JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

2/2002

Quality of service in packet networks: experiments and theoretical studies

Special issue edited by Wojciech Burakowski

AQUILA network architecture: first trial experiments
A. Bąk et al. — *Paper* — 3

Relations between ATM traffic parameters and quality of multimedia transmission - an experimental approach
A. Sochan — *Paper* — 14

Multistage optical switching networks
S. Kaczmarek — *Paper* — 21

Greedy randomised adaptive search procedures for topological design of MPLS networks
A. Mysłek — *Paper* — 26

On effectiveness of conditional admission control for IP QoS network services with REM scheme
M. Dąbrowski, W. Burakowski, and A. Bęben — *Paper* — 33

Regular papers

Efficient procedure for capacitance matrix calculation of multilayer VLSI interconnects using quasi-static analysis and Fourier series approach
H. Ymeri, B. Nauwelaers, and K. Maex — *Regular paper* — 40

Editorial Board

Editorial Advisory Board

ISSN 1509-4553

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

Preface

This special issue is dedicated to the quality of service (QoS) multi-service networks, based on packet switched technologies like Internet protocol (IP) or asynchronous transfer mode (ATM). Introduction of QoS mechanisms and appropriate traffic control rules in such networks is of special interest since they allow us for effective transfer of information corresponding to a variety of applications, like voice telephony, computer data and multimedia. These applications differ in (1) type of traffic profile emitted to the network (from constant bit rate to variable bit rate), (2) volume of requested bandwidth (from kbit/s up to Mbit/s), and (3) QoS requirements (some applications are more sensitive on delay while others are more sensitive on loss rate). For this purpose, a set of network services is defined and implemented in multi-service networks, each of them designed for transferring traffic generated by adequate types of applications. A network service should have its own QoS mechanisms protecting target QoS objectives, commonly expressed in the form of guaranteed packet loss rate, maximum packet transfer delay, throughput etc. Current research activities are mainly focused on enhancement of IP-based networks toward multi-service networks.

The papers, designated for this special issue were selected from the papers presented at the annual Polish Teletraffic Seminar, which was held in Zakopane on September 1–3, 2001. Traditionally, the Seminar gathers the researchers form Universities of Technologies and Polish Academy of Science working on traffic modeling, methods for network dimensioning, network mechanisms supporting QoS, mechanisms for communication protocols, queuing models for telecommunications and computer networks etc.

Wojciech Burakowski
Guest Editor

On May 22nd, 2002 we lost one of our dearest Colleagues,

Professor Władysław Majewski, Ph.D.

an outstanding Scientist and Telecommunications Expert, a long-term Scientific Worker at the National Institute of Telecommunications, Director and Chairman of the Scientific Council of the Institute,
an unforgettable University Teacher and Scientific Worker at the Warsaw University of Technology, Director of the Institute of Telecommunications at the Faculty of Electronics and Information Technology and Minister of Communications,
Member of Editorial Advisory Board of our magazine, Author of numerous scientific books and articles, and an incomparable Reviewer and Advisor in many editorial efforts.

We part with Him in sorrow. He will remain in our memory forever.

Editorial Board

Paper

AQUILA network architecture: first trial experiments

Andrzej Bąk, Andrzej Bęben, Wojciech Burakowski, Marek Dąbrowski, Monika Fudała, Halina Tarasiuk, and Zbigniew Kopertowski

Abstract — The paper presents trial experiments with IP QoS network services (NS) defined and implemented in the AQUILA pilot installation. The AQUILA NSs (premium CBR, premium VBR, premium multimedia and premium mission critical) provide a framework for supporting a variety of applications generating both streaming and elastic traffic. The measurement experiments confirm that AQUILA architecture differentiates the QoS offered to these NSs. The presented numerical results were obtained in the test network installed in the Polish Telecom (Warsaw) consisting of 8 CISCO routers.

Keywords — IP QoS, measurements.

1. Introduction

The IP QoS networks have gained recently a significant attention. The future Internet architecture will have to offer QoS as well as keep its current advantages of scalability and simplicity. The differentiated service (DiffServ) architecture [1, 2] is considered to be a promising solution for IP QoS networks. The DiffServ architecture is not a complete solution rather it is a framework that defines basic mechanisms and rules for providing QoS. The AQUILA project aims at developing complete QoS architecture for IP networks that enhances the original DiffServ concept. The main innovations of AQUILA include the definition of network services, specification of admission control rules and specification of signalling mechanisms for dynamic resource reservation. The AQUILA project adds new element to the DiffServ enabled network by implementation of resource control layer (RCL) that can be treated as a distributed bandwidth broker. The RCL implements all logic necessary for providing end-to-end dynamic QoS guarantees to the end-users.

The AQUILA project is on going activity. Currently the first phase of RCL development was finished, that resulted in the implementation of the AQUILA prototype in 3 different test-beds. The main AQUILA test-bed is located in Polish Telecom in Warsaw. The paper presents comprehensive description of basic AQUILA functions and additionally presents some results of the first trial experiments curried out in the Warsaw test-bed.

The paper is organized as follows. In Chapter 2 the basic concepts of AQUILA network are briefly described. The Chapter 3 introduces reader to the AQUILA test-bed used in the trial experiments. The trial results are presented in Chapter 4. Finally, the Chapter 5 concludes and summarizes the paper.

2. AQUILA network and service architecture

The AQUILA network architecture distinguishes two types of network elements (similarly as in DiffServ): the edge devices (ED) and core routers (CR). The EDs connect end users to the network and implement all per flow oriented traffic control functions e.g. traffic policing, marking etc. The CRs perform only packet forwarding according to the assigned per hope behaviour (PHB). The CRs do not distinguish individual flows rather they operate on behaviour aggregates (the collections of packet that belongs to the same traffic class).

In AQUILA architecture (Fig. 1), additional layer, named resource control layer (RCL), was added on the top of DiffServ enabled network.

This layer is responsible for controlling and managing network resources (the RCL represents the control plane while the DiffServ network represents data pane of the AQUILA network). The RCL is subdivided into two sublayers implemented by software-based components i.e. resource control agent (RCA) and admission control agents (ACA). The ACA agent is responsible for performing admission control, while the RCA manages the resource provisioning and distribution (or redistribution) between ACA agents.

For providing QoS differentiation to users traffic a set of four NSs was defined and implemented. The particular NSs differ with respect to QoS objectives, traffic characterization and supported traffic type (streaming or elastic). Each NS is designed for effective transmission of traffic produced by applications with similar QoS requirements (e.g. voice, video or data applications). The NSs are mapped into a set of traffic classes (TCL) that define appropriate traffic handling rules for meeting NS objectives, i.e. traffic conditioning mechanisms, admission control algorithms and PHB rules.

2.1. AQUILA network services

The following QoS NSs were defined in AQUILA: premium CBR (PCBR), premium VBR (PVBR), premium multimedia (PMM) and premium mission critical (PMC) in addition to standard best effort Internet service. The first two services are for streaming traffic while the other are for elastic traffic.

The PCBR is intended for applications that require constant bit rate (CBR traffic) with very low packet transfer delay and packet losses e.g. voice applications. The user

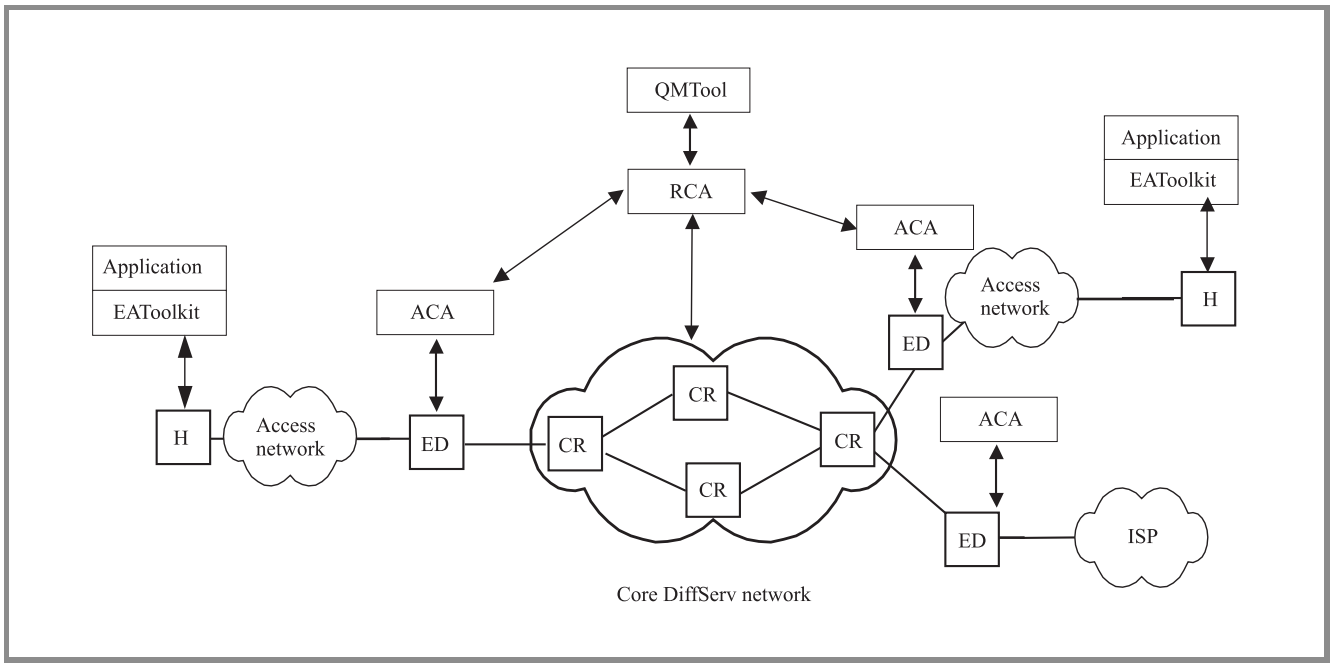


Fig. 1. General AQUILA network architecture.

traffic in this case is characterized by peak bit rate, only. The next NS (PVBR) is designed for applications requiring variable bit rate (VBR traffic). This service provides similar QoS guarantees as PCBR. Typical application of this NS is video transmission. The user traffic is characterized by sustainable and peak bit rates.

The remaining NSs are intended for elastic traffic i.e. TCP-controlled. The PMM service uses single rate characterization (sustainable bit rate). This service allows for much higher burst tolerances comparing to PCBR. Therefore, the PMM service can be seen as a special case of VBR service with the peak bit rate equal to the link bit rate. This service is mainly intended for long life greedy TCP connections. On the contrary, the PMC service is designed for short life non-greedy TCP connections. In this service, the user traffic is characterized, similarly as in case of PVBR service, by the peak and sustainable bit rates.

2.2. AQUILA traffic handling mechanisms

The AQUILA NSs are mapped to appropriate traffic classes TCL (currently one NS is mapped to one traffic class). Depending on the TCL, the traffic conditioning (policing) is performed by single or dual token bucket algorithm. The TCL1 class (employed by PCBR service) uses single token bucket algorithm (peak bit rate) while the TCL2 (used by PVBR service) uses dual token algorithm (peak and sustained bit rates). In both cases the excess traffic is dropped (only conforming packets are allowed to enter the network). In the case of TCL3 traffic class (used by PMM) the user traffic is characterized by single token bucket algorithm while in TCL4 (used by PMC) by dual token bucket. In both cases the excess traffic is marked as "out of the profile" and entered to the network.

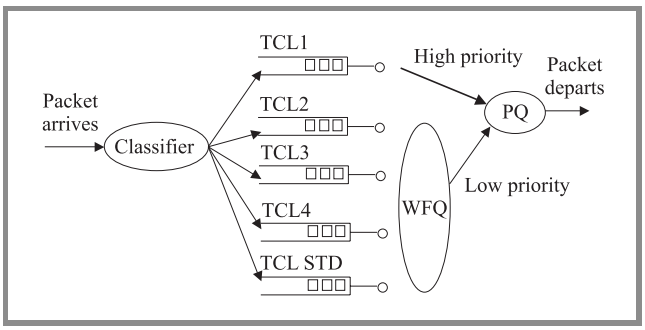


Fig. 2. Routers output port architecture.

After passing conditioning block, the packets are buffered in separate queues according to the traffic class they belong (the conditioning is done only in the ED). The TCL1 and TCL2 queues use simple drop tail queue management mechanism while TCL3 and TCL2, since they are designed for elastic traffic, employ the WRED (weighted random early discard) algorithm to tailor the access of "in profile" and "out of profile" packets to the output link. The TCL1 queue has non-preemptive priority over the other queues in the access to the output link (priority queuing scheduler – PQ). The remaining four queues (including the queue for best effort traffic) are served by WFQ (weighted fair queuing) scheduler. The AQUILA scheduling algorithm is an equivalent of CBWFQ (class based WFQ) scheduler available on CISCO routers (Fig. 2).

2.3. Resource management

The resource distribution is done on the per traffic class basis, i.e. resources assigned to a particular class cannot be used by other classes (there is no dynamic resource

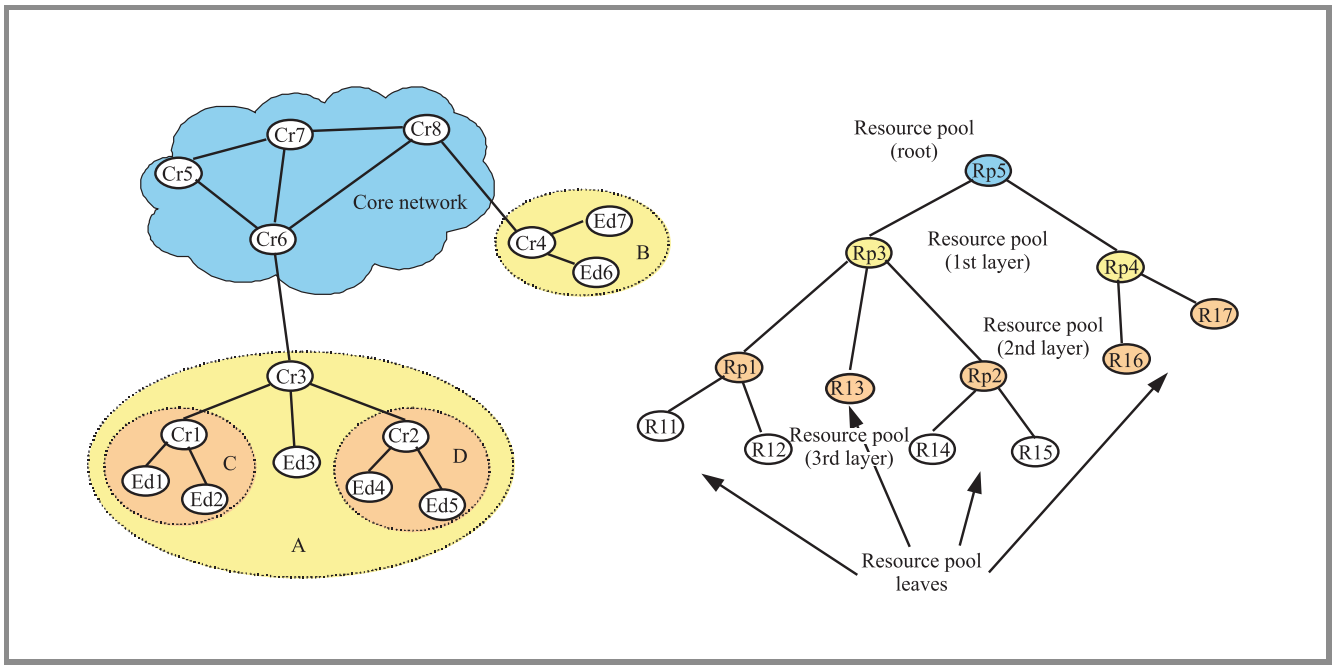


Fig. 3. Resource pools.

sharing between AQUILA traffic classes). As a result the network is divided into 5 layers each representing virtual networks related to particular traffic class. Each ACA agent represents one ED. The ACA agent accepts request for resources up to a given limit. To calculate the admission limits the RCL takes into account the prediction of traffic between EDs, network routing, etc. The resources assigned to a given ED represent total traffic that can be accepted by its ACA agent irrespectively of the traffic destination (the resources are assigned on per ACA basis).

Usually precise prediction of traffic demands in the Internet is almost impossible. Therefore, additional adaptive resource management mechanism was developed in AQUILA, which aims to improve overall network resource utilization. This mechanism allows shifting resources between ACA agents if the resource utilization at certain point in the network is lower that assumed threshold. To facilitate this, resources assigned to ACA agents are grouped into, so called, resource pools. The resource pool represents a part of network resources that can be shared by its ACA agents. The resource pool can be further aggregate creating higher level of resource pools and so on. Any level of resource pool hierarchy can be defined depending on network topology. The topology structure where resource pools can be easily defined is the tree topology. In the simplest scenario, one resource pool can be created for each tree node (the network nodes on the lowest level represents ACA agents, and are called resource pool leaves). This approach can be generalized for tree like structures of network sub-areas.

The exemplary resource pool hierarchy is shown in Fig. 3. The root of the resource pool hierarchy represents overall network resources. In the exemplary network from Fig. 3 we have two network sub-areas (A and B) for which the resource pool Rp3 and Rp4 are created. The resource pools Rp3 and Rp4 form the first level of the resource pool hierarchy. The Rp3 and Rp4 are further shared by the resource pools of the second hierarchy level. This process is continued until the lowest level of resource pool hierarchy is reached. The lowest resource pool in each tree branch represents the resource pool leave, which is associated with ACA.

2.4. Admission control

Typical IP network offers only the best effort service. Because this service does not provide any QoS guarantees no admission control is necessary. Consequently, there is no control over the number of flows that are handled by the network. In AQUILA project new QoS services were defined with differentiated QoS requirements. Therefore, the admission control methods, capable to provide required QoS guarantees, had to be defined. In AQUILA the QoS guarantees are provided to the flow, which is defined as an unidirectional flow of IP packets that can be distinguished at the network entry based on the contents of packets headers (source, destination IP addresses, protocol type etc.). This definition is essentially compliant with the concept of the flow used in DiffServ architecture. Generally, two types of admission control algorithms can be distinguished i.e. the methods based on the declaration submitted by the user and methods that additionally take into account some traffic measurement (apart the declarations). The first approach was taken by AQUILA project. A request for network resources is accepted or rejected based on the traffic descriptors provided by the user (special signalling procedure was developed for this purpose). The objective of the AQUILA project was to ver-

ify if similar approach (declaration based admission control) as used in ATM network could also be applied to IP networks.
In the AQUILA architecture the admission decision is perform only at the network ingress and egress. This makes the admission control algorithm more difficult (or uncertain) as link-by-link verification of resource availability is not possible. To perform the admission control at the ingress the single link model was assumed. It means that, for each traffic class, the admission control algorithm operates as if there was only one link with capacity (C) and buffer size (B) in the network. The parameter (C) corresponds to the amount of network resources assigned to the particular ACA agent where as the parameter (B) is related to the buffering capacity provided by the network (currently it is set to the smallest buffer size among all routers devoted to the particular traffic class). Whenever below parameters (C) or (B) are mentioned they correspond to the capacity or buffer space dedicated to serve the given traffic class.

2.4.1. Admission control algorithm for streaming traffic characterized by single token bucket

The TCL1 class uses peak bit rate allocation scheme [4]. Within this class the flows are characterized by parameters of single token bucket algorithm (this parameter correspond to the peak bit rate (PBR) and peak bit rate tolerance (PBRT). In the AQUILA architecture, the TCL1 traffic is served with the highest priority. It means that the TCL1 traffic is not mixed with other traffic in the network. Small variation in the TCL1 traffic rate can be still present due to the *PBRT* parameters and influence of residual packet transmission time of low priority traffic. Taking this into account, it can be assumed that the TCL1 streams have negligible packet delay variation [11]. Consequently, the worst-case traffic pattern for the superposition of a number of TCL1 flows takes the form of poissonian stream (with the mean rate equal to the sum of the PBR parameters of the particular flows).
New flow characterized by PBR and $PBRT$ parameters, is admitted if the following condition is satisfied:

$$PBR + \sum_i PBR_i \leq \rho C\,. \tag{1}$$

Parameter ρ ($\rho < 1$) specifies the admissible load of capacity allocated to the TCL1 class. The value of ρ can be calculated from the analysis of M/D/1/B system taking into account the target packet loss ratio and the buffer size [5].

2.4.2. Admission control algorithm for streaming traffic characterized by dual token bucket

In the case of TCL2 traffic class rate envelope multiplexing (REM) scheme is assumed for guaranteeing low packet delay [4]. Therefore, the only QoS parameter that requires concern is the packet loss rate. In the REM multiplexing the buffer has to be dimensioned for absorbing, so called, packet scale congestion (simultaneous arrival of packets from different sources). For this purpose the N*D/D/1 queuing system analysis can be used.
In the TCL2 class, each flow is characterized by parameters of dual token bucket algorithm: the PBR together with the $PBRT$ and the sustainable bit rate (SBR) together with the burst size (BSS). It is commonly assumed that the worst-case traffic pattern for given values of PBR, SBR and BSS is of the ON/OFF type.
The proposed admission method for TCL2 is based on the notion of effective bandwidth. One may find a number of methods for calculating effective bandwidth [4]. For its simplicity the methods proposed in [12] was chosen for AQUILA. In this method the value of effective bandwidth, $Eff(.)$, is calculated on the bases of PBR, SBR and BSS parameters taking into account the target packet loss rate. The new flow with $Eff(.)$ is admitted if the following condition is satisfied:

$$Eff(.) + \sum_i Eff(i) \leq C\,. \tag{2}$$

The effective bandwidth usually depends on the link capacity. In the case of AQUILA admission concept (admission performed only at the ingress) this requires to use some reference link capacity (not the admission limit C). In AQUILA the rate of the first link connecting the ED to the CR is used.

2.4.3. Admission control algorithm for elastic traffic characterized by single token bucket

In the case of TCL3, each flow is characterized by parameters of single token bucket algorithm that correspond to the SBR and the BSS. Usually, the BSS values are rather large what allows high variability of submitted traffic. Comparing to TCL2 or TCL4, the declaration of the PBR parameter is omitted. In fact, the PBR is only limited by the link bit rates.
Taking into account that the traffic flows submitted to TCL3 class are TCP-controlled, only rough QoS guarantees are assumed to be provided. Therefore, an admission control method that maximizes the network utilization can be employed in this case. The new flow characterized by SBR, BSS and MTU (maximum transfer unit) is admitted only if the following conditions are satisfied:

$$SBR + \sum_i SBR_i \leq C\,, \tag{3}$$

$$MTU + \sum_i MTU_i < B\,. \tag{4}$$

Remark, that depending on the value of the buffer size (B) the multiplexing in this class can follow the REM

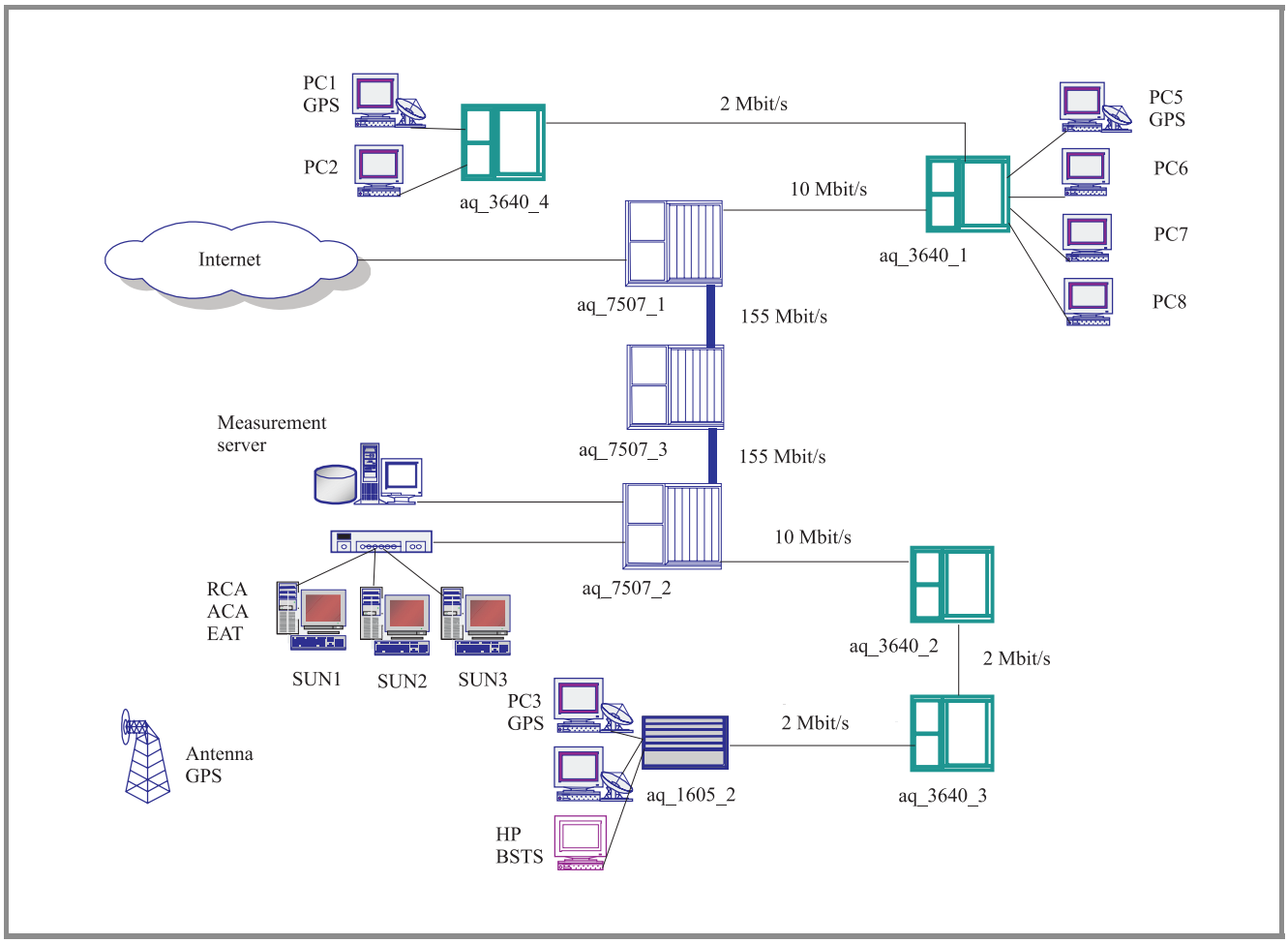


Fig. 4. AQUILA test-bed configuration: PC1-8 – terminals, SUN1-3 – sun workstations with implemented RCA, ACA and EAT modules, aq_1605, aq_3640, aq_7507 – CISCO routers.

or the rate sharing multiplexing (RSM) scheme [4]. The RSM multiplexing corresponds to the case of large buffers capable to absorb temporary traffic fluctuations (to absorb packet burst scale congestion). In this case, the second condition (related to buffer) can be omitted.

2.4.4. Admission control algorithm for elastic traffic characterized by dual token bucket

In the TCL4, a flow is characterized by parameters of dual token bucket algorithm, similarly as in the case of TCL2 class. Because this class does not guarantee any delay parameters the RSM multiplexing scheme was assumed. The proposed admission control algorithm provides the service rate that will guarantee no losses for packets that are within traffic contract. New flow characterized by PBR, SBR and BSS is accepted if the following condition is satisfied:

$$Eff(.) + \sum_i Eff(i) \leq C\,. \tag{5}$$

The effective bandwidth in this case is calculated as follows [13]:

$$Eff(.) = \max\left\{SBR,\ \frac{PBR \cdot T}{B/C + T}\right\}, \tag{6}$$

where, $T = \frac{BSS}{PBR-SBR}$.

3. Description of the AQUILA test-bed

The experiments were carried out in the AQUILA test-bed (Fig. 4), installed in the Polish Telecom in Warsaw.
The test-bed network consisted of 8 CISCO routers (of different types) connected in the form of the chain to emulate large number of hops. The end terminals were connected to the edge routers by Ethernet ports. The access links between the ED and the first CR had capacity of 2 Mbit/s. The core routers were connected by higher capacity links (10 and 155 Mbit/s). The following types of routers were installed in the test-bed: two edge routers – 1605 (aq1605_2) and 3640 (aq3640_4), 6 core routers – 3640 (aq3640_1, aq3640_2, aq3640_3,) and 7507 (aq7507_1, aq7507_2, aq7507_3). The core and edge routers were configured according to the

AQUILA architecture design. The AQUILA software was installed on 3 workstations. Each workstation was used for different type of RCL components (ACA, RCA, EAT).

4. Measurement results

4.1. Separation of streaming and elastic flows

The first set of experiments shows the rationale for defining separate NSs for traffic generated by streaming and elastic applications. The test traffic was generated between the end-stations PC4 and PC2 (Fig. 4) as the 200 kbit/s constant bit rate stream controlled by UDP transport protocol. Additionally, a number of greedy TCP connections was established between the end-stations PC3 and PC1. Notice that the 2 Mbit/s link connecting ED router to the core is the bottleneck link in the considered topology.
Two test cases were taken into account. First, the FIFO scheduling discipline was set-up in the ED. As a consequence, traffic produced by UDP and TCP connections was handled inside ED in the same way. In the second case, the PQ scheduler was selected in ED and UDP traffic was served in isolation from TCP traffic. The highest priority was assigned for UDP traffic because of its QoS requirements. Note that it is not reasonable to assign the highest priority for greedy TCP traffic since it can lead to bandwidth starvation for the lower priority traffic. In both tested scenarios, the performances of transferring the UDP traffic were checked by measuring the one-way packet delay and packet loss ratio QoS parameters.

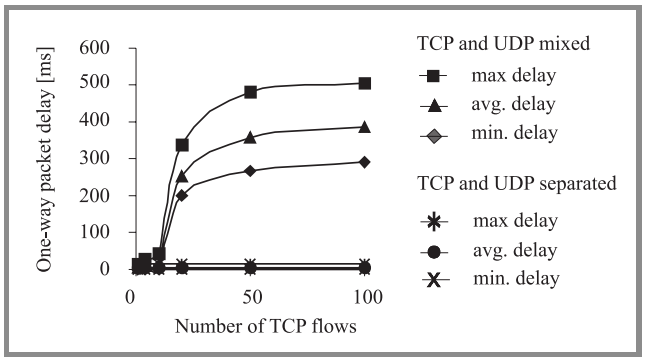


Fig. 5. One-way delay of UDP packets mixed with TCP traffic and separated from TCP traffic.

The obtained results are shown in Figs. 5 and 6. One can observe, that when TCP and UDP traffic is mixed together, the one-way delay and packet loss ratio of the UDP packets strongly depend on the number of active TCP connections competing for the bottleneck link capacity. This is the case now observed in IP networks, where the number of TCP connections is not limited by any admission control mechanism. When 100 TCP connections were set up in the network, the aggressive TCP congestion control scheme caused that the buffer in ED associated with the bottleneck link was heavy overloaded. Therefore, the delays experienced by tested UDP packets was about 500 ms.

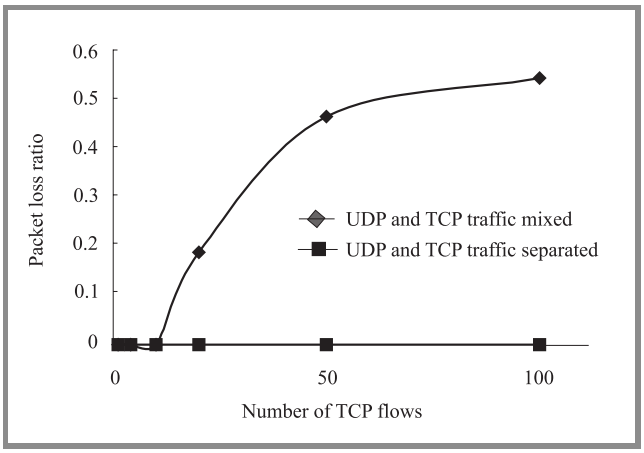


Fig. 6. Packet loss ratio of UDP packets mixed with TCP traffic and separated from TCP traffic.

The observed values of packet loss ratio were also considerably high (up to 50%).
This non desirable effect can be eliminated when UDP packets are served in the network in isolation, with highest assigned priority. In this case, no packet losses were observed and the maximum packet delay was close to 12 ms, regardless the number of active TCP connections. This result confirms that streaming and elastic traffic should be separated in the network, otherwise the QoS guarantees for streaming traffic are hard to satisfy. In AQUILA, this requirement is met by appropriate definitions of PCBR and PVBR services for streaming traffic and PMM and PMC services for elastic traffic.

4.2. Premium CBR service

The carried out experiments corresponding to PCBR service were aimed for measuring effectiveness of CBR flows transmission. For the test scenario was assumed that the QoS parameters of carrying traffic were measured under the maximum submitted traffic allowed by the admission control of the TCL1 class. The test traffic was modelled by a Poisson stream with constant packet lengths. Notice that such traffic represents the worst case of the superposition of large number of CBR streams. To take into account the impact of other classes on the TCL1 packets (residual transmission time of low priority packets) the background traffic was also added to the STD class. The background traffic filled the rest of link capacity not dedicated to TCL1.
In the following experiments, 200 kbit/s of the access link capacity was reserved for TCL1. Furthermore, the TCL1 queues in the routers were set to 5 packets to guarantee low packet delay. The performance of TCL1 was validated assuming target packet loss ratio (Ploss) equal 10^{-2}. According to the specified admission control algorithm, the maximum admissible load in this case (acc. to the M/D/1 system analysis [4]) is $r = 0.685$, what is equivalent to 137 kbit/s.

The foreground traffic was transmitted between PC1 and PC3 terminals (Fig. 4) while the background traffic was generated only to the access link. The background traffic was created as a mix of packets with different lengths: 7% with 44 bytes, 21% with 256 bytes and 72% with 1280 bytes. Both foreground and background traffic was transmitted using UDP protocol.
The characteristic of packet loss rate as a function of the TCL1 traffic volume is presented in Fig. 7.

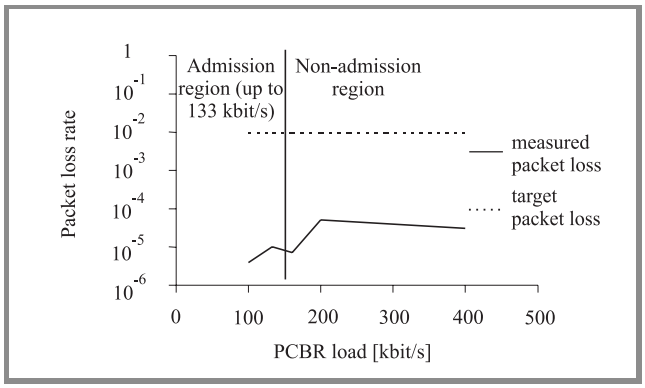


Fig. 7. Packet loss rate versus TCL1 traffic load.

One can observe that the measured value of packet loss rate are in the range of 10^{-5} and they are significantly below the assumed target value 10^{-2} even for the load above the admission region (133 kbit/s). This is rather expected result since the TCL1 traffic is served with higher priority with the effective service rate of 2 Mbit/s (instead of 200 kbit/s as was assumed for the admission control algorithm). Anyway, increasing TCL1 traffic significantly above the assumed limit (133 kbit/s) can degrade the quality experienced by packets carried in low priority classes (e.g. TCL2). The recommended admissible load of TCL1 traffic is approximately 10% of total link capacity.

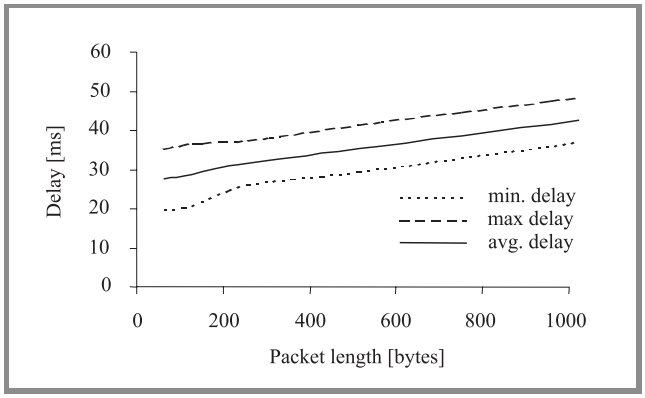


Fig. 8. One-way delay versus TCL1 packet length.

The characteristics of one-way packet delay as a function of TCL1 packet length are depicted in Fig. 8. These curves were measured assuming that Poisson traffic with the rate equal to 133 kbit/s (up to admission limit) was submitted to TCL1. The background traffic of ON/OFF type was submitted independently to each intermediate link (Fig. 4), with the peak bit rate equal to the appropriate link rate. The maximum observed delay was below 50 ms. This value is acceptable for most voice applications, that tolerate delay in the order of 150 ms (with codec and packetization delay).

4.3. Premium VBR service

The PVBR service was designed for real time applications producing variable bit rate traffic. The QoS objectives for this service are low packet transfer delay (≤ 150 ms) and low packet loss ratio ($\leq 10^{-4}$). In the followed experiments, the target packet loss ratio was fixed to 10^{-2}. The presented measurement experiments correspond to validation of packet loss ratio as well as end-to-end delay measured for artificial flows. Traffic served by TCL2 was generated between PC3 and PC1 while traffic served by TCL1, TCL3, TCL4, and TCL5 was generated between HP BSTS and aq_3640_3 router. The bandwidth dedicated for TCL2 class (on 2 Mbit/s access link) was 300 kbit/s. The following bandwidths were allocated to the other traffic class: TCL1: 200 kbit/s, TCL3: 600 kbit/s, TCL4: 100 kbit/s, and TCL5: 800 kbit/s. Output buffer in ED dedicated for TCL2 was assumed to be 5 packets long. The aggregated traffic of TCL2 class was assumed to be superposition of ON/OFF sources and was modelled by Markov modulated deterministic process (foreground traffic). Flows of constant bit rate sufficient to load the network resources dedicated to the other traffic classes were used as background traffic flows.
In the first experiment the packet loss ratio and the end-to-end packet delay were obtained as a function of the number of admitted flows (Fig. 9). In this case each ON/OFF flow was described by the following parameters: $PBR = 32$ kbit/s, $SBR = 16$ kbit/s, $BSS = 15000$ bytes, packet size $= 500$ bytes. Effective bandwidth for each flow was equal to 27.29 kbit/s. Number of admitted flows, according to the AC algorithm, was 10. Values of traffic descriptors used in the following experiments correspond to the traffic generated by a voice applications.

Table 1
The measured values of QoS parameters

Min. delay [ms]	Max delay [ms]	Avg. delay [ms]	Packet loss ratio
15	44	26	10^{-5}

The second experiment was done for the superposition of 30 ON/OFF flows served by TCL2 (to achieve greater multiplexing gain). Each ON/OFF flow was characterized by: $PBR = 18.5$ kbit/s, $SBR = 5.6$ kbit/s, $BSS = 526$ bytes, and packet size $= 256$ bytes. Effective bandwidth of each flow was 9.97 kbit/s. The sum of the peak bit rates of 30 flows was 555 kbit/s while the sum of the effective bandwidths over all flows was 299.1 kbit/s. The obtained results are presented in the table bellow. The measured values of QoS parameters satisfy requirements and

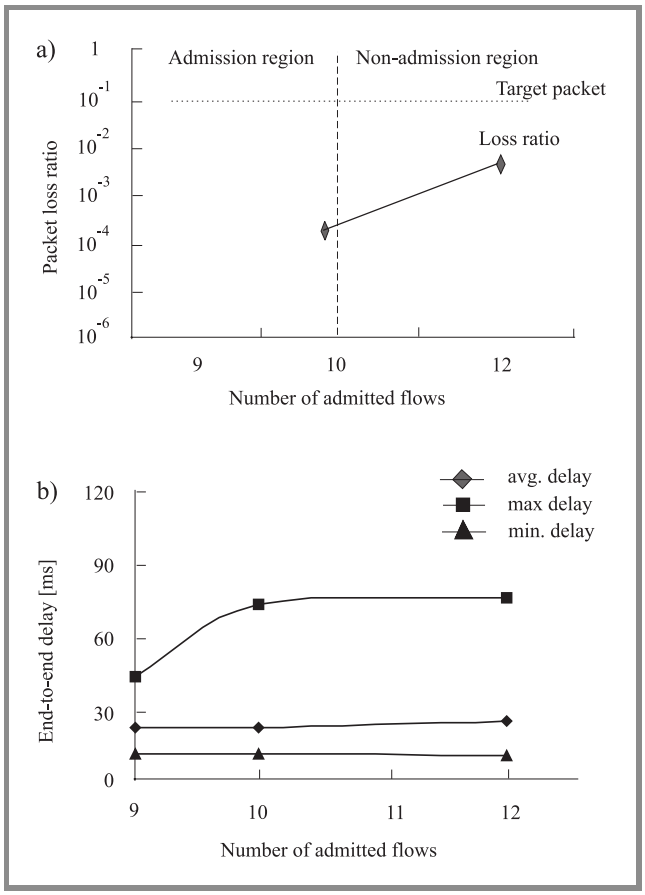


Fig. 9. Packet loss ratio (a) and end-to-end (b) delay versus number of admitted flows.

objective of TCL2 (Table 1). The values of packet loss ratio ($10^{-5} - 10^{-3}$) are smaller than the assumed target value (10^{-2}).
Some trials were also carried out with real video application (e.g. NetMeeting). Unfortunately, it appeared that it is very difficult to tune traffic descriptors for real application, like NetMeeting. This suggest that traffic declarations should be limited to the peak rate while admission control should get additional information about carried traffic from measurements.

4.4. Premium multimedia

The PMM service was mainly designed for effective support of greedy TCP-controlled flows. The objective of this service is to guarantee a minimum throughput to the TCP flows. The measurement results presented below were obtained assuming that the traffic generated by a number of TCP greedy sources (between PC1 and PC3) was submitted to the TCL3 class while the CBR background traffic was submitted to other NSs with the maximum possible rates (according to the assigned capacity for given service). In this way the capacity available for the TCP connections was limited to the capacity allocated to TCL3 class. In the considered case this capacity was equal to 600 kbit/s with target utilization factor set to 0.9 (this parameter represents target utilization for "in-profile" packets) what gives 540 kbit/s of available capacity for setting up flows. Notice that the "out-of-profile" packets in case of TCL3 (and also TCL4) are not dropped. Consequently, the considered greedy TCP flows can achieve higher rate than requested by the SBR parameters.

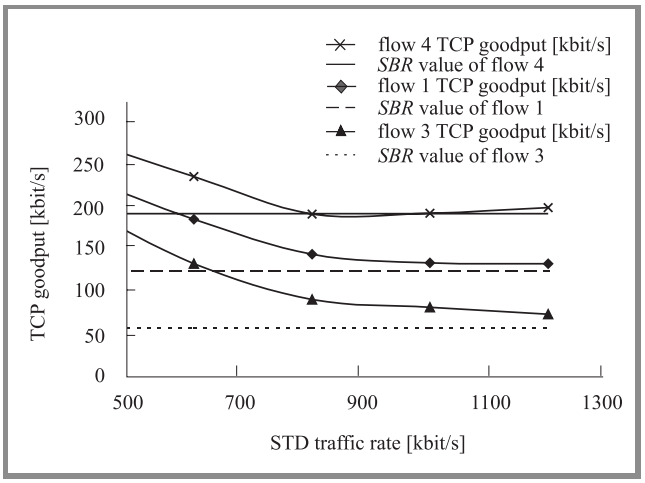


Fig. 10. TCP goodput in PMM service versus STD traffic rate.

The goodput characteristics of 4 TCP flows, with different declared SBR values, as a function of best effort traffic volume are depicted in Fig. 10. The assigned SBR values were as follows: 135, 135, 70 and 200 kbit/s. One can observe that the available capacity for PMM service is effectively used by the considered TCP connections. When the background traffic exceeds a certain threshold (in this case 800 kbit/s) the available capacity for TCL3 is limited to assumed AC limit (600 kbit/s). This capacity is shared by TCP flows roughly in proportion to their declared SBR rates. For example, the connections with $SBR = 200$ gets 206 kbit/s while the connection with $SBR = 70$ gets only 85 kbit/s.
Figure 11 shows the proportion of packets, marked as "in-profile" and "out-of-profile", contributing to the throughput of 2 exemplary TCP connections. The traffic submitted into the network in excess of the traffic profile requested in the reservation is marked as "out-of-profile". In the underload network conditions, the "out-of-profile" traffic obtains the capacity unused by the other NSs, but in time of network congestion, throughput of the TCP flow in PMM service is guaranteed roughly up to the requested SBR value.
By applying the WRED mechanism, one can expect that the "in-profile" packet will obtain better service, than "out of profile" packets. In AQUILA, WRED mechanism controls the access of "in-profile" and "out-of-profile" packets to the TCL3 buffer. When the average queue size exceeds a predefined threshold, incoming packets can be dropped with a certain probability. By setting different threshold values for "in" and "out" packets, one can differentiate the loss ratio of packets marked as "in" and "out of profile". Figure 12 shows the received packet loss characteristic for "in" and "out" packets as a function of the number of TCP

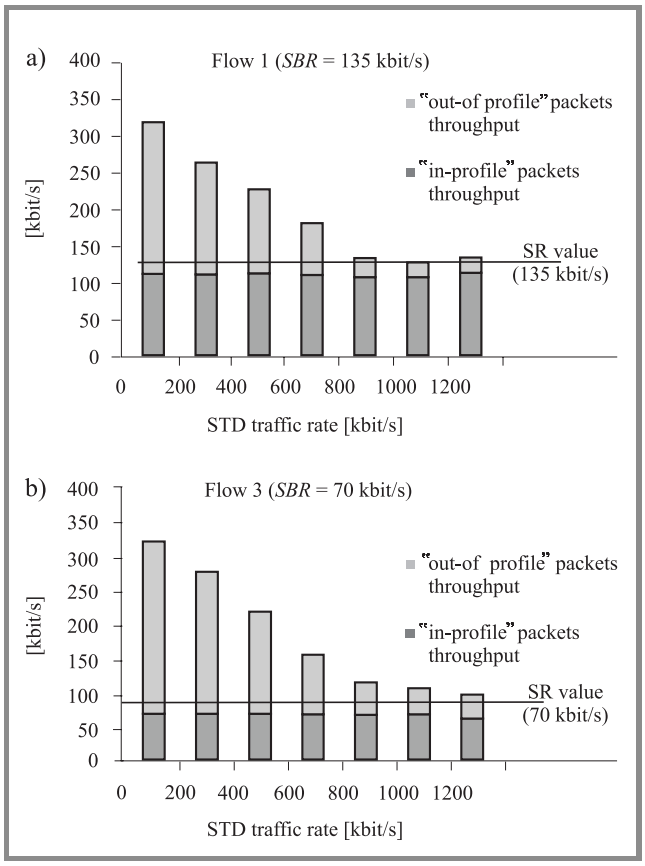


Fig. 11. Throughput of "in-profile" and "out-of-profile" packet streams within flow 1 (a) and flow 3 (b).

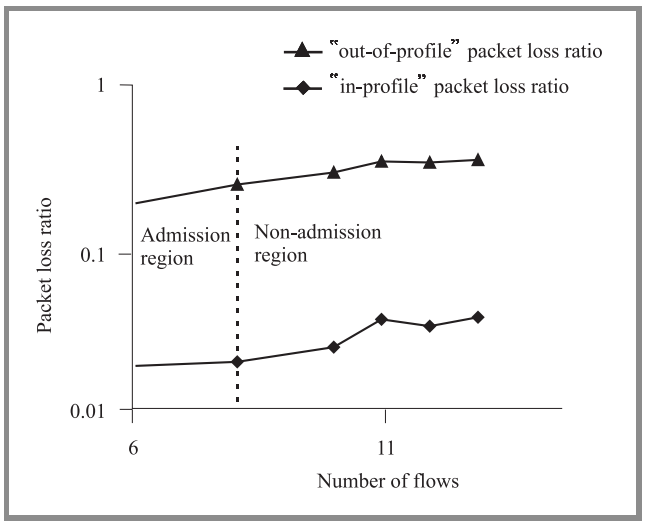


Fig. 12. "In-profile" and "out-of-profile" packet loss ratio versus number of TCP flows in PMM service.

connections. As it was expected, the packets conforming to the assumed traffic profile are delivered by the network with a significantly higher probability than the non-conforming packets.

Note that the proper configuration of WRED mechanism is crucial for flow differentiation inside the TCL3 class with respect to the guaranteed throughput. Since there is high probability that the "out of profile" packets will be lost during network congestion, the TCP source sending "out-of-profile" traffic will likely be forced to decrease its sending rate. On the other hand, when the source transmits only conforming traffic, its packet loss rate will be much lower than in the previous case. This TCP source will not be forced to decrease its sending rate. Consequently, the average throughput of each PMM flow will be close to the SBR declaration.

4.5. Premium missioncritical service

The PMC service was designed for short-lived non-greedy TCP flows, typically generated by transaction-oriented applications, e.g. database queries. The objective of this service is to provide high guarantee of packet delivery (low packet loss rate and relatively low packet transfer delay). In the following experiment, the test traffic was produced non-greedy TCP flows, transmitted between end-stations PC3 and PC1 (additionally for comparison this experiment was repeated with greedy TCP flows). Similarly as in the PMM service tests, the background traffic of CBR type was submitted to other NSs with the maximum possible rate. The capacity available for the tested TCL4 class was equal to 100 kbit/s.

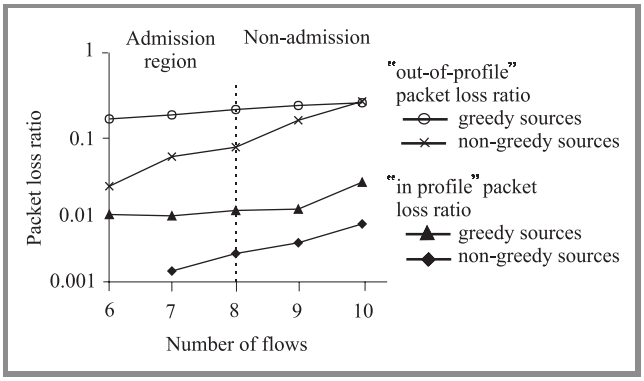


Fig. 13. "In-profile" and "out-of-profile" packet loss ratio versus number of greedy and non-greedy TCP flows in PMC service.

Figure 13 shows characteristic of packet loss ratio for "in-profile" and "out-of-profile" traffic as a function of number of running TCP connections (greedy or non-greedy). As one could expect, in the case of greedy TCP sources the packet loss ratio is significantly higher for both "in" and "out" packets. On the other hand, the observed packet loss ratio for non-greedy TCP flows is greater than the target value (fixed to 10^{-6}). Therefore, some modifications of queue management algorithm are required.

5. Conclusions

In the paper the approach for traffic engineering in the AQUILA QoS IP network was described. The measure-

ment experiments confirm that IP network with appropriate traffic handling mechanisms is capable to offer QoS guarantees with dynamic reservation scheme. The experiments curried out in AQUILA trial demonstrated the need for separation of streaming and for elastic traffic. Currently, AQUILA project defines 4 QoS NSs with different QoS objectives. Such approach allows efficient support of wide range of applications with different requirements. The obtained measurement results confirm that such approach can be implemented in practice, and that the AQUILA NSs generally provide the assumed QoS guarantees.

To guarantee QoS objectives appropriate AC algorithms should be employed. AQUILA project shows that the declaration based AC algorithms could guarantee the QoS requirements on the per flow basis in the IP network based on the DiffServ architecture. On the other hand, AQUILA AC algorithms assume the worst-case traffic patterns that are rarely met in practice. This sometimes can lead to low network resource utilization. Future work is focus on adding some measurements in the resource control algorithm for improving network utilization.

References

[1] S. Blake *et al.*, "An architecture for differentiated services". Internet RFC 2475", Dec. 1998.

[2] Y. Bernet *et al.*, "Differentiated services". Internet Draft, draft-ietf-diffserv-framework-0.2.txt, Feb. 1999.

[3] Y. Bernet *et al.*, "A conceptual model for DiffServ routers". Internet Draft, draft-ietf-diffserv-model-02.txt, Sept. 2000.

[4] "Final report COST 242, Broadband network teletraffic: Performance evaluation and design of broadband multiservice networks", J. Roberts, U. Mocci, and J. Virtamo, Eds. Lectures Notes in Computer Science 1155, Springer 1996.

[5] "Final report COST 257, Impact of new services on the architecture and performance of broadband networks". COST 257 Management Committee, Phuoc Tran-Gia, Norbert Vicari, Eds. Wuerzburg, Germany, 2000.

[6] Deliverble D1201, "System architecture and specification for the first trial, AQUILA project consorcium", June 2000, http://www-st.inf.tu-dresden.de/AQUILA

[7] Deliverable D1301, "Specification of traffic handling for the first trial, AQUILA project consortium", Sept. 2000, http://www-st.inf.tu-dresden.de/AQUILA

[8] Deliverable D2301, "Report on the development of measurement utilities for the first trial, AQUILA project consortium", Sept. 2000, http://www-st.inf.tu-dresden.de/AQUILA

[9] Deliverable D3101, "First trial integration report, AQUILA project consortium", March 2001, http://www-st.inf.tu-dresden.de/AQUILA

[10] Deliverable D3201, "First trial report, AQUILA project consortium", Draft version, Apr. 2001.

[11] F. Brichet *et al.*, "Stochastic ordering and the notion of negligible CDV", in *Proc. 15th Int. Teletraf. Congr.*, Washington, USA, 1996.

[12] K. Lindberger, "Dimensioning and design methods for integrated ATM networks", in *Proc. 14th Int. Teletraf. Congr.*, Antibes, 1994.

[13] A. Elwalid *et al.*, "A new approach for allocating buffers and bandwidth to heterogeneous, regulated traffic in an ATM node", *IEEE J. Selec. Areas Commun.*, pp. 1115–1127, 1995.

[14] A. Bak, W. Burakowski, F. Ricciato, S. Salsano, and H. Tarasiuk, "Traffic handling in AQUILA QoS IP network", in *2nd Int. Workshop Qual. Fut. Internet Serv.*, Coimbra, Portugal, Sept. 2001.

Andrzej Bąk was born in Warsaw, Poland in 1968. He studied automatic at Warsaw University of Technology where he received the M.Sc. degree in 1993. In 1998 he received the Ph.D. degree in telecommunications from Warsaw University of Technology. Since 1998 he has been an Assistant Professor at the Institute of Telecommunications, Warsaw University of Technology. His research interest is mainly focus on IP QoS and ATM technologies.
email: bak@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Andrzej Bęben was born in Poland in 1974. He received both M.Sc. and Ph.D. degrees in telecommunications from Warsaw University of Technology in 1998 and 2001, respectively. Now he is senior researcher at Military Communication Institute. His research interests include ATM and IP networks.
e-mail: abeben@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Wojciech Burakowski was born in Poland in 1951. He received the M.Sc., Ph.D. and D.Sc. degrees in telecommunications from Warsaw University of Technology, in 1975, 1982 and 1992, respectively. He is now a Professor at the Institute of Telecommunications, Warsaw University of Technology. Since 1990, he has been involved in the European projects COST 224, COST 242, COST 257 and 279. His research interests include ATM and IP network design as well as traffic control mechanisms.
e-mail: wojtek@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Marek Dąbrowski was born in Poland in 1976. He received the M.Sc. degree in telecommunications from Warsaw University of Technology in 2000. He is now a Ph.D. student. His research interest focus mainly on traffic control in IP QoS networks.

e-mail: mdabrow5@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Halina Tarasiuk was born in Poland in 1972. She received the M.Sc. degree in computer science from Szczecin University of Technology in 1996. In 1998 she started Ph.D. studies at Warsaw University of Technology. Since 2000 she has kept assistant position. Her research activities cover traffic control and architectural aspects of QoS IP and ATM technologies.

e-mail: halina@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Monika Fudała was born in Poland in 1976. She received the M.Sc. degree in telecommunications from Warsaw University of Technology in 2000. She is now a Ph.D. student. Her research interests mainly focus on traffic handling in QoS IP networks.

e-mail: mkrol@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Zbigniew Kopertowski was born in 1969 in Poland. He received the M.Sc. and Ph.D. degrees in telecommunications from Warsaw University of Technology in 1993 and 1997, respectively. He participated in European project COST 242 and Copernicus. His research interests mainly include ATM and IP QoS network design.

e-mail: zet@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Paper

Relations between ATM traffic parameters and quality of multimedia transmission – an experimental approach

Arkadiusz Sochan

Abstract — In the paper, we present the results of research of MPEG-2 streams transmission through CBR and VBR channel in ATM network. We applied not only experimental and measurement methods but analytical ones as well. A description of the proposed measurement station includes: installed hardware, created software and prepared experimental data. Next, we present the applied measurement method. In the following section, we define a term of the quality of multimedia transmission. The measurement results were compared with the results of the analytical station model. We proposed a functional description of dependence between ATM traffic parameters, video stream bit rate, and quality of multimedia transmission.

Keywords — multimedia transmission, quality of service, traffic parameters, MPEG-2, ATM, CBR and VBR.

1. Introduction

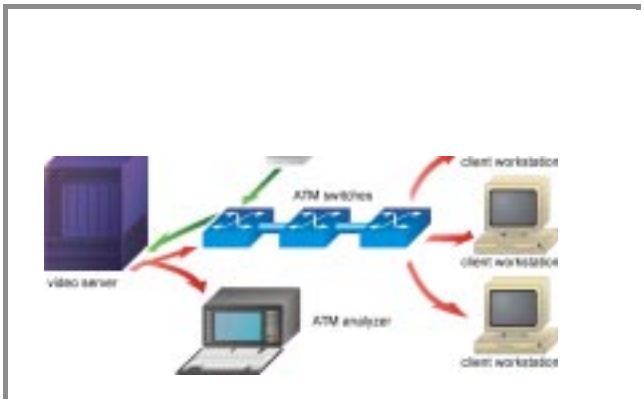
Applications of real-time multimedia services like videoconferences or video on demand has been lately widely discussed. Such services require much more bandwidth than traditional data transmissions. In the case of multimedia transmissions, besides the bandwidth, the important factors are: low loss rate, bounds of end-to-end delay, and delay variation. On the other hand, a lot of users should share network resources. Network operators ought to find a compromise between a number of users and quality of transmission.

Network technology employed to real-time multimedia transmission has to guarantee quality of service (QoS). Currently, the only one native technology assuring QoS on a large topographically area is asynchronous transfer mode (ATM). There are two categories of services in the ATM for real-time communication: constant bit rate (CBR) and variable bit rate (VBR). The both categories are described by combination of traffic parameters and QoS parameters called a traffic descriptor. The CBR traffic descriptor is simpler than the VBR traffic descriptor but the VBR category is potentially more efficient than its CBR counterpart in terms of network resources utilization. Additionally, the traffic descriptors are not always easy to understand clearly for end users who pay operators for their services. Moreover, it is difficult to predict traffic descriptors of encoded multimedia streams.

In this paper we propose an application of relative simply functional dependencies for efficiently bandwidth management in ATM networks. The obtained dependencies were drawn from experimental results. In our experiments we transmitted MPEG-2 streams both in CBR and in VBR ATM channels. The MPEG-2 streams were classified into various categories. Moreover, we made an analytical model of the station to compare with measurement results. The found dependencies can be an alternative for complex analytical and statistical models.

2. Testbed and samples of streams

The experimental research was performed on a testbed configured in the Laboratory of Computer Multimedia Systems at the Institute of Theoretical and Applied Informatics of PAS. The station included the server providing video on demand service, three workstations as clients of this service with prepared measurement software and an ATM network analyser. The components were connected through an ATM network built with three switches. Samples of multimedia streams we prepared on the encoding workstation with a built-in real-time MPEG-2 coder. The schema of the station is shown in Fig. 1.



Paper

Multistage optical switching networks

Sylwester Kaczmarek

Abstract — The backbone networks evolution to high-speed DWDM networks generates new problems for switching. This function element must be also based on optical technology. For large capacity this cannot be realized by a single matrix, but by multistage switching networks. In the paper three types of optical matrices have been described: fiber switch *FX*, wavelength fiber switch *WSX* and wavelength interchanging fiber switch *WIX*. Based on these matrices, four switching network constructions were considered. The connection properties of these switching networks were evaluated, determining the electronic equivalent switching network for which these properties are well known.

Keywords* — *optical matrix, optical switching networks, electronic switching networks equivalent, connection properties.

1. Introduction

The DWDM technology has increased the transmission capacity of fiber links and transmission systems. Thus the bandwidth bottleneck problem in transmission was overcome. But the backbone transport networks must fulfil two main functions: transmission and switching. The last one is necessary to make connections between users of this backbone network. It is realized by switching nodes of the backbone network. As the capacity of this backbone grows very strongly, the processing power of switching function must keep up with that growth. Thus the connection capability and power processing of nodes must also grow. To reach one's aim some problems must be solved: kinds of switched medium (space, wavelength, time), structure of matrices type, scalability of connection capacity. The solution all of these problems is difficult, as they very strongly depend on each other and great influence on that has the state of technology [1–4].

Now two kinds of medium can be switched: space (fibers) and wavelength. Based on this different type of switching matrices can be constructed. The connection capacity of these matrices is limited by the technology and, for that reason multistage optical switching networks must be used to overcome the capacity problem and realize the switching function. Their connection properties depend on the used matrix types and switching network structure.

In this paper the consideration about matrix types and network structures is presented. For the matrix we account two dimensions: fiber (space) and wavelength. Based on this assumption three types of optical matrix (element, fabric) were introduced: FX, WSX, and WIX. In Section 2 these elements are defined and described from the point of view of connection. The optical switching network can be constructed on these types of matrices and, we can do that in many ways. In Section 3 we have presented four constructions limited only to two side of multistage switching networks. The structures of these constructions are named GS_{FX}, GS_{WSX}, $GS_{WIX\_WSX\_WIX}$ and $GS_{WIX\_WIX\_WIX}$ and are described. To evaluate the connection properties we transform the optical switching network structure into electronic switching network structure. The last one is very well recognized and theoretically described. The results of these considerations are presented in Section 4. Finally Section 5 concludes the paper.

2. Optical switching matrix

We have limited our consideration to three types of optical switching matrix (element, fabric) that will be used to construct the multistage switching networks. The matrices are named: fiber switch FX, wavelength fiber switch WSX and wavelength interchanging fiber switch WIX. Each of these elements is described by the structure parameters, the number a of input fibers and number b of output fibers. The input and output fiber guide n optical waves denoted as $n\lambda$. The general notation of this structure is presented in Fig. 1. We introduce the following sets: fiber input set $A = \{1, 2, \ldots, a\}$, fiber output set $B = \{1, 2, \ldots, b\}$ and wave set $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. The connection between an input x and an output y we denote as $c = <x, y>$, where $x = (f^x, \lambda^x)$, $f^x \in A$, $\lambda^x \in \Lambda$ and $y = (f^y, \lambda^y)$, $f^y \in B$, $\lambda^y \in \Lambda$. All of the possible connections form a set $C = \{c : x \in X,\ y \in Y\}$. For each type of element considered in this paper we assume that

$$\underset{c \neq c_1}{\forall} \underset{(f,\lambda)}{\forall} \left((f, \lambda) \in c \Rightarrow (f, \lambda) \notin c_1 \right). \tag{1}$$

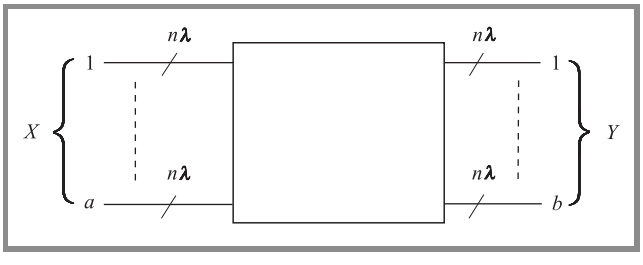


Fig. 1. General structure of optical switching element (matrix).

The FX matrix can connect only the input fibers with the output fibers, but that can be done only one_to_one fiber and has no influence on the n guide waves in these fibers.

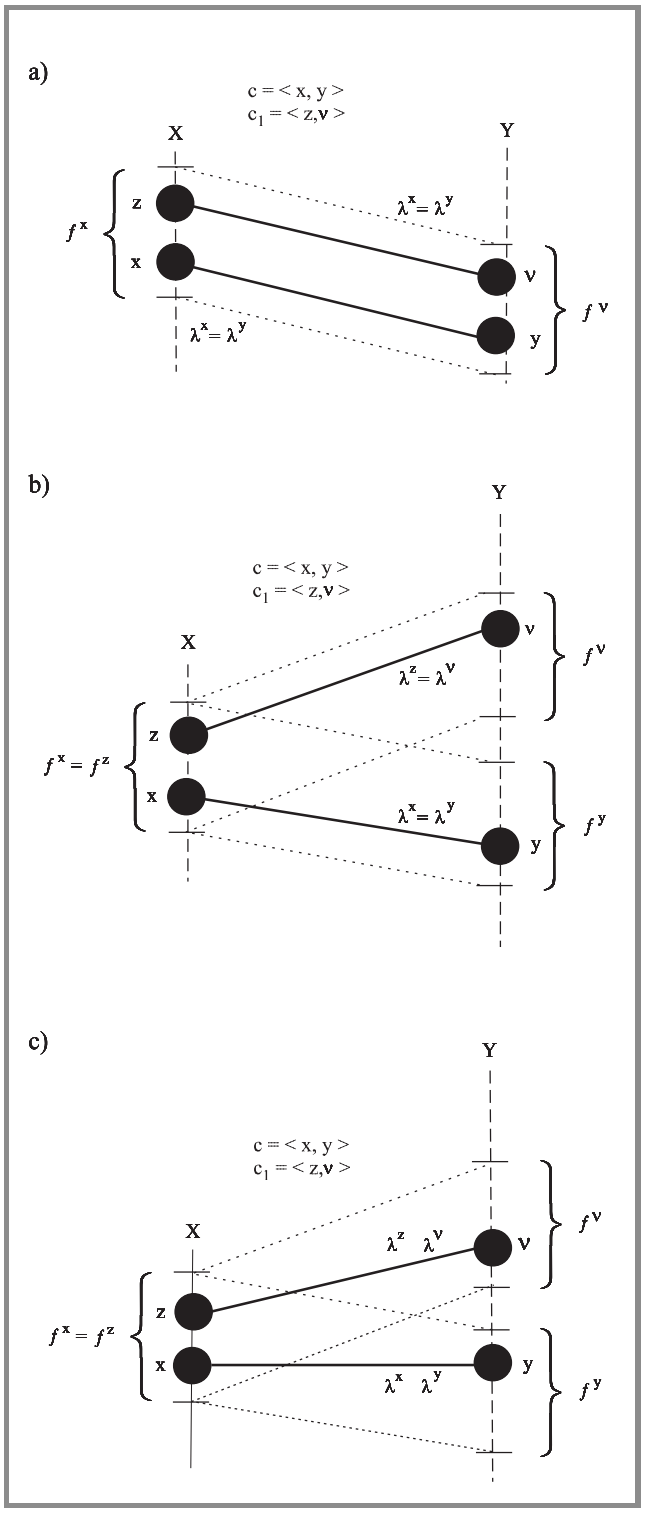


Fig. 2. Connection set C element illustration for matrix type: (a) FX; (b) WSX; (c) WIX.

We say that the connection is wavelength transparent. Let $c = <x, y>$ and $c_1 = <z, v>$ then the set C of the connections can be described as

$$C = \left\{ c : \underset{x}{\forall} \underset{\lambda}{\forall} \left(\lambda^x = \lambda^y = \lambda \right) \wedge \left((f^x = f^z) \Rightarrow (f^y = f^v) \right) \right\}. \quad (2)$$

Properties of this set are presented graphically in Fig. 2a.
For WSX matrix we assume that this element can additionally switch any wave between any input and any output fibers without changing the wavelength. The set C of the connection can be written as

$$C = \left\{ c : \underset{x}{\forall} \underset{\lambda}{\forall} \left(\lambda^x = \lambda^y = \lambda \right) \right\} \quad (3)$$

and this is presented graphically in Fig. 2b.
The WIX matrix has additional feature towards the last one. It can change the wavelength of connection between input and output fibers. Connection set C has elements, which fulfil

$$C = \left\{ c : \underset{x}{\exists} \underset{\lambda^x}{\exists} \left(\lambda^x \neq \lambda^y \right) \right\}. \quad (4)$$

This feature is presented graphically in Fig. 2c.
All of these matrices are nonblocking. The maximum number $|C|$ of connections is given by formula

$$|C| = n \cdot \min(a, b). \quad (5)$$

In this paper we have not get into consideration the technology of realization the optical switching matrix. It was assumed that the defined matrices, with above described connection properties, are available and based on these matrices the connection capabilities of the multistage optical switching networks were evaluated.

3. Multistage switching networks

Based on switching matrices presented in Section 2 we can construct many types of multistage switching networks. The most interesting are two side switching networks. For these ones we can distinguish four constructions:

- in each stage only FX matrices are used,
- in each stage only WSX matrices are used,
- in the first and last stage only WIX matrices are used and in the remaining stages WSX matrices are used,
- in each stage only WIX matrices are used.

We will later prove that each of these constructions has various connection capabilities, the first one has the worst and the last one has the best.
The switching network structure G is described by four parameters: left W_l and right W_r dimensions of matrices, set W_λ of waves and construction Alg of the connection between matrices belonging to the neighbouring stages.
All of these parameters can be denoted as follows:

- $G = < W_l, W_r, W_\lambda, Alg >$,
- $W_l = < a_1, a_2, \ldots, a_s >$,
- $W_r = < b_1, b_2, \ldots, b_s >$,

– $W_\lambda = < n\lambda_0, n\lambda_1, \ldots, n\lambda_s >$, where $n\lambda_i = \Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$,

– *Alg* is an algorithm given by Cantor, Clos, Benes or others [5–7],

where s is the G structure stage number. The switching network capacity can be counted in number N of fiber ports and, in maximum number $|C|$ of connections.
These numbers are given by the following formulas

$$N = \min\left\{\prod_{i=1}^{s} a_i, \prod_{i=1}^{s} b_i\right\}, \quad |C| = n \cdot N. \tag{6}$$

Typical practical switching networks have $GS = GG^{-1}$ structure with $(2s-1)$ stages, where G^{-1} is reversed (mirrored) structure to the G structure. We will now describe this *GS* structure for four above mentioned switching network constructions.
When all matrices are of FX type then GS_{FX} structure can only connect input fiber with output fiber. The switching network for the wavelength is transparent. The combinatorial power of connection realization depends on W_l, W_r and W_λ parameters. The switching network can be blocking, rearrangeable and nonblocking in strict and wide sense. Theoretical results are given and are just the same as for the analog space switching networks [5–7].

Conclusion 1. The GS_{FX} switching network structure has the same connection properties like the analog space switching network.

In the case when all matrices are of WSX type, the GS_{WSX} structure can be divided into n parallel *GS* switching networks, each one for different wavelength. Each of these switching networks has similar properties like the GS_{FX} structure and, is described by the same theorems. We must only remember that in the *GS* structure wave corresponds to fiber in GS_{FX} structure and is an input and output. Each of these structures can be blocking, rearrangeable and nonblocking in strict and wide sense. The connection features of *GS* structure depend on W_l and W_r parameters. The equivalent switching network of GS_{WSX} structure is presented in Fig. 3.

Conclusion 2. The GS_{WSX} switching network structure has the same connection properties like the n parallel separated analog space switching network.

The third construction $GS_{WIX\_WSX\_WIX}$ based on WIX and WSX type of matrices can be divided into two parts. The first part consists of the first and the last stage and, the second part consists of b_1 GS_{WSX} switching networks with structure parameters as follows: $W_l = < a_2, \ldots, a_s >$, $W_r = < b_2, \ldots, b_s >$, $W_\lambda = < n\lambda_1, \ldots, n\lambda_s >$, where $n\lambda_i = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Each of these GS_{WSX} can be divided into n parallel switching networks like those presented in Fig. 3. Thus the second part of $GS_{WIX\_WSX\_WIX}$ has $n \times b_1$ this one switching network. For this reason when we compare the former described two constructions with the same capacity N and $|C|$ we can say that the combinatorial power of connection realization is greater for $GS_{WIX\_WSX\_WIX}$ than for GS_{FX} and GS_{WSX} structures. $GS_{WIX\_WSX\_WIX}$ structure can be blocking, rearrangeable and nonblocking in strict and wide sense. The connection features depend on W_l, W_r and W_λ parameters.

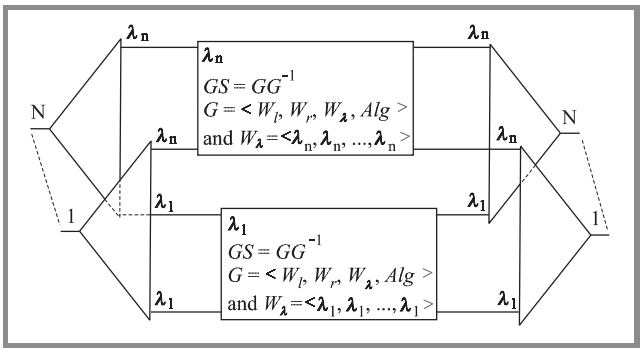


Fig. 3. Equivalent of $GS_{WSX} = GG^{-1}$ structure: $G = <W_l, W_r, W_\lambda, Alg>$ and $W_\lambda = < n\lambda_0, n\lambda_1, \ldots, n\lambda_s >$.

Conclusion 3. The $GS_{WIX\_WSX\_WIX}$ switching network structure has better connection properties than GS_{FX} and GS_{WSX} switching network structures.

The last one, $GS_{WIX\_WIX\_WIX}$ construction is based only on WIX matrices and each matrix can change any input wavelength to any output wavelength. For this reason this construction has the combinatorial power of connection greater than other above-mentioned constructions. This construction can be also blocking, rearrangeable and nonblocking in strict and wide sense. The connection features depend on W_l, W_r and W_λ parameters.

Conclusion 4. The $GS_{WIX\_WIX\_WIX}$ switching network structure has better connection properties than GS_{FX}, GS_{WSX} and $GS_{WIX\_WSX\_WIX}$ switching network structures.

Apart from the two side optical switching networks the one side optical switching networks can be constructed based on the same set matrix types. The most interesting is the structure based only on WIX matrices. This structure has the best connection properties. But all the one side constructions have disadvantages in transmission domain because of different path length in switching network – the number of matrices in connection path is variable. And for that reason these constructions will not be described in the paper.

4. Optical and electronic switching network similarity

Between optical and electronic matrices technology we can indicate similarity with respect of connection capability. The electronic matrices are analog or digital. For analog

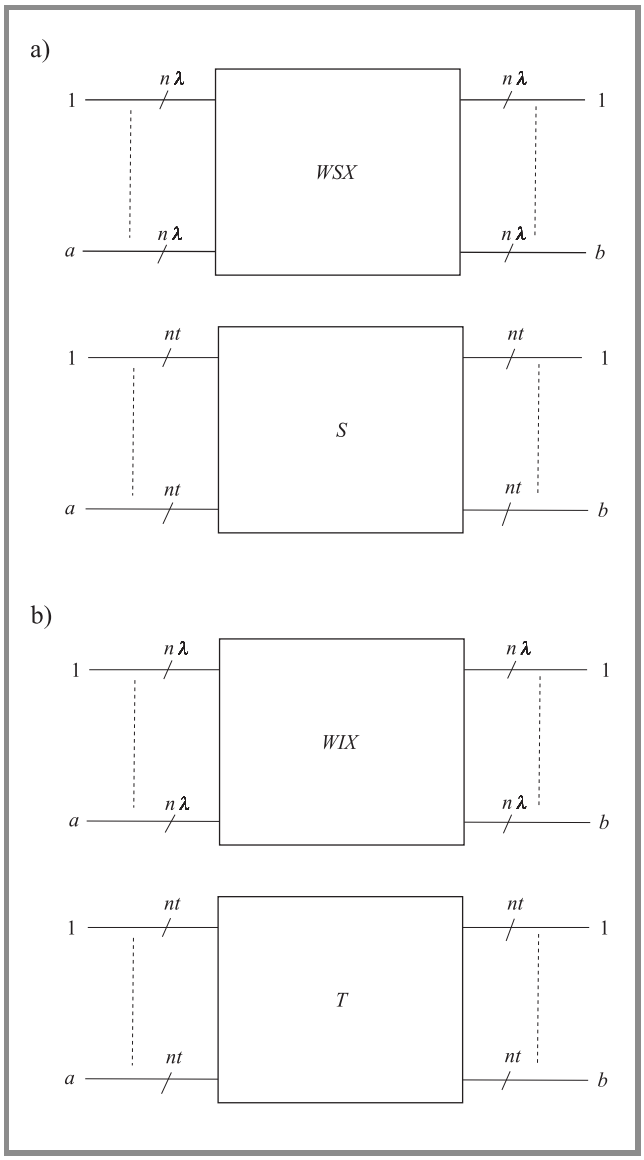


Fig. 4. Electronic equivalent of optical *WSX* (a) and *WIX* (b) matrices.

matrices we have only space connection between matrix input and output. The same situation can be observed for *FX* fiber matrix that connects input and output fiber and for that reason we can say that the analog space matrix is equivalent for fiber matrix. From this it results that the analog space switching network is an equivalent for GS_{FX} switching network (Conclusion 1).

Digital matrix is constructed based on time slot division multiplexing, however the optical matrix on wavelength division multiplexing. When we, for optical matrix, replace wavelength with time slot, then we obtain time slot division matrix as an equivalent. That replacing can be made for two types of optical matrix presented in the paper, *WSX* and *WIX* type of matrix. In the first one we obtain as an equivalent digital space matrix *S* with frame length equal to *n* and for that reason GS_{WSX} switching network has equivalent *S*-*S*-...-*S*-*S* switching network (Conclusion 2). For the second one the time switching matrix *T* is the equivalent matrix. These two cases are presented in Fig. 4. The *nt* symbol denotes *n* time slots in the frame.

The two, above described optical switching network constructions with *WIX* matrices, have two different equivalent electronic switching networks. For $GS_{WIX\_WSX\_WIX}$ the switching network structure the *T*-*S*-...-*S*-*T* structure is an equivalent however for $GS_{WIX\_WIX\_WIX}$ is the *T*-*T*-...-*T*-*T* structure.

The electronic *S*-*S*-...-*S*-*S*, *T*-*S*-...-*S*-*T* and *T*-*S*-...-*S*-*T* switching network structures are very well recognized and theoretically described [8, 9]. All these results can be transferred to the optical switching network.

Conclusion 5. Connection properties of above considered optical switching networks are the same as their electronic switching network equivalents and thus are very well described.

Only the theoretical results of optimalisation problems must be reformulated because of different criteria for optical and electronic switching networks. For the optical switching networks very important are the number of interstage links and the number of stages and the crosstalk [10], on the contrary to the electronic switching networks are the memory capacity and dimension of matrix.

We must emphasize that these optical switching networks because of limitation to switching in space (fiber) and wavelength do not cover all needs, which must be realized by switching function. These considered optical switching network constructions are transparent for flows aggregated in and transmitted by one wavelength. For the backbone network on the upper plane that will be done but not for the lower one. In the lower plane must be process of individual flows. This can be solved only when the optical switching function will be realized also in time not only in space and wavelength dimensions. Switching in time dimension can be realized in two ways, that is by assumption time or packet division multiplexing. But to meet this an appropriate optical matrix type must be designed.

5. Conclusions

In this paper connection properties of optical switching networks based on presented three types of matrices were analysed. Each of these optical matrices has an equivalent in the electronic domain. The multistage optical switching network constructed on these matrices type can be transformed into an electronic equivalent from the connection properties point of view. The four described main structures of two side optical switching networks have the same properties as the electronic switching networks. Wavelength is replaced by time slot by transformation from optical to electronic equivalent. All theoretical results for electronic space and time switching networks can be transferred to optical switching networks. Because of wavelength transparency, the optical switching networks can connect only

aggregated streams in the wave. To have switching possibility of individual connections and streams transported by the wave, an optical matrix with time dimension must be designed.

References

[1] ITU-T, G.872, "Architecture of optical transport networks", 02/1999.

[2] F. X. Ollivier, C. Zugno, and S. Thompson, "Evolution of high speed DWDM backbone networks", *ALCATEL Telecommun. Rev.*, 3th quart., pp. 181–193, 2000.

[3] P. Perrier and S. Thompson, "Optical cross-connects: the newest element of the optical backbone network", *ALCATEL Telecommun. Rev.*, 3th quart., pp. 195–201, 2000.

[4] N. A. Jackman, S. H. Patel, B. P. Mikkelson, and S. K. Korotky, "Optical cross connects for optical networking", *Bell Labs Techn. J.*, Jan.–March, pp. 262–281, 1999.

[5] C. Clos, "A study of non-blocking switching networks", *BSTJ*, vol. 32, no. 2, pp. 406–424, 1953.

[6] V. E. Benes, *Mathematical Theory of Connecting Networks and Telephone Traffic*. New York: Academic Press, 1965.

[7] D. G. Cantor, "On non-blocking switching networks", *Networks*, vol. 1, no. 4, pp. 367–377, 1972.

[8] A. Jajszczyk, *Switching Networks Composed of Time-Space Switches*. Rozprawy nr 160. Poznań: Politechnika Poznańska, 1985 (in Polish).

[9] W. Kabaciński, *Selected Problems of Constructing Broadband Switching Networks*. Rozprawy nr 342. Poznań: Politechnika Poznańska, 1999 (in Polish).

[10] T-S. Wong and C-T. Lea, "Crosstalk reduction through wavelength assignment in WDM photonic switching networks", *IEEE Trans. Commun.*, vol. 49, no. 7, pp. 1280–1287, 2001.

Sylwester Kaczmarek received the M.Sc./B.Sc. in electronics engineering, Ph.D. and D.Sc. in switching and teletraffic science from the Technical University of Gdańsk, Gdańsk, in 1972, 1981 and 1994, respectively. His research interests include: broadband ATM and IP networks, switching, routing, teletraffic and quality of service. He has published more than 120 papers.

e-mail: kasyl@eti.pg.gda.pl
Faculty of Electronics, Telecommunications and Informatics
Technical University of Gdańsk
G. Narutowicza st 11/12
80-952 Gdańsk, Poland

Paper

Greedy randomised adaptive search procedures for topological design of MPLS networks

Andrzej Mysłek

Abstract — In this paper, the IP/MPLS network cost optimisation problem of selecting localisation of nodes and links, combined with link's dimensioning, is discussed. As the considered problem is hard, we discuss and propose greedy randomised adaptive search procedure (GRASP) based solution method. GRASP is an iterative randomised sampling technique which combines adaptive randomised greedy function in constructing initial solution with local search optimisation. The effectiveness of the method is illustrated by means of a numerical study. We compare the GRASP results with results for both exact and heuristic methods obtained in previous research concerning topological design problem.

Keywords — network design, optimisation, MPLS, GRASP, local search.

1. Introduction

New communications systems, such as ATM or IP backbone networks, have to be deployed in a short time due to user's demand for provision of new services. Furthermore, ensuring the high quality of supplied services is required. Whereas the IP network does not possess built-in capabilities to ensure the quality of service (QoS), the multiprotocol label switching (MPLS) introduced to the IP network can add mechanisms for QoS assurance.

Effective and fast network deployment requires good network design tools, incorporating fast network optimisation algorithms for diverse optimisation tasks. To identify the problem described in this paper in the range of different network optimisation problems, we can use the following classification, proposed in [1]:

- **access network topology optimisation** problems are usually looking for hierarchical network structures, assuming a simple traffic demand pattern;
- **backbone network topology optimisation** problems require (arbitrary) mesh topology with traffic demands between each pair of end nodes given.

The optimisation problem, called the transit node and link localisation problem (TNLLP), is a general instance of the backbone network topology optimisation problem.

The TNLLP is characterised as follows.

For a given set of access nodes and the demand between each pair of access nodes find:

- number and locations of the actually installed transit nodes (in these nodes non traffic is originated, they only switch the traffic streams between access node pairs);
- capacity of links connecting access nodes to transit nodes;
- capacity of links interconnecting transit nodes;

such that the demand is realised at a minimum total network cost, which is composed of:

- fixed installation cost of each transit node,
- fixed installation cost of each link,
- capacity-dependent cost of each link.

The considered problem can be applied to MPLS-capable IP networks, with the access nodes representing ingress/egress label edge routers (LER), and the transit nodes – label switching routers (LSR). Since the evolving IP/MPLS networks will comprise large numbers of LERs and LSRs to be placed in many possible sites, the problem of optimal node and link location becomes important, especially for economical network extension. The problem definition is general enough to describe topology design problem without node localisation (cf. ONDP in [1]).

A subproblem of the TNLLP, known as optimal network design problem (ONDP) was widely studied in [2–5] using greedy and branch-and-bound algorithms. In [6] the multicommodity capacitated network design problem, an arc-based formulation with variable cost of flow and installation cost of multiple facilities installed on the arc, is solved combining the cutting plane method with the Lagrangean relaxation and heuristics. In Ref. [7] the dualascent procedure in the Lagrangean relaxation to solve an uncapacitated network design problem is used. Other references can be found in [1].

Since a problem very similar to the considered one is known to be NP-complete [8], one cannot expect to find time efficient algorithms for exact solving of the TNLLP. The branch-and-bound approach [9], although exact, is in general too time-consuming, especially for large networks. Hence, we have to use heuristic methods. Some heuristics were already studied in [1], and some of them (H4B, SAL, SAN) supply quite good suboptimal solutions. In this paper we inspect another heuristic method called the greedy randomised adaptive search procedure, that comprises greedy,

but randomised construction of the initial solution, and the local search method. GRASP has already been applied to many different problems (cf. [10–12]), and seems to be a good metaheuristic method, which may be adapted to solve TNLLP.

This paper is a continuation of [1] and [9], and it is organised as follows. In Section 2 we give a formal statement of the considered problem. Section 3 consists of GRASP method's presentation. A discussion of GRASP implementation for TNLLP is provided in Section 4. In Section 5 we illustrate efficiency of the GRASP implementation using numerical examples. Conclusions are drawn in Section 6.

2. Problem formulation

Let us consider two disjoint sets of nodes: access nodes labelled with $w = 1,\ldots,W$, and transit nodes labelled with $v = 1,\ldots,V$. The nodes are connected by undirected links, labelled with $e = 1,\ldots,E$. The links connecting access nodes to transit nodes are called access links, and transit links interconnect transit nodes. There are no links between the access nodes. The incidence of links and transit nodes is given by the binary incidence coefficients $b_{ev} : b_{ev} = 1$ if link e is incident with transit node v, and $b_{ev} = 0$ otherwise.

The access nodes are installed (fixed) and cost nothing[1]. A transit node v can be provided or not. If it is installed, it costs l_v. Any transit or access link can also be provided or not. An installed link costs $c_e y_e + k_e$ (where y_e is the capacity of the link e), otherwise it costs nothing. Since the network must be consistent, if a link is provided then its end nodes must be provided as well.

The demands, which are imposed on network and labelled with $d = 1,\ldots,D$, are realised by means of flows allocated to admissible paths. With each demand d there is associated its origin node s_d, destination (target) node t_d and demand volume h_d.

There are two main ways of formulating the flow problem for such a graph: link-path formulation and node-link formulation. The former requires to define for each demand a set of admissible paths labelled with $j = 1,\ldots,J_d$. Each such path begins at node s_d, then traverses (a non-empty) subset of transit nodes, and ends at node t_d. The paths are defined by the binary incidence coefficients a_{edj}, where $a_{edj} = 1$ if link e belongs to path j of demand d, and $a_{edj} = 0$ otherwise.

Transit nodes and links localisation problem can be formally stated in its link-path formulation as the following mixed-integer programme (MIP).

[1]Introduction of the installation cost of the access nodes would change an objective function just for a constant value, because provision of access links is determined by imposed demands. Hence, it does not change our optimisation problem.

TNLLP1 (link-path formulation)

indices

$d = 1,2,\ldots,D$	demands
$j = 1,2,\ldots,J_d$	paths for realising flows of demand d
$v = 1,2,\ldots,V$	transit nodes
$e = 1,2,\ldots,E$	links

constants

h_d	volume of demand d
a_{edj}	$= 1$ if link e belongs to path j of demand d, 0 otherwise
c_e	cost of capacity unit of link e
k_e	fixed cost of installing link e
Y_e	upper bound for the capacity of link e
b_{ev}	$= 1$ if link e is incident with transit node v, 0 otherwise
l_v	fixed cost of installing transit node v
Y_e	upper bound for the capacity of link e
G_v	upper bound for the degree of transit node v

variables

x_{dj}	flow realising demand d allocated to path j (non-negative continuous variable)
y_e	capacity of link e (non-negative continuous variable)
σ_e	$= 1$ if link e is provided, 0 otherwise (binary variable)
ε_v	$= 1$ if node v is provided, 0 otherwise (binary variable)

objective

$$\text{minimise} \quad C = \Sigma_e c_e y_e + \Sigma_e k_e \sigma_e + \Sigma_v l_v \varepsilon_v \tag{1}$$

constraints

$$\Sigma_j x_{dj} = h_d \quad d = 1,2,\ldots,D \tag{2}$$

$$\Sigma_d \Sigma_j a_{edj} x_{dj} = y_e \quad e = 1,2,\ldots,E \tag{3}$$

$$y_e \le Y_e \sigma_e \quad e = 1,2,\ldots,E \tag{4}$$

$$\Sigma_e b_{ev} \sigma_e \le G_v \varepsilon_v \quad v = 1,2,\ldots,V. \tag{5}$$

In TNLLP1 constraints (2) assure that the demands are realised, and constraints (3) and (4) that the capacity of a non-provided link is equal to 0. Constraints (5) assure that if a link is provided then also its end nodes are. The objective is to minimise the cost of all links and of transit nodes. Parameters Y_e and G_v are high enough not to limit the capacity of edges and the number of edges outgoing from node respectively.

The set of admissible paths for each demand d may be defined as the set of all paths between s_d and t_d, or it may be somehow limited (e.g. to at most two-hop paths). In the former case, the values of J_d can be very large, leading to an excessive amount of flow variables in the link-path problem formulation given above. For such a problem, we can use the alternative node-link problem formulation.

The node-link formulation of TNLLP uses link flows instead of path flows and requires the flows to be directed. Hence, with each link e we associate two directed arcs traversing the link in two opposite directions. Note that also the demands are directed. To suppress the number of flow variables we use flow variables aggregated by originating node (non-negative continuous variables).

TNLLP2 (node-link formulation)
indices

$v = 1,2,\ldots,V$ transit nodes
$w = 1,2,\ldots,W$ access nodes
$e = 1,2,\ldots,E$ links
$t = 1,2,\ldots,T$ directed transit arcs (between transit nodes)
$f = 1,2,\ldots,F$ directed access arcs (between access and transit nodes)

constants

$h_{ww'}$ demand originating at access node w and destined for access node w
H_w $= \Sigma_{w'} h_{ww'}$ total demand originating at access node w
b_{ev} $= 1$ if node v is incident with link e, 0 otherwise
b_{tv} $= -1$ if transit arc t is incoming to transit node v
$= 1$ if transit arc t is outgoing from transit node v
$= 0$ otherwise
b_{fv} $= -1$ if access link f is incoming to transit node v
$= 1$ if access link f is outgoing from transit node v
$= 0$ otherwise
b_{fw} $= -1$ if access link f is incoming to access node w
$= 1$ if access link f is outgoing from access node w
$= 0$ otherwise
a_{et} $= 1$ if transit arc t is realised on link e, 0 otherwise
a_{ef} $= 1$ if access arc f is realised on link e, 0 otherwise
Y_e upper bound for the capacity of link e
G_v upper bound for the degree of transit node v

variables

x_{tw} flow realising all demands originating at access node w on transit arc t
x_{fw} flow realising all demands originating at access node w on access arc f
y_e capacity of link e
σ_e $= 1$ if link e is installed, 0 otherwise (binary variable)
ε_v $= 1$ if node v is installed, 0 otherwise (binary variable)

objective

minimise (1)

constraints

$$\Sigma_t a_{et} \Sigma_w x_{tw} + \Sigma_f a_{ef} \Sigma_w x_{fw} = y_e \quad e = 1,2,\ldots,E \tag{6}$$

$$\Sigma_f b_{fw'} x_{fw} = H_w \quad w = 1,2,\ldots,W \tag{7}$$

$$\Sigma_f b_{fw'} x_{fw} = h_{ww'} \quad w = 1,2,\ldots,W, \quad w' = 1,2,\ldots,W \tag{8}$$

$$\Sigma_t b_{tv} x_{tw} + \Sigma_f b_{fv} x_{fw} = 0 \quad v = 1,2,\ldots,V, \quad w = 1,2,\ldots,W \tag{9}$$

$$y_e \le Y_e \sigma_e \quad e = 1,2,\ldots,E \tag{10}$$

$$\Sigma_e b_{ev} \sigma_e \le G_v \varepsilon_v \quad v = 1,2,\ldots,V. \tag{11}$$

In the formulation of the constraints in TNLLP2 only the following flow variables are used:

- x_{tw} for all pairs (t,w) such that $t = 1,2,\ldots,T$, $w = 1,2,\ldots,W$
- x_{fw} for all pairs (f,w) such that $f = 1,2,\ldots,F$, $w = 1,2,\ldots,W$ and either access link f is outgoing from access node w (i.e. $b_{fw} = 1$) or link f is incoming to some other access node w' with $h_{ww'} > 0$ ($b_{fw} = -1$ and $h_{ww'} > 0$).

In TNLLP2 each link e supports two directed arcs (either both access or both transit) whose numbers are specified by coefficients a_{ef} and a_{et} (respectively) equal to 1. Constraint (7) forces the total demand H_w generated in access node w to flow out, constraint (8) – that the portion $h_{ww'}$ of the flow originated at node w and destined for node w' stays at w', and constraint (9) – that no flow stays in a transit node.

The savings in the number of variables and constraints while using TNLLP2 instead of TNLLP1 are illustrated in [1].

Note that in the optimal solution of TNLLP, the demands can be routed on single paths, so if a demand corresponds – to an MPLS tunnel, the solution guarantees that the tunnel is realised on one path. This follows from the linearity of the capacity-dependent cost used in the cost function: for each fixed configuration of nodes and links, the capacity dependent part of (1) is minimised by assigning the entire demand volume h_d to one of its shortest paths with respect to the link metrics c_e. This property was used in [9] to construct a lower bound for the branch-and-bound algorithm.

3. Greedy randomised adaptive search procedure

A greedy randomised adaptive search procedure [10, 11] is one of neighbourhood search methods, like local search and tabu search methods. The GRASP is metaheuristic approach, that was applied to various optimisation problems: graph, scheduling, assignment and other problems, and for different domain-specific problems like aircraft routing or network planning (for the full bibliography of the GRASP see [12]).

A general GRASP heuristic is a two-phase iterative process. In each GRASP iteration there are two phases – the construction of a new greedy randomised solution, which consequently is used as the starting point for the second phase, and a local search algorithm that is run afterwards. This procedure is executed repeatedly until some termination criterion is met. The best solution over all iterations is the result. Pseudo-code for a generic GRASP method is given in Source 1.

Source 1. Pseudo-code for GRASP

```
procedure GRASP (var bestSolution);
begin
  repeat
     ConstructGreedyRandomisedSolution(solution);
     LocalSearch(solution);
     UpdateBestSolution(solution,bestSolution);
  until TerminationCriterion();
end;
```

In the construction phase we are looking for a solution, building it up from smaller parts. If we describe the solution as a vector of, for example, binary decision variables, constructing a solution means to choose variables one by one, and decide about their values. Randomly choosing and setting one of the best variables is the probabilistic component of the GRASP method.

To make construction process more formal and applicable to different problems, in Ref. [11], the concept of a list of candidates, called restricted candidate list (RCL) is introduced. The RCL is a list of the best candidates (parts of a solution) that can be used to form a new solution. It is restricted, because instead of looking for just the best candidate we allow worse candidates to be chosen to form the solution. Candidates are chosen randomly, and after adding a candidate to the solution, we adapt greedy function, i.e. we calculate the impact of our choice for the greedy function. This procedure is presented in Source 2.

Source 2. Pseudo-code for the GRASP construction phase

```
procedure ConstructGreedyRandomisedSolution
  (var solution);
begin
  solution = Ø;
  while not SolutionConstructed(solution) do
  begin
    MakeRCL(RCL);
    s = SelectRandom(RCL);
    solution = solution ∪ {s};
    AdaptGreedyFunction(s);
  end;
end;
```

The solution obtained in the construction phase is not guaranteed to be locally optimal with respect to simple neighbourhood definition. Hence, we try to improve the solution, applying a local search loop. This algorithm works in an iterative fashion by moving to a better solution in the neighbourhood of the current solution. It terminates when no better solution can be found (in the neighbourhood). A neighbourhood $N(s)$ of the solution s relates this solution to a set of solutions. Choosing the best solution from the neighbourhood we run the greedy local search.

The two phases of GRASP give foundation for the good optimisation technique. A greedy choice of candidates forms random (but good) solution, which is then improved by a local search.

4. GRASP implementation for TNLLP

Greedy randomised adaptive search procedure uses a two-phase iterative approach to solve the problem. In the first phase, the construction phase, construction of restricted candidate list is repeated. In each iteration of the construction phase we choose a part of the solution to form initial solution for the second phase of GRASP.

In TNLLP a solution is a set of edges and nodes provided, and the set of paths used to realise demands (each demand on a single path). Note, that if we decide which edge and node should be provided, we can easily find paths for demands using the shortest path according to variable cost of edges. Hence, our construction phase would consist of repeatedly choosing one edge and adding it to the solution. We could select, for example, edges with the highest flow belonging to the shortest path according to the modified weights from [9] (incorporating the installation cost of nodes and links, and the variable cost of edges).

Selecting edges to form a solution seems to be a good method. We can rank edges according to their flow or the number of demands realised on and we have termination criterion for construction (when all the edges not chosen do not bare the flow). However, the solutions generated by this method are usually poor.

The better starting solution can be constructed using demand's flows allocation. This method was used in [1] to construct starting solution for heuristics, but here we can add some randomisation. Let each demand's flow be routed through the shortest paths, according to the current state of the network. We calculate the length of the path as the sum of the variable cost of the demand and the fixed cost of not provided links (and optionally nodes) on that path. We adapt our greedy function by setting the edges (and nodes) as provided after the shortest path is found. Our modified construction phase is shown in Source 3.

Source 3. Pseudo-code for the GRASP modified construction phase (for TNLLP)

```
procedure ConstructGreedyRandomisedSolution
  (var solution);
begin
  solution* = Ø;
  MakeRCL(demandRCL);
  repeat
    d = SelectRandom(demandRCL);
    path = FindShortestPath(solution,d);
    solution = solution ∪ NodesAndEdges(path);
  until size(demandRCL)= 0;
end;
```

* solution consists of edges and nodes

In the construction phase we first construct a demand RCL. Then, selecting in each iteration one demand at random, we route this demand through the shortest path, updating the solution. When all the demands are routed, the construction phase is done. Our initial solution is good enough to apply efficiently the local search phase. Moreover, our initial solutions are well randomised, since the order of demands routed is random in each run of the construction phase.

Next, we try to improve GRASP initial solution obtained in the construction phase applying a local search. To identify the local search procedure for TNLLP we have to choose the neighbourhood model (the solution model was determined by the construction phase). Certainly, we could use the model from the construction phase (the demand reallocation), but to make the search more exhaustive we select nodes and links switching. An iteration of the local search will comprise switching selected edge (or node), i.e. making edge unavailable if it is provided and vice versa.

Switching an edge off requires rerouting demands which use the edge. It is rather simple, since the shortest paths of other demands do not change and there are not so many edges provided. Alternatively, switching edge on should involve rerouting all the demands, because any shortest path can change[2].

Finally, a GRASP iteration is composed of the construction phase, where we obtain the initial solution randomly routing demands on the shortest paths, and the local search, that improves the initial solution. We repeat the GRASP iteration until some arbitrary termination criterion is met, for example maximum iteration limit is reached or there was no improvement for last n iterations.

The GRASP for TNLLP, as it was described above, has some variants. We can use the installation cost of transit nodes during construction phase combined with the variable cost of demand and installation cost of edges, or just the variable cost of demand and installation cost of edges to determine the shortest path for a demand. In the local search phase the nodes can be switched first, and then edges instead of switching edges only.

5. Numerical results

We have considered three artificially generated network structures (N7, N14 and N28) determined by the geographical locations of nodes, and one realistic network (PL49) reflecting Polish public backbone network. The networks were used in [1] and are available on the web site [13]. The basic parameters of the networks are given in Table 1. All links are potentially available.

The unit cost c_e of link e is in all cases proportional to its geographical length. The fixed installation cost is given by $k_e = c_e \cdot 10^n$ (n is a parameter in computations; the fixed cost of access links is additionally multiplied by 3, and for the transit links by 2). The fixed installation cost of a transit node is the same for all nodes and is given by $l_v = 10^k$ (k is another parameter in computations).

[2]Finding the shortest path is the most time-consuming activity in the demand's routing.

Table 1
Test networks parameters

Network	W	V	F	T	D	$\min h_d$	$\max h_d$	$\Sigma_d h_d$
N7	7	5	70	20	42	240	1920	34 320
N14	14	11	308	110	182	120	7560	172 320
N28	28	15	840	210	756	120	30 240	892 492
PL49	49	12	1178	132	2352	36	53 572	2 788 073

Applied GRASP methods vary in network installation cost factors used in the solution construction, node switching and edge adding. G1 operates on the fixed cost of nodes and the cost of flow (variable cost of edges times demand volume) during the construction phase and removes edges in the local search. The fixed cost of nodes was added as a factor in the construction phase of G2. Methods G3 and G4 are respectively similar to G1 and G2, but in a local search phase they try to switch off transit nodes first, and then remove edges. All the four methods G1–G4 have their counterparts G1$^+$–G4$^+$, which in the local search phase not only remove edges, but can add them too.

Table 2
Results for TNLLP (cost in units of 10^6)

Network	n	k	H4B*	SAN*	SAL*	EX*	G1	G2
N14	4	4	58.93	81.38	57.61	57.61	58.54	58.54
N14	4	5	59.92	81.83	58.60	58.60	59.53	62.24
N14	4	6	69.82	84.07	69.53	67.14	69.43	75.28
N14	5	4	266.23	426.88	265.32	—	276.39	276.39
N14	5	5	267.22	427.24	267.22	—	277.38	277.38
N14	5	6	277.12	430.84	277.88	—	287.28	287.28
N28	4	4	253.21	274.45	262.46	—	262.32	262.32
N28	4	5	254.56	277.21	256.34	—	263.67	263.67
N28	4	6	268.06	283.51	284.04	—	277.17	318.72
N28	5	4	751.60	1369.97	742.94	—	781.85	781.85
N28	5	5	752.95	1370.78	769.64	—	783.20	795.81
N28	5	6	761.37	1378.88	784.93	—	796.70	809.31
PL49	4	7	1199.75	1186.80	1207.50	—	1272.84	1361.90
* Results for the best methods published in [1].								

Table 3
Results for TNLLP – continued (cost in units of 10^6)

Network	n	k	G3	G4	G1$^+$	G2$^+$	G3$^+$	G4$^+$
N14	4	4	58.54	58.54	57.61	57.61	57.61	57.61
N14	4	5	59.53	62.24	58.60	58.60	58.60	58.60
N14	4	6	69.43	75.28	68.50	75.28	68.50	75.28
N14	5	4	276.39	276.39	271.25	271.25	271.25	271.25
N14	5	5	277.38	277.38	272.24	272.24	272.24	272.24
N14	5	6	287.28	287.28	282.14	282.14	282.14	282.14
N28	4	4	262.32	262.32	250.02	250.02	250.02	250.02
N28	4	5	263.67	263.67	251.37	251.86	251.37	251.86
N28	4	6	277.17	318.72	264.87	277.15	264.87	277.15
N28	5	4	781.85	781.85	756.25	756.25	756.25	756.25
N28	5	5	783.20	795.81	757.60	757.60	757.60	757.60
N28	5	6	796.70	809.31	771.10	771.10	771.10	771.10
PL49	4	7	1272.84	1361.90	1197.89	1361.90	1197.89	1361.90

Table 4
Relative cost difference for TNLLP with respect to the optimal solution [%]

Network	n	k	H4B*	SAN*	SAL*	G1	G2	G3	G4	G1$^+$	G2$^+$	G3$^+$	G4$^+$
N14	4	4	2.28	41.24	0	1.61	1.61	1.61	1.61	0	0	0	0
N14	4	5	2.24	39.63	0	1.59	6.21	1.59	6.21	0	0	0	0
N14	4	6	3.98	25.21	3.55	3.41	12.12	3.41	12.12	2.03	12.12	2.03	12.12
* Results for the best methods published in [1].													

Results for TNLLP are given in Tables 2–4 together with results for specialised heuristic H4B, simulated annealing (SAN) and simulated allocation (SAL) from [1]. Results for G1–G4 were obtained after 100 iterations of the algorithm, except PL49, for which results were obtained after 50 iterations. Methods G1$^+$–G4$^+$ were run for 20 iterations.
Table 5 presents convergence of methods. Running times are shown in Tables 6 and 7. We can notice that methods that employ the node cost in the construction phase, i.e. search the shortest path using the installation cost of nodes, the installation cost of edges and the flow cost, give worse results then, one might think less reasonable, method of constructing a solution using just the installation cost of edges and the flow cost.

Table 5
Convergence to the optimal solution [%]

Network	n	k	G1, G3			G2, G4			G1$^+$		
			20*	50*	100*	20*	50*	100*	5*	20*	100*
N14	4	4	10.57	6.32	1.61	10.57	6.32	1.61	1.61	0.00	0.00
N14	4	5	10.39	6.21	1.59	10.39	6.21	6.21	1.59	0.00	0.00
N14	4	6	11.10	7.45	3.41	12.12	12.12	12.12	3.41	2.03	2.03
* Number of iterations.											

We can also find that methods using node switching in their local search give the same results as their counterparts, and run for a slightly longer time. Nevertheless, these methods can be much faster for networks that have high fixed cost of nodes, because they remove in one local search iteration a bunch of edges. Certainly, they can miss the better solution then.

Table 6
Running times for TNLLP [s] (H4B, SAN, G1–G3: Sun Sparc Ultra-4; SAL: PC 800 MHz)

Network	n	k	H4B*	SAN*	SAL*	G1	G2	G3
N14	4	4	3.71	326.74	21.36	18.55	18.69	24.18
N14	4	5	3.77	327.14	3.14	18.40	19.20	23.98
N14	4	6	3.59	318.09	64.08	18.55	7.75	24.26
N14	5	4	3.77	317.61	5.15	13.11	13.06	16.44
N14	5	5	3.78	316.90	19.45	13.01	13.07	16.55
N14	5	6	3.77	324.74	7.46	13.33	12.95	16.84
N28	4	4	35.90	1287.99	78.08	245.01	241.91	300.66
N28	4	5	36.61	1272.05	193.14	242.51	258.77	300.99
N28	4	6	38.65	1271.74	58.66	243.09	84.70	299.33
N28	5	4	35.92	1321.07	58.64	131.13	135.81	160.54
N28	5	5	35.82	1314.30	44.83	130.53	150.39	159.83
N28	5	6	35.00	1304.07	89.97	131.78	133.38	160.58
PL49	4	7	135.46	265.19	146.11	452.67	94.69	514.30
* Results for the best methods published in [1].								

Table 7
Running times for TNLLP [s] – continued (G4, G1$^+$–G4$^+$: Sun Sparc Ultra-4)

Network	n	k	G4	G1$^+$	G2$^+$	G3$^+$	G4$^+$
N14	4	4	23.89	143.55	145.70	151.23	152.01
N14	4	5	25.18	143.01	142.10	150.92	149.89
N14	4	6	11.53	146.95	40.63	163.52	42.81
N14	5	4	16.68	23.51	24.09	24.97	25.01
N14	5	5	16.57	23.76	24.05	24.66	25.03
N14	5	6	16.75	24.20	23.79	24.90	24.95
N28	4	4	297.97	4062.85	4047.29	4164.93	4176.44
N28	4	5	305.95	4080.50	4183.02	4150.94	4289.03
N28	4	6	116.61	4154.86	1961.39	4155.54	2005.43
N28	5	4	160.71	1431.43	1427.85	1522.83	1485.44
N28	5	5	162.96	1435.02	1430.90	1474.59	1506.80
N28	5	6	163.65	1438.98	1491.40	1471.11	1499.58
PL49	4	7	114.45	25439.00	562.95	25736.20	573.21

Finally, edge switching, instead of edge removing, is much better in terms of the final result of GRASP, but it takes a lot of time, because usually there are not as many edges chosen for initial solution (thus not as many to check for removing) as edges not included in the initial solution. In edge switching all the edges of a network must be checked (whether to be removed or included), what can change the method running time for an order of magnitude or even more.
G1 and G3 methods in one case give better results than other heuristic methods from [1] (N14, $n = 4$, $k = 6$), but generally results are 3–5% worse. Their running time is usually longer than the running time for H4B and SAL, but shorter than for SAN. G1$^+$ and G3$^+$ methods give results in 2% neighbourhood ($+/-$) of other heuristic methods from [1], but their running times are very long (cf. PL49).

6. Conclusions

In the paper we proposed a GRASP method implementation to solve topological node and link localisation problem. The problem, consisting in optimal locating of links and nodes under demand constraints, is *NP*-hard, and hence heuristic methods are needed. Heuristics become a **must** for larger networks, for which the running time of exact algorithms becomes infinite.
Numerical examples show that some of GRASP implementations drawn here, which use the demand routing through the shortest path and the local search, are quite efficient and supply good suboptimal solutions. Our study shows that during the construction phase (when we construct the initial solution) it is better not to use installation cost of nodes. Using edges switching in the local search phase instead of edge removing can give better results after fewer iterations, but the time of one iteration is large, and the overall running time of modified method can be order of magnitude larger.
Certainly, the proposed method can still be improved, for example by incorporating more advanced procedures to the construction phase to prepare better initial solution. RCL concept of GRASP can be further deployed by intro-

duction of a function of ranking demands. The promising way might be employing simulated allocation instead of the local search.

References

[1] M. Pióro, A. Jüttner, J. Harmatos, Á. Szentesi, P. Gajowniczek, and A. Mysłek, "Topological design of telecommunication networks. Nodes and links localization under demand constraints", in *17th Int. Teletraf. Congr.*, 2001.

[2] M. Minoux, "Network synthesis and optimum network design problems: models, solution methods and application", *Networks*, vol. 19, pp. 313–360, 1989.

[3] H. H. Hoang, "A computational approach to the selection of an optimal network", *Manag. Sci.*, vol. 19, pp. 488–498, 1973.

[4] D. E. Boyce, A. Farhi, and R. Weischedel, "Optimal network problem: a branch and bound algorithm", *Envir. Plan.*, vol. 5, pp. 519–533, 1973.

[5] R. Dionne and M. Florian, "Exact and approximate algorithms for optimal network design", *Networks*, vol. 9, pp. 37–59, 1979.

[6] B. Gendron, T. G. Crainic, and A. Fragnioni, "Multicommodity capacitated network design", in *Telecommunication Network Planning*, B. Sanso and P. Soriano, Eds. Boston: Kluwer, 1996.

[7] A. Balakrishnan, T. L. Magnanti, and R. T. Wong, "A dual-ascent procedure for large-scale uncapacitated network design", *Oper. Res.*, vol. 37, no. 5, pp. 726–740, 1989.

[8] D. S. Johnson, J. K. Lenstra, and A. H. G. Rinnoy Kan, "The complexity of the network design problem", *Networks*, vol. 8, pp. 279–285, 1978.

[9] M. Pióro, A. Mysłek, A. Jüttner, J. Harmatos, and Á. Szentesi, "Topological design of MPLS networks", in *Globecom*, San Antonio, 2001.

[10] T. A. Feo, M. G. C. Resende, and S. H. Smith, "Greedy randomized adaptive search procedure for maximum independent set", *Oper. Res.*, vol. 42, no. 5, pp. 860–887, 1994.

[11] T. A. Feo and M. G. C. Resende, "Greedy randomized adaptive search procedures", *J. Glob. Opt.*, vol. 6, pp. 109–133, 1995.

[12] M. G. C. Resende, "GRASP bibliography", http://www.research.att.com/ mgcr/doc/graspbib.pdf

[13] Examples of TNLLP networks, http://www.tele.pw.edu.pl/networks/TNLLP/

Andrzej Mysłek was born in Poland in 1973. He received M.Sc. degrees in telecommunications and in management from the Warsaw University of Technology in 1997 and 2000, respectively. He is now preparing his Ph.D. theses, concerning topological design of telecommunications networks, at the Warsaw University of Technology, in the Institute of Telecommunications.

e-mail: amyslek@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Paper

On effectiveness of conditional admission control for IP QoS network services with REM scheme

Marek Dąbrowski, Wojciech Burakowski, and Andrzej Bęben

Abstract — Future IP QoS (quality of service) networks are aiming at differentiating transfer quality of packets belonging to different flows. For this purpose, a set of network services (NS) with different QoS objectives is defined and implemented in the network. To a NS a certain amount of network resources, i.e. dedicated link capacity with associated buffer size, is allocated. Moreover, the resources dedicated for one NS are not available for other NSs. Traditional approach for admission control algorithm corresponding to given NS takes into account current traffic conditions inside considered NS. This can lead to the situation, due to traffic fluctuations, that temporary overloaded NS cannot use the spare bandwidth from underloaded in this time other NSs. This paper describes a conditional admission control algorithm (C-AC), allowing us to admit new packet flow conditionally in the case where no available capacity inside a given NS. For conditionally accepted flow currently unused capacity, dedicated to other NS, is allocated. This can be done only in the case when QoS requirements for both the conditionally accepted flow and the flows in progress are satisfied. The conditions for effective using of C-AC algorithm are discussed in the paper, like characteristics of NS borrowing and lending capacity and their current traffic load. To show potential benefits of the approach, exemplary numerical results are included, corresponding to hypothetical NSs using REM (rate envelope multiplexing) scheme.

Keywords — QoS IP network, conditional admission control.

1. Introduction

For the development of the future IP-based network, called IP QoS, two network architectures are discussed by the IETF: IntServ [4] and DiffServ [2, 3]. Despite that these architectures differ in many points, each of them offers a possibility for defining a set of network services with different QoS objectives. The NSs can be similar to these supported by ATM, like CBR (constant bit rate) and VBR (variable bit rate), or can be arranged for transferring packet stream associated with specific application (like WWW – world wide web). Implementation of a NS is possible thanks to QoS mechanisms available in IP routers. These mechanisms correspond to classification, policing, scheduling and buffer management. The excellent example of a new NS for IP network is the premium service for transmitting voice traffic [6].

Each NS is designed to offer specific QoS objectives, usually expressed in terms of maximum allowed packet transfer delay, packet transfer delay variation, packet loss ratio, etc. A certain amount of network resources, i.e. dedicated link capacity with associated buffer size, is assigned for each NS. Access to this capacity can be assured e.g. by setting appropriate weight value in the WFQ (weighted fair queuing) packet scheduler in the output port of the router. The maximum volume of traffic allowed inside a given NS is controlled by appropriate admission algorithm. In the case when there is not enough capacity available inside a given NS, new flow request is simply blocked. Let us remark that strict partitioning of network resources between NSs limits multiplexing gain only to the capacity dedicated for a single NS.

This paper describes a conditional admission control algorithm (C-AC), allowing us to admit new packet flow conditionally in the case where no available capacity inside a given NS. For conditionally accepted flow currently unused capacity, dedicated to other NS, is allocated. This can be done only in the case when QoS requirements for both the conditionally accepted flow and the flows in progress are satisfied. The conditions for effective using of C-AC algorithm are discussed in the paper, like characteristics of NS borrowing and lending capacity and current traffic load. To show potential benefits of the approach, exemplary numerical results are included, corresponding to hypothetical NSs using REM multiplexing scheme.

The paper is organized as follows. Section 2 discusses the implementation of differentiated NSs in IP networks. Section 3 describes the proposed C-AC algorithm and discusses its application and implementation aspects. Numerical examples are included in Section 4. Finally, Section 5 summarizes the paper.

2. Supporting differentiated NSs in IP networks

Conditional admission control algorithm can be engaged in the case when capacity of transmission link is strictly partitioned among a number of NSs. Each NS supports different QoS requirements corresponding to the packet transfer characteristics. The studied system with n NS is depicted

in Fig. 1a. Dedicated capacity in the link is assigned for each NS, adequate to the fixed value of the weight in the WFQ scheduler. Exemplary link partitioning is shown in Fig. 1b, where the ith NS has dedicated capacity equal to $C_i = w_i C$ (w_i – value of the weight for the ith NS, $i = 1, \ldots, n$, C-link capacity). The maximum allowed carried traffic inside the ith NS is limited by the C_i value (and the length of the associated buffer) and is controlled by appropriate admission algorithm. The type of applied admission rules directly depends on the type of multiplexing scheme assumed for the considered NS.

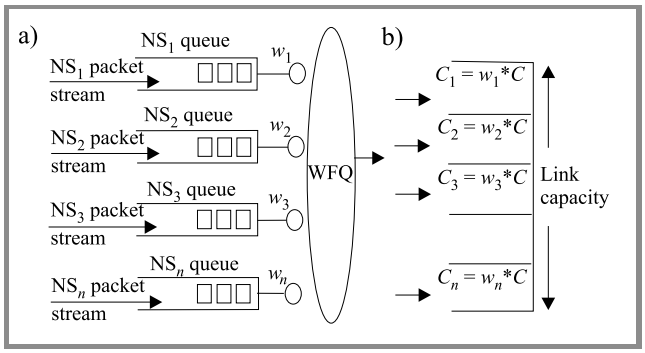


Fig. 1. Exemplary structure of the output port in the IP router supporting N different NSs: (a) studies system; (b) exemplary link partitioning.

We distinguish two types of multiplexing which are REM and RSM (rate sharing multiplexing). Let us recall that the REM scheme is dedicated for traffic with rigorous requirements with respect to packet delay characteristics. Usually, for this scheme a small buffer is dedicated for absorbing packets arriving to the system in the same time. On the contrary, the RSM multiplexing scheme is for bursty traffic and it requires relatively large buffer for absorbing traffic fluctuations in time. The differences between these schemes are important in the case of the discussed conditional admission.

The above system is in fact partitioned into N subsystems, each corresponding to different NS. High overall link utilization is reached only in the case the traffic load submitted to each NS is heavy at the same time. However, assuming that fluctuations in time of the traffic submitted for a given NS follow a stochastic process, there is a chance that a high percentage of new flows is blocked despite of spare capacity on the link. Better link utilization (lower flow request blocking) can be achieved by "borrowing" the resources from the NS that is temporarily under-utilized to the one that is actually overloaded. This requires changing the WFQ weights on the link, which has two serious drawbacks:

- Updating the values of weights in the WFQ scheduler in a dynamic way can cause uncontrolled traffic oscillations [5].
- Repartitioning of link resources may require adequate changes on all the subsequent links in the network (see Fig. 3). In a network based on the DiffServ architecture [2], this affects the scalability of AC mechanism. In an ideal case AC can be applied locally, taking into account traffic conditions in a particular Edge Router [6].

Therefore, changing WFQ weights is an operation, which can be performed in rather long time scale (e.g. hours). Repartitioning of resources between the NSs in the whole network can be done for example as a result of a long-term analysis of the traffic demands. New weights can be calculated in an off-line process of network reprovisioning, based on the observed changes of the traffic matrix.

If the network should be able to react quickly to traffic fluctuations in a shorter time scale, e.g. minutes, changing WFQ weights is not a reasonable solution. Therefore, we propose, so called conditional admission as a mechanism, which could significantly decrease the probability of call blocking in the case of short-time fluctuations of traffic offered to the network.

3. Conditional admission control

The proposed C-AC assumes that in the case of blocking, new flow can be accepted and submitted to other NS guaranteeing requested by this flow packet transfer quality. This is illustrated in Fig. 2. The conditionally accepted flow is submitted to the queue associated to other NS. It can take place only when this NS is under light load and available capacity is sufficient. It is obvious that the volume of available capacity limits the number of conditionally accepted flows. The flow is turned back to its own NS when capacity will be sufficient for its service. Anyway, the service of the considered flow can be terminated successfully even if the conditional status of this flow will not be changed. Of course, the service of the conditionally admitted flows is not always finished with success. This is due to the fact that new flows from the NS lending the capacity can arrive to the system. These flows are admitted with the highest priority and, as a consequence, can interrupt the service of the conditionally accepted flows.

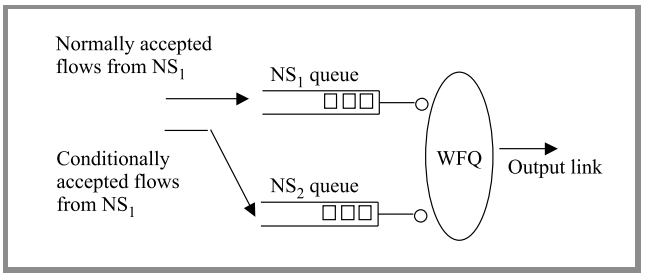


Fig. 2. Illustration of the conditional admission. Conditionally accepted new flow from the NS_1 borrows the capacity originally dedicated to the NS_2.

Effective application of the proposed C-AC scheme requires satisfactory solutions for the following questions:

- What about co-operation level of different NSs? It means, which NS can lend/borrow capacity to/from other NSs?
- When co-operation of the NSs can give expected profit?
- What about the implementation complexity level?
- What is the level of risk that the conditionally accepted flow will be terminated before the proper time?

3.1. Co-operating NSs

Basic requirement for co-operating NSs is that a flow "belonging" to a given NS can efficiently transfer its packets inside co-operating NS. Let us consider two NSs both working under REM (or RSM) scheme. Furthermore, let us assume that the first NS provides stronger QoS guarantees than the second. In order to assure appropriate QoS, the first NS has potential to lend the capacity to the second NS without essential limitations. On the contrary, the second NS can lend its resources only when it is currently under very light traffic conditions. In the next section, an example of co-operation between two NSs, both using REM multiplexing scheme, is more deeply discussed.
The rules of the co-operation between two NSs with different multiplexing schemes, i.e. REM and RSM, are not so clear. We can deduce that one NS can lend a limited volume of its capacity only when current traffic load conditions are rather low. Anyway, the detailed studies for specific NSs are required.

3.2. Expected profit

The potential profit that we can reach from the application of C-AC depends on the degree of the flow level (not the packet level) traffic fluctuations inside the co-operated NSs. It is obvious that no profit is possible when these NSs are overloaded at the same time. Otherwise, one can expect better traffic service when the traffic submitted to the considered NSs will alternate in time.

3.3. Risk assessment

The fundamental question is *when conditional admission of a new flow is reasonable?* Let us recall that a flow admitted in such a way can be terminated before the proper time. Note, that the probability of such event depends on the flow level traffic conditions and the current system state. The decision whether a new flow is conditionally accepted or not, should take into account the above probability. It is reasonable to admit a new flow conditionally only when this probability is relatively low, e.g. on the level comparable to the flow blocking probability.

3.4. Complexity level

Two essential factors influence the complexity of C-AC schema, which are:

- Modifications to the packet handling mechanisms in edge and core routers.
- Additional complexity of AC and user-network signaling.

A comparison of the proposed scheme with the method based on changing the WFQ weights is presented in Fig. 3. Consider a simple network, consisting of 2 links. Both links are equally partitioned between NS_1 and NS_2. Let us assume, that NS_1 is temporary overloaded (marked as light grey area) while NS_2 is at the same time underloaded. Then, the call blocking is observed in NS_1 despite there are unused resources in NS_2.

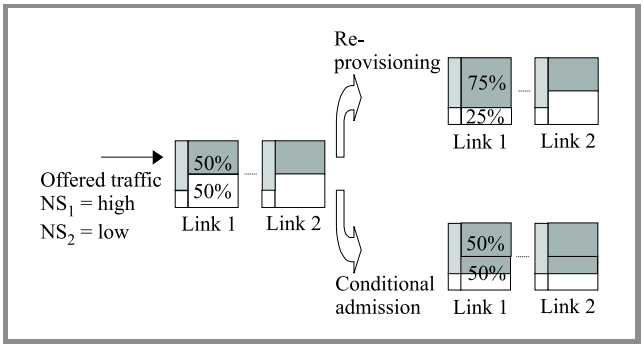


Fig. 3. Comparison of two methods of adapting network resources to fluctuating traffic load: changing WFQ weights (reprovisioning) and conditional admission.

As it was stated before, two solutions are possible for increasing the network utilization. One of them is to update the WFQ weights for NS_1 to serve the higher load. These weights must be adequately changed on all the consecutive links. Otherwise, as it is depicted in the Fig. 3, the NS_1 capacity on the first link (dark-grey area) is sufficient for serving the submitted higher load (light-grey area), while in the case of unchanged weights on the second link, an overload for NS_1 is still observed. If we admit the excess traffic from NS_1 conditionally, it must be served within capacity allocated for NS_2 on both links. In the framework of DiffServ architecture, conditional acceptance requires only marking the packets in the edge router with the code point corresponding to the PHB (per-hop behaviour) associated with NS_2. Then, packets are served in the same way on all subsequent routers in a given domain. Therefore, one can conclude that WFQ weights can be changed as a result of re-calculation process taking into account all links in the network. This imposes a significant complexity level and can be done only in a long time scale. On the contrary, conditional admission applies modifications to the packet marking locally at the edge router, not requiring changes in the core network.

Anyway, introducing conditional admission increases complexity of the control algorithm and user-network signaling implementation. This is caused by:

- Conditionally accepted flow should submit its traffic to another queue in the WFQ scheduler. If sufficient capacity will become available to accept this flow in the original NS, this flow should be switched to the proper WFQ queue, loosing its status of a conditionally accepted flow.
- Conditionally accepted flow may be terminated before the proper time by the new flows arriving to the NS, which lends its capacity to conditionally accepted flows.

4. Numerical example

In this section, the effectiveness of the proposed C-AC algorithm will be illustrated by considering an exemplary system with two NSs, say NS_1 and NS_2, both working under the REM multiplexing scheme.

Example 1. Co-operation of NSs designed for CBR traffic

Consider a system with NS_1 and NS_2, both designed for serving CBR traffic and working under typical AC algorithm (not C-AC). Guaranteed packet loss ratio (PLR) by NS_1 is 10^{-4} and by NS_2 is 10^{-2}. The associated buffer sizes, say B_1 and B_2, for NS_1 and NS_2 are of 10 packets each. The maximum allowed value of the link utilization for the NS_1, say ρ_1, can be calculated by [1, 5]:

$$\rho_1 = \frac{2B_1}{2B_1 - \ln(PLR_1)}. \qquad (1)$$

The value of $\rho_1 = 0.68$. Similarly, for NS_2 we calculate $\rho_2 = 0.81$. (Notice that ρ_1 and ρ_2 do not depend on the link capacity). The traffic submitted to the NS_1 (NS_2) and corresponding to the flow (call) level follows Poissonian process with parameter λ_1 (λ_2) while the holding times are negative-exponentially distributed with parameter μ_1 (μ_2). Additionally, each flow submitted to NS_1 or NS_2 requests for the same amount of bandwidth, fixed to 1. The total link capacity is 100.

Three scenarios are considered, which differ in link partitioning between NS_1 and NS_2. They are the following:

Scenario 1: $NS_1 = 50$, $NS_2 = 50$.
Scenario 2: $NS_1 = 70$, $NS_2 = 30$.
Scenario 3: $NS_1 = 30$, $NS_2 = 70$.

Furthermore, we investigate the system assuming that the offered traffic (at the call/flow level) is such that the resulting call/flow blocking probabilities for both NS_1 and NS_2 are 10^{-2}. Therefore, the corresponding to the considered scenarios values of the offered traffic to NS_1 and NS_2 (calculated from Erlang formula), in the case when flow holding times are 1 ($1/\mu_1 = 1/\mu_2 = 1$), are:

Scenario 1: $\lambda_1 = 23.8$, $\lambda_2 = 29.5$.
Scenario 2: $\lambda_1 = 35.2$, $\lambda_2 = 15.3$.
Scenario 3: $\lambda_1 = 12.0$, $\lambda_2 = 43$.

The admission regions for scenario 1 are depicted in Fig. 4. The curve corresponding to typical AC shows that the maximum number of admitted flows from the NS_1 and NS_2

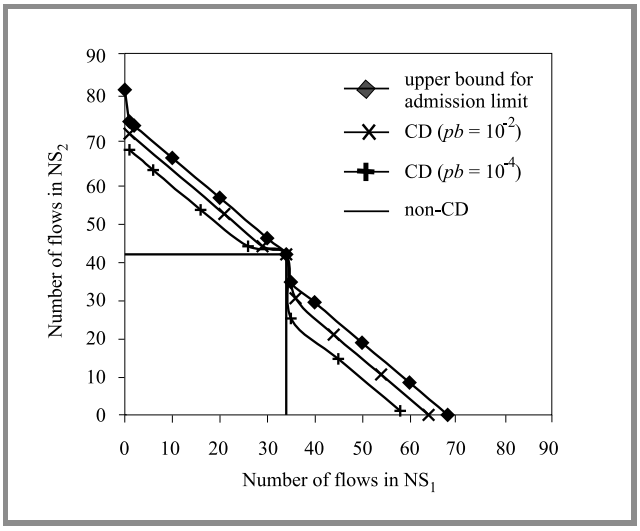


Fig. 4. Scenario 1: admission regions for the C-AC and typical AC for CBR traffic (CD – conditional admission; non-CD – typical admission, *pb* – level of risk).

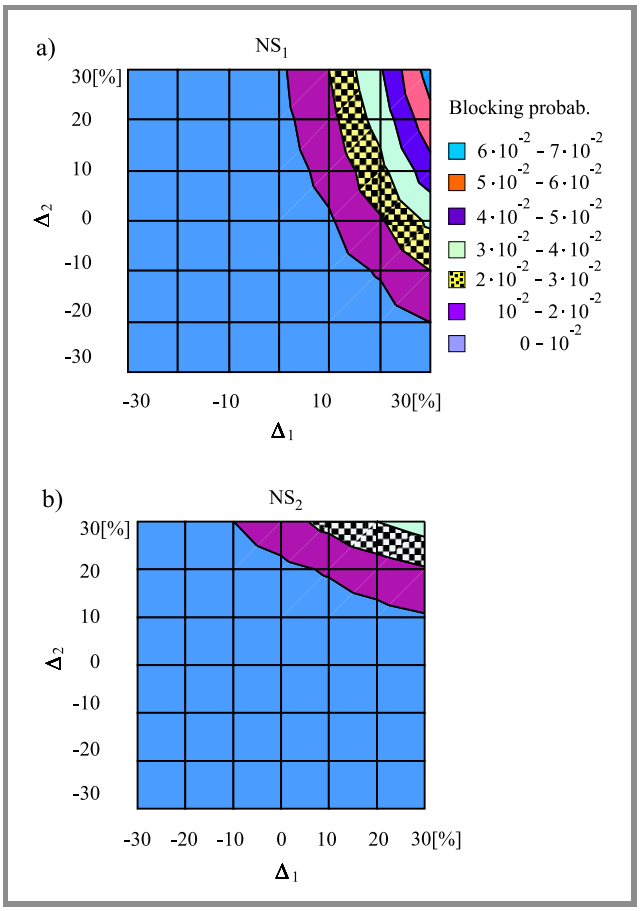


Fig. 5. Scenario 1: blocking probability in NS_1 (a) and NS_2 (b) versus the fluctuations of traffic offered to NS_1 and NS_2 for CBR traffic.

is 34 and 40, respectively (non-CD curve). Note that the upper bounds with applied the C-AC algorithm are in this case significantly greater. Now, the admission region is limited by 68 flows for NS_1 and 80 flows for NS_2 (!). Anyway, this result is too optimistic since this curve does not take into account any level of risk, expressed by the probability, say pb ($pb > 0$), that a conditionally admitted flow can be terminated before its proper finish time. The remaining two curves from Fig. 4 (CD curves) show the resulting admission regions in the cases when $pb = 10^{-2}$ and 10^{-4}. One can observe that the curve corresponding to $pb = 10^{-2}$ is very close to the upper bound curve. For the case of $pb = 10^{-4}$, we observe smaller admission region, but still significantly greater than this obtained by typical AC. Summarizing, we can conclude that the C-AC scheme can radically improve the admission region.

Figure 5 illustrates the effectiveness of the C-AC algorithm when the offered traffic to the NS_1 and NS_2 deviates from the assumed. Now, the call/flow arrival rates are $\lambda_1 = \lambda_1 + \Delta_1 \cdot \lambda_1$ and $\lambda_2 = \lambda_2 + \Delta_2 \cdot \lambda_2$. The factor

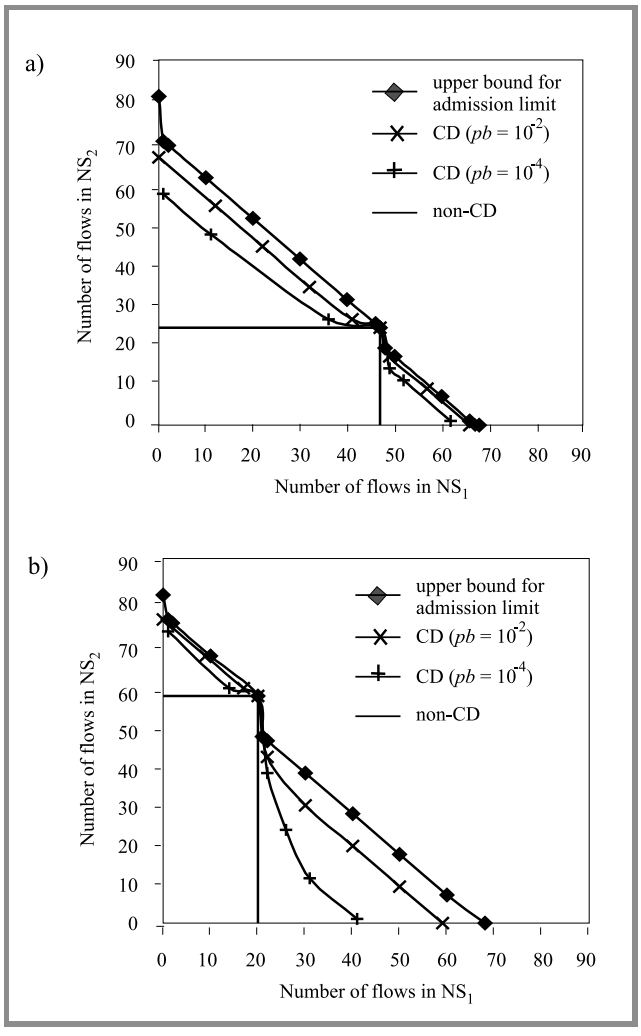


Fig. 6. Scenario 2 (a) and scenario 3 (b): admission regions for the C-AC and typical AC for CBR traffic (explanations – see Fig. 4).

Δ_1 (Δ_2) expresses a bias coefficient of the offered traffic to NS_1 (NS_2).

Let us recall that, when $\Delta_1 = 0\%$ and $\Delta_2 = 0\%$, the assumed blocking probabilities (with typical AC) for NS_1 and NS_2 are 10^{-2}. The curves from Fig. 5 show that, thanks to the C-AC algorithm, the call/flow blocking probabilities for NS_1 are still below 10^{-2} even when $\Delta_1 = 10\%$ and $\Delta_2 = 0\%$. As it was expected, more profit of using C-AC is observed when $\Delta_2 < 0$. For instance, when $\Delta_2 = -20\%$ and $\Delta_1 = 30\%$, call/flow blocking probability in NS_1 is still below 10^{-2}.

Figure 6 shows the obtained admission regions in scenario 2 and 3. One can observe that for scenario 2 the C-AC algorithm is more effective for flows submitted to NS_2 than NS_1. This is caused by the fact that in this case the capacity allocated for NS_2 is smaller than for NS_1. Therefore, one can expect that more NS_2 flows can be conditionally accepted within NS_1 than in reverse.

In scenario 3 more capacity is allocated to NS_2 than NS_1. The obtained upper bound for admission region in the case of NS_1 with the C-AC algorithm is now essentially larger comparing to the system with typical AC. However, the gained profit is now much less than expected. This is caused by the fact that NS_1 provides more rigorous QoS (at the packet level) than NS_2. When NS_1 flow is conditionally accepted to the capacity assigned for NS_2, the maximum link utilization in NS_2 must be decreased from 0.81 to 0.68.

Example 2. Co-operation of NSs designed for VBR and CBR traffic

Co-operating NSs from example 1 were both designed for CBR traffic, although with different target PLR. In this example we consider the case when NS_1 and NS_2 serve VBR and CBR traffic, respectively. Both NSs guarantee PLR value not greater than 10^{-4}. VBR flows are characterized by parameters of dual token bucket, e.g. peak bit rate (PBR) and sustainable bit rate (SBR). For non-conditionally accepted flows, we use the AC algorithm based on calculation of, so called, effective bandwidth, following Lindeberger formula [1, 5]. Therefore, new flow can be admitted only if the sum of effective bandwidths of all multiplexed flows is not greater than the capacity assigned for NS_1. In the NS_2 case, AC is performed as in the example 1. Maximum link utilization factor for NS_2 is $\rho = 0.68$. For conditionally accepted flows, the rules for AC differ for both NS_1 and NS_2 class. In the case when a flow originally submitted to NS_1 is rejected and as a consequence is re-submitted to NS_2, the admission control takes into account only PBR values of the flow in question as well as the flows being in progress and served by NS_2. The same rule is kept when flows are conditionally accepted in NS_1. The above algorithm seems to be a bit restrictive, since at least theoretically more flows could be admitted conditionally when effective bandwidth for VBR flows instead of PBR was taken into account.

In the considered example, CBR flows submitted to NS_2 request for 1 unit of link capacity, while VBR flows sub-

mitted to NS_1 request for a certain amount of effective bandwidth (EB), calculated assuming PBR= 2, SBR= 1. The total link capacity is 100. Three network scenarios are again considered, with the link partitioning as in example 1. Therefore, a single VBR flow requires 1.6, 1.45 and 1.9 units of EB, in scenario 1, 2 and 3, respectively.

Furthermore, we investigate the system assuming that the offered traffic (at the call/flow level) is such that the resulting call/flow blocking probabilities for both NS_1 and NS_2 are 10^{-2}. Therefore, the corresponding to the considered scenarios values of the offered traffic to NS_1 and NS_2 (calculated from Erlang formula), in the case when flow holding times are 1 ($1/\mu_1 = 1/\mu_2 = 1$), are:

Scenario 1: $\lambda_1 = 20, \quad \lambda_2 = 23.8.$
Scenario 2: $\lambda_1 = 31.6, \quad \lambda_2 = 12.$
Scenario 3: $\lambda_1 = 8.7, \quad \lambda_2 = 35.7.$

The admission regions for scenario 1 are depicted in Fig. 7. The curve corresponding to typical AC shows that the maximum number of admitted flows from the NS_1 and NS_2

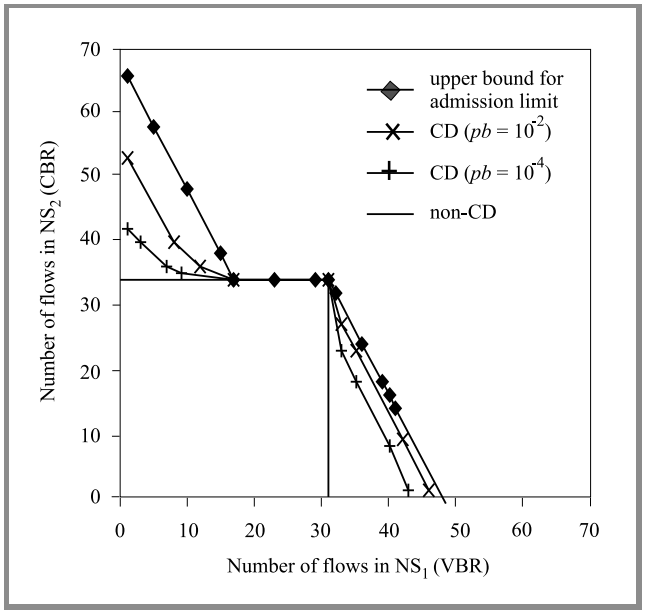


Fig. 7. Scenario 1: admission regions for the C-AC and typical AC for VBR and CBR traffic (explanations – see Fig. 4).

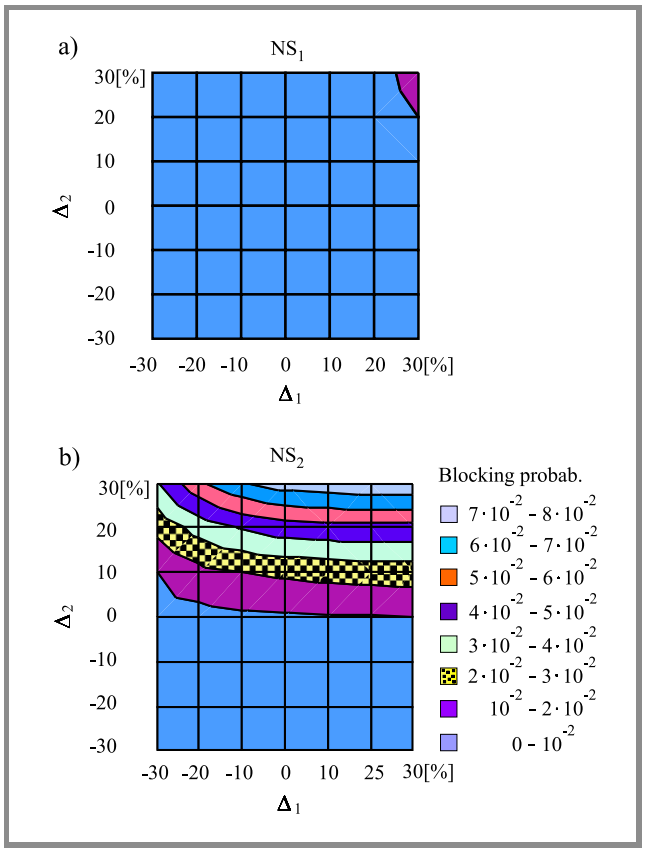


Fig. 8. Scenario 1: blocking probability in NS_1 (a) and NS_2 (b) versus the fluctuations of traffic offered to NS_1 and NS_2 for VBR and CBR traffic.

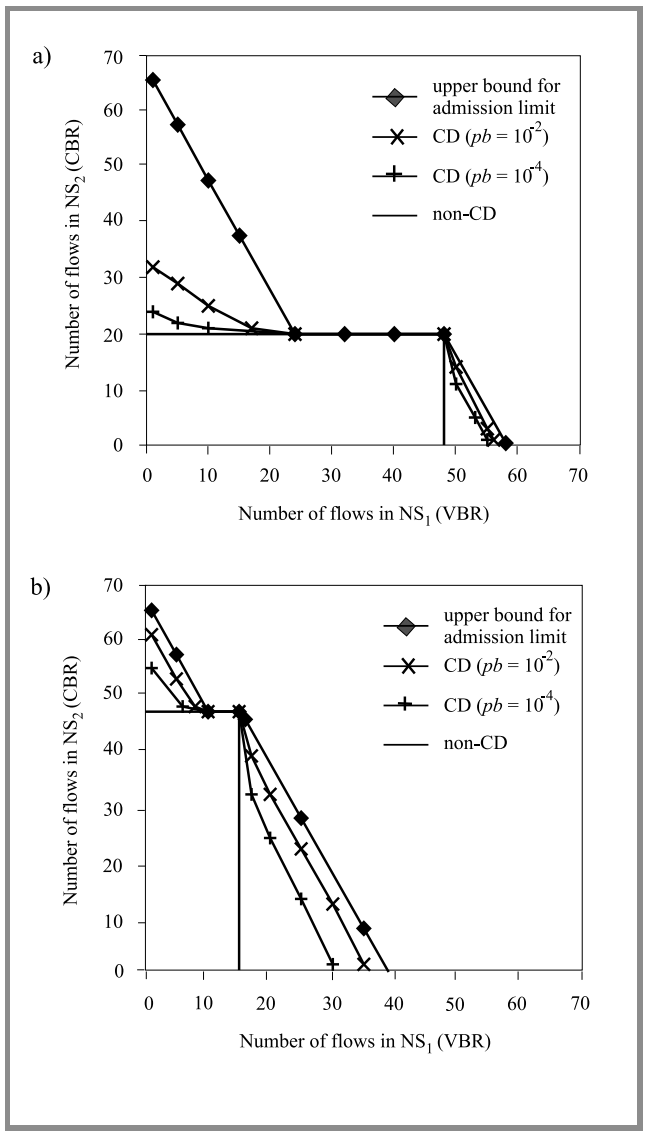


Fig. 9. Scenario 2 (a) and scenario 3 (b): admission regions for the C-AC and typical AC for VBR and CBR traffic (explanations – see Fig. 4).

is 31 and 34, respectively (non-CD curve). Note that, as in example 1, the upper bounds with applied the C-AC algorithm are in this case significantly greater. Now, the admission region is limited by 48 flows for NS_1 and 66 flows for NS_2. One can observe that, comparing to the scenario 1, the possibility of shifting resources between NSs is now limited. This is caused by the fact, that when CBR and VBR flows are mixed, the admission rules take into account only values of PBR of submitted flows, which does not allow for a multiplexing gain within VBR service. As a consequence, CBR flows can be conditionally admitted to the NS_1 only when it is under very low traffic conditions.

Figure 8 illustrates the effectiveness of the C-AC algorithm when the offered traffic to the NS_1 and NS_2 deviates from the assumed. The presented curves show that, thanks to the C-AC algorithm, the call/flow blocking probabilities for NS_1 are still below 10^{-2} even when $\Delta_1 = 25\%$ and $\Delta_2 = 25\%$. As it was expected, much less profit of using C-AC is observed in the case of NS_2. Now, Δ_2 can be increased only when $\Delta_1 = -30\%$ and below.

Admission regions obtained in the case of scenario 2 and scenario 3 are depicted in Fig. 9. One can observe, that the risk related with conditional admission of CBR flows within the capacity allocated for NS_1 is quite high in scenario 2. This is caused by the fact that in this case the conditional admission is allowed only when the current traffic load carried by NS_1 is very low.

5. Conclusions

The concept of conditional admission of new calls/flows was presented and discussed in the paper. The proposed approach assumes that new flows, which would normally be blocked, are conditionally accepted. This is possible by using spare at this moment capacity dedicated for other NSs. The preliminary numerical results confirm that the proposed approach is reasonable, leading to admitting larger number of flows and higher overall resource utilization, with a low probability that conditionally accepted flows will be terminated before the proper finish time. The admission regions for particular NSs can be, in some cases, radically extended. The proposed conditional admission can be especially attractive for the QoS IP networks where network resources are strictly partitioned between supported NSs.

References

[1] J. Roberts, U. Mocci, and J. Virtamo, Eds., "Broadband network teletraffic. Performance evaluation and design of broadband multiservice networks". Final Report of Action COST 242, 1996.

[2] S. Blake *et al.*, "An architecture for differentiated services", IETF RFC 2475 Document, Dec. 1998.

[3] Y. Bernet *et al.*, "A framework for differentiated services", Internet Draft, draft-ietf-diffserv framework-0.2.txt, Feb. 1999.

[4] R. Braden, D. Clark, and S. Shenker, "Integrated services in the Internet architecture: an overview", IETF RFC 1633 Document, June 1994.

[5] Deliverable D1301, "Specification of traffic handling for the first trial", AQUILA project consorcium, Sept. 2000, http://www-st.inf.tu-dresden.de/aquila/

[6] A. Bąk, W. Burakowski, F. Ricciato, S. Salsano, and H. Tarasiuk, "Traffic handling in AQUILA QoS IP network", to appear in Quality of Future Internet Services, Qofis2001, Portugal 2001.

[7] W. Burakowski, M. Fudała, and H. Tarasiuk, "Performance of premium service in QoS IP network", in *Proc. IEEE Workshop IP-oriented Oper. & Manag., IPOM '2000*, Cracow, Poland, 2000.

[8] J. Roberts, "Traffic theory and the Internet", *IEEE Commun. Mag.*, Jan. 2001.

Wojciech Burakowski, Andrzej Bęben – for biography, see this issue, p. 12.

Marek Dąbrowski – for biography, see this issue, p. 13.

Regular paper

Efficient procedure for capacitance matrix calculation of multilayer VLSI interconnects using quasi-static analysis and Fourier series approach

Hasan Ymeri, Bart Nauwelaers, and Karen Maex

Abstract — In this paper, we present a new approach for capacitance matrix calculation of lossy multilayer VLSI interconnects based on quasi-static analysis and Fourier projection technique. The formulation is independent from the position of the interconnect conductors and number of layers in the structure, and is especially adequate to model 2D and 3D layered structures with planar boundaries. Thanks to the quasi-static algorithms considered for the capacitance analysis and the expansions in terms of convergent Fourier series the tool is reliable and very efficient; results can be obtained with relatively little programming effort. The validity of the technique is verified by comparing its results with on-surface MEI method, moment method for total charges in the structure, and CAD-oriented equivalent-circuit methodology, respectively.

Keywords — lossy IC interconnect, Fourier projection method, line capacitance.

1. Introduction

Calculation of the capacitance matrix in multilayer IC interconnects is a well-known problem that can be solved by many analytical and numerical techniques [1–9]. Often these procedures were based on the integral equation formulation, differential equation formulation, or have been the results of extensive numerical simulations using adequate empirical corrections.

This letter proposes a new and more general formulation for computation of capacitance matrix of the most common 2D interconnect structures using quasi-static analysis and Fourier projection approach.

2. Background of the method

In the formulation, 2D L-layered interconnect structures with planar boundaries are considered. Each layer is linear, homogeneous, and isotropic, and has permittivity $\varepsilon^{(l)}$ and conductivity $\sigma^{(l)}$, where $l = 1, \ldots, L$. For lossy medium the complex permittivity is $\underline{\varepsilon}^{(l)} = \varepsilon^{(l)} - j\sigma/\omega$. The point charge source is located along $y = 0$, $x = x_s$ and $z = z_s$, respectively (see Fig. 1).

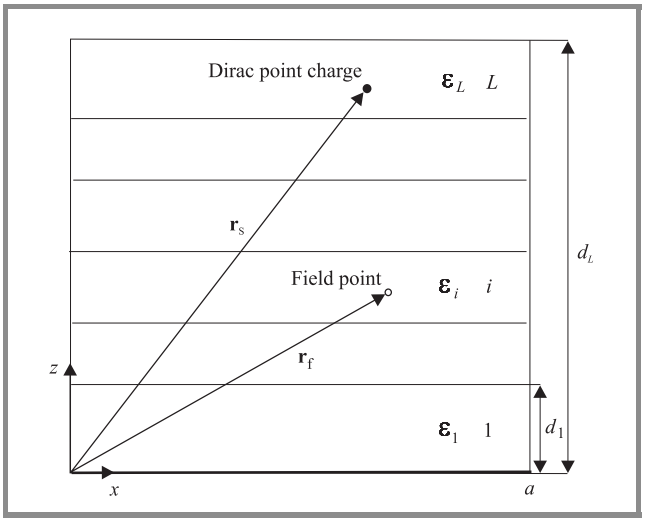


Fig. 1. Geometry of a layered structure for multilayer Green's function determination.

Inside each layer l and excluding the source point layer, the potential function $\varphi^{(l)}$ satisfies

$$\nabla^2 \varphi^{(l)} = 0 \tag{1}$$

and the induction vector $\mathbf{D}^{(l)}$ is obtained from

$$\mathbf{D}^{(l)} = -\underline{\varepsilon}^{(l)} \nabla \varphi^{(l)} . \tag{2}$$

Here, the problem is solved by developing each potential $\varphi^{(l)}$ as a Fourier series. In the source layer, the general solution to Eq. (1) can be written as

$$\varphi^{(s)}(x_f, y_f, z_f) = \varphi_P(x_f, y_f, z_f) + \varphi_H(x_f, y_f, z_f) , \tag{3}$$

where φ_P is the source term given by

$$\varphi_P(x_f, y_f, z_f) = \frac{Q}{4\pi \underline{\varepsilon}^{(s)}} \left\{ \frac{1}{\left[(x_f - x_s)^2 + y_f^2 + (z_f - z_s)^2 \right]^{\frac{1}{2}}} \right\} \tag{4}$$

and

$$\varphi_H(x_f, y_f, z_f) = \sum_{n,m\geq 0} \left[C_{nm}^{(s)} \exp(K_{nm} z_f) + \right.$$
$$\left. + D_{nm}^{(s)} \exp(-K_{nm} z_f)\right] \cos(k_n x_f) \cos(k_m y_f)\,, \qquad (5)$$

where $k_n = n\pi/a$, $k_m = m\pi/b$, $K_{nm} = (k_n^2 + k_m^2)^{\frac{1}{2}}$, (x_f, y_f, z_f) are field point coordinates, and a and b are dimensions of the structures in x and y direction.
Considering Eq. (2), $\mathbf{D}^{(s)}$ is given by

$$\mathbf{D}^{(s)}(x_f, y_f, z_f) = \mathbf{D}_P(x_f, y_f, z_f) + \mathbf{D}_H(x_f, y_f, z_f) \qquad (6)$$

with

$$\mathbf{D}_P(x_f, y_f, z_f) = \frac{Q}{4\pi} \left\{ \frac{(x_f - x_s)\mathbf{1}_x + y_f \mathbf{1}_y + (z_f - z_s)\mathbf{1}_z}{\left[(x_f - x_s)^2 + y_f^2 + (z_f - z_s)^2\right]^{\frac{3}{2}}} \right\} \qquad (7)$$

and

$$\mathbf{D}_H(x_f, y_f, z_f) = \underline{\varepsilon}^{(s)} \left\{ \sum_{n,m\geq 0} k_n \left[C_{nm}^{(s)} \exp(K_{nm} z_f) + \right.\right.$$
$$\left.\left. + D_{nm}^{(s)} \exp(-K_{nm} z_f)\right] \sin(k_n x_f) \cos(k_m y_f) \right\} \mathbf{1_x} +$$
$$+ \underline{\varepsilon}^{(s)} \left\{ \sum_{n,m\geq 0} k_m \left[C_{nm}^{(s)} \exp(K_{nm} z_f) + \right.\right.$$
$$\left.\left. + D_{nm}^{(s)} \exp(-K_{nm} z_f)\right] \cos(k_n x_f) \sin(k_m y_f) \right\} \mathbf{1}_y +$$
$$- \underline{\varepsilon}^{(s)} \left\{ \sum_{n,m\geq 0} K_{nm} \left[C_{nm}^{(s)} \exp(K_{nm} z_f) + \right.\right.$$
$$\left.\left. - D_{nm}^{(s)} \exp(-K_{nm} z_f)\right] \cos(k_n x_f) \cos(k_m y_f) \right\} \mathbf{1}_z\,. \qquad (8)$$

In the other layers, the solutions are

$$\varphi^{(l)}(x_f, y_f, z_f) = \sum_{n,m\geq 0} \left[C_{nm}^{(l)} \exp(K_{nm} z_f) + \right.$$
$$\left. + D_{nm}^{(l)} \exp(-K_{nm} z_f)\right] \cos(k_n x_f) \cos(k_m y_f) \qquad (9)$$

and

$$\mathbf{D}^{(l)}(x_f, y_f, z_f) = \underline{\varepsilon}^{(l)} \left\{ \sum_{n,m\geq 0} k_n \left[C_{nm}^{(l)} \exp(K_{nm} z_f) + \right.\right.$$
$$\left.\left. + D_{nm}^{(l)} \exp(-K_{nm} z_f)\right] \sin(k_n x_f) \cos(k_m y_f) \right\} \mathbf{1_x} +$$
$$+ \underline{\varepsilon}^{(l)} \left\{ \sum_{n,m\geq 0} k_m \left[C_{nm}^{(l)} \exp(K_{nm} z_f) + \right.\right.$$
$$\left.\left. + D_{nm}^{(l)} \exp(-K_{nm} z_f)\right] \cos(k_n x_f) \sin(k_m y_f) \right\} \mathbf{1}_y +$$
$$- \underline{\varepsilon}^{(l)} \left\{ \sum_{n,m\geq 0} K_{nm} \left[C_{nm}^{(l)} \exp(K_{nm} z_f) + \right.\right.$$
$$\left.\left. - D_{nm}^{(l)} \exp(-K_{nm} z_f)\right] \cos(k_n x_f) \cos(k_m y_f) \right\} \mathbf{1}_z\,. \qquad (10)$$

The potential function distribution $\varphi^{(l)}$ and the normal component of electric induction vector $\mathbf{D}^{(l)}$ are expressed by series expansions in terms of solutions of the Laplace Eq. (1). One such expansion is written down for each homogeneous region of the layered structure in Fig. 1. The expansion coefficients $C_{nm}^{(l)}$ and $D_{nm}^{(l)}$ of the different series are related to each other and to the charge density distribution on the interconnect conductors via boundary conditions. Then, coefficients $C_{nm}^{(l)}$ and $D_{nm}^{(l)}$ are determined recursively. In this way we have found the multilayer Green's function $G(\mathbf{r}_f; \mathbf{r}_s)$ of the problem. By deriving the Green's function over a multilayer dielectric region and allowing evaluation of potential distribution in any layer, we can place interconnect conductors anywhere in the multilayer structure, and therefore solve for the capacitance per unit length matrix for an arbitrary number of conductors.

3. Capacitance matrix calculation

In the following the complex capacitance calculation procedure will be treated in more detail. In an equivalent circuit, the value of a capacitance is the ratio of the free charge associated with a voltage difference between two interconnect conductors or between a interconnect conductor and the reference (e.g. the ground plane or the point at infinity), and that voltage difference. The values of these capacitances are known as network capacitances.
According to the equivalent source principle for the electromagnetic field, we can replace the rectangular conductor (c) (see Fig. 2a) with a piece of surface charge density distribution $\sigma_c(\mathbf{r}_s)$ around the surface S_c, as shown in Fig. 2b. Using a Green's function of the medium $G(\mathbf{r}_p; \mathbf{r}_s)$ that incorporates all boundary conditions in the structure in Fig. 2b (see Sec. 2), the voltage at any point $\mathbf{r}_p$ is generated by the charge density $\sigma_c(\mathbf{r}_s)$ on all conductors ($c = 1, \ldots, N$)

$$V(\mathbf{r}_p) = \sum_{c=1}^{N} \oint_{(c)} \sigma_c(\mathbf{r}_s) G(\mathbf{r}_p; \mathbf{r}_s) dS_c\,. \qquad (11)$$

Element C_{cj} of the capacitance matrix [C] may be calculated as the charge Q_c per unit length on conductor (c) when the voltage on conductor (j) is 1 and 0 V on all other conductors. The charge per unit length on conductor (c) is the integral of the surface charge density $\sigma_c(\mathbf{r}_s)$ over the circumference of conductor (c): $Q_c = \oint_{(c)} \sigma_c(\mathbf{r}_s) dS_c$. The charge distribution on every conductor (c) may be approximated by a number N_b of well-chosen basis functions $\sigma_{c,r=1,\ldots,N_b}(\mathbf{r}_s)$ along the contour of the conductor: $\sigma_c(\mathbf{r}_s = \sum_{r=1}^{N_b} W_{c,r} \sigma_{c,r}(\mathbf{r}_s)$. The problem has been reduced to the computation of the discrete charge constants $\{W_{c=1\ldots N, r=1\ldots N_b}\}$. As the result we obtain a series of simultaneous equations and represent them as follows:

$$\sum_{c=1}^{N} \sum_{r=1}^{N_b} W_{c,r} p_{j|c,r|t} = V_{j=1\ldots N}\,, \qquad (12)$$

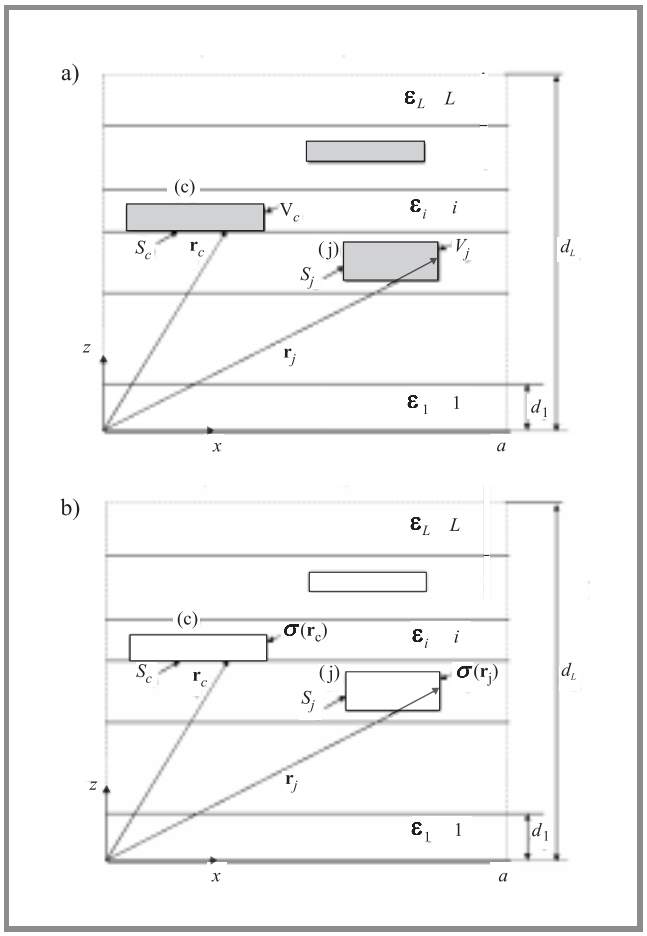


Fig. 2. Geometry of a layered structure with (a) embedded conductors, and (b) charge density distribution on the discretized surface of the conductors.

where $V_{j=1...N}$ is the voltage on any conductor (j), with

$$p_{j|c,r|t} = \frac{\oint_{(j)} \oint_{(c)} \sigma_{j,t}(\mathbf{r}_j) G(\mathbf{r}_j;\mathbf{r}_c) \sigma_{c,r}(\mathbf{r}_c) dS_c\, dS_j}{\oint_{(j)} \sigma_{j,t}(\mathbf{r}_j) dS_j} \qquad (13)$$

as potential coefficients of the Galerkin matrix. Solving the matrix Eq. (12) on a computer, we can determine the constants $\{W_{c,r}\}$ and then the capacitance per unit length C_{cj} can be obtained in the form:

$$C_{cj} = Q_c(V_j = 1,\ V_{c\neq j} = 0) = \sum_{r=1}^{N_b} W_{c,r} \oint_{(c)} \sigma_{c,r}(\mathbf{r}_c) dS_c. \qquad (14)$$

The lossy semiconducting substrate is taken into account by the complex permittivity

$$\varepsilon_{cs} = \varepsilon_s - j\frac{\sigma}{\omega}, \qquad (15)$$

where ε_s is the permittivity and σ conductivity of the semiconducting substrate (silicon).

Due to the quasi-TEM character of the electromagnetic fields in the examined structure the frequency dependent distributed admittance per unit length Y can be calculated as

$$Y = G + j\omega C = j\omega \frac{Q}{\Delta V}, \qquad (16)$$

where Q is the total charge per unit length, ΔV denote the voltage difference between the conductors, G is the conductance per unit length (losses) and C is the capacitance per unit length.

4. Discussion of the results

In this section we apply the new procedure to calculate some examples. In these examples we use multilayer IC interconnects whose strip conductors are infinitely thin (zero-thickness) or of rectangular cross-section and very thick (as usually in on-chip interconnets).

Example 1. Let us consider the system of four strip conductors embedded in a two-layered dielectric region with structure as shown in Fig. 3, where the conductors are numbered from left to right and upper to lower as 1, 2, 3 and 4, respectively. Numerical values for the ca-

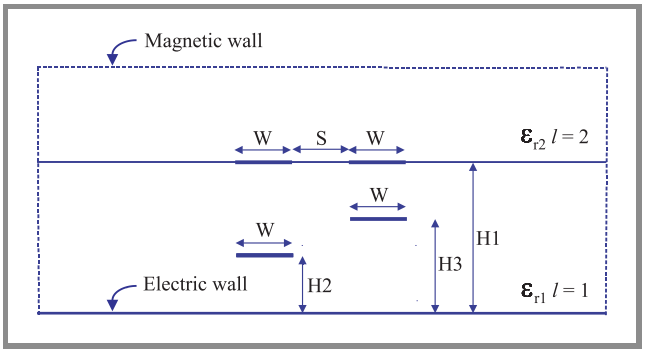


Fig. 3. Geometry of the structure from example with four strips (W/H1 = S/H1 = H2/H1 = 1/3, H3/H1 = 2/3, $\varepsilon_{r1} = 5$ and $\varepsilon_{r2} = 1$).

Table 1
Capacitance matrix of the structure of Fig. 3

Capacitance [pF/m]	MoM [5, 7]	MEI [1]	This letter
C_{11}	70.158	69.514	70.158
C_{12}	–12.842	–12.832	–12.839
C_{13}	–12.960	–13.110	–12.967
C_{14}	–22.240	–23.014	–22.230
C_{22}	87.327	87.028	87.227
C_{23}	–54.195	–55.462	–54.234
C_{24}	–4.052	–3.988	–4.049
C_{33}	133.935	128.86	128.50
C_{34}	–14.16	–14.93	–14.21
C_{44}	135.70	141.31	135.94

pacitance matrix elements, generated by the proposed approach (the moment method has been used) and by the on-surface MEI procedure [1] and the moment method with total charge in structure [5, 8], respectively, are given in Table 1. Note that the discrepancies between the values generated by our approach and one by [5, 8] are practically smaller than 0.2% over a wide range of physical dimensions and material parameters (all treated cases are not reported in this letter).

Example 2. In order to prove the validity of the given approach self and mutual shunt admittance per unit length (capacitance and conductance per unit length) calculated using our procedure are compared with the results of the full-wave analysis (spectral domain approach) in conjunction with equivalent circuit modeling technique [9]. In Fig. 4, an asymmetric coupled interconnect structure is depicted with the following electrical and geometrical parameters:

$$t_{\mathrm{Si}} = 500\ \mu\mathrm{m},\ t_{\mathrm{ox}} = 2\ \mu\mathrm{m},\ w_1 = 4\ \mu\mathrm{m},\ w_2 = 1\ \mu\mathrm{m},$$
$$T_1 = T_2 = 1\ \mu\mathrm{m},\ \varepsilon_{\mathrm{si}} = 11.8,$$
$$\rho_{\mathrm{Si}} = 0.01\ \Omega\mathrm{cm},\ \varepsilon_{\mathrm{ox}} = 3.9 \ \text{ and } \ \mathrm{s} = 4\ \mu\mathrm{m}.$$

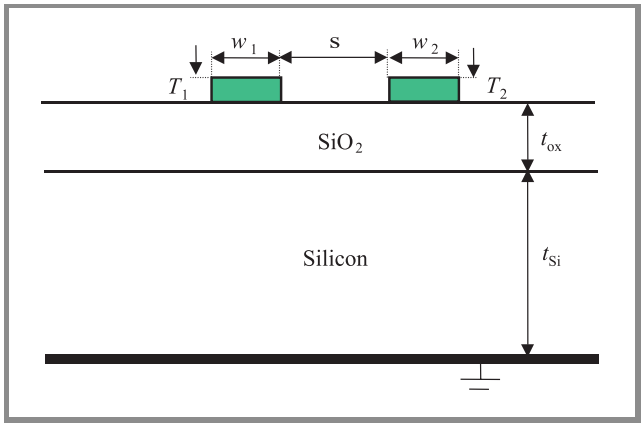


Fig. 4. Asymmetric coupled interconnects on lossy silicon substrate.

Figure 5a shows the variation in the distributed self and mutual capacitance per unit length $C_{11}(\omega)$, $C_{22}(\omega)$, and $C_{12}(\omega)$, as a function of frequency. Similarly, Fig. 5b shows the variation of the distributed self and mutual conductance per unit length $G_{11}(\omega)$, $G_{12}(\omega)$, and $G_{22}(\omega)$ as a function of frequency. The solid lines are computed using the new multilayer Green's function procedure and the dashed lines are the results from the equivalent-circuit model approach [9]. It is observed that the values of the self and mutual capacitance and conductance per unit length, respectively, are in good agreement with those of [9]. As expected, the lossy silicon semiconducting substrate has significant impact on the frequency-dependence of the capacitance and conductance per unit length as compared to the lossless or low loss dielectric substrate.

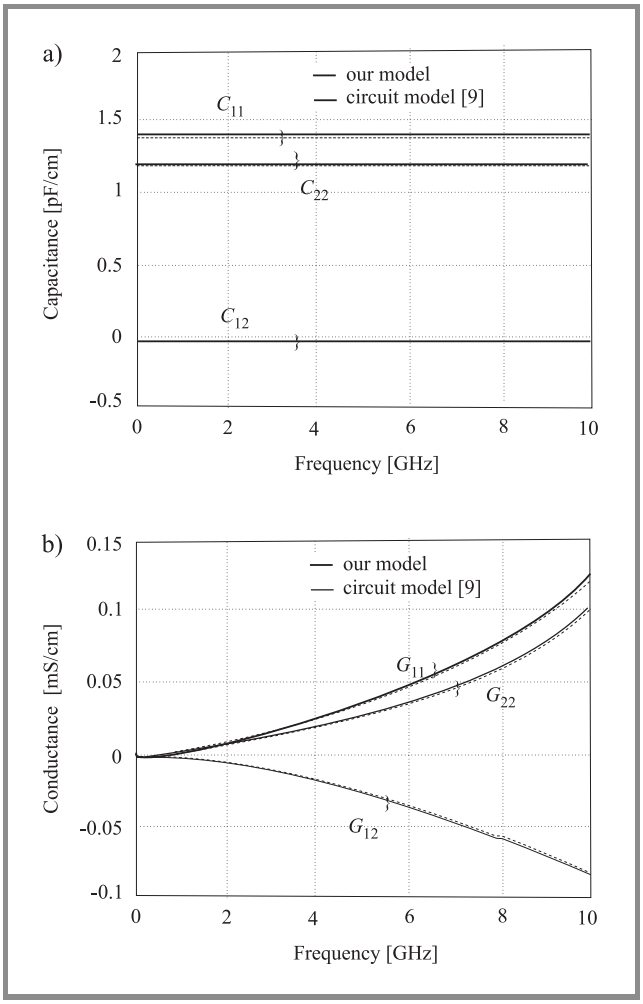


Fig. 5. (a) Self and mutual capacitance per unit length of asymmetric coupled interconnects on lossy silicon substrate; (b) self and mutual conductance per unit length of asymmetric coupled interconnects on lossy silicon substrate.

5. Conclusion

In this paper, we have discussed a technique for capacitance matrix extraction over a multilayer Si substrate. We derived the appropriate Green's function using quasi-static analysis and Fourier projection method. The potential function and electric induction vector components are defined as series expansions in terms of the Laplace equation which are periodic in the direction parallel to the plane of interconnect conductors. The proposed semi-analytical procedure allows us: first, to assess in an analytical and simple way the integral equations of the problem, and second, to obtain a fast convergence of the numerical results due to the averaging technique used in the Galerkin approach which leads to better accuracy in the numerical calculations. This method results in a very simple formulation of the problem that is well suited for computer solutions with relatively little programming effort.

References

[1] W. Y. Liu, K. Lan, and K. K. Mei, "Computation of capacitance matrix for integrated circuit interconnects using on-surface MEI method", *IEEE Trans. Microw. Guid. Wave Lett.*, vol. 9, pp. 303–304, 1999.

[2] Z. Zhu, W. Hong, Y. Chen, and Y. Wang, "Electromagnetic modeling and transient simulation of interconnects in high speed VLSI circuits", *IEEE Proc. Microw. Anten. Propag.*, vol. 143, pp. 373–378, 1996.

[3] E. Groteluschen, L. S. Dutta, and S. Zaage, "Full-wave analysis and analytical formulas for the line parameters of transmission lines on semiconductor substrates", *INTEGRATION, VLSI J.*, vol. 16, pp. 33–58, 1993.

[4] F. Stellari and A. L. Lacaita, "New formulas of interconnect capacitances based on results of conformal mapping method", *IEEE Trans. Electron. Dev.*, vol. ED-47, pp. 222–231, 2000.

[5] C. Wei, R. F. Harrington, J. R. Mautz, and T. K. Sarkar, "Multiconductor transmission lines in multilayered dielectric media", *IEEE Trans. Microw. Theory Tech.*, vol. MTT-32, pp. 439–449, 1984.

[6] T. Sakurai, "Simple formulas for two- and three-dimensional capacitances", *IEEE Trans. Electron. Dev.*, vol. ED-40, pp. 118–124, 1993.

[7] C. P. Yan and T. N. Trick, "A simple formula for the estimation of the capacitance of two-dimensional interconnects in VLSI circuits", *IEEE Trans. Electron. Dev. Lett.*, vol. EDL-3, pp. 391–393, 1982.

[8] A. R. Djordjevic, M. B. Bazdar, T. K. Sarkar, and R. F. Harrington, *LINPAR: Matrix Parameters for Multiconductor Transmission Lines*. New York: Artech House, 1999.

[9] J. Zheng, Y.-C. Hahm, A. Weisshaar, and V. K. Tripathi, "CAD-oriented equivalent circuit modeling of on-chip interconnects for RF integrated circuits in CMOS technology", in *Proc. Int. Microw. Symp.*, Anaheim, USA, June 1999, pp. 35–38.

Hasan Ymeri was born in Druar near Mitrovicë, Kosovë on 24 October 1957. He received the Dipl. Ing. (M.Sc.) degree in electronic engineering from University of Prishtina, Prishtinë, Kosovë, in 1983, the M.Sc. degree in electrical engineering from the University of Ljubljana, Slovenia, in 1988, and the D.Sc. degree in electronic engineering from the Polytechnic University, Tirana, Albania, in 1996, respectively. Since 1985 he has been with the University of Prishtina, first as Lecturer and later as Senior Lecturer, where has been involved in research on electromagnetic theory, microwave integrated circuits and digital electronics. At present he is with the Katholieke Universiteit of Leuven, Belgium, as a Researcher in the field of silicon IC interconnects.
e-mail: Hasan.Ymeri@esat.kuleuven.ac.be
Department of Electrical Engineering (ESAT)
Division ESAT-TELEMIC
Katholieke Universiteit Leuven
Kasteelpark Arenberg 10
B-3001 Leuven-Heverlee, Belgium

Bart Nauwelaers was born in Niel, Belgium on 7 July 1958. He received the M.Sc. and Ph.D. degrees in electrical engineering from the Katholieke Universiteit Leuven, Leuven, Belgium in 1981 and 1988, respectively. He also holds a Mastere degree from ENST, Paris, France. Since 1981 he has been with the Department of Electrical Engineering (ESAT) of the K.U. Leuven, where he has been involved in research on microwave antennas, microwave integrated circuits, MMICs, and wireless communications. He teaches courses on microwave engineering, on analog and digital communications, on wireless communications and on design in electronics and in telecommunications.
e-mail: Bart.Nauwelaers@esat.kulueven.ac.be
Department of Electrical Engineering (ESAT)
Division ESAT-TELEMIC
Katholieke Universiteit Leuven
Kasteelpark Arenberg 10
B-3001 Leuven-Heverlee, Belgium

Karen Maex received the M.Sc. degree in electrical engineering in 1982 and the Ph.D. degree in 1987 both from the Katholieke Universiteit Leuven, Leuven, Belgium. From 1982 until 1987 she has been a Research Assistant of the Belgian National Fund for Scientific Research (FWO). At present she is continuing her research at the Interuniversity Microelectrics Center (IMEC) as a Research Director of the Fund for Scientific Research Flanders. She is a Professor at the E.E. Department of the Katholieke Universiteit Leuven. At Imec she is Director of the Interconnect Technologies and Silicides (ITS) department within the Silicon Process Technology Division. Her interests are in materials science and technology for deep-sub-micron semiconductor devices.
e-mail: maex@imec.be
The Interuniversity
Microelectronics Center (IMEC)
Kapeldreef 75
B-3001 Leuven, Belgium

Regular paper

On a method to authenticate and verify digital streams

Beata J. Wysocki, Yejing Wang, and Reihaneh Safavi-Naini

Abstract — Recently, digital streams have become widely used to make audio, video, and other media available in real-time over the Internet. As with other transmission methods, the recipient needs to have a possibility to verify the source and authenticity of the received information. Several techniques have been proposed to deal with this issue. Most of them are vulnerable to packet losses or they introduce unacceptable computational and/or communication overheads. Some of the graph-based techniques provide immunity to burst losses of certain length. However, these techniques are not immune to the loss of packets containing signatures or occasional burst of lengths greater than the assumed one. In the paper, we propose a modification to one of the graph-based techniques that introduces immunity to the loss of packets containing signatures, without introducing any additional overheads.

Keywords — authentication algorithms, hash chains, signing digital streams, Markov modelling, Gilbert-Elliot channel.

1. Introduction

Streamed data or a stream of data packets is generated by a specialized application on a server machine, and then send in a form of autonomous packets over the Internet (or other packet switched network). The process of splitting a large block of information into the autonomous packets by a server application needs to be distinguished from normal splitting of a large file into packets before sending them through the packet switched network. In the later case, a file to be transmitted is divided into packets by a transport protocol (e.g. TCP [1]). The packets are then transmitted to the receiving host. Because they are numbered, they can be reassembled at the receiver in the desired order, and in case of packets missing or erroneous ones, they can be retransmitted. Only after correctly reassembling the packets (some integrity is checked through the error control coding), the full file is passed by the transport protocol to the application layer of the receiving host. The application can then execute this file (or perform other operation). This is illustrated in Fig. 1 where it is shown that for the application layers in normal data transmission over the Internet, the transmitted file is one contiguous block, not much different from a file read from disk.

In case of streaming, a file is divided into packets at the application layer. These packets are of the size, which would fit into the size of a packet in the underlying packet network. In the case of the Internet, it is usually not more than 1000–1500 bytes [2]. Every packet of a stream is than treated like a full file by the transport layer, and transmitted to the receiving host independently. As soon as such a packet arrives at the transport layer of the receiving host it is passed to the application layer, as there is no need to reassemble anything. For the transport protocol, this packet

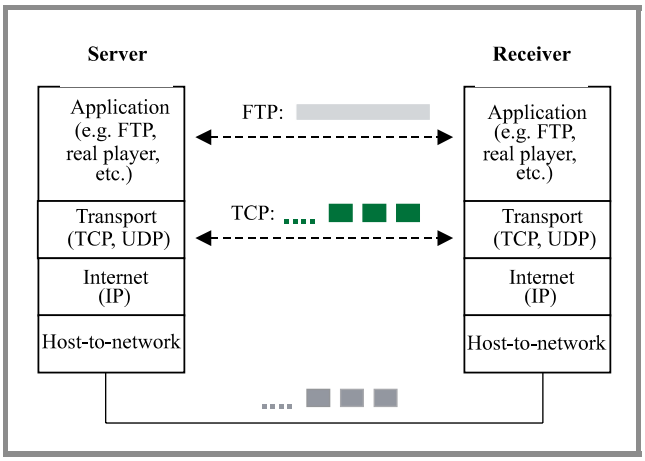


Fig. 1. Transmission of non-streamed data during a file transfer.

constitutes the whole file. The client application can than execute or play the received packet as it comes. It is the application task, to discard packets coming out of order. This is illustrated in Fig. 2.

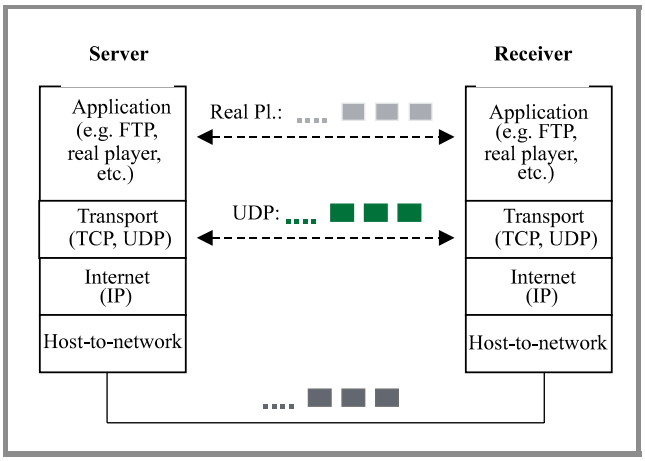


Fig. 2. Transmission of streamed data during a real player session.

As the streaming involves mainly real time applications [3] (e.g. real player) the UDP [1] transport is usually used, which means that there are no retransmissions or integrity checks at the transport layer. This must be catered for by the application. The advantage of UDP over TCP is its speed and lower communication overhead but there is no guarantee for packet delivery or for its correctness.

Another feature of streams, which distinguish them from normal messages, is that the receiver utilizes data and receives at more or less the input rate. Hence, it cannot buffer large amount of unused data [4].

The importance of the need to authenticate streamed data manifests itself in a fact that a recipient would like to have a possibility to verify the source and authenticity of the received information. For example, another feature of streams, which distinguish them from normal messages is important to the listeners of an Internet radio station that the audio stream they receive was really broadcasted by the station they listen to. On the other hand, it is equally important to that station that only the content it broadcasts is attributed to it. Malicious parties should be prevented from injecting commercials or offensive material into the stream.

Moreover, those guarantees must be non-repudiable, which means that a positive verification has to be enough to hold the transmitter responsible for the contents. An answer to those problems is the use of digital signatures. However, the digital signature technology has been developed for single messages, not for a continuous stream of autonomous packets [5–7].

Authentication of a single message or a file transferred from a server to a remote client can be done using one of the standard schemes, like digital signature. The signature or other authentication features are generated by the application and appended in some form to the file. The file is then passed to the transport layer, where it is packetized. The part that contains the signature is not treated in anyway different to the other contents of the message (file). At the remote client machine, the transport layer reassembles the received packets, checks the message (file) integrity and passes the whole message (file) to the application layer, where verification of the message (file) is performed.

For the streamed data, this approach is not possible to use. First of all, the application layer at the server generates autonomous packets itself, and the transport layer considers those packets as separate messages. At the remote client the packets are used as they arrive. This means that they are not assembled back at the application layer. In the optimal solution, they are even not buffered. Hence, it is not possible to have the approach used to authenticate and verify messages (file) used for non-streamed data directly applied to the streamed data. The nature of a digital stream forces the need to authenticate and verify each of the received packets. As a result, each packet must carry some features verifiable by a remote client.

Several different authentication/verification schemes have been proposed in literature, e.g. [5–7, 9, 10, 14]. Most of them can be classified into two major groups:

- The schemes where every packet carries the full information necessary to verify the packet [6, 7], which in some sense resembles signing of every packet individually.
- The schemes where only one packet in the stream (usually the first or last) is signed and the other packets are connected with the signature by a chain of hashes [5, 9, 14].

There are, of course, several modifications to the second group, e.g. [14]. However, all of them suffer from the fact that once the verification chain is broken, there is no means to verify packets incoming after the break. The way to avoid this drawback, and somehow combine the benefits of both groups has been proposed by Golle and Modadugu in [10].

They proposed to divide the stream into sequences of N packets, with each of the sequence being individually verifiable. Hence, a break in the verification chain would result in rejection of a maximum of one sequence of packets but not all packets coming after the break. The technique proposed in [10] is resistant to bursty losses with bursts of up to a predefined length. However, even an isolated single packet loss, when the packet containing a signature is involved, results in a loss of the whole sequence of packets. In this paper, we propose a simple but efficient way to mitigate this problem.

The paper is organized as follows. In Section 2, we briefly introduce some of the techniques proposed to authenticate digital streams, concentrating on the method proposed in [10]. Section 3 describes the Gilbert-Elliott channel model [15, 16] and shows the level of vulnerability of Golle and Modadugu method to the loss of signed packets. In Section 4, we introduce our modification and calculate the probability that the modified authentication chain can be broken for a given bursty channel, while Section 5 concludes the paper.

2. Some of the proposed schemes to authenticate digital streams

2.1. Block signatures

Signing every packet of the data stream can be considered as a direct extension of the message signing technique into authentication of digital streams. This approach, however, involves unacceptable overheads, both a computational and a communication one. Some reduction in those overheads without compromising the benefits of having every packet independently verifiable has been proposed in [6].

One of the stream signing techniques proposed by Wong and Lam in [6] is a star chaining technique. During the authentication phase, a block of "m" packets is formed and hash functions are calculated for every packet in a block, and then for a sequence of these individual hashes a block hash is calculated. After that, the block hash is signed using a standard digital signature algorithm. For packets to be individually verifiable, each packet needs the full authentication information, called a packet signature. The packet signature consists of the block signature, the packet number in the block, and the hashes of all other packets in the block. The authentication process and the transmission of the resulting packet block are illustrated in Fig. 3.

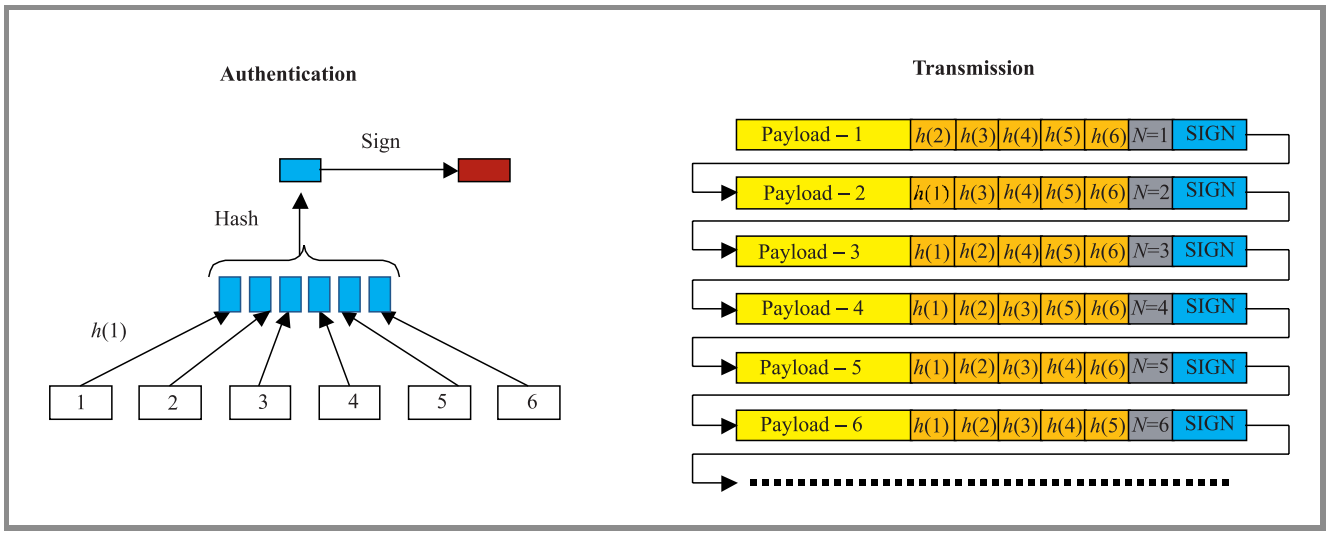


Fig. 3. Authentication process in star chaining scheme [13], and transmission of the resulting packet block.

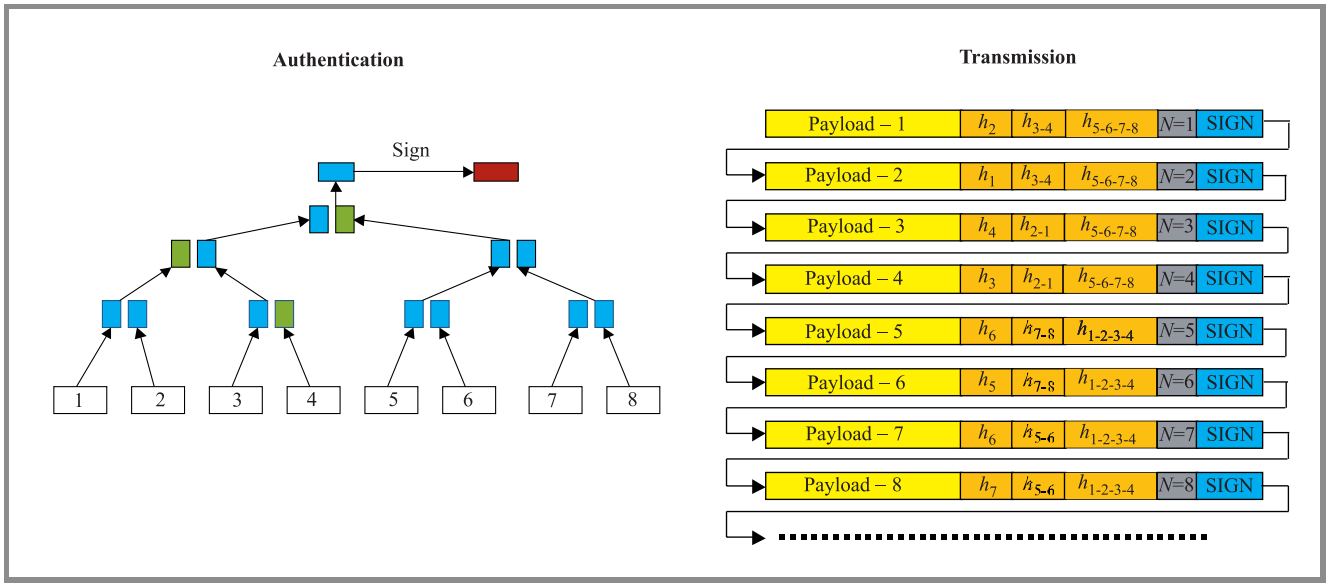


Fig. 4. Tree chain authentication process [13], and transmission of the resulting packet block.

The verification procedure is as follows:

- For the first packet received from the block, the verifier calculates the packet's hash.
- Based on the calculated packet hash, the packet number, and the hashes of other packets in the block contained in the packet signature, the verifier computes the block hash.
- Using the calculated block hash, receiver verifies the block signature, if it is correct, the packet is accepted, if not it is rejected.

For all other packets from the block, the verifier needs only to calculate the new packet hash and compare it to the hash contained in the packet from this block previously positively verified. If they agree, the new packet is verified.

Another scheme proposed by Wong and Lam in [6], is a generalization of their star chaining technique. In this generalized scheme, the block hash is computed as a root node of an authentication tree (see Fig. 4). In such a tree, the packets' hashes are the leaf nodes of the second-degree authentication tree, with other nodes of the tree computed as hashes of their children. For example, in Fig. 4, the parent of leaves D_3 and D_4 is $D_{3-4} = h(D_3, D_4)$. The block signature is calculated on the block hash.

In tree chaining, a packet signature (packet overhead) consists of:

- The block signature.
- The packet number in a block.
- Siblings of each node in the packet's path to the root.

From Fig. 4, it is visible that the communication overhead can be reduced compared to the star chaining scheme. However, this is paid by the increase in computational overhead.

During the verification process the packet's path to the root is verified, i.e. all nodes on that path. The procedure is similar to that for the star-chaining scheme.
The main advantage of the both schemes is the ability to independently verify each of the received packets, so there is no problems with packets lost or tampered with that can be discarded independently. On the other hand, they both involve high computational overhead and quite high communication overhead [10].

2.2. Hash chains

The simplest scheme using a hash chain was proposed by Gennaro and Rohatgi in [5]. It involved the use of just one full digital signature for the whole stream and hashes for each block of "c" packets. The receiver required a buffer of size "c". The receiver first received the signature of the 20-byte hash of the first block and the hash itself. After verifying the signature, the first hash should be verified. He then started receiving the first block and calculated the hash for this block. If it matched the verified hash, it could then play the block. Otherwise the whole stream was rejected. The first block carried the hash for the second block, and so on. At the sender, the whole stream had to be known in advance, as the hash for the block "$i+1$" was appended to the block "i". Reducing the block size "c" to a single packet meant no need for a buffer at the receiver side.
The main advantages of the scheme were very low computational and communicational overheads. However, the scheme was very vulnerable to packet losses, and even a single packet loss would result in a break in the verification chain and rejection of all successive packets. Moreover, the scheme was suited for the off-line applications only. Some modification to the scheme was also proposed in [5] for on-line applications. Unfortunately, it involved significant complication of the scheme while maintaining its vulnerability to packet losses.
Miner and Staddon in [14] propose to introduce additional connections in the authentication chain in order to achieve immunity against lost packets. As the result, the verification chain can tolerate even burst losses of up to the predefined number of packets. However, once the chain is broken due to the longer than assumed burst, the verification cannot be continued. Another disadvantage of the scheme is the fact that it is suited for the off-line authentication only, as the sender needs to know all packets in the stream to calculate the desired sequence of hashes
A breakthrough in design of the hash chain based authentication schemes for streamed data has been the technique proposed by Golle and Modadugu in [10]. It follows from the fact that if a collision-resistant hash of packet P_1 is appended to several modifications to the scheme were proposed in literature, and well generalized by packet P_2 before signing P_2, then the signature on P_2 guarantees the authenticity of both P_2 and P_1. In general, a hash h_1 of P_1 is appended to P_2 before calculating a hash h_2 for P_2, and h_2 is appended to P_3 before calculating h_3 for P_3, and so on. The final packet in a sequence, P_n is then signed after appending h_{n-1} to it. If then the sequence of packets $P_1, \ldots, P_n$ is received without tampering or losses, all packets in that sequence can be authenticated. The process of creating such a simple authenticating graph is presented in Fig. 5.

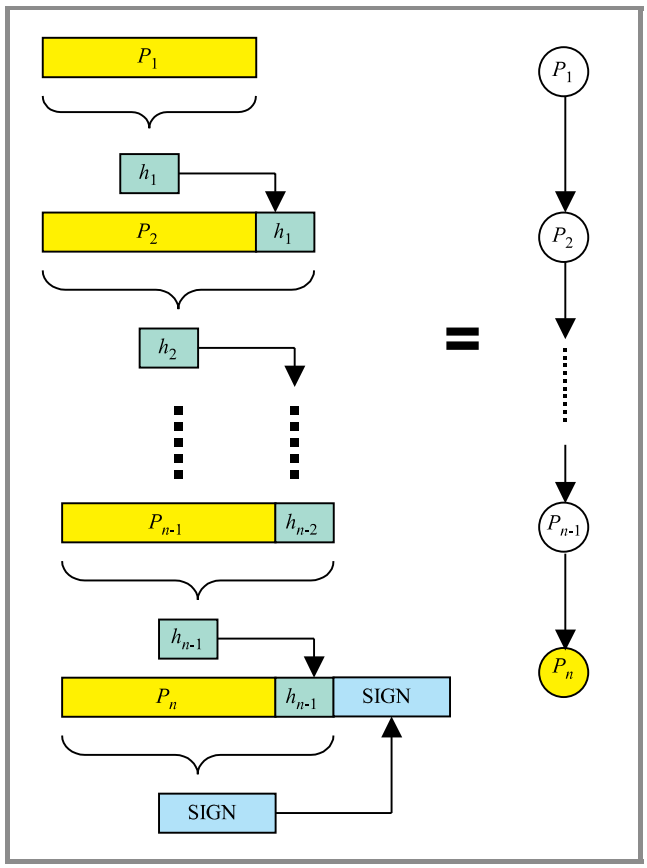


Fig. 5. Basic one-way hash chain of Golle and Modadugu.

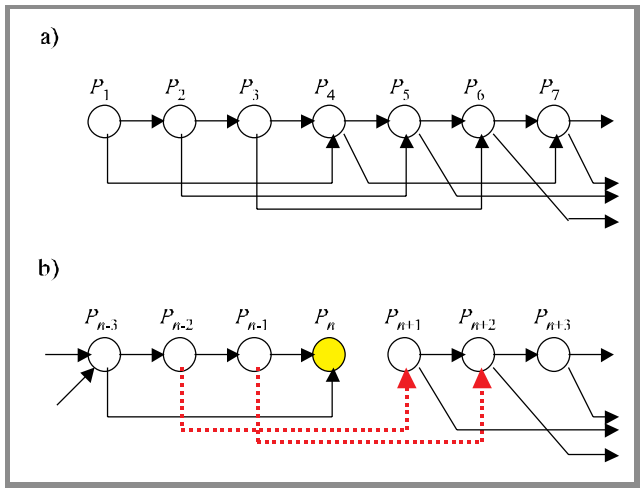


Fig. 6. Explanation of the proposed modification method: (a) the original authentication chain of order 3 [10]; (b) continuation of the chain beyond the signed packet P_n.

The simple chain presented in Fig. 5 can be modified to include supplementary connections to prevent it from being broken in a case of packet losses. Depending on the

type of construction of those additional connections, immunity from burst losses can be achieved. As an example, in Fig. 6a, the chain immune to a loss of two consecutive packets is presented. Several different constructions are presented in [10] and analyzed.

The main advantage of the scheme is the fact that by splitting the stream into smaller sequences of packets and reversing the order of the authentication chain, Golle and Modadugu achieved a scheme suited for on-line applications where a break in the verification chain means rejection of usually just one sequence of packets (maximum two sequences if the break includes a signed packet). Other advantages of the technique are immunity to bursts of up to a given length, low communication overhead, and low computational overhead [10]. The drawbacks of the scheme are the delayed verification and susceptibility to a loss of signed packets.

3. Model of a bursty channel

In his fundamental paper [15], Gilbert introduced a two state Markov chain to model a transmission channel with burst errors. The model has been later refined by Elliott in [16], and is generally known in telecommunications literature as the Gilbert-Elliott channel. The Gilbert-Elliott model has been introduced to analyze channel at bit level. However, we can use the same approach to perform analysis at the packet level.

Gilbert in [15] has shown that a Markov chain with two states can be used to generate bursts. The model is shown in Fig. 7, where the states G and B denote the "good" channel state and the "bad" channel state, respectively. In the "good" state the probability of packet loss approaches zero, while in the "bad" state it can take any arbitrary value greater than zero. In the original model [15], that probability was set to 0.5, as the author dealt with bursts at the bit level, i.e. in the physical channel, where bursts contain good bits interspersed with the errors. In this paper, to simplify considerations, we will assume that at the packet level it is very high, approaching 1. However, more general approach can be taken at the later stage.

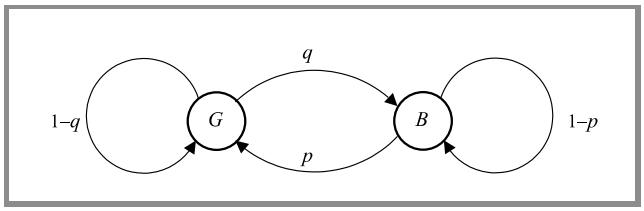


Fig. 7. Gilbert-Elliot channel model.

The model is described by the probability transition matrix $\mathbf{P}_T$ given by:

$$\mathbf{P}_T = \begin{bmatrix} 1-q & q \\ p & 1-p \end{bmatrix} \tag{1}$$

and the corresponding graph is presented in Fig. 7.

For this Markov chain, the stationary probability vector $\mathbf{P}_S = [p_1,\, p_2]$ can be calculated using the formula:

$$\mathbf{P}_S = \mathbf{P}_S \mathbf{P}_T \tag{2}$$

and the normalization condition:

$$p_1 + p_2 = 1\,. \tag{3}$$

From Eqs. (1) and (2), we get a set of two simultaneous equations:

$$\begin{cases} p_1 = p_1(1-q) + p_2 p \\ p_1 + p_2 = 1 \end{cases} \tag{4}$$

that give us:

$$p_1 = \frac{p}{1+p-1+q} = \frac{p}{p+q} \tag{5}$$

and:

$$p_2 = \frac{q}{p+q}\,. \tag{6}$$

The stationary probabilities p_1 and p_2 are the probabilities that at any discrete time instant the channel is in the "good" state or the "bad" state, respectively. For the transmitted packets, it translates on the probabilities that the packet is either received correctly or lost. In addition, the exact values of the probabilities p_{12} and p_{21} fully determine the probabilities of bursts occurrences and their lengths.

Let now denote by $P_g\{M\}$ the probability that out of M signed packets in the stream all packets have been transmitted in "good" state, and that subsequently no such a packet has been lost. Following the above analysis of the Gilbert-Elliott model, we can write that:

$$P_g\{M\} = p_1^M\,. \tag{7}$$

Hence, the probability $P_b\{1\}$ that at least one signed packet is transmitted in the "bad" state and subsequently lost is given by:

$$P_b\{1\} = 1 - P_g\{M\} = 1 - p_1^M\,. \tag{8}$$

To illustrate the problem of the authentication method proposed in [10], let us consider the following example.

Example 1. We consider here the chain construction proposed in [10] that provides immunity against bursts of length $l = 5$ and the stream consists of 50 sequences of 50 packets each. The packets are transmitted through the Gilbert-Elliott channel having the following parameters[1]:

Probability of a transition from G to B: $q = 0.01$

Probability of a transition from B to G: $p = 0.3$

[1]Even though the exact channel parameters q and p are difficult to find for virtual transport channels, i.e. where packet loss is considered, the values used in the example correspond to one of the best channels reported on in [17].

The stationary probabilities p_1 and p_2 can be calculated using Eqs. (5) and (6), and are equal to:

$$p_1 = \frac{0.3}{0.3+0.01} = 0.9677 \text{ and } p_2 = \frac{0.01}{0.3+0.01} = 0.0323$$

and the probability that at least one signature is lost equals to:

$$P_b\{1\} = 1 - 0.9677^{50} = 0.8059\,.$$

The presented example indicates that even for quite a modest stream (50 sequences of 50 packets of 1000 bytes correspond to about 2.5 MB), the probability of loosing at least one signed packet is very high, and with the increased number of sequences it approaches 1.

4. Modification method

Here we propose a simple method to overcome the problem of losing a signed packet without introducing any additional overheads or delays in packet verification, when no signature is lost. An extra delay in verification time will be only introduced in the case of a burst containing the signed packet. However, in the original scheme described in [10], this situation would result in the whole considered sequence of packets being rejected.

The proposed method is based on extending the authentication chain of [10] beyond the packet containing the signature, i.e. the packet, on which the authentication graph should end for the given sequence. In other words, we propose to consider the signed packet in a similar way to any other packet, with the exception that no hash is generated for it. Suppose a stream is divided into sequences $S_1, S_2, \ldots$, where each S_k consists of n packets. Let

$$S_k = \{P_{(k-1)n+1}, P_{(k-1)n+2}, \ldots, P_{kn}\}\,. \qquad (9)$$

For each sequence Golle and Modadugu [10] proposed an optimal authentication chain to provide resistance for bursty loss of packets with the assumption that the signed packet P_{kn} is not lost. We modify their authentication chain by introducing extra edges between S_k and S_{k+1} and removing the assumption of no signed packet lost.

Our authentication chain is defined as follows. Let a be a positive integer. The hash of a packet P_i is appended to two packets P_{i+1} and P_{i+a} for all i, $(k-1)n < i < kn$, and for all $k \geq 1$. This authentication chain sustains bursts of $a-1$, packets, even including the packet with signature. If the signature packet is lost, the authentication has a delay of $2n$.

An example application of the proposed solution together with the authentication chain of [10] is given in Fig. 6. The diagram of Fig. 6a shows the original authentication chain, while the diagram of Fig.6b presents what happens around the signed packet P_n. As the authors of [10] point out, their authentication chain (Fig. 6a) is immune to bursts of 2 packets, but the signed packet P_n must be delivered correctly. In our case, if the burst of 2 packets contains the signed packet, e.g. packets P_{n-1} and P_n are lost, the authentication chain is not broken, and verification can be performed after receiving the next signed packet, i.e. P_{2n}. Moreover, if the bursts are no longer than the length for which the original chain has been designed, any received signature can be used to verify all previously received packets, no matter how many signed packets have been lost.

There are no any additional computational overheads involved, and there is almost no (the hashes from the previous sequence packets are now attached to the current sequence packets, as in Fig. 6b) additional communication overhead involved in the proposed scheme compared to the original scheme described in [10]. All chain constructions proposed in [10] can be used together with the proposed method. This does not seem to be the case of constructions proposed in [14], as the signatures are there transmitted with the first packet in sequences. However, this requires some further investigations.

Let us now analyze the probability that the authentication chain constructed in accordance with our modification will be broken. This can only happen when the burst containing more than b packets occur, while the construction is immune to bursts of no more than b packets.

The burst of length $l = b$ occurs when channels changes state from G to B, stays there for b consecutive steps and returns to G afterwards. Because the verification chain is resistant to burst of length $l = b$, the break of the verification chain occurs at the state $s^{(n)} = B$ if, $s^{(n-1)} = B$, $s^{(n-2)} = B$, $\ldots$, $s^{(n-b)} = B$, and $s^{(n-b-1)} = G$. Hence, we can model the verification chain breaking process by a sequential stochastic machine of $b+2$ states. We denote the states of this machine by $\sigma_1, \sigma_2, \ldots, \sigma_{b+2}$.

The state σ_1 occurs when the channel is in the "good" state and the state σ_{b+2} when the verification chain is broken. The states $\sigma_2, \sigma_3, \ldots, \sigma_{b+1}$ are the states corresponding to the bursts of length $1, 2, \ldots, b$, respectively. Using the original Gilbert-Elliot model, we can find the respective transition probabilities. They are given by the transition probability matrix:

$$\mathbf{P}_{\sigma\sigma} = \begin{bmatrix} 1-q & q & 0 & 0 & 0 & \cdots & 0 \\ p & 0 & 1-p & 0 & 0 & \cdots & 0 \\ p & 0 & 0 & 1-p & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p & 0 & 0 & 0 & 0 & \cdots & 1-p \\ p & 0 & 0 & 0 & 0 & \cdots & 1-p \end{bmatrix}. \qquad (10)$$

The corresponding graph is given in Fig. 8.

As the resulting $(b+2)$ – state Markov chain is a stationary one, (the matrix $\mathbf{P}_{\sigma\sigma}$ does not depend on the number of a currently transmitted packet) we can find the vector of stationary probabilities $\mathbf{P}_\sigma$ by solving the set of $(b+2)$ linear equations:

$$\begin{cases} \mathbf{P}_\sigma = \mathbf{P}_\sigma \mathbf{P}_{\sigma\sigma} \\ \sum_{i=1}^{b+2} p_{\sigma i} = 1 \end{cases}. \qquad (11)$$

The probability $p_{\sigma(b+2)}$ is the stationary probability that during the transmission of a packet, the verification chain is broken.

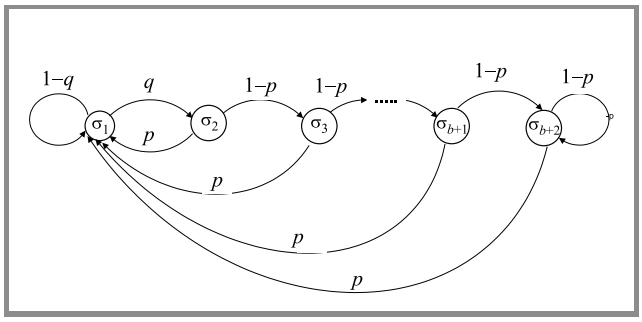


Fig. 8. Graph of the Markov chain given by matrix (9).

Having found the probability $p_{\sigma(b+2)}$, we can calculate the probability $P_K\{break\}$ that during the transmission of K packets the verification chain is broken at least once. Those K packets are assumed as transmitted in the steady state, i.e. they do not include first b packets in the stream. It is given by:

$$P_K\{break\} = 1 - (1 - p_{\sigma(b+2)})^K . \tag{12}$$

This will be illustrated by the following example.

Example 2. Let us consider here the same stream transmitted as in Example 1 through the same channel. We will find the probabilities that the verification chain will be broken during the transmission of a single sequence and the whole stream of 50 sequences. We compare the results for the cases when the authentication chain is created to be immune against bursts of lengths 3, 5, and 8.

To find answers to the problem, we first need to construct the relevant transition probability matrices and find the stationary probabilities of the states corresponding to the break in the verification process. While for $b = 3$, this is still a feasible task to find the stationary probabilities manually, as

$$\mathbf{P}_{\sigma\sigma} = \begin{bmatrix} 1-q & q & 0 & 0 & 0 \\ p & 0 & 1-p & 0 & 0 \\ p & 0 & 0 & 1-p & 0 \\ p & 0 & 0 & 0 & 1-p \\ p & 0 & 0 & 0 & 1-p \end{bmatrix} =$$

$$= \begin{bmatrix} 0.99 & 0.01 & 0 & 0 & 0 \\ 0.30 & 0 & 0.70 & 0 & 0 \\ 0.30 & 0 & 0 & 0.70 & 0 \\ 0.30 & 0 & 0 & 0 & 0.70 \\ 0.30 & 0 & 0 & 0 & 0.70 \end{bmatrix} . \tag{13}$$

However, it becomes more cumbersome for the cases when $b = 5$ or $b = 8$. Any standard software package can be used to provide solution to the set of linear equations. We used here MATLAB, and the results are given in Table 1.

After that, we can use formula (12) to find the probabilities of breaking the verification stream for a single sequence of 50 packets, and for the whole stream of 50 sequences. The results are given in Table 2.

Table 1
Stationary probabilities of breaking the verification chain for the case considered in Example 2

Verification chain immune against bursts of length b	Stationary probability of a break in the verification chain
3	0.0111
5	0.0054
8	0.0019

Table 2
Probabilities of the verification chain being broken during the transmission of a single sequence of 50 packets and the whole stream of 50 sequences for the Example 2

Verification chain immune against bursts of length b	Probability of breaking verification chain in a single sequence	Probability of breaking verification chain in a whole stream
3	0.4277	1.0000
5	0.2372	1.0000
8	0.0907	0.9913

The results from Example 2 indicate that the stationary probabilities of breaking the verification chain are quite low, as shown in Table 1, and decrease with an increase in the maximum length of burst to which the verification chain is immune to. However, the probabilities of having at least one break in the verification chain for the transmission of whole stream are very high, approaching 1 in all considered cases. This is due to the fact that the considered channel is characterized with the high value of a transition from the "good" to the "bad" state and equal to 0.01. This value is more likely to be experienced in wireless channels. However, with the proliferation of multimedia services to the wireless environment such channels need to be considered while designing authentication algorithms for streamed data.

5. Conclusions

In the paper we presented a modification to the Golle and Modadugu authentication chain [10] that results in a chain resistant to losses of packets containing sequence signatures. The modification is achieved without any additional computational or communication overheads. In addition, we have introduced here a Markov chain method to analyze performance of authentication schemes. It draws from the fact that channel with burst losses can be modeled by a two state Markov process, referred to as the Gilbert-Elliot model. The method can be very useful in analyzing performance of different authentication schemes under realistic channel conditions.

In the future, we expect to perform such an analysis for other hash chain based schemes taking into account multiple bursts. This will of course result in more complicated Markov chains of the verification process.

References

[1] W. Stalings, *Data and Computer Communications*. 6th ed. Upper Saddle River: Prentice Hall, 2000.

[2] J. Lu, "Signal processing for Internet video streaming", *Proc. SPIE, Image and Video Commun. Proc.*, Jan. 2000.

[3] G. Pipkin, "Video on the web", http://www.itc.virginia.edu/atg/techtalks/powerpoint/video/sld001.htm (accessed on 23/11/01).

[4] "Utah education network", http://www.uen.org/technical/html/streamingfaq.html (accessed on 23/11/01).

[5] R. Gennaro and P. Rohatgi, "How to sign digital streams", in *Cypto '97*. Springer-Verlag, 1998, pp. 180–197.

[6] C. K. Wong and S. S. Lam, "Digital signatures for flows and multicasts", *IEEE/ACM Trans. Netw.*, vol. 7. no. 4, pp. 502–513, 1999.

[7] M. J. Moyer, J. R. Rao, and P. Rohatgi, "A survey of security issues in multicast communications", *IEEE Network*, pp. 12–23, Nov./Dec. 1999.

[8] D. R. Stinson, *Cryptography Theory and Practice*. CRC Press.

[9] A. Perrig, R. Canetti, J. D. Tygar, and D. Song, "Efficient authentication and signing of multicast streams over lossy channels", in *Proc. IEEE Secur. Priv. Symp.*, May 2000, pp. 56–73.

[10] P. Golle and N. Modadugu, "Authenticating streamed data in the presence of random packet loss", in *Proc. Netw. Distrib. Syst. Secur. Symp.*, 8–9 Feb. 2001.

[11] P. Golle and N. Modadugu, "Authenticating streamed data in the presence of random packet loss", http://www.isoc.org/isoc/conferences/ndss/01/2001/papers/golle.pdf (accessed on: 12/02/02).

[12] W. Stalings, *Cryptography and Network Security: Principles and Practice*. 2nd ed. Upper Saddle River: Prentice Hall, 1998.

[13] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*. CRC Press, 1996.

[14] P. Golle, "Authenticating streamed data in the presence of random packet loss", power point presentation, http://crypto.stanford.edu/p̃golle/papers/auth.html (accessed on: 10/12/01).

[15] S. Miner and J. Staddon, "Graph-based authentication of digital streams", http://www-cse.ucsd.edu/users/sminer/abstracts/GraphAuth.html (accessed on: 11/01/02).

[16] E. N. Gilbert, "Capacity of burst-noise channel", *Bell Syst. Techn. J.*, pp. 1253–1265, Sept. 1960.

[17] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels", *Bell Syst. Techn. J.*, pp. 1977–1997, Sept. 1963.

[18] J. M. Boyce and R. D. Gaglianello, "Packet loss effects on MPEG video sent over the public Internet", in *ACM Multimedia '98, Electron. Proc.*

Beata Joanna Wysocki graduated from Warsaw University of Technology receiving her M.Eng. degree in electrical engineering in 1991. In 1994, she started the Ph.D. study in the Australian Telecommunications Research Institute at Curtin University of Technology. In March 2000, she was awarded her Ph.D. for the thesis: "Signal Formats for Code Division Multiple Access Wireless Networks". During the Ph.D. studies, she was involved in a research project "Wireless ATM Hub" at the Cooperative Research Centre for Broadband Telecommunications and Networking, and worked as a research assistant at Edith Cowan University, within the ARC funded "CDMA with enhanced protection against frequency selective fading", and "Reliable high rate data transmission over microwave local area networks". Since October 1999 she has been with the Telecommunications & Information Technology Research Institute at the University Wollongong as a research fellow. Her research interests include sequence design for direct sequence (DS) code division multiple access (CDMA) data networks and error control strategies for broadband wireless access (BWA) systems.
e-mail: beata@elec.uow.edu.au
University of Wollongong
Northfields Avenue
Wollongong, NSW 2522, Australia

Yejing Wang received a Ph.D. in computer science in 2001 and is currently a research fellow in the School of Information Technology and Computer Science of the University of Wollongong. Her research interest include cryptography and algebraic coding.
University of Wollongong
Northfields Avenue
Wollongong, NSW 2522, Australia

Reihaneh Safavi-Naini is a Professor of Computer Science in University of Wollongong. She received a Ph.D. in electrical and computer engineering from University of Waterloo in Canada. Her research interests include, cryptography, computer and communication security and digital right management.
University of Wollongong
Northfields Avenue
Wollongong, NSW 2522, Australia

Regular paper

Simple analysis of the impact of packet loss and delay on voice transmission quality

Krzysztof Perlicki

Abstract — In this paper a simple analytical expression based on E-model is applied to analyse the impact of packet loss and delay on voice transmission quality. The relationship between overall transmission quality rating, packet loss and delay G.711, G.723 and G.729 codecs is presented.

Keywords — E-model, packet network, packet loss, delay, transmission quality rating.

1. Introduction

The IP network will become an ubiquitous infrastructure, which is used by numerous applications with various requirements and with different characteristics of generated traffic. In particular, it concerns to data and voice applications. In the future, the voice communication is expected to migrate from the public switched telephone network (PSTN) to the packet (IP) transmission. Due to constant improvements within the years, traditional voice communications over the PSTN is characterized today by low delay, high availability and adequate voice quality. For the packet transmission to compete with the PSTN, it should provide the same level of quality, which implies stringent delay, packet loss and also packet delay variation (jitter). The packet delay variation is bounded by the dejitter buffer, which removes small amounts of variation by increasing delay and/or by the increasing packet loss [1].

The E-model [2, 3] is a network planning tool used in the design of networks for carrying voice applications. The model estimates the relative impairments to voice quality. It combines the impairments caused by transmission parameters into R factor (overall transmission quality rating). The factor R is given by [3]:

$$R = R_o - I_s - I_d - I_e + A\,. \tag{1}$$

The factor R_o expresses the basic signal-to-noise ratio. The factor I_s represents all impairments which occur more or less simultaneously with the voice signal, such as: too loud speech level, non-optimum sidetone, quantization noise, etc. The delay impairment factor (I_d) sums all impairments due to delay and echo effects. The equipment impairment factor (I_e), represents impairments which are caused by low bit-rate codecs used in special equipment; this factor can also be used to take into account the influence of packet loss. The advantage factor A represents an advantage of access which certain systems may provide in comparison to conventional systems (for wirebound communication system factor A is equal zero). Typically, the values of the R factor are categorized as shown in Table 1. This table also contains equivalent transformed values of R into *MoS* (mean opinion score).

Table 1
R factor, *MoS* and the user satisfaction measure [3]

R	*MoS*	User satisfaction
$90 < R < 100$	$4.34 < MoS < 4.5$	Very satisfied
$80 < R < 90$	$4.03 < MoS < 4.34$	Satisfied
$70 < R < 80$	$3.60 < MoS < 4.03$	Some users dissatisfied
$60 < R < 70$	$3.10 < MoS < 3.60$	Many users dissatisfied
$50 < R < 60$	$2.58 < MoS < 3.10$	Nearly all users dissatisfied

An interesting aspect of the E-model is that terms R_o, I_s, I_d, I_e and A are additive and that delay and packet loss contributions are isolated into I_d and I_e, respectively. The others factors, like noise (R_o) and a connection that is too loud (I_s) of a packetized voice call are not fundamentally different from traditional switched telephone network.

2. Delay in packet network – I_d factor

The delay for packetized voice call, where the most important contributions are encoding, packetization, propagation, queuing, service, dejittering and decoding delay is larger than for traditional switched telephone network where the end-to-end delay is equal to propagation delay and switching delay. Figure 1 shows the relationship between the I_d factor and the absolute end-to-end (mouth-to-ear) delay (T_a). The results are derived from E-model (all E-model input parameters of default values, which are shown in [3]) and calculated using following expression:

$$I_d = 0.65 + (0.1T_a - 15.93) \cdot \delta(T_a - 165)\,, \tag{2}$$

where: $\delta(x) = 0$ for $x < 0$, $\delta(x) = 1$ for $x \geq 0$.

The absolute end-to-end delay in packet connections is given by:

$$T_a = T_{enc} + T_p + T_{dec} + T_n\,, \tag{3}$$

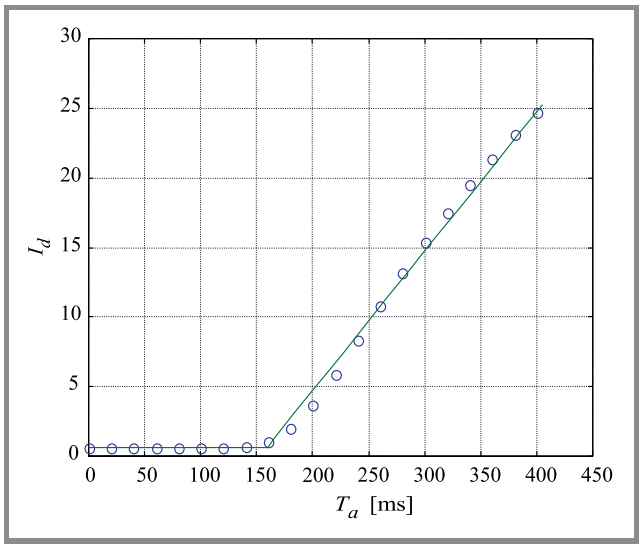


Fig. 1. The relationship between the I_d factor and the T_a delay; ∘ – based on E-model, curve fit – Eq. (3).

where: T_{enc} is the encoding delay, T_p is the packetization delay, T_{dec} is the decoding delay and T_n is the so-called network delay, which is equal to the sum of the propagation, queuing, service and dejittering delay [4].

3. Packet loss – I_e factor

No analytical expressions for the equipment impairment factor (I_e) are directly available today. Instead, the I_e factor must be obtained from subjective tests of voice quality for particular codec and various operating conditions (e.g. packet loss, number of voice frames per packet) [5]. ITU-T Rec. [6] gives measured value of the I_e factor only for G.711, G.723 and G.729 codecs.

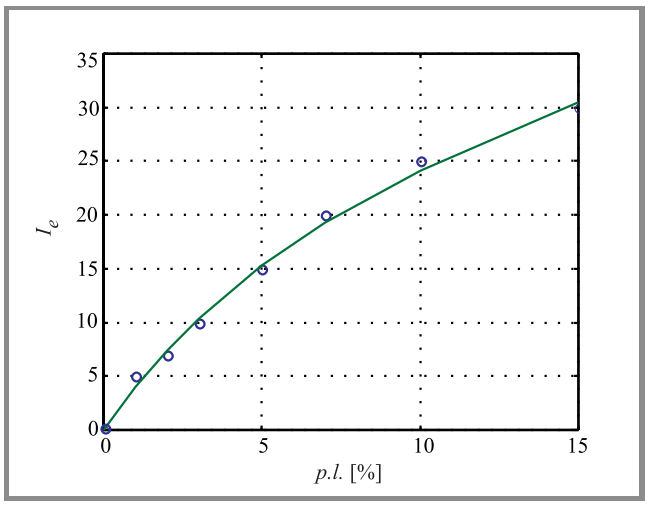


Fig. 2. The equipment impairment factor (I_e) versus the packet loss for G.711 codec (random packet loss); ∘ – based on [6], curve fit – Eq. (4).

Figure 2 shows results for the I_e factor for G.711. Also shown in Fig. 2 is a curve fit, which is derived from an expression of the form:

$$I_e = 22 \ln (1 + 0.2 \cdot p.l.), \tag{4}$$

where $p.l.$ is the packet loss in percent.

The results for the I_e factor and G.723 codec are shown in Fig. 3. The curve fit in this figure is derived from the following expression:

$$I_e = 15 + 33 \ln(1 + 0.15 \cdot p.l.). \tag{5}$$

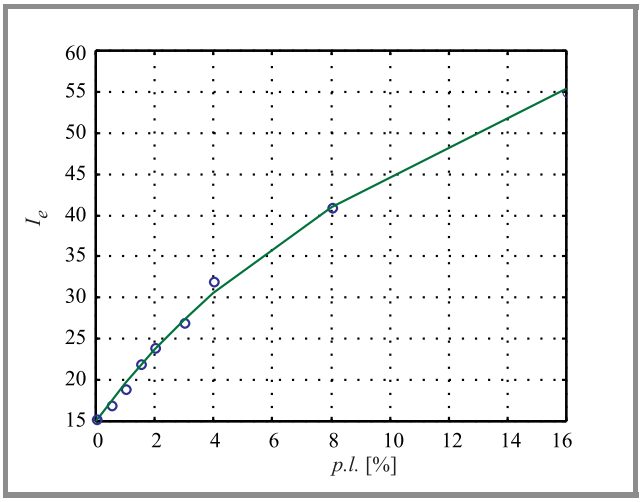


Fig. 3. The equipment impairment factor (I_e) versus the packet loss for G.723 codec (number of voice frames per packet: 1); ∘ – based on [6], curve fit – Eq. (5).

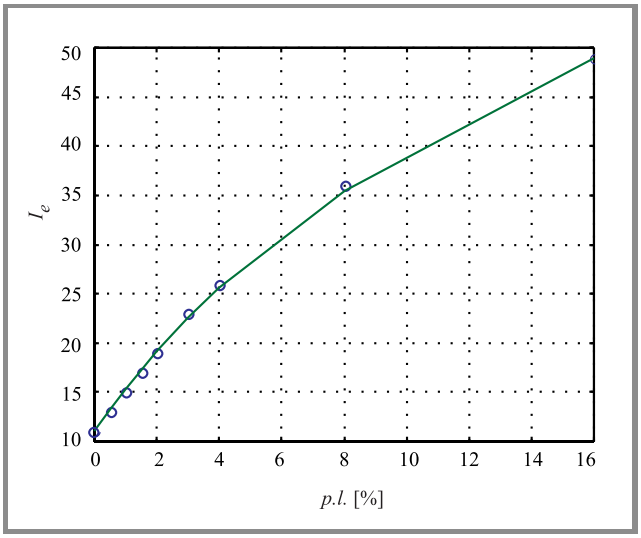


Fig. 4. The equipment impairment factor (I_e) versus the packet loss for G.729 codec (number of voice frames per packet: 2); ∘ – based on [6], curve fit – Eq. (6).

The results for the I_e factor and G.729 codec are shown in Fig. 4. The curve fit in this figure is derived from the following expression:

$$I_e = 11 + 31 \ln(1 + 0.15 \cdot p.l.). \tag{6}$$

4. Impact of packet loss and network delay on the factor R

Choosing E-model default values [3] we can reduce the expression (1) to:

$$R = 93.33 - I_d - I_e\,. \tag{7}$$

Substituting Eqs. (2)–(6) into (7), the factor R can be rewritten as:

- for G.711 codec:

$$\begin{aligned} R = {} & 92.68 - (0.1\,T_{n\max} - 15.90) \times \\ & \times \delta(T_{n\max} - 164.75) + \\ & - 22 \ln(1 + 0.2 \cdot p.l.), \end{aligned} \tag{8}$$

- for G.723 codec:

$$\begin{aligned} R = {} & 77.68 - (0.1\,T_{n\max} - 9.18) \times \\ & \times \delta(T_{n\max} - 97.50) + \\ & - 33 \ln(1 + 0.15 \cdot p.l.)\,, \end{aligned} \tag{9}$$

- for G.729 codec:

$$\begin{aligned} R = {} & 81.68 - (0.1\,T_{n\max} - 12.43) \times \\ & \times \delta(T_{n\max} - 130) + \\ & - 31 \ln(1 + 0.15 \cdot p.l.)\,, \end{aligned} \tag{10}$$

where $T_{n\max}$ is the maximum of the network delay.

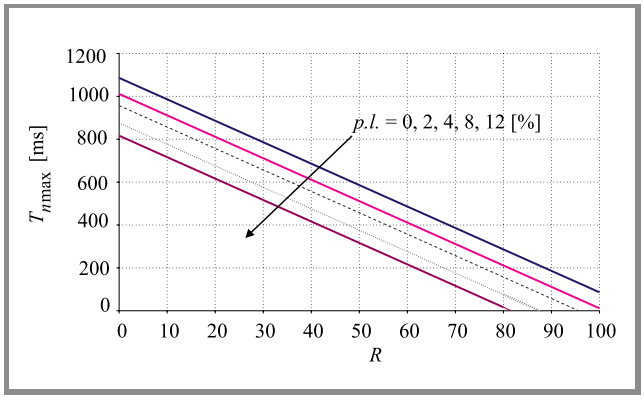


Fig. 5. Maximum of the network delay versus the factor R (G.711 codec).

It is assumed in Eqs. (8)–(10) that the minimum values of the sum of the encoding delay, the packetization delay and the decoding delay are [7]:

- for G.711 codec: 0.25 ms,
- for G.723 codec: 67.5 ms (number of voice frames per packet: 1),
- for G.729: 35 ms (number of voice frames per packet: 2).

Figures 5, 6, 7 show the relationship between the maximum of the network delay ($T_{n\max}$) and the R factor and G.711, G.723 and G.729 codecs, respectively.

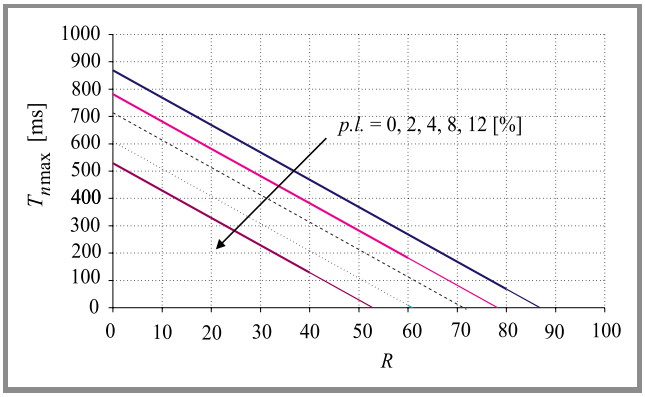


Fig. 6. Maximum of the network delay versus the factor R (G.723 codec).

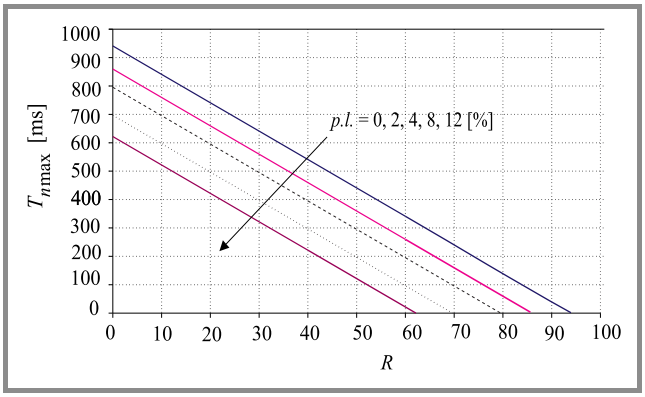


Fig. 7. Maximum of the network delay versus the factor R (G.729 codec).

Table 2
The value of $T_{n\max}$ for $R = 50,\ 60,\ 70,\ 80$ and 90 (G.711 codec)

Maximum of the network delay $T_{n\max}$ [ms]					
R	Packet loss [%]				
	0	2	4	8	12
50	585.8	511.8	456.5	375.6	316.6
60	485.8	411.8	356.5	275.6	216.6
70	385.8	311.8	256.5	175.6	116.6
80	285.8	211.8	156.5	75.6	16.6
90	185.8	111.8	56.5	—	—

The value of $T_{n\max}$ for $R = 50,\ 60,\ 70,\ 80$ and 90 under conditions of packet loss for G.711, G.723 and G.729 codecs are given in Tables 2, 3, and 4, respectively.

Table 3
The value of $T_{n\max}$ for $R = 50,\ 60,\ 70,\ 80$ and 90 (G.723 codec)

Maximum of the network delay $T_{n\max}$ [ms]					
R	Packet loss [%]				
	0	2	4	8	12
50	368.6	282.0	213.5	108.4	28.8
60	268.6	182.0	113.5	8.4	—
70	168.6	82.0	13.5	—	—
80	68.6	—	—	—	—
90	—	—	—	—	—

Table 4
The value of $T_{n\max}$ for $R = 50,\ 60,\ 70,\ 80$ and 90 (G.729 codec)

Maximum of the network delay $T_{n\max}$ [ms]					
R	Packet loss [%]				
	0	2	4	8	12
50	441.1	359.8	295.4	196.7	121.9
60	341.1	259.8	195.4	96.7	21.9
70	241.1	159.8	95.4	—	—
80	141.1	59.8	—	—	—
90	41.1	—	—	—	—

5. Conclusions

The E-model has been used to study the quality of packetized voice calls. The overall transmission quality rating (user satisfaction) decreases as the packet loss and the delay increases. The quality "Very satisfied" ($R > 90$) is reached only for G.711 and G.729 codecs. In the case of G.711 codec $R > 90$ is reached at: $p.l. = 0$ and $T_n \leq 185.5$ ms, $p.l. = 2$ and $T_n \leq 111.8$ ms, $p.l. = 4$ and $T_n \leq 56.5$ ms. In the case of G.729 codec $R > 90$ is only reached at $p.l. = 0$ and $T_n \leq 41.1$ ms. It can be seen that for G.723 codec the greatest quality is "Satisfied" ($R > 80$). It is reached at $p.l. = 0$ and $T_n \leq 68.6$ ms. For the packet loss 12% the G.711 codec has a rating R equal to 80 ($T_{n\max} = 16.6$ ms). For this packet loss the G.723 codec has the R rating equal to 50 ($T_{n\max} = 28.8$ ms), and G.729 codec has $R = 60$ ($T_{n\max} = 21.9$ ms).

References

[1] ETSI TS 101 329-2 V1.1.1, "Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON)"; "End to End Quality of Service in TIPHON Systems"; Part 2: "Definition of Quality of Service (QoS) Classes", Sophia Antipolis Cedex – France, ETSI, 7/2000.

[2] ETSI EG 201 050 V1.2.2, "Speech Processing, Transmission and Quality Aspects (STQ)"; "Overall Transmission Plan Aspects for Telephony in a Private Network", Sophia Antipolis Cedex – France, ETSI, 2/1999.

[3] ITU-T Rec. G.107, "The E-model, a computational model for use in transmission planning", Geneva, ITU, 5/2000.

[4] D. De Vleeschauwer and J. Janssen, "Quality bounds for packetized voice transport", *Alcatel Telecommun. Rev.*, no. 1, pp. 19–23, 2000.

[5] ITU-T Rec. P.833, "Methodology for derivation of equipment impairment factors from subjective listening – only tests", Geneva, ITU, 2/2001.

[6] ITU-T Rec. G.113, "Transmission impairments", Appendix I: "Provisional planning values for the equipment impairment factor I_e", Geneva, ITU, 9/1999.

[7] ITU-T Rec. G.114, "One-way transmission time", Geneva, ITU, 5/2000.

Krzysztof Perlicki
e-mail: perlicki@tele.pw.edu.pl
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

The author's fee was financed by the Association for Authors Rights Collective Administration of Scientific and Technical Works KOPIPOL with a seat in Kielce from the remuneration collected on the basic of Art. 20 of the Law on Author Right and Related Rights.

INFORMATION FOR AUTHORS

The *Journal of Telecommunications and Information* Technology is published quarterly. It comprises original contributions, both regular papers and letters, dealing with a broad range of topics related to telecommunications and information technology. Items included in the journal report primary and or experimental research results, which advance the base of scientific and technological knowledge about telecommunications and information technology.

The *Journal is* dedicated to publishing research results which advance the level of current research or add to the understanding of problems related to modulation and signal design, wireless communications, optical communications and photonic systems, speech devices, image and signal processing, transmission systems, network architecture, coding and communication theory, as well as information technology. Suitable research-related manuscripts should hold the potential to advance the technological base of telecommunications and information technology. Tutorial and review papers are published by invitation only.

Papers published by invitation and regular papers should contain up to 15 and 8 printed pages respectively (one printed page corresponds approximately to 3 double-space pages of manuscript, where one page contains approximately 2000 characters).

Manuscript: An original and two copies of the manuscript must be submitted, each completed with all illustrations and tables attached at the end of the papers. Tables and figures have to be numbered consecutively with Arabic numerals. The manuscript must include an abstract limited to approximately 100 words. The abstract should contain four points: statement of the problem, assumptions and methodology, results and conclusion, or discussion, of the importance of the results. The manuscript should be double-spaced on only one side of each A4 sheet (210 x 297 mm). Computer notation such as Fortran, Matlab, Mathematica etc., for formulae, indices, etc., is not acceptable and will result in automatic rejection of the manuscript. The style of references, abbreviations, etc., should follow the standard IEEE format.

References should be marked in the text by Arabic numerals in square brackets and listed at the end of the paper in order of their appearance in the text, including exclusively publications cited inside. The **reference entry** (correctly punctuated according to the following rules and examples) **has to contain**.

From journals and other serial publications: initial(s) and second name(s) of the author(s), full title of publication transliterated into Latin characters in case it is in Russian, possibly preceded by the title in Russian characters), appropriately abbreviated title of periodical, volume number, first and last page number, year. E.g.:

[1] Y. Namihira, „Relationship between nonlinear effective area and modefield diameter for dispersion shifted fibres", *Electron. Lett.*, vol. 30, no. 3, pp. 262-264, 1994.

From non-periodical, collective publications: as above, but after title – the name(s) of editor(s), title of volume and/or edition number, publisher(s) name(s) and place of edition, inclusive pages of article, year. E.g.:

[2] S. Demri, E. Orłowska, „Informational representability: Abstract models versus concrete models" in *Fuzzy Sets, Logics and Reasoning about Knowledge*, D. Dubois and H. Prade, Eds. Dordrecht: Kluwer, 1999, pp. 301-314.

From books: initial(s) and name(s) of the author(s), place of edition, title, publisher(s), year. E.g.:

[3] C. Kittel, *Introduction to Solid State Physics*. New York: Wiley, 1986.

Figure captions should be started on separate sheet of papers and must be double-spaced.

Illustration: Original illustrations should be submitted. All line drawings should be prepared on white drawing paper in black India ink. Drawings in Corel Draw and Postscript formats are preferred. Colour illustrations are accepted only in exceptional circumstances. Lettering should be large enough to be readily legible when drawing is reduced to two- or one-column width – as much as 4:1 reduction from the original. Photographs should be used sparingly. All photographs must be gloss prints. All materials, including drawings and photographs, should be no larger than 175 x 260 mm.

Page number: Number all pages, including tables and illustrations (which should be grouped at the end), in a single series, with no omitted numbers.

Electronic form: A floppy disk together with the hard copy of the manuscript should be submitted. It is important to ensure that the diskette version and the printed version are identical. The diskette should be labelled with the following information: a) the operating system and word-processing software used, b) in case of UNIX media, the method of extraction (i.e. tar) applied, c) file name(s) related to manuscript. The diskette should be properly packed in order to avoid possible damage during transit.

Among various acceptable word processor formats, TEX and LATEX are preferable. The *Journal's* style file is available to authors.

Galley proofs: Proofs should be returned by authors as soon as possible. In other cases, the article will be proofread against manuscript by the editor and printed without the author's corrections. Remarks to the errata should be provided within two weeks after receiving the offprints.

The *copy of the „Journal"* shall be provided to each author of papers.

Copyright: Manuscript submitted to this journal may not have been published and will not be simultaneously submitted or published elsewhere. Submitting a manuscript, the authors agree to automatically transfer the copyright for their article to the publisher if and when the article is accepted for publication. The copyright comprises the exclusive rights to reproduce and distribute the article, including reprints and also all translation rights. No part of the present journal may be reproduced in any form nor transmitted or translated into a machine language without permission in written form from the publisher.

Biographies and photographs of authors are printed with each paper. Send a brief professional biography not exceeding 100 words and a gloss photo of each author with the manuscript.

On a method to authenticate and verify digital streams
B. J. Wysocki, Y. Wang, and R. Safavi-Naini — *Regular paper* — 45

Simple analysis of the impact of packet loss and delay on voice transmission quality
K. Perlicki — *Regular paper* — 53

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Editorial Office
tel. +48(22) 872 43 88
tel./fax:+48(22) 512 84 00
e-mail: redakcja@itl.waw.pl
http://www.itl.waw.pl/jtit