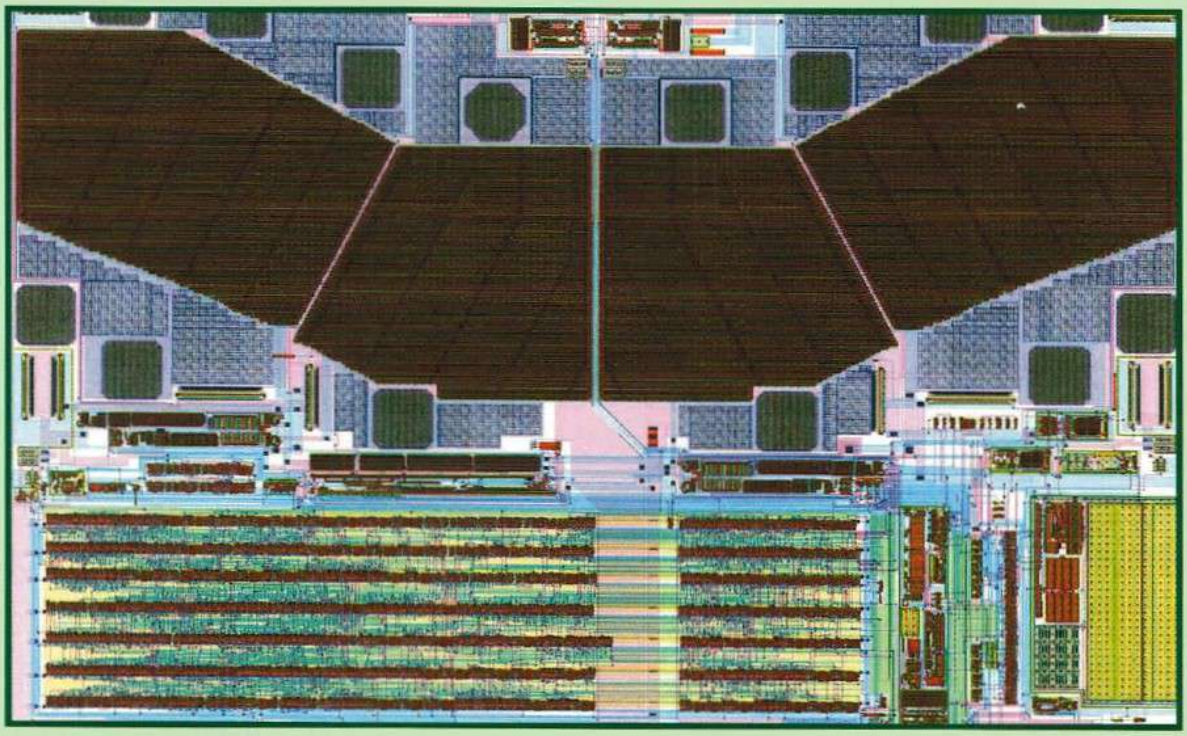
# híradástechnika

1945 VOLUME LX. 2005

## info-communications-technology



**Queueing models**

**Data mining**

**Protocols for wireless networks**

**Personal area communications**

Selected Papers 2005/12

Journal of the Scientific Association for Infocommunications with co-operation with the National Council of Hungary for Information and Communications Technology

# Contents



*Cover photo: Philips*

# Welcome

*szabo@hit.bme.hu*

This is the first English language issue that was put together by the new Editorial Board appointed in July this year. We thought that this issue would be a good opportunity to introduce our colleagues to the international readers. We hope that most names sound familiar to those working in the respective fields, not only in Hungary but at international level, as well.

- *István Bartolits* (National Communications Authority) – telecommunication management and regulation
- *István Bársony* (Research Institute for Technical Physics and Material Science of the Hungarian Academy of Sciences) – new communication devices and technologies
- *Levente Buttyán* (Budapest Institute of Technology and Economics – BUTE, Dept. of Telecommunications) – information security
- *Erzsébet Győri* (BUTE, Dept. of Telecommunications and Media Informatics) – telecommunication software
- *Sándor Imre* (BUTE, Dept. of Telecommunications) – mobile comunications and computing
- *Csaba Kántor* (Hungarian Telecom) – satellite and space communications
- *László Lois* (BUTE, Dept. of Telecommunications) – multimedia communications
- *Géza Németh* (BUTE, Dept. of Telecom. and Media Inform.) – speech processing, service automation
- *Géza Paksy* (BUTE, Dept. of Telecom. and Media Inform.) – optical telecommunications
- *Gergő Prazsák* (National Council for Communications and Information Technology) – society-related issues
- *István Tétényi* (Computer and Automation Institute of the Hungarian Academy of Sciences) – research networks and testbeds
- *Gyula Veszely* (BUTE, Dept. of Broadband Com.) – propagation, antennas, electromagnetic compatibility
- *Lajos Vondervisztt* (National Communications Authority) – the Internet and WWW

The new Editorial Board would like to pay tribute to György Lajtha who was editing our journal during the last 5 years. Professor Lajtha, the always youthful “great senior” of telecommunications, has played a critical role in renewing our journal and maintaining its quality over these years. His unprecedented professional experience and never ceasing interest in all areas of the wide scope of our journal made it possible to develop its technical level and bring its content to a wide comunity of readers. We can only hope that Prof. Lajtha will stay with us by giving valuable advice, and – needless to say – we’ll gladly publish his papers!

The present issue contains English versions of reviewed research papers, selected from the preceeding five Hungarian issues. We intend to continue with this practice, but we also welcome submissions intended directly for the English issues. We also hope we will be able to gradually increase their number from two per year to maybe four per year, yielding to shorter waiting times for those wishing to submit their results for these issues. Now let us briefly introduce the papers selected for this English issue.

Horváth and Telek investigate the class based weighted fair queueing (WFQ) used to model a number of computer and communication systems and present a very simple approach that provides a fast approximation for the queue length and waiting time measures.

Data mining is an area of increasing importance, with it’s goal to extract implicit, unknown and useful information from large data sets. Neural networks can be effectively used for nonlinear function approximation (regression) problems. The paper of Valyon and Horváth deals with some special types of networks, namely Support Vector Machines.

QoS requirements toward VoIP services necessitate adequate Operations and Maintenance (OAM) support, as well as a specialized fault management system. The study of Varga et al describes a complex fault management system customized for VoIP service providers.

In the paper by Buchholcz et al, we propose a new TCP variant (called TCP-ELN) which is capable to considerably improve the transfer rate over radio channels.

The paper of Hottmar deals with modelling of double-processor system by closed service network, presents a queueing system as a method of modelling and analyses the performance and quantifies the time characteristics of the processor systems.

Rónai et al propose a middleware called MAIPAN that provides a uniform computing environment for creating dynamically changing personal area networks (PANs). In this solution session transfers and dynamic session management are tightly integrated with strong and intuitive access control security.

The paper by Lois et al deals with multi-view video encoding and presentation. Based on the OpenGL system, the authors present a Depth Image-base Representation (DIBR) method to render an image from several existing reference pictures in a multi-view environment.

The provision of the sophisticated mobile services over IPv6 requires efficient transport protocols. The paper by Gál et al investigates the effect of mobility on the TCPv6 and UDPv6 protocols using comparative measurements.

Lastly, we would like to use this opportunity to wish a Happy and Prosperous New Year to our authors, reviewers and readers!

*László Zombory*
*President of the Editorial Board*

*Csaba A. Szabó*
*Editor-in-Chief*

# An Approximate Analysis of Two Class WFQ Systems

GÁBOR HORVÁTH, MIKLÓS TELEK

*Budapest University of Technology and Economics, Department of Telecommunications*
*{ghorvath, telek}@hit.bme.hu*

***Keywords: WFQ service, 2D Markov chain, expected value and deviation of delay***

*The class based weighted fair queueing (WFQ) is applied in a lot of computer and communication systems. It is a popular way to share a common resource. Its efficient analytical solution is an open question for a long time. Different solutions were proposed, using complex analysis or numerical techniques. But all of these methods have their limits in usability. In this paper we present a very simple approach that provides a fast approximation for the queue length and waiting time measures. Although it is simple and looks rough, the comparison with simulation shows that it provides reasonable accuracy with an execution time less than a second.*

## 1. Introduction

Class based weighted fair queueing is a service policy in multi-class systems. Consider a service station that provides its resources for customers belonging to different classes. The customers inside a class are waiting in an FCFS manner for their service. There are weights assigned to each class. The ratio of server capacity available for a class is given by the ratio of weights of the classes that are "active" (there are customers waiting in the queue belonging to that class). So the "importance" of the customers is regulated by the weight assigned to their class.

If the customer arrivals are according to Poisson processes, and service times are exponentially distributed, the system can be modeled by a "two dimensional" Markov chain.

There were many methods proposed to give the solution of this Markov chain. First we list numerical solutions. In [1],[2] the authors consider the same problem, but they call this kind of system *Coupled Processor Model.* They express the steady state probabilities of the two dimensional Markov chain as a power series of the load, and give a recursive way to compute the coefficients of the powers. With that approach only a few number of classes can be handled (2-3), and as the load tends to 1 a large number of coefficients have to be computed to reach a given accuracy.

An other approach ([3]) approximates the infinite model with a finite one, and uses a kind of Gaussian elimination to solve the finite Markov chain. During the Gaussian elimination the structure of the system is exploited, and reasonable speedup is achieved. But again, if the load is high, the reduced finite Markov chain has too many states and speed decreases fast.

In [4] the authors provide the generating function of the two dimensional Markov chain. The result is a two-variable complex (actually analytical) function, where the problem is to determine the one dimensional boundary generating functions. This gives the difficulty of this approach, since it needs Wiener-Hopf factorization.

In this paper we consider a two-class system, however the approach itself can be extended to more classes as well. Contrary to the solutions mentioned above, the arrival intervals and service times are given with two moments. Using this input two moments of the waiting time are approximated.

## 2. Concept of the Approximation

The concept is to approximate the 2-queue system as if the queues were separated, and construct a service process for each that approximately imitates the behaviour of the original server *(see Figure 1)*.

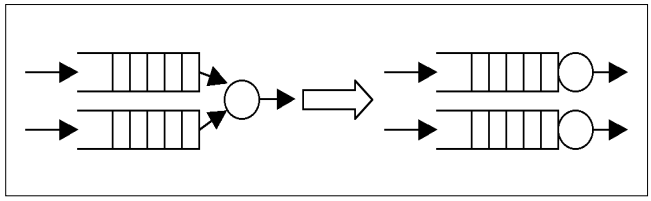


*Figure 1.Separation of the customer classes*

For example looking at queue 1 the server capacity is changing between the full capacity (*C*) and decreased capacity (according to the ratio of weights) depending on whether queue 2 is idle or busy. So the idea is simple: let's characterize the busy period process of queue 2, and construct a modulated server process for queue 1 where the modulation of the server is given by the busy period process of queue 2 *(see Figure 2)*.

As soon as the servers are separated, the queues are modeled by quasi birth-death processes (QBDs), and solved using matrix geometric techniques. In the Markov chain that models a queue the states are duplicated, corresponding to the idle or busy state of the other queue. In one state the customers can use the

full server capacity, in the other they receive reduced service rate according to weights. *Figure 3* shows the macro-structure of the Markov chain.

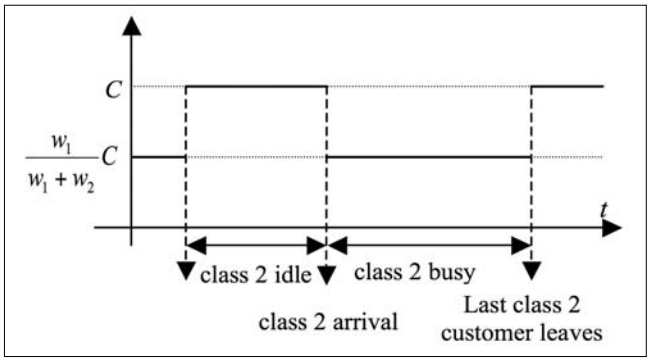



*Figure 2. The modulation in the service process*

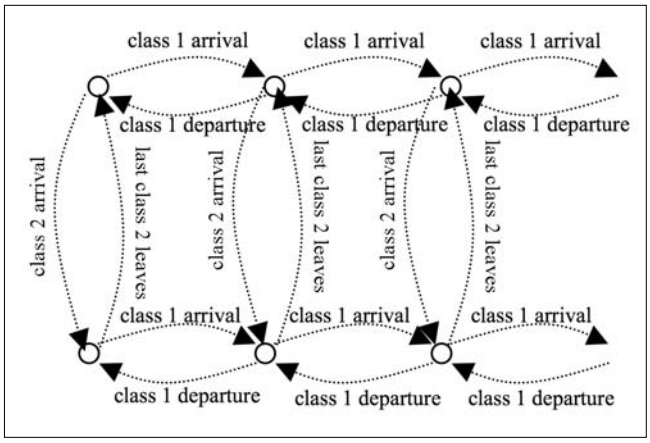



*Figure 3. Structure of the approximating Markov chain*

Before going into technical details, we summarize the steps of the algorithm to help understanding its structure.

- Construct PH representation of the arrival and service processes. The PH representation enables the use of matrix geometric methods.
- Compute the length of the busy periods. This step will produce the parameters of 4 PH random variables, since the computation detailed in that section has to be performed for queue 1 and queue 2, beside full capacity service and beside reduced capacity service.
- Construct the QBD that models queue 1 (depicted by Figure 3) and compute the waiting time parameters. Do the same for queue 2.

### A. Arrival and Service

There are two classes. Measures and notations associated to class *i* are denoted by superscript $^{(i)}$. The arrival process is characterized by the arrival intensity $\lambda^{(i)}$ and the squared coefficient of variation of the inter-arrival times $c_A^{2(i)}$. Based on these two moments a second order acyclic phase type distribution (APH, [6]) is constructed having the same moments. This PH random variable is described by its transient generator matrix $D^{(i)}$, the vector of absorbing transitions $d^{(i)}$, and initial probability vector $\delta^{(i)}$. It is easy to check that the following PH random variable has the proper moments:

$$D^{(i)} = \begin{bmatrix} -\frac{\lambda^{(i)}}{c_A^{2(i)}} & \frac{\lambda^{(i)}}{c_A^{2(i)}} \\ 0 & -2\lambda^{(i)} \end{bmatrix}, \quad d^{(i)} = \begin{bmatrix} 0 \\ 2\lambda^{(i)} \end{bmatrix}$$

$$\delta^{(i)} = \begin{bmatrix} \frac{1}{2c_A^{2(i)}} & 1-\frac{1}{2c_A^{2(i)}} \end{bmatrix}$$

The length of the job brought by the customers is characterized by its mean $m_l^{(i)}$ and squared coefficient of variation $c_l^{(i)}$. Instead of these measures we compute and use the service rate and its squared coefficient of variation of the queues, assuming full capacity (other queue is idle):

$$\mu_f^{(i)} = \frac{C}{m_l^{(i)}}, \quad c_{Sf}^{2\ (i)} = c_l^{2(i)},$$

and assuming reduced capacity (other queue is busy):

$$\mu_r^{(i)} = \frac{w_i}{\sum_i w_i}\frac{C}{m_l^{(i)}}, \quad c_{Sr}^{2\ (i)} = c_l^{2(i)},$$

where $C$ denotes the server capacity and $w_i$ denotes the weight assigned to class *i*.

Again, as above, a PH distribution is constructed from these parameters:

$$S_f^{(i)} = \begin{bmatrix} -\frac{\mu_f^{(i)}}{c_{Sf}^{2\ (i)}} & \frac{\mu_f^{(i)}}{c_{Sf}^{2\ (i)}} \\ 0 & -2\mu_f^{(i)} \end{bmatrix}, \quad s_f^{(i)} = \begin{bmatrix} 0 \\ 2\mu_f^{(i)} \end{bmatrix}$$

$$\sigma_f^{(i)} = \begin{bmatrix} \frac{1}{2c_{Sf}^{2\ (i)}} & 1-\frac{1}{2c_{Sf}^{2\ (i)}} \end{bmatrix}.$$

For the reduced capacity case ($S_r^{(i)}$, $s_r^{(i)}$, $\sigma_r^{(i)}$) are constructed similarly.

### B. Busy Period

Let us look at Figure 3 again. The transitions related to arrivals and services are already described, by PH distributions (see the last section). What is still missing is the busy period computation. We will compute two moments of the busy period of the queues as if they were isolated (without the impact of the presence of the other queue), beside full and reduced server capacity.

We have PH arrival and service process, so the Markov chain model of the queues in isolation has a special block-tri-diagonal, a so-called QBD structure:

$$Q = \begin{bmatrix} A_1' & A_0' & & & \\ A_2' & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \\ & & \ddots & \ddots & \ddots \end{bmatrix}$$

The blocks of the generator of queue i with full server capacity are the following (see [6]):

$$A_{0f}^{(i)} = d^{(i)}\delta^{(i)} \otimes I_{2x2}$$
$$A_{1f}^{(i)} = D^{(i)} \oplus S_f^{(i)}$$
$$A_{2f}^{(i)} = I_{2x2} \otimes s_f^{(i)}\sigma_f^{(i)}$$
$$A_{0f}^{(i)\prime} = d^{(i)}\delta^{(i)} \otimes \sigma_f^{(i)}$$
$$A_{1f}^{(i)\prime} = D^{(i)}$$
$$A_{2f}^{(i)\prime} = I_{2x2} \otimes s_f^{(i)}$$

The blocks assuming reduced capacity are obtained similarly. To compute the $k^{\text{th}}$ moment of the busy period ($m_{Bf}^{k\ (i)}$) generated by an arrival the following equation has to be solved:

$$m_{Bf}^{k\ (i)} = (-1)^k(\delta^{(i)} \otimes \sigma_f^{(i)})\frac{d^k}{ds^k}G_f^{(i)}(s)|_{s=0}\ h,$$

where $G_f^{(i)}(s)$ satisfies the following matrix equation:

$$sG_f^{(i)}(s) = A_{2f}^{(i)} + A_{1f}^{(i)}G_f^{(i)}s + A_{0f}^{(i)}(G_f^{(i)}(s))^2.$$

The *0*th derivative of $G_f^{(i)}(s)$ at *s=0* leads to the *fundamental matrix* geometric equation (for matrix *G*), see for example [6]. For the first derivative we have the following implicit equation:

$$\frac{d}{ds}G_f^{(i)}(s)|_{s=0} = \left(A_{1f}^{(i)} - A_{0f}^{(i)}G_f^{(i)}(0)\right)^{-1} \cdot \cdot\left(I - A_{0f}^{(i)}\frac{d}{ds}G_f^{(i)}(s)|_{s=0}\right)G_f^{(i)}(0)$$

which we solved by a fix point iteration.

In our approximation only two moments are utilized, they are the following (from the definition, after some algebra):

$$m_{Bf}^{1\ (i)} = -(\delta^{(i)} \otimes \sigma_f^{(i)})\left(A_{2f}^{(i)} + A_{0f}^{(i)}G_f^{(i)}(0)\right)^{-1}h$$
$$m_{Bf}^{2\ (i)} = 2(\delta^{(i)} \otimes \sigma_f^{(i)})\left(A_{2f}^{(i)} + A_{0f}^{(i)}G_f^{(i)}(0)\right)^{-1} \cdot \cdot\left(A_{0f}^{(i)}\frac{d}{ds}G_f^{(i)}(s)|_{s=0} - I\right)\cdot m_{Bf}^{1\ (i)}$$

Having these two moments the same PH fitting is performed like before, with the parameters of the obtained PH distribution denoted by $B_f^{(i)}$, $b_f^{(i)}$, $\beta_f^{(i)}$.

To build up the Markov chain of Figure 3, we will also need the phase probability vector of the arrival process at the moment when the busy period finishes. This is computed as:

$$\alpha_f^{(i)} = (\delta^{(i)} \otimes \sigma_f^{(i)}) \cdot G_f^{(i)}(0) \cdot (h_2 \otimes I_{2x2}).$$

### C. Queue Model

Now, we construct the block matrices of the QBD modeling queue *i* (its structure is depicted in Figure 3). The index of the other queue will be denoted by *j* (thus, if *i=1* then *j=2* and vice versa). As seen in Figure 3, the state space is divided into two parts.

In the upper part, where the other queue is idle, queue *i* receives full service capacity. In this part of the state space, one has to keep track (1) the phase of the arrival process of queue *i*, (2) the phase of service process of queue *i*, and (3) the phase of arrival process of queue *j*.

When a class *j* customer appears, the capacity of the server is shared between the two classes according to the weights. Thus, the arrival of a class *j* customer drives the Markov chain to the lower part. In this part phases (1) and (2) have to be kept track too, but the phase of the busy period of queue *j* has to be kept track instead of the phase of its arrival process. This behavior is reflected in the definition of the block matrices of the QBD:

$$C_0^{(i)} = \begin{bmatrix} d^{(i)}\delta^{(i)} \otimes I \otimes I & 0 \\ 0 & d^{(i)}\delta^{(i)} \otimes I \otimes I \end{bmatrix}$$
$$C_1^{(i)} = \begin{bmatrix} D^{(i)} \oplus D^{(j)} \oplus S_f^{(i)} & I \otimes d^{(j)}\beta_r^{(j)} \otimes I \\ I \otimes b_r^{(j)}\alpha_r^{(j)} \otimes I & D^{(i)} \oplus B_r^{(j)} \oplus S_r^{(i)} \end{bmatrix}$$
$$C_2^{(i)} = \begin{bmatrix} I \otimes I \otimes s_f^{(i)}\sigma_f^{(i)} & 0 \\ 0 & I \otimes I \otimes s_r^{(i)}\sigma_r^{(i)} \end{bmatrix}$$

The irregular matrices at level 0 are the following:

$$C_0^{(i)\prime} = \begin{bmatrix} d^{(i)}\delta^{(i)} \otimes I \otimes \sigma_f^{(i)} & 0 \\ 0 & d^{(i)}\delta^{(i)} \otimes I \otimes \sigma_r^{(i)} \end{bmatrix}$$
$$C_1^{(i)\prime} = \begin{bmatrix} D^{(i)} \oplus D^{(j)} & I \otimes d^{(j)}\beta_f^{(j)} \\ I \otimes b_f^{(j)}\alpha_f^{(j)} & D^{(i)} \oplus B_f^{(j)} \end{bmatrix}$$
$$C_2^{(i)\prime} = \begin{bmatrix} I \otimes I \otimes s_f^{(i)} & 0 \\ 0 & I \otimes I \otimes s_r^{(i)} \end{bmatrix}$$

The number of phases is 16, so the classical QBD solver algorithms can provide the steady state probabilities and waiting times quickly (see [6]).

## 3. Numerical Results

Below, we evaluate two examples to demonstrate the algorithm. In case 1 the moments of the job size of the two classes are similar and in the second case, the job size of customers of class 2 is 10 times longer.

The plots contain the results of both the analysis and discrete event simulation to allow the evaluation of the accuracy of the presented approximation.

**Example 1.**

On the first pair of plots the waiting time moments are depicted as a function of the load of class 1 *(Figure 4 and 5)*. As it is expected we obtained that the mean waiting time increases while its squared coefficient of variation decreases with increasing load.

The second scenario investigates the influence of job size variance. In the next two plots the squared coefficient of variation of the inter arrival and service times are changed, and again two moments of the waiting times are captured *(Figures 6, 7, 8 and 9)*. According to the results (both simulation and analysis) the increasing variance does only have a very small impact on the waiting time of the other queue.

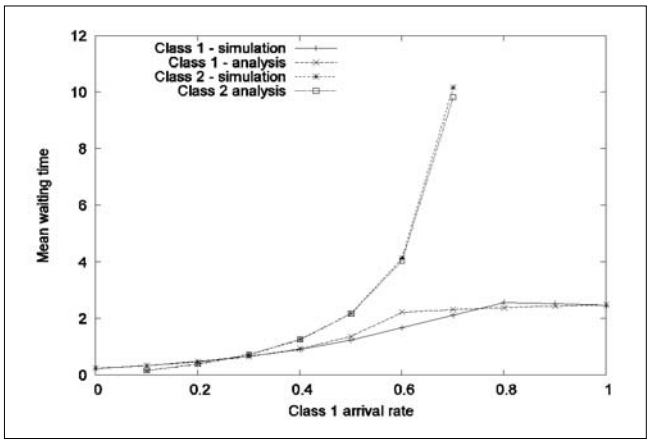



*Figure 4. Mean waiting time vs. class 1 arrival rate*

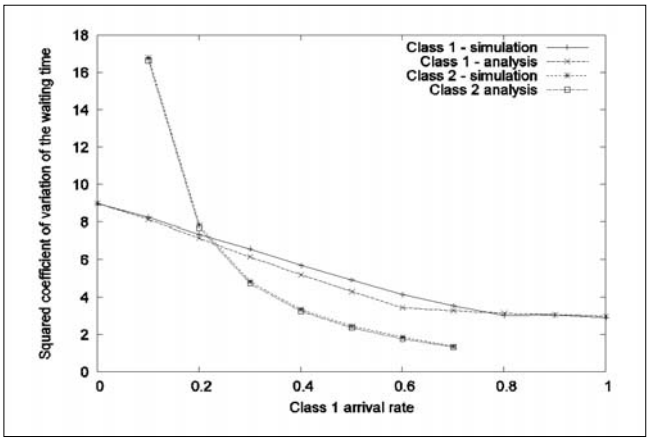



*Figure 5. Squared coefficient of variation of the waiting time vs. class 1 arrival rate*

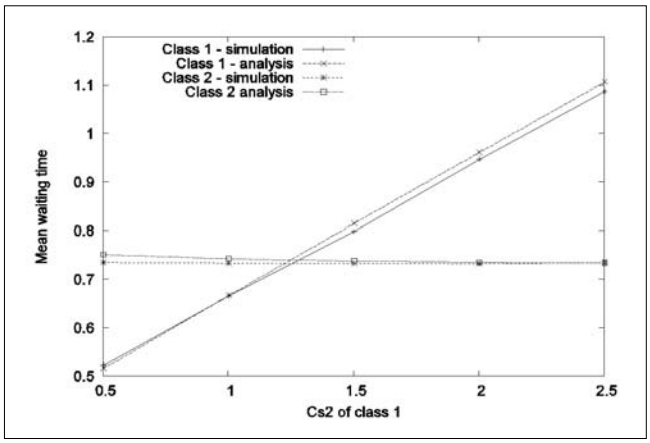



*Figure 6. Mean waiting time vs. the squared coefficient of variation of the class 1 service time*

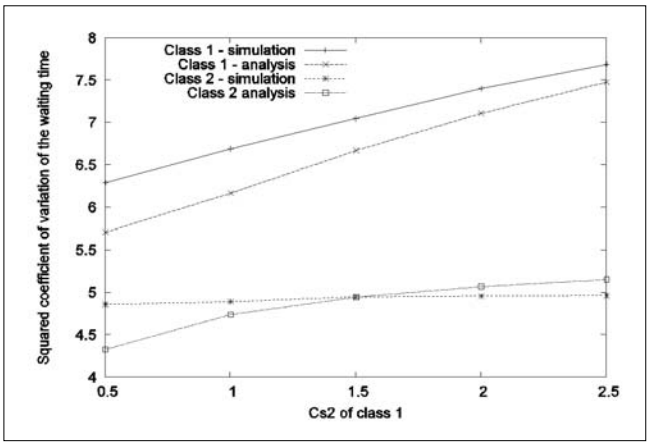



*Figure 7. Squared coefficient of variation of the waiting time vs. the squared coefficient of variation of the class 1 service time*

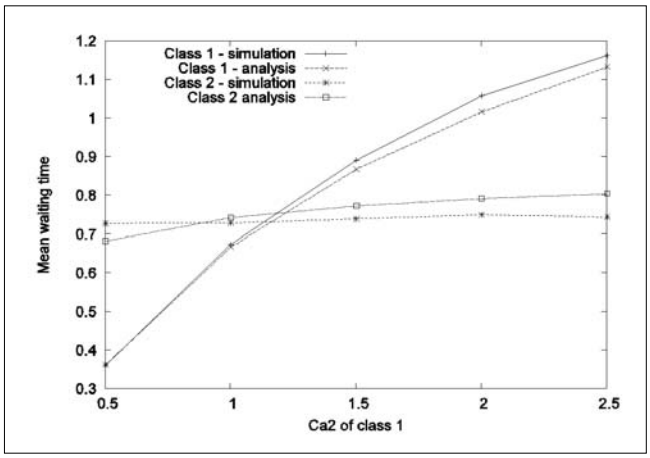



*Figure 8. Mean waiting time vs. the squared coefficient of variation of the class 1 inter-arrival time*

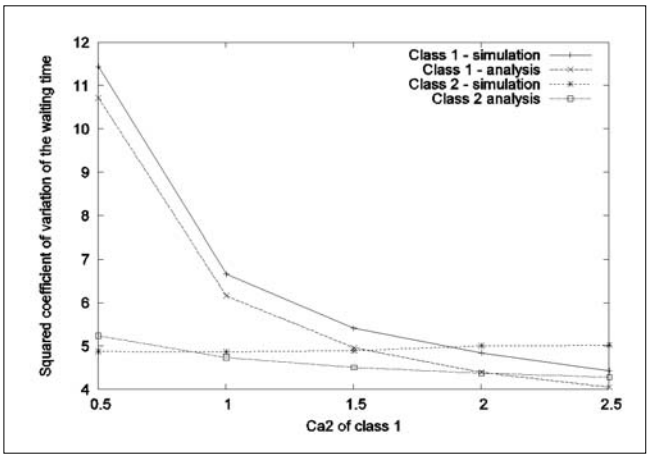



*Figure 9. Squared coefficient of variation of the waiting time vs. the squared coefficient of variation of the class 1 inter-arrival time*

Finally, the waiting time is depicted as a function of the weight *(Figures 10 and 11)*. Weight 0 means that the other queue has preemptive priority over the corresponding one. The results reflect this behavior.

**Example 2.**

In this case the mean job size of class 2 customers is 10 times larger. The first figure *(Figure 12)* shows the waiting time as a function of the load. Even when the 'large job' (class 2) queue is overloaded, customers in class 1 get their guaranteed service, as it is indicated by its low waiting time.

*Figures 14, 15, 16 and 17* show the effect of the variance on the waiting time. The simulation results show again that the waiting time of queue 2 does not get worse when increasing the variance of the inter arrival

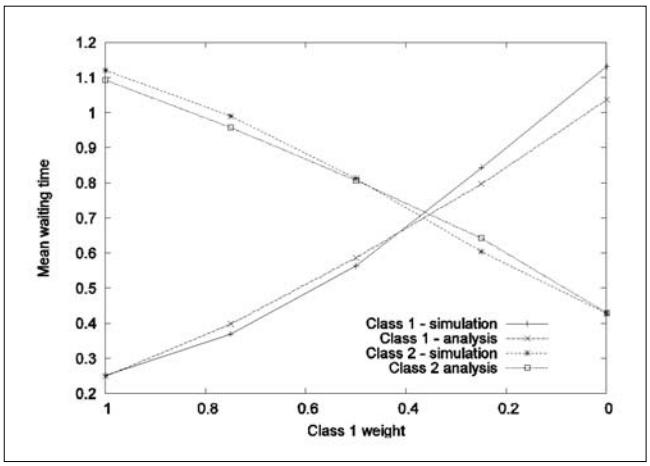



*Figure 10. Mean waiting time vs. the weight of class 1*

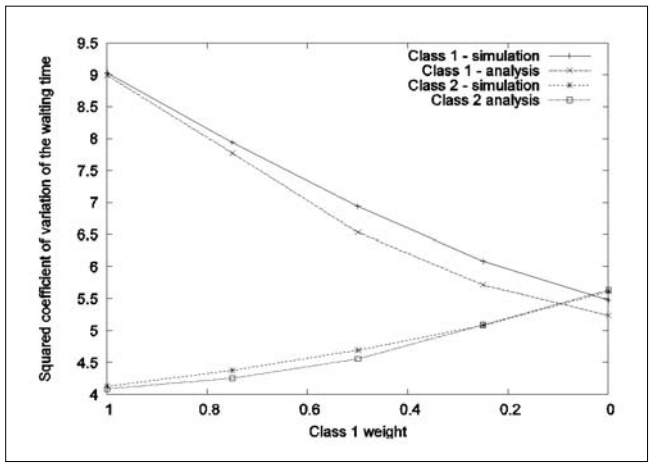



*Figure 11. Squared coefficient of variation of the waiting time vs. the weight of class 1*

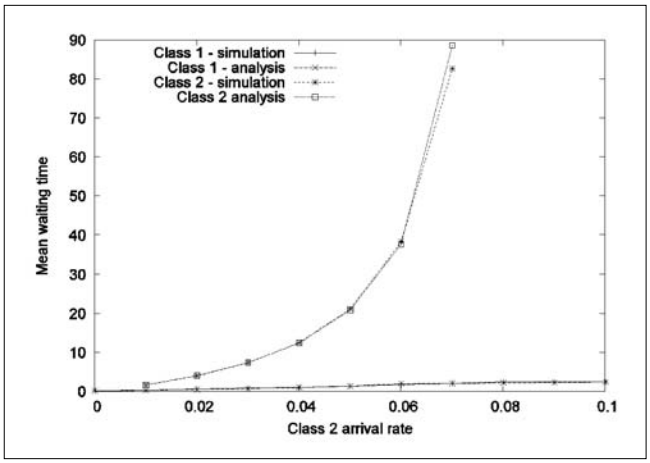



*Figure 12. Mean waiting time vs. class 2 arrival rate*

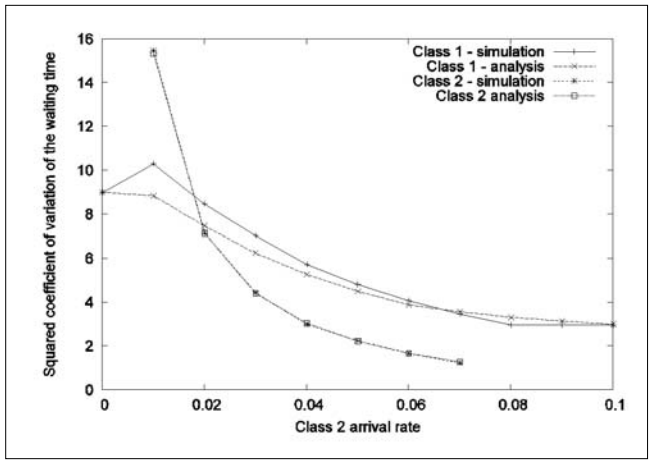



*Figure 13. Squared coefficient of variation of the waiting time vs. class 1 arrival rate*

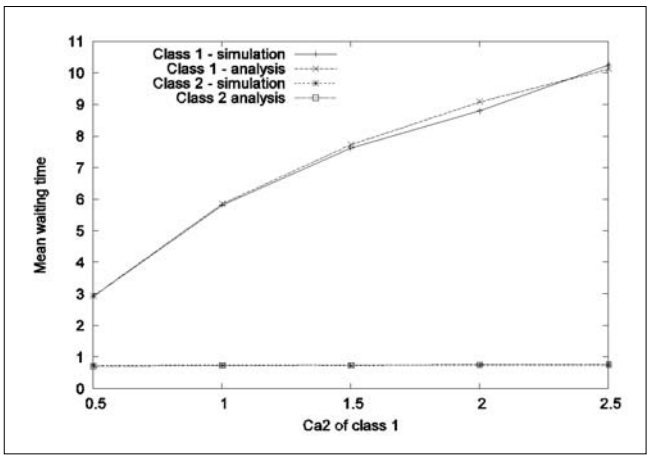



*Figure 14. Mean waiting time vs. the squared coefficient of variation of the class 1 inter-arrival time*

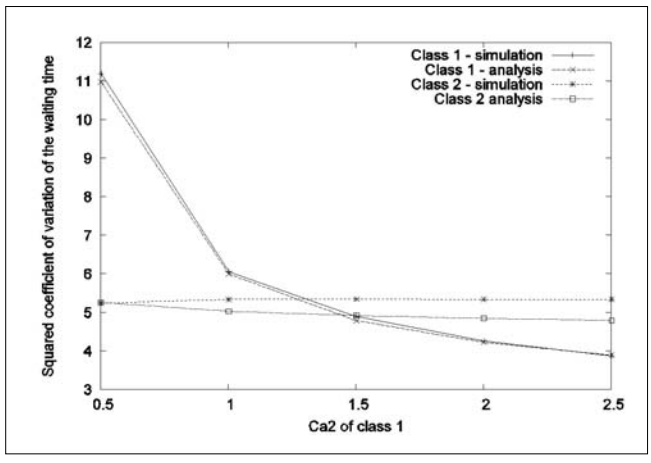



*Figure 15. Squared coefficient of variation of the waiting time vs. the squared coefficient of variation of the class 1 inter-arrival time*

or the service time of queue 1. The analysis shows a little correlation, but the difference from simulation remains reasonable.

In *Figures 18 and 19*, the weight of class 1 is changed. This case provides the worst approximation. We obtain reasonable differences with respect to the mean waiting time, but the squared coefficient of variation difference grows to 20%.

## 4. Conclusion

In this paper we presented an approximate performance analysis method for the two class weighted fair queueing system.

After the simplification of the structure of the Markov chain, we solved the upcoming queueing problem by matrix geometric methods. The advantage of our me-

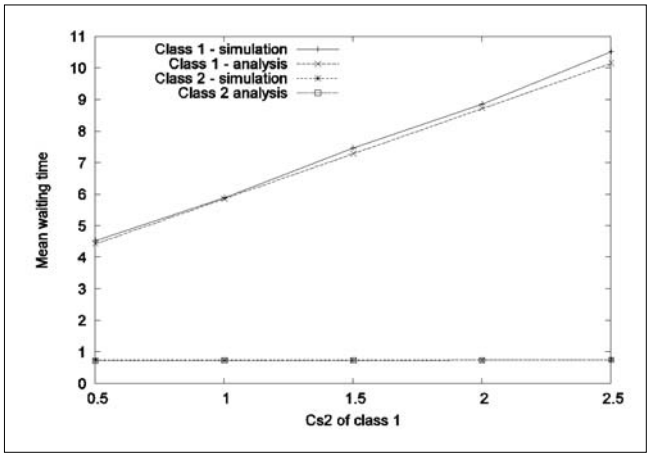



*Figure 16. Mean waiting time vs. the squared coefficient of variation of the class 1 service time*

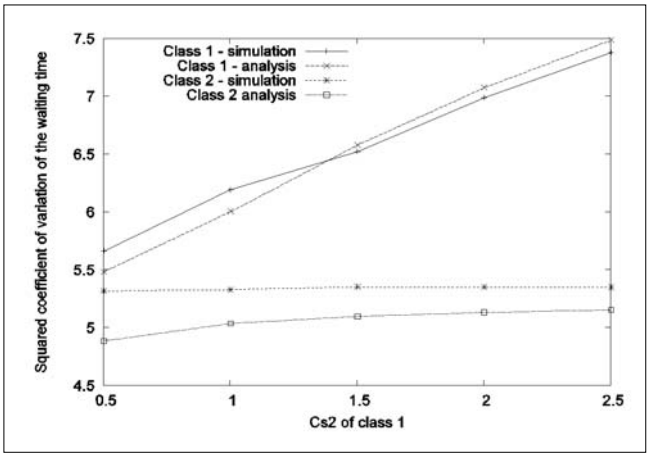



*Figure 17. Squared coefficient of variation of the waiting time vs. the squared coefficient of variation of the class 1 service time*

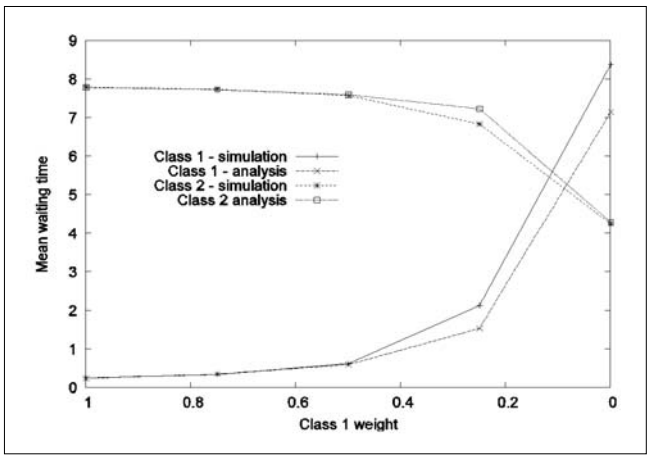



*Figure 18. Mean waiting time vs. the weight of class 1*

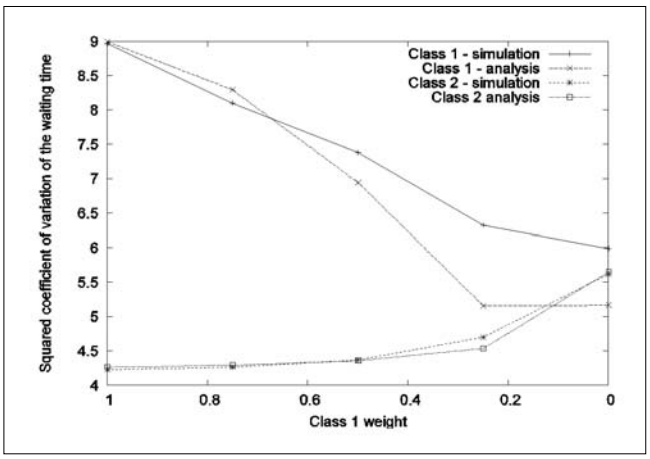



*Figure 19. Squared coefficient of variation of the waiting time vs. the weight of class 1*

thod is that its computation complexity is very low; and it is more general than most of the methods published so far since it also takes the variance of the arrival and service times in consideration.

We evaluated two numerical examples to examine the accuracy of the approximation exhaustively. In most of the cases the results were very close to the ones obtained by simulation.

The largest gap (15-20%) was experienced in the squared coefficient of variation of the waiting time, however we got reasonably good approximation for the mean waiting time in all of the cases.

## References

[1] G. Koole,
"On the power series algorithm", Technical Report, Centrum voor Wiskunde en Informatica, 1994.

[2] J. P. C. Blanc,
"A numerical study of the coupled processor model", in Computer Performance and Reliability, 1988.

[3] Leslie D. Servi,
"Algorithmic solutions to two-dimensional birth-death processes with application to capacity planning", Telecommunication Systems, Vol. 21, no.2, pp.205–212., 2002.

[4] F. Guillemin, R. Mazumdar, A. Dupuis, J. Boyer,
"Analysis of the fluid weighted fair queueing system", J. Appl. Probab., Vol. 40, no.1, pp.180–199., 2003.

[5] Guy Fayolle, Roudolf Iasnogorodski, Vadim Malyshev, Random Walks in the Quarter Plane, Springer-Verlag New York, 1999.

[6] G. Latouche and V. Ramaswami,
Intr. to Matrix Analysic Methods in Stochastic Modeling, American Statistical Association and the Society for Industrial and Applied Mathematics, 1999.

# Least Squares Support Vector Machines for Data Mining

JÓZSEF VALYON, GÁBOR HORVÁTH

*Budapest University of Technology and Economics,*
*Department of Measurement and Information Systems*

*{valyon, horvath}@mit.bme.hu*

***Keywords: Support Vector Machines, Least Squares Support Vector Machines, function approximation, time series prediction***

*Due to the extensive use of databases and data warehouses; it is very easy to collect large quantities of data from any aspect of life. The analysis of these data may lead to many useful or interesting conclusions. Data mining is a collection of methods and algorithms that can effectively be used to discover implicit, previously unknown, hidden trends, relationships, patterns in masses of data. Classical data mining uses many methods from different fields, like the foundations of linear algebra, graph theory, database theory, machine learning and artificial intelligence. In this paper, we address the problem of black-box modeling, where the model is built based on the analysis of input-output data. These problems usually cannot be addressed analytically, so they require the use of soft computing methods. We focus on the use of a special type of Neural Network (NN), the least squares version of Support Vector Machines (SVMs), the LS-SVM. We also present an extension of this method, the LS$^2$-SVM, which enables us to deal with large quantities of data. Using this method we present solutions for function approximation, time series analysis, and time series prediction.*

## 1. Introduction

Most of the real-life problems concern very complex systems, whose exact properties, inside functioning and operation are unknown. Such problems are often found in the field of medical research and diagnosis, in industrial and economic problems, etc. During the operation of such systems a large number of input and output data samples may be collected, which can be processed to extract knowledge or to create a model of the system. This is called black-box modeling.

The goal of data mining is to extract implicit, previously unknown and useful information from large data sets accumulated in databases or data warehouses.

The tools most commonly used in data mining contain many different soft computing techniques, inclusive of neural networks [1],[2]. Neural networks can be effectively used for nonlinear function approximation (regression) problems. This paper deals with some special types of networks, namely Support Vector Machines [3],[4]. The main idea of support vector machines is to map the complex nonlinear primal problem – or to be exact, the data samples – with the use of nonlinear transformations into a higher dimensional space, where a linear solution can be found.

The main advantage of this method is that it guarantees an upper limit on the generalization error of the resulting model. Another important property of the method is that the training algorithm seeks to minimize the model size and creates a sparse model. This is a trade-off problem between model complexity and the approximation error that can be is controlled by a hyper parameter.

The biggest problem with the traditional SVM is its high algorithmic complexity and memory requirement, caused by the use of a quadratic programming. This shuts out large datasets and data mining. Many different solutions have been introduced to overcome this problem. These are mostly iterative solutions, breaking down the large optimization problem into a series of smaller tasks. The different “chunking” algorithms differ in the way they decompose the problem [5-7].

Another possibility to use the Least Squares Support Vector Machine, where the algorithmic problems are tackled by replacing quadratic programming by a simple matrix inversion. The price paid for this simplicity is the loss of sparseness, thus all samples are embodied in the resulting – and, therefore, large – model.

Data mining problems incorporate very large data sets; therefore it is extremely important that – in contrast to traditional LS-SVM – the size of the resulting model must be independent of the training set size.

In the sequel, we propose some modifications for the LS-SVM that allows us to control the network size, while at the same time they simplify the formulations, speed up the calculations and/or provide better results.

In the field of black-box modeling, there are two general problem types that must be discussed: (a) *function approximation* (regression), and in case of dynamic systems (b) *time series prediction*.

#### Function approximation

The available data set is analyzed to find, and determine mathematical relations among them. In case of $N$ $p$-dimensional samples one must find how one of the variables (which is mostly the known output) depends on the $p$-1 others (or any subset of them). In this case there is no ordering (sequence – e.g. time) in the data, the system is expected to be static, thus the output depends only on the actual inputs.

**Time series prediction**

It is assumed, that the system has a memory (e.g. it contains feedback connections), so it is dynamic. This means that in time series prediction problems, the input and output data must have an ordering and the model constructed should represent the dynamics of the process. This way a proper model can continue a data series by predicting the future values.

This problem can be generalized, since the prediction may be done along any variable (not only the time). Most of the real life systems are dynamic, where the output depends not only on the inputs, but the current state of the system. This case can also be handled as a regression (function approximation), but the input variables are extended to include earlier inputs and/or outputs.

Section 2 shows a common way to convert a time series prediction problem to a function approximation one. Section 3 describes the LS-SVM regression, and the proposed LS-LS-SVM (the LS$^2$-SVM). Section 4 of this paper contains some experiments and then finally the conclusions are drawn.

## 2. Creating a data set for dynamic problems

In a time series prediction problem, a series of output values changing in time must be predicted, based on some earlier values and sometimes some other inputs. In cases like this, the approximating model must also have some dynamics. The easiest way to achieve this is to take a (usually nonlinear) static model and extend it with some dynamic components (e.g. delays or feedback paths). Probably the most common solution for adding external dynamic components is the use of tapped delay lines, as shown below.

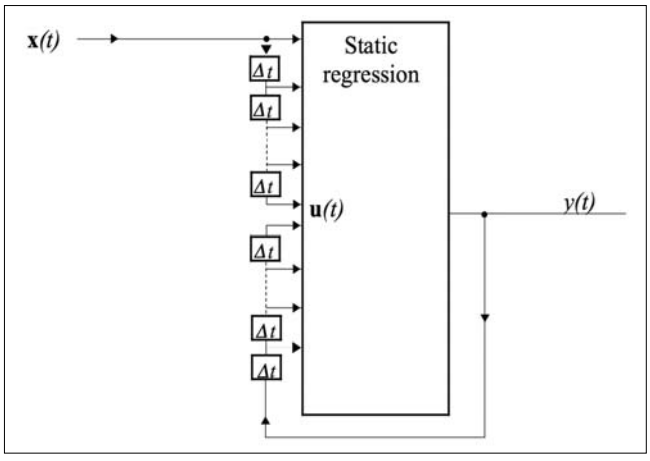



*Figure 1. A static system that is made dynamic by delays*

As it can be seen in the figure, the static system can be made dynamic, by extending the inputs with delays. The $N$ dimensional $\mathbf{x}$ vectors are expanded to $N+K$ dimensions, by incorporating certain past input and output values in the new model input.

$$\mathbf{u}^T(t) = [\, x_1(t), x_1(t - T_{11}), \ldots, x_1(t - T_{K_1 1}), \ldots, x_i(t), x_i(t - T_{1i}), \ldots, x_i(t - T_{K_i i}), \ldots, y(t - T_{1Y}), \ldots, y(t - T_{K_Y Y})\,] \quad (1)$$

Where

$x_i(t)$ is the $i$-th input in the $t$-th time step,
$x_i(t - T_{1i})$ is the $i$-th input in the $t$-$T_{1i}$ time step,
$[T_{1i} \ldots T_{K_i i}]$ is the collection of delays ($i$=1...$N$) for the $i$-th input,
$K_i$ is the number of delays for the $i$-th input,
$\mathbf{u}^T(t)$ is the $N$+$K$ dimensional vector at time step $t$,
$[T_{1Y} \ldots T_{K_Y Y}]$ is the collection of delays for the $y$ output,
$K_i$ is the number of delays for the $y$ output,
$y$ is the output.

The method described above is very general, since it allows different delays for all inputs and the output. In practice this is usually simplified:

– by the use of the same delays for all input components,
– by using a sliding window (e.g. the last n samples) instead of custom delays,

With the use of this extended input vector, the originally dynamic problem can be handled by a static regression. After training in the recall phase the model's output is also calculated based on an extended input vector. This means that the result must be calculated iteratively, since the previous results may be needed as an input to predict the next output.

## 3. Function approximation

There are a large number of different methods for function approximation. These methods can be organized by many different aspects, but based on the field of use and the implementation, the following characterization is emphasized:

- Linear regression
- Non-linear regression

This categorization is especially important, because the kernel based methods discussed here transform the nonlinear regression problem to address it in a linear manner as a much easier linear regression.

The problem is the following in both cases:

Given the $\{\mathbf{x}_i, d_i\}_{i=1}^N$ training data set, where $\mathbf{x}_i \in \Re^p$ represents a $p$-dimensional input vector and $d_i \in \Re$ is the scalar target output, our goal is to construct a $y = f(\mathbf{x})$ function, which represents the dependence of the output $d_i$ on the input $\mathbf{x}_i$.

Besides the well known analytic methods (linear/polynomial regression, splines etc.), the function approximation problem can also be solved by neural networks [1],[2]. The described LS-SVM can be considered as a special kind of neural solution [8].

As mentioned earlier, Support Vector Machines use nonlinear transformations to transform the input sam-

ples to a higher dimensional space, where a linear solution is constructed. Whilst doing this, the SVM selects a subset of the samples – the support vectors – as relevant for the solution and discards the rest. This property is called sparseness [9].

In order to simplify the calculations, the LS-SVM sacrifices this, therefore its model is often unnecessarily large. In this paper a reduction method is proposed, which enables us to achieve a sparse solution, while at the same time the algorithmic complexity is further reduced [10],[11].

### 3.1. LS-SVM regression

The Least Squares SVM (LS-SVM) is a modification of Vapnik's traditional Support Vector Machines (SVMs) [3]. In this formulation the solution is obtained by solving a linear set of equations, instead of solving a quadratic programming problem involved by standard SVM. The main advantage of this is that algorithmic complexity is reduced.

The solution is formed $y(\mathbf{x}) = \mathbf{w}^T \varphi(\mathbf{x}_i) + b$,

where $\varphi(.): \Re^n \rightarrow \Re^{n_h}$ is a mostly non-linear function, which maps the input data into a higher – possibly infinite – dimensional feature space. The data is transformed into a higher dimensional space, where the solution is calculated according to the (**w** and *b*) parameters resulting from the linear solution of the training. To minimize the generalization error, an additional constraint – the minimization of the length of **w**, that is $w^T w$ – is introduced.

The optimization problem and the inequality constraints are defined by the following equations ($i$ = 1,...,$N$):

$$\min_{\mathbf{w},b,e} J_p(\mathbf{w},e) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\frac{1}{2}\sum_{i=1}^{N} e_i^2 \qquad (2)$$

$$\text{with constraints: } d_i = \mathbf{w}^T\varphi(\mathbf{x}_i) + b + e_i \,.$$

The $C \in \Re^+$ is the trade-off parameter between a smoother solution, and training errors. From this, a Lagrangian is formed (3):

$$L(w,b,e;\alpha) = J_p(w,e) - \sum_{i=1}^{N} \alpha_k \left\{ \mathbf{w}^T \varphi(\mathbf{x}_i) + b + e_i - d_i \right\}$$

The solution concludes in a constrained optimization, where the conditions for optimality lead to the following overall solution:

$$\begin{bmatrix} 0 & \vec{\mathbf{1}}^{\mathbf{T}} \\ \vec{\mathbf{1}} & \boldsymbol{\Omega} + C^{-1}\mathbf{I} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{d} \end{bmatrix},$$
$$d = [d_1, d_2, ..., d_N]., \ \boldsymbol{\alpha} = [\alpha_1, \alpha_2, .., \alpha_N] \,, \qquad (4)$$
$$\vec{\mathbf{1}} = [1, ..., 1] \,, \ \Omega_{i,j} = K(\mathbf{x}_i, \mathbf{x}_j) = \varphi^T(\mathbf{x}_i)\varphi(\mathbf{x}_j)$$

where $K(\mathbf{x}_i,\mathbf{x}_j)$ is the kernel function, and $\boldsymbol{\Omega}$ is the kernel matrix.

The result is:

$$y = \sum_{i=1}^{N} \alpha_i K(\mathbf{x}, \mathbf{x}_i) + b \qquad (5)$$

Where $\alpha_k$ and $b$ come from the solution of Eq. 5. It can be seen that the achieved solution is linear but the nonlinear mapping is replaced by a new $K(.,x_i)$ kernel function, which is obtained as the dot product of the φ(.)-s.

The training of a support vector machine is a series of mathematical calculations, but the equation used for determining the machine's answer for a given input represents similar calculations as those of a one hidden layer neural network. Although in practice SVMs are rarely formulated as actual networks, this neural interpretation is important, because it provides an easier discussion framework than the purely mathematical point of view. *Figure 2* illustrates the neural interpretation of an SVM.

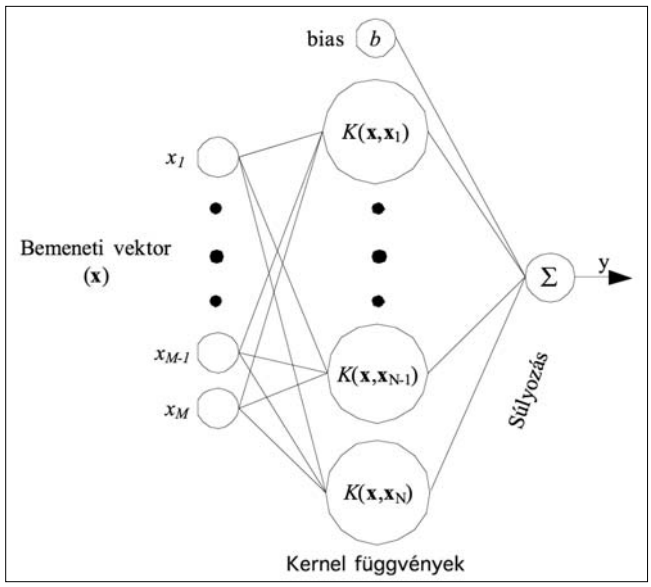



*Figure. 2. A neural network that corresponds to an SVM*

The input is an *M*-dimensional vector. The hidden layer implements the mapping from the input space into the kernel space, where the number of hidden neurons equals to the number of selected training samples, the support vectors. In this interpretation, the network size of a standard SVM – because of the sparseness of this solution – is determined by the number of support vectors, which is usually much smaller than the number of all training samples (in case of LS-SVM the number of neurons equals to the number of training samples *N*).

The response of the network (y) is the weighted sum of the hidden neurons' outputs, where the $\alpha_i$ weights are the calculated Lagrange multipliers.

### 3.2. LS²-SVM regression

The main goal of an LS²-SVM is to reduce model complexity, to reduce the number of neurons in the hidden layer. As a side effect, the algorithmic complexity of the required calculation is also reduced.

The starting point of the new method is Eq. 4 which is modified by the following two steps: *(i)* The *first step* reformulates the LS-SVM solution to use only a subset of the training samples as "support vectors'. We also

show that the resulting overdetermined system can still be solved. (ii) In the *second step* an automatic "support vector" selection method is proposed.

**Using an overdetermined equation set**

If the training set consists of $N$ samples, then our original linear equation set will have ($N$+1) unknowns, the $\alpha_i$-s and $b$, ($N$+1) equations and ($N$+1)$^2$ coefficients. These factors are mainly the values of the $K(\mathbf{x}_i,\mathbf{x}_j)$ $i,j$ = 1,...,$N$ kernel function calculated for every combination of the training input pairs. The cardinality of the training set therefore determines the size of the kernel matrix, which plays a major part in the solution, as algorithmic complexity; the complexity of the result etc. depends on this.

Let's take a closer look at the linear equation set describing the regression problem. The first row means:

$$\sum_{k=1}^{N}\alpha_k = 0\,, \tag{6}$$

and the $j$-th row stands for the:

$$b+\alpha_1 K(\mathbf{x}_j,\mathbf{x}_1)+\ldots+\alpha_k\left[K(\mathbf{x}_j,\mathbf{x}_k)+C^{-1}\mathbf{I}\right]+\ldots$$
$$\ldots+\alpha_N K(\mathbf{x}_j,\mathbf{x}_N)=d_j \tag{7}$$

condition.

To reduce the equation set, columns and/or rows may be omitted.

- If the $k$-th column is left out, then the corresponding $\alpha_k$ is also deleted. therefore the resulting model will be smaller. The $\sum_{i=1}^{N}\alpha_i = 0$ condition automatically adapts,
  since the remaining $\alpha$-s will still add up to zero.
- If the $j$-th row is deleted, then the condition defined by the $(\mathbf{x}_j,d_j)$ training sample is lost, because the $j$-th equation is removed.

The most important component of the main matrix is the $\mathbf{\Omega}+C^{-1}\mathbf{I}$ kernel matrix; its elements are the results of the kernel function for pairs of training inputs ($\Omega_{i,j}$= $K(\mathbf{x}_i,\mathbf{x}_j)$).To reduce the number of elements in $\mathbf{\Omega}$, one column, one row, or both (a column and a corresponding row) may be eliminated. The rows, however, represent the constraints (input-output relations) that the solution must satisfy. The following two reduction techniques can be used on the regularized $\mathbf{\Omega}+C^{-1}\mathbf{I}$ matrix:

*Traditional full reduction* – a training sample $(\mathbf{x}_k,d_k)$ is fully omitted, therefore both the column and the row corresponding to this sample are eliminated. The next equation demonstrates how the equation changes by fully omitting some training points. The deleted elements are colored grey.

$$\begin{bmatrix} 0 & \vec{1} \\ \vec{1} & \mathbf{\Omega}+C^{-1}\mathbf{I} \end{bmatrix}\begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix}=\begin{bmatrix} 0 \\ \mathbf{d} \end{bmatrix} \tag{8}$$

In this case, however, reduction also means that the knowledge represented by the numerous other samples are lost. This is exactly the case in traditional LS-SVM pruning since pruning iteratively omits some training points. The information embodied in these points is entirely lost. To avoid this information loss, one may use the technique referred here as partial reduction.

*Proposed partial reduction* – a training sample $(\mathbf{x}_j,d_j)$ is only partially omitted, by eliminating the corresponding $j$-th column, but keeping the $j$-th row. The corresponding input-output relation is still in effect, which means that the weighted sum of that row should still meet the $d_j$ (regression) goal, as closely as possible.

By selecting some (e.g. $M$, $M < N$) vectors as *"support vectors"*, the number of columns is reduced, resulting in more equations than unknowns. The effect of this reduction is shown in the next equation, where the removed elements are colored grey.

$$\begin{bmatrix} 0 & \vec{1} \\ \vec{1}^T & \mathbf{\Omega}+C^{-1}\mathbf{I} \end{bmatrix}\begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix}=\begin{bmatrix} 0 \\ \mathbf{d} \end{bmatrix} \tag{9}$$

As a consequence of partial reduction, our equation set becomes overdetermined, which can be solved as a linear least-squares problem, consisting of only ($M$+1)x($N$+1) coefficients. Let's simplify the notations of our main equation as follows:

$$\mathbf{A}=\begin{bmatrix} 0 & \vec{1}^{\mathrm{T}} \\ \vec{1} & \mathbf{\Omega}+C^{-1}\mathbf{I} \end{bmatrix},\quad \mathbf{u}=\begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix},\quad \mathbf{v}=\begin{bmatrix} 0 \\ \mathbf{d} \end{bmatrix}. \tag{10}$$

then the solution will be

$$\mathbf{u}=\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{v}\,. \tag{11}$$

The omission of columns with keeping the rows means that the network size is reduced; still all the known constraints are taken into consideration. This is the key concept of keeping the quality, while the equation set is simplified.

The modified, reduced LS-SVM equation set is solved in a least squares sense, therefore we call this method Least Squares LS-SVM or shortly LS$^2$-SVM. This proposition resembles to the basis of the Reduced Support Vector Machines (RSVM) introduced for standard SVM in [12]. The RSVM also selects a subset of the samples as possible delegates to be support vectors, but the selection method, the solution applied and the purpose of this reduction differs from the propositions presented. Since SVM is inherently sparse, the purpose of this selection is to reduce the algorithmic complexity, while our main goal is to achieve a sparse LS-SVM.

**Selecting support vectors**

To achieve sparseness by the above described partial reduction, the linear equation set has to be redu-

ced in such a way, that the solution of this reduced (overdetermined) problem is the closest to what the original solution would be.

As the matrix is formed from columns, we can select a linearly independent subset of column vectors and omit all others, which can be formed as linear combinations of the selected ones. This can be done by finding a "basis" (the quote indicates, that this basis is only true under certain conditions defined later) of the coefficient matrix, because the basis is by definition the smallest set of vectors that can solve the problem. The linear dependence discussed here, does not mean exact linear dependence, because the method uses an adjustable tolerance value when determining the "resemblance" (parallelism) of the column vectors. The use of this tolerance value is essential, because none of the columns of the coefficient matrix will likely be exactly dependent (parallel).

The tolerance parameter indirectly controls the number of resulting basis vectors ($M$). This number does not really depend on the number of training samples ($N$), but only on the problem, since $M$ only depends on the number of linearly independent columns. In practice it means that if the complexity of a problem requires $M$ neurons, then no matter how many training samples are presented, the size of the resulting network does not change.

The reduction is achieved as a part of transforming the $\mathbf{A}^T$ matrix into reduced row echelon form, using a slight modification of Gauss-Jordan elimination with partial pivoting [13],[14].

The algorithm uses elementary row operations:

- Interchange of two rows.
- Multiply one row by a nonzero number.
- Add a multiple of one row to a different row.

The algorithm goes as follows:

1. Work along the main diagonal of the matrix starting at row one, column one ($i$-row index, $j$-column index).
2. Determine the largest element $p$ in column $j$ with row index $i \geq j$.
3. **If** $p \leq \varepsilon'$ (where $\varepsilon'$ is the tolerance parameter) then zero out the elements in the $j$-th column with index $i \geq j$ ; **else** *remember the column* index ($j$) because we found a basis vector (support vector). If necessary move the row, to have the pivot element in the diagonal and divide the row with the pivot element $p$. Subtract the right amount of this row from all rows below this element, to make their entries in the $j$-th column zero.
4. Step forward to the next diagonal element ($i = i+1$, $j = j+1$). Go to step 2.

This method returns a list of the column vectors which are linearly independent form the others considering a tolerance $\varepsilon'$.

The problem of choosing a proper $\varepsilon'$ resembles the selection of Vapnik's SVM hyper-parameters, like of $C$, $\varepsilon$ and the kernel parameters. One possibility is to use cross-validation. With a larger tolerance value, we can achieve smaller networks, but consequently the error of the estimation grows.

It is easy to see that the selection of the $\varepsilon'$ tolerance is a trade-off problem between network size and performance. It is also important to emphasize that by using 0 tolerance, LS$^2$-SVM and LS-SVM are equivalent, since all of the input samples will be kept by the described selection method.

## 4. Experiments

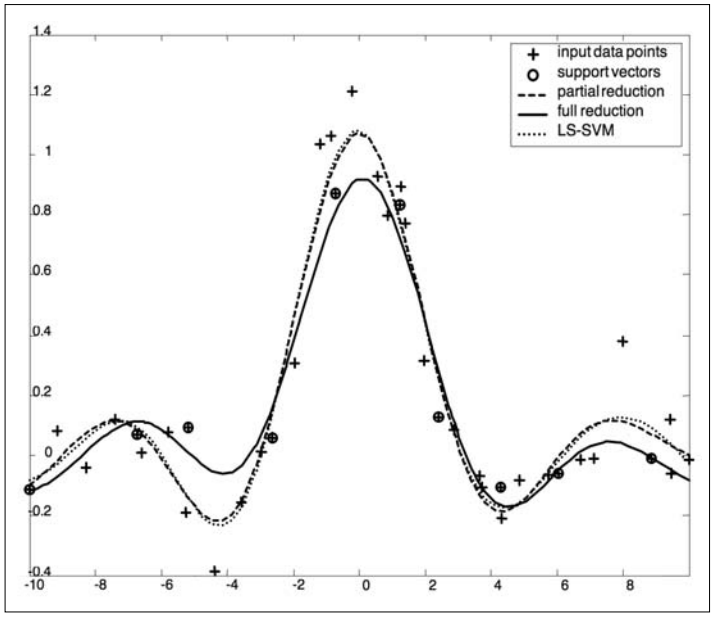
The results will be demonstrated on the most commonly used benchmark problem in the literature of LS-SVM. Most of the experiments where done with the *sinc(x)* function in the [-10,10] domain. The kernel is Gaussian like (RBF kernel), where $\sigma = \pi$. The tolerance ($\varepsilon'$) is set to 0.2 and $C$ =100.

The samples are corrupted by additive Gaussian output noise of zero mean and standard deviation of 0.1.

First the effects of partial reduction are examined. This is extended with the automatic selection method. The same problem is used to compare the described methods to the original solutions (LS-SVM and pruned LS-SVM), but a more complex time series prediction problem – the Mackey-Glass chaotic time series prediction problem – is also described.

To compare the reduction methods first an extremely simple support vector selection method is applied: every forth input is chosen from the 40 samples (the 40 sample points are randomly selected from the domain). *Figure 3* shows the result of the partial- and the full reduction plotted together, along with the original – unreduced – LS-SVM.

*Figure 3.*
*The different reduction methods plotted together*



It can be seen that the partial reduction gave the same quality result as the original LS-SVM, while in this case the complexity is reduced to its one fourth. The fully reduced solution is only influenced by the "support vectors", which can be easily seen on the figure. In this case the resulting function is burdened with much more error.

The original unreduced LS-SVM almost exactly covers the partial reductions' dotted line ($MSE_{partial\ red.}$: $1.04x10^{-3}$, $MSE_{full\ red.}$: $6.43x10^{-3}$, $MSE_{LS\text{-}SVM}$: $1.44x10^{-3}$).

The "support vectors" may be selected automatically, by the use of the proposed selection method. *Figure 4* shows a solution that is based on the automatically selected support vector set.

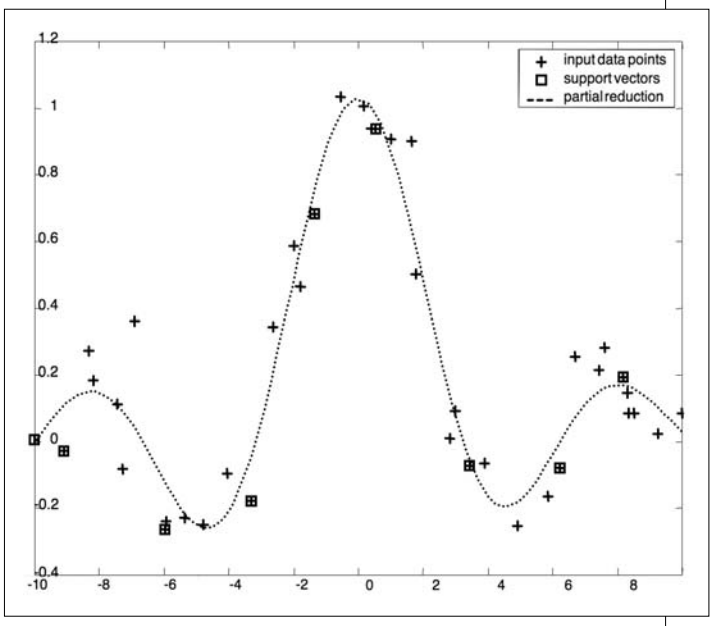



*Figure 4.*
*A partially reduced LS-SVM, where the support vectors were selected by the proposed method (ε'= 0.2)*

Since 40 samples were provided, the original network would have 40 nonlinear neurons, while the reduced net incorporates only 9. An even more appealing property of the proposed solution is that the cardinality of the support vector set is indeed independent from the number of training samples. If the problem can be solved with *N* neurons in the hidden layer, then no matter how many inputs are presented, the network size should not change.

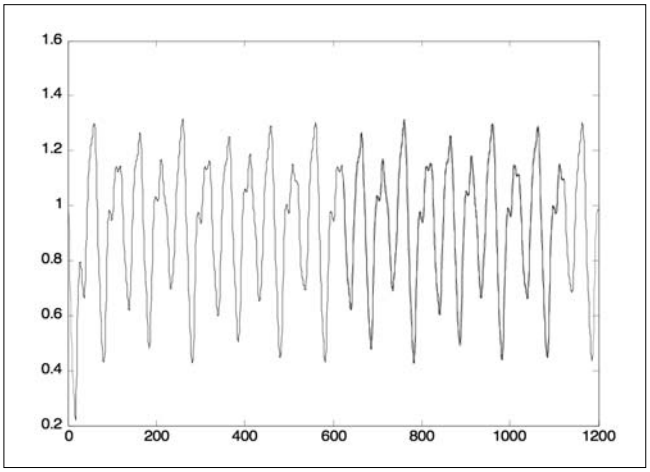


*Figure 5.*
*An LS-SVM approximation of the Mackey-Glass time series prediction problem*

The *Table* shows the number of "support vectors" calculated by the algorithm for different training set sizes of exactly the same problem (*sinc*(*x*) with matching noise etc. parameters). The mean square errors tested for 100 noise-free samples are also shown. It can be seen that by increasing the number of training samples, the error decreases – as expected –, but the network size does not change significantly.

The next figure *(Figure 5)* shows a solution to the widely used Mackey-Glass time series prediction problem. In the prediction we have used the [-6,-12,-18,-24] delays, thus the **x**(*t*) value of the *t*-th time instant is approximated by four past values (in Mackey-Glass process, there is no input – the output only depends on the past values). In this experiment the training is done by using 500 training samples.

This experiment shows that the above/described solution along with LS-SVM regression is applicable to solve time series prediction problems.

## 5. Summary

In this paper, the possibilities of data mining applications of LS-SVMs are investigated. We have shown the original formulation of LS-SVM and proposed some

*Table:*
*The number of support vectors, and the mean squared error calculated for different training set sizes of the same problem using the proposed methods (the tolerance was set to 0.25)*

| Number of training samples *(This would also be the network size in case of standardLS-SVM)* | Number of "support vectors" *(Network size using the proposed reduction method)* | The mean square error *(The approximation error for the proposed method)* |
|---|---|---|
| 40 | 8 | $1.890 \times 10^{-3}$ |
| 80 | 9 | $0.877 \times 10^{-3}$ |
| 800 | 9 | $0.155 \times 10^{-3}$ |
| 1600 | 9 | $0.029 \times 10^{-3}$ |

modifications to extend it, in order to achieve a small, sparse model, even in the case of large datasets. This paper also shows a simple method to solve a time series prediction problem through simple static regression, which can be solved by analytic methods, or by the neural model (e. g. the described LS-SVM).

The described methods can be used for a wide range of problems; therefore they provide an efficient tool for a large number of research or real-life problems.

**References**


[1] Horváth G. (ed.), "Neural Networks and their Applications", Műegyetem kiadó, Budapest, 1998. (in Hungarian)

[2] S. Haykin, "Neural networks. A comprehensive foundation", Prentice Hall, N. J., 1999.

[3] V. Vapnik, "The Nature of Statistical Learning Theory", New York: Springer Verlag, 1995.

[4] E. Osuna, R. Freund, F. Girosi, "Support vector machines: Training and applications", Technical Report AIM-1602, MIT A.I. Lab., 1996.

[5] C. J. C. Burges, B. Schölkopf, "Improving the accuracy and speed of support vector learning machines", In: M. Mozer, M. Jordan, T. Petsche (ed.), Advances in Neural Information Processing Systems 9, pp.375–381., Cambridge, MA, MIT Press, 1997.

[6] E. Osuna, R. Freund, F. Girosi, "An improved training algorithm for support vector machines" In: J. Principe, L. Gile, N. Morgan, E. Wilson (ed.), Neural Networks for Signal Processing VII – Proceedings of the 1997 IEEE Workshop, pp.276–285, New York, 1997.

[7] Thorsten Joachims, "Making Large-Scale SVM Learning Practical", Advances in Kernel Methods-Support Vector Learning', MIT Press, Cambridge, USA, 1998.

[8] J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, "Least Squares Support Vector Machines", 2002. World Scientific, Singapore, (ISBN 981-238-151-1)

[9] F. Girosi, "An equivalence between sparse approximation and support vector machines," Neural Computation, 10(6), pp.1455–1480., 1998.

[10] J. Valyon, G. Horváth, "A generalized LS-SVM", SYSID'2003 Rotterdam, 2003, pp.827–832.

[11] J. Valyon, G. Horváth, "A Sparse Least Squares Support Vector Machine Classifier", Proceedings of the International Joint Conference on Neural Networks IJCNN 2004, pp.543–548.

[12] Yuh-Jye Lee, Olvi L. Mangasarian, "RSVM: Reduced support vector machines", Proc. of the First SIAM International Conference on Data Mining, Chicago, April 5-7, 2001.

[13] W. H. Press, S. A. Teukolsky, W. T. Wetterling, B. P. Flannery, "Numerical Recipes in C", Cambridge University Press, Books On-Line, www.nr.com, 2002.

[14] G. H. Golub, Charles F. Van Loan, "Matrix Computations", Gene Johns Hopkins University Press, 1989.


# Complex Fault Management Solution for VoIP Services

PÁL VARGA, ISTVÁN MOLDOVÁN
*Budapest University of Technology and Economics,*
*Department of Telecommunications and Media Informatics, {pvarga, moldovan}@tmit.bme.hu*
GERGELY MOLNÁR
*Ericsson Hungary, gergely.molnar@ericsson.com*

***Keywords: VoIP, quality of service, fault management***

*The more widespread the quality VoIP solutions are getting, the more users utilize these services – without even noticing it. High QoS requirements necessitate adequate Operations and Maintenance (OAM) support, as well as a fault management system specialized for VoIP services. Our current study describes a complex fault management system customized for VoIP service providers. This system covers the functions of detecting and processing of alarms, moreover, it features a unique root cause analysis subsystem, also. The solution integrates the classic passive, rule based event processing approach with a method operating by active, fault localizing checks. The appropriate elementary checks are scheduled by an approach novel in the Fault Management field: Petri nets. On one hand the eye-catching nature of this solution makes root cause analysis steps easy to understand, and on the other hand it is fast, due to the simultaneous submission of active checks.*

## 1. Introduction

Voice over IP (VoIP) services are employed not only in cases when we run the VoIP client on our Internet-capable computer to make a phone call. There are numerous hidden or unnoticed cases of VoIP usage – even when we initiate a simple call from our desk phone. Achieving a certain quality with the calls carried by VoIP networks require careful network planning and – once the service is operational – effective execution of operations and maintenance tasks.

The operation of VoIP services with high availability over a carrier network requires a complex Fault Management System (FMS) to be running. This paper describes an FMS supporting VoIP *services-related* Operations and Maintenance (OAM), covering VoIP *network-related* OAM issues, too. Overcoming faults related to QoS of the VoIP service is the task of the service provider, whereas handling carrier errors is the task of the network operator.

During the continuous detection and collection of events arriving from the managed objects, the FMS is able to filter the alarming events, to find their original sources (root cause) and to suggest corrective actions. This way the job of the operating personnel gets significantly simplified, although it never gets superfluous, since the detection and elimination of complex faults still remains their major role.

QoS measures of a VoIP service can be degraded by the faults originated from the network element errors, the misbehavior of the IP network as a managed object and the malfunction of VoIP-related applications, too. Any status-change of the above entities can generate “normal” or “alarming” event reports, which arrive to the FMS. Evaluating the standalone event reports would not allow the FMS to grade the event in question. In most fortunate cases the report could be explicitly alarming, describing the exact root cause, but in the majority of the cases it only reports about propagation or a side effect of an error – not to mention the cases where event reports describe normal and harmless status changes. Since managed objects can report events about different representations of the very same fault, the data to be analyzed by event processing is redundant. Finding the root cause of the fault from this redundant data is quite a challenging job.

Fault Management itself is the process of eliminating the alarms through the steps of event detection, alarm processing and correlation, root cause analysis and fault correction. During event detection the reports of the various event sources get collected together, and transformed to a common format that is easy to process. This set of events gets processed by filtering algorithms (tailored to reduce the number of alarms), event correlation and trend analysis methods (introduced to provide new, more descriptive alarms). The outputs of event processing are alarm notifications *(alarms)*. Each of these alarms describes a specific error, being a subject of the root cause analysis (RCA). The automated Fault Management process results a suggestion for corrective actions. In the sequel, we specify the requirements set toward the mentioned Fault Management steps, together with their design considerations.

In the telecommunications field one of the classic, best documented, complex Fault Management framework is TMN (Telecommunications Network Management) [11]. This has been enriched with some novel ideas while being applied to Hungarian conditions. This TMN adaptation is described in [12]. The drawback of these systems is their hierarchical setup, which makes them inflexible and ineffective for networks suffering from frequent topology changes. This disadvantage puts them out of the picture when choosing the FMS solution for

rapidly changing networks. A brief description of traditional Fault Management frameworks as well as modern root cause analysis methods can be found in [6].

The prototype of our VoIP Fault Management framework has been run at the Department of Telecommunications and Media Informatics of BUTE. The system planning, development and installation has been carried out with the kind help of Ericsson Hungary and Kovax'95 Ltd. The call data record (CDR) evaluation theories had been verified by using anonymized CDRs from the VoIP network of NIIF, Hungary. The description of our FMS framework and the corresponding event processing methodology was presented at the workshop [10], too.

## 2. Fault Management

A complex FMS should cover the following functionalities:
- detects and stores the event and alarm notifications,
- supports filtering the alarm notifications,
- initiates diagnostic checks to find out root causes of errors, suggests corrective actions.

The FMS continuously watches the state changes of the network and the various managed objects, "hunting" for suspicious events and clearly alarming reports. The appearing errors are handled through the steps of
- event detection,
- alarm processing and correlation,
- root cause analysis and
- fault elimination (see *Figure 1*).

The outputs of this process from the FMS point of view are the result of diagnostic checks carried out by the system and the list of the suggested corrective actions. Based on this information, the operator can start eliminating the fault. Since the FMS may not be able to provide accurate details of the root cause, the operator can *analyze the diagnostic results* and continue refining the details of the original error without having to initiate (duplicate) the diagnostic checks already finished.

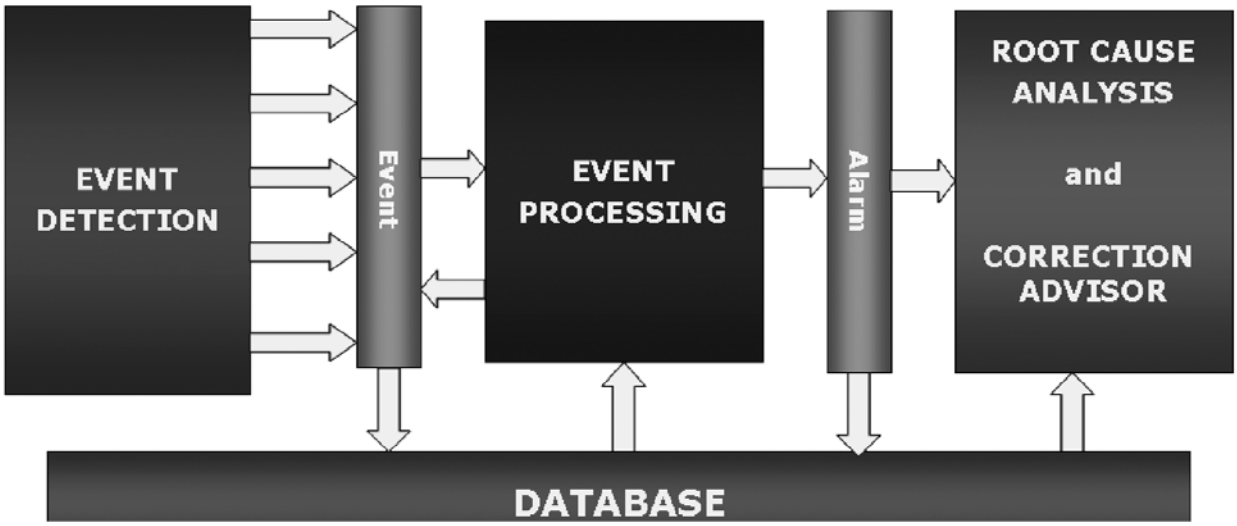
To ease the referencing of the process elements in this article, we should clarify the difference between *events* and *alarms*. We use the notion "event" as a short form of *event notification*, the output of event detection sub-process. This covers the status-reporting events sent by the network elements (including, but not limited to alarming events of direct error notifications) and the events reported by the active elements of the FMS (call generator, active monitor, etc.). On the other hand, we use the term "alarm" as a short form of *alarm notification*, which is the actual output of the event processing subsystem. These cover all the alarming entities that become the input of the RCA, being serious enough to be taken special care of. The operator should analyze each of these alarms separately.

## 3. Event Detection

The aim of event detection is to perceive the errors endangering VoIP service availability as early as possible. To support this, the main function of this subsystem is forwarding events (arriving from various sources) to event processing in a standardized format. Another task of this subsystem is feeding these events into the event database. Building a database with standard fields requires the database records to be in a similar format. This is another reason for the event detection subsystem to standardize the event notifications before storing and forwarding them. Events arriving to the processing subsystem in a general format ease their handling – both to the automatic FMS algorithms and the operator.

During the fault management of VoIP services the following types of events appear to be interesting:
- reports sent by error detecting entities installed in the network (Syslog, QoS monitor (automatic call generating and evaluating system) [2],[3]),

*Figure 1. The elements of Fault Management System and the interaction among them*



– reports on call data records generated for VoIP calls (RADIUS records [4],[5] – e.g. frequent appearance of release cause codes classified as erroneous or alarming can indicate faulty system elements),
– events generated by the active monitor (checking the availability of key network nodes and processes),
– errors logged by operators in the HelpDesk system (e.g. triggered by a subscriber complaint).

Current FMS is designed to evaluate the erroneous cases to be corrected by the VoIP service provider. The system is capable of indicating some types of carrier grade errors; suggesting detailed lists of corrective actions for these kinds of faults, however, is out of its scope. To satisfy such demands it is recommended to apply special fault management applications focusing on lower, network level issues.

## 4. Event processing

As mentioned before, messages of the event detection subsystem get stored in the event database (Figure 1). These event-description types of records form the input data set of the event processing subsystem. Manual processing of this database is not feasible, since it grows in a rate of tens of records per second – not to mention that it contains redundant information (record contents can be multiplicated both physically and logically).

Event processing should be considered as an automated data processing method, resulting in the presentation of alarms in front of the operator. The general process is depicted in *Figure 2*.

Events are stored in the database for further utilization. Their processing starts in the correlator module. The trend analysis module works on the event database, searching for alarming trends in the data set. These two modules actually generate new kinds of events, as opposed to the filter modules, which aim to reduce the number of events by removing redundancies and "highlighting the essence" of an event-pattern. These modules are detailed in the following sections.

### 4.1. Filters

Event filtering is the simplest among the event processing algorithms. We have chosen *rule based filtering* [7] out of the practical filtering approaches used in current FM systems [6]. During event processing, the filter module checks if the current event meets or fits into the description of any of the active rules. If there is no such filtering rule, the event should not be filtered, hence an alarm notification will be generated (propagation by default). If the event satisfies a rule, it either gets suppressed or promoted to be an alarm – depending on the nature of the rule. The filter criteria include parameters such as validity period, source type, event code, threshold value, etc. We have found that defining merely four types of filters (namely: suppress, counter, redundancy and dominance) are enough to address all raising issues of event suppressing [1]. In brief, our filter types provide the following functionalities:

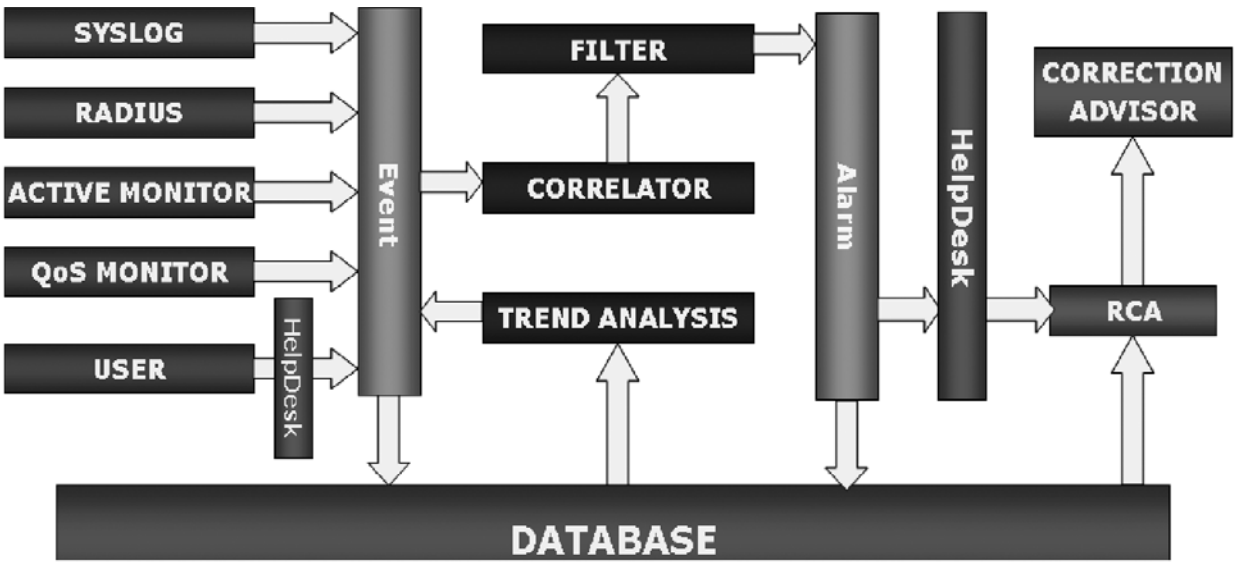
- *Suppress:* complete suppressing of event types
- *Counter:* supressing events of a kind if their amount is under a given threshold
- *Redundancy:* passing through only one event of a kind during a given period – and suppressing all the others arriving meanwhile
- *Dominance:* more serious events override others with lower priority or less descriptive power

The operator can easily change the active rule-set at any time: as rules are defined in a human readable format, it is as easy to create new alarms, as to modify or delete them.

### 4.2. Event correlation

Event processing is often referred to as "event correlation". In our terminology event correlation does not include filtering (suppressing), but it enriches the event

*Figure 2. The flow of event processing*



space with information found by correlating various event types. This way we have the opportunity to extract information from events arriving more or less at the same time and provide an essential alarm with high descriptive power, featuring parameters taken from several (different) events. Our event correlation module does not suppress events, but passes all incoming events to the filtering modules together with the specific, correlated ones. Supressing the events successfully exploited by event correlation is a typical task of dominance filters.

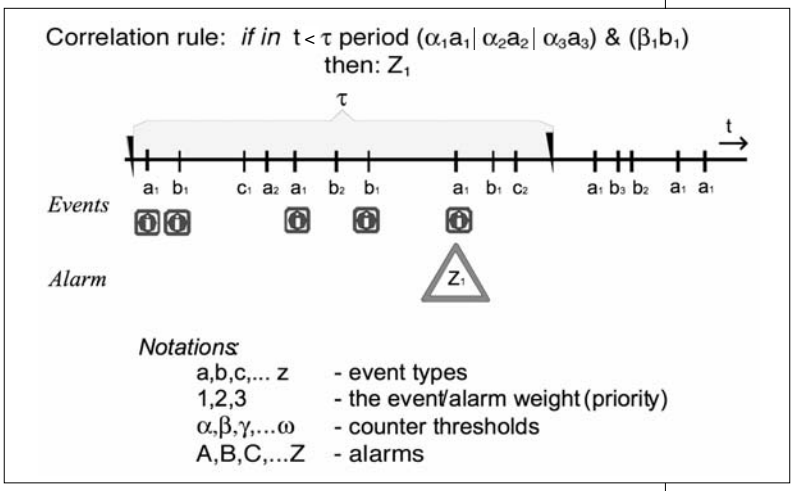



*Figure 3. Rule-based event correlation*

Among the practical correlation methods used in FM systems [6] we have chosen the rule-based approach. Our event correlation rules can be defined in the same manner as described for the filters. The effect of these rules is depicted in *Figure 3*. The example shown in this figure describes how a new, correlated alarm ($Z_1$) is generated when the correlation criteria (several number of $a_1$ events and a given amount of $b_1$ events arrive in the τ timeframe) gets satisfied [8].

The ultimate reason of event correlation is providing events with the highest descriptive power, gathering many elementary fault-descriptions into a more specific description of the root cause.

### 4.3. Trend analysis

Utilizing trend analysis methods allows us to predict misbehavior from event patterns of a longer time-scale. To solve trend analysis problems in general, the task is to find an appropriate trend function for the prediction. This method is hard to be applied for our case, since the "trend functions" could use only a few samples of suspicious events before the actual fault happens. This leads to another approach, namely: we should search for exact event patterns predicting future faults. The method should notify about possible future faults when the observed variable (e.g. a parameter inside specific types of event notifications) gets higher than a predefined threshold. The order of the predictive events is indifferent, however, they should arrive inside a given timeframe (otherwise there could be no correlation between them).

We have found that trend analysis methods can only be powerful in predicting accumulative type of faults (where the fault grows out of the continuous degradation of some system resource). Suddenly hitting, catastrophic faults are impossible to be predicted with trend analysis methods, since there are no patterns to predict from.

## 5. Root cause analysis and advise of corrective actions

We have implemented a novel Root Cause Analysis (RCA) approach during the definition of the FM framework focusing on VoIP services. Traditional FM systems operate with merely passive processing of the events. Some types of alarms generated by such methods were descriptive enough to provide the root cause of the fault, whereas others were too complex. Processing of these complex alarms were still left to the operator, who runs some active checks searching for specific proofs of the misbehavior, this way getting closer and closer to the root cause.

Whereas *passive* RCA systems try to find root causes by taking merely event notifications as input, the *active* approach is to carry out diagnostic checks continuously. Our system integrates the advantages of *passive*, model based event correlation systems that are able to follow topology changes in a flexible way with the automatic scheduling of *active* checks that belong to the every day routine of the operator personnel. The result is a complex fault management system, utilizing both passive and active (checking/intrusion) algorithms.

One of the principles of the passive, rule based event correlation systems is the following: using high complexity rule-sets allow the system to provide alarm descriptions featuring the place of the error and its probable cause. Extracting these descriptive parameters from the complex, correlated alarm report is the simplest RCA task: these alarm parameters are the actual output of the "root cause analysis".

This method, however, can only be used for root cause analysis using passively correlated events. The greatest drawback utilizing this passive rule-set is the inability of actively requesting status and configuration information from the nodes. Without the possibility of fetching key pieces of information from the nodes themselves, it is impossible to automatically provide root cause information for all types of alarm reports. The key questions when solving these problems are when to request the information and how to fetch the required data. Our algorithmic root cause analysis approach – described in the following section – aims to answer this, as well as other issues raised by these basic questions.

### 5.1. Algorithmic root cause analysis

During the RCA process we utilize all kinds of information (topological, node configuration, node status) to determine the most probable cause of the fault. The result of this process is a list of suggested corrective actions. Taking this list into consideration can greatly ease the job of the network operator in charge.

We have defined the following set of requirements against the fault management system specialized for VoIP services. The system should be capable of

– searching in the event database,
– initiating active checks and scheduling these initiations,
– running these checks and database searches at the same time (simultaneously),
– handling the topology changes in a dynamic way.

Beside these key principles we have born modular system design and short implementation period in mind.

RCA methods in the literature are grouped into three categories: *alarm correlation based systems, statistical methods* and *model based systems*. We have defined a framework capable of initiating active checks and evaluate their result. To meet all the above-listed requirements, we have integrated this framework into a model based RCA system.

One of the advantages of model based RCA systems is their ability of describing network topology in a flexible way. When a new node gets introduced to the system, the correlation and filtering rules can be automatically generated, rather than having to manually redefine them. Considering VoIP architecture, the topology description includes the IP addresses assigned to the physical nodes, and the identifiers of the first hops of the particular node.

The rule-set follows this flexibility as well. It handles the topology information dynamically; hence topology changes can be introduced without having to reconstruct the rules. Event hierarchy is also taken into consideration in the model.

### 5.2. Petri net of checking routines

In fortunate cases event filtering and correlation provide alarm reports describing the features and the place of the fault relatively well. The parameters of these alarms include the ID, type, and – beside others – the set of node-IDs related to these alarms. The aims of alarm report evaluation are identifying the root cause(s) and suggesting a list of corrective actions. *Figure 4* depicts this process.

Event notifications arrive to the RCA and correction advisor module from the event processing sub-module. There is an *RCA descriptor graph* for each alarm type. These descriptors are standalone Petri nets. We have chosen the principle of Petri nets because it controls the RCA processing by taking into account what pieces of information are available to run specific tests. RCA descriptor Petri nets depict the connection of basic checking routines and the parameters required to initiate these checks. Such a graph is shown by *Figure 5*, which will be explained later in this section in detail.

The *execution scheduler* determines the execution order of elementary checks. Once all the predefined checks have returned with some result, and there are no other executable tasks, the scheduler passes the RCA result to the advisor module. This assigns appropriate corrective actions to the result and presents this list to the operator.

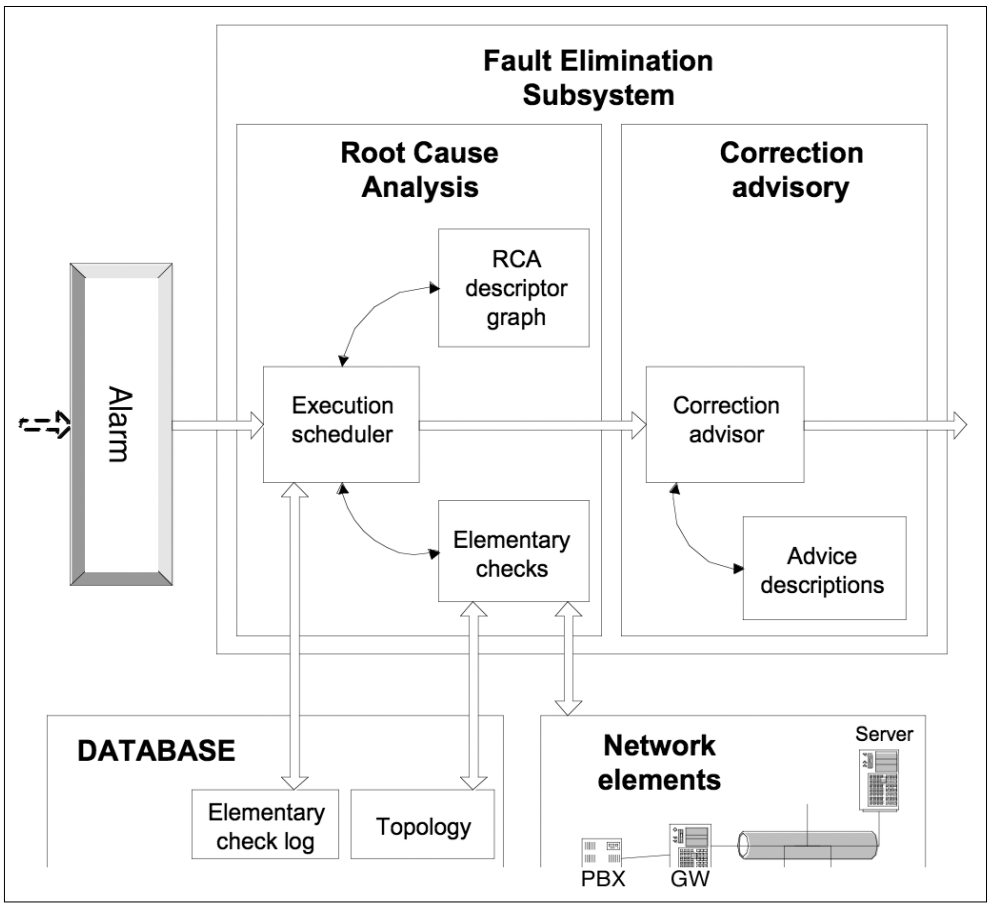



*Figure 4. The architecture of the Fault Elimination Subsystem*

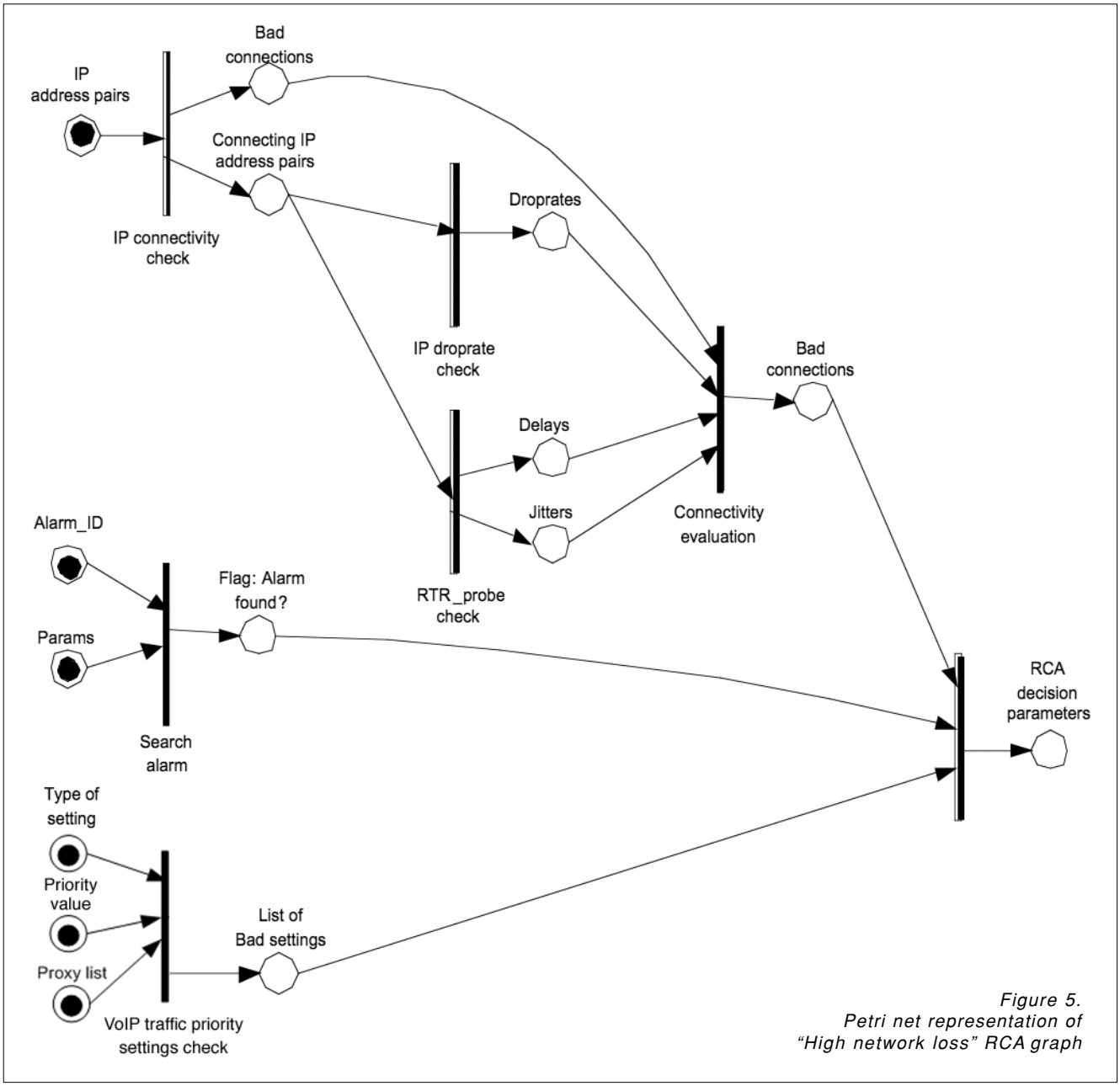



*Figure 5.*
*Petri net representation of*
*"High network loss" RCA graph*

The central element of the above process is the RCA descriptor graph. We have implemented this by utilizing Petri nets, a framework novel to fault management [9]. Using this framework allows the system to simultaneously execute elementary checks (the *transitions* of the Petri net), and schedule new checks in the order of the availability of their input data (the *nodes* of the Petri net). Once the data is available the Petri net node representing that data becomes "tokened". An elementary check gets triggered (the representing transition fires) when all its input data available (all its input nodes are tokened). As a result of the elementary checks, the corresponding output data become available – the nodes representing these data get tokened. The graph can include transitions representing data-request type of functions. These allow the system to fetch data from any kind of source (e.g. interface configuration file, topology database). Fetching such data is an important part of the process, since these complete the input data set of the elementary checks.

Fault management of VoIP services requires various elementary routines to be initiated at some point. These fall into the following categories:

– functions requesting interface status information,
– configuration data fetching,
– active checking routines,
– other
(e.g. active search of alarm patterns).

Let us take the alarm notification of "high packet loss" as a case study for demonstrating the RCA process. Figure 5 depicts the corresponding Petri net. The execution scheduler activates this RCA descriptor when the corresponding alarm arrives to the RCA module. At the first execution step all the *transitions* (checking and data fetching routines) having completely tokened input nodes *fire* (get triggered).

After completing an elementary routine, its output node gets tokened. The data associated with this node could be an input of other elementary routines. The execution scheduler "spends some time" evaluating each active RCA descriptor, in a round robin fashion. During its stay at a designated Petri net, it evaluates if there is a transition able to fire – if there are elementary functions having all their input parameters available.

For every firing transition the scheduler triggers the corresponding elementary function. Once the output of the last (typically the final decision-maker) transition gets tokened, the scheduler passes the results to the advisor module and ceases the RCA entity.

Due to the Petri net based RCA description the execution time of the RCA processes can be cut to half [8], moreover, the execution order does not need to be determined in advance. The task of assigning checking routines to alarm types cannot be eliminated, but still this should be determined only once for each alarm-type. The checks will be executed following the connections of the Petri net, in order of the data availability.

## 6. Summary

The amount of events, their diversity and arrival rate are all factors making the job of the VoIP network operator personnel extreamly hard.

Using appropriate filtering rule-set the alarm flooding phenomenon – when extream amount of alarms get generated and presented to the operator – can be avoided.

The arriving events can be clustered, counted, prioritized, temporarily suppressed or passed through. Applying proper event correlation and filtering we can generate alarm reports describing the fault better than merely analyzing standalone events. The job of the operator can be eased very much by presenting these verbose alarm notifications rather than the event floods themselves. Being in possession of longer term event and alarm information, we can predict some types of errors by utilizing trend-analysis mechanisms. In the FM case the trend-analysis method covers pattern matching algorithms applied to the event database for predicting saturation-type of faults.

Employing merely passive event processing methods does not always lead close enough to the root cause of the fault. The FM system should initiate the active root cause analysis. During the RCA process the FM can check the possible fault sources by initiating active verification routines. After evaluating the results the FM puts forward a proposal for the place and nature of the root cause, moreover, it suggests corrective actions.

To schedule and evaluate the active tests we have developed a method novel to the FM field: we used Petri nets to describe the connection between active RCA steps. This data-driven framework allows simultaneous submission and evaluation of active checks. Due to this method the RCA process is easy to understand and fast to execute.

The research and development activities described in this study was kindly supported by the Hungarian Ministry of Education, under the project identifier IKTA-00092-2002.

**References**

[1] Management System for Integrated Voice-Data Networks (Technical Appendices, Phase I. and II.) – BUTE-TMIT, Ericsson Hungary, Kovax'95, IKTA-00092-2002 report, 2003 (in Hungarian).
[2] ETSI TS 101 329-5 V1.1.2; Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 3; End-to-end Quality of Service in TIPHON systems; Part 5: Quality of Service (QoS) measurement methodologies – ETSI, 2000.
[3] ITU-T Recommendation P.862; Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs – ITU-T, 2001.
[4] RFC 2866 – RADIUS Accounting, 06/2000.
[5] RFC 2865 – Remote Authentication Dial in User Service (RADIUS), 06/2000.
[6] Dilmar Malheiros Meira, "A Model For Alarm Correlation in Telecommunications Networks", 1997.
[7] Roy Sternitt, "Discovering Rules for Fault Management", Eighth Annual IEEE International Conference and Workshop on the Engineering of Computer Based Systems (ECBS '01), 2001.
[8] Tamás Szíjártó, "Fault Management Considerations for VoIP Networks" Scientific Student Conference at BUTE-FIEE, 2004 (in Hungarian).
[9] James Lyle Peterson, Petri Net Theory and the Modeling of Systems, Prentice-Hall, 1981.
[10] Pál Varga, István Moldován, Gergely Molnár, "Fault Management for VoIP Services", Networkshop 2005, Szeged (in Hungarian).
[11] Recommendation M.3000 Series – TMN: Telecommunications Management Network, ITU-T, 1992-1997.
[12] Roxán Reznák, OSS Concept of the Hungarian Telecom

# TCP-ELN: Efficient Communication over Wireless Channels

GERGŐ BUCHHOLCZ, TIEN VAN DO
*BME Department of Telecommunications, {buchholcz, do}@hit.bme.hu*
THOMAS ZIEGLER
*Telecommunications Research Center Vienna, ziegler@ftw.at*

***Keywords: TCP protocol, ARQ, explicit loss notification***

*As wireless technologies becoming more wide-spread, the importance of efficient communications over radio channels increases. TCP versions applied currently do not take into account the special characteristics of radio channels, thus their performance can be significantly lower compared to the available resources. In this paper, we propose a new TCP variant (called TCP-ELN) which is capable to considerably improve the transfer rate. The performance of the new solution is evaluated using simulations.*

## 1. Introduction

TCP congestion control has been designed for the application in wired networks that have negligible packet losses due to bit error on the channel. Consequently, standard TCP versions do not distinguish between losses due to congestion and losses due to bit-error on the channel and simply reduce the congestion window for both types of packet losses.

In the recent years, TCP/IP has been used in wireless networks to support Internet access through radio interfaces. In a wireless environment, where bit errors on the channel happen with high probability, reduction of the congestion window without distinguishing losses due to bit error and congestion at the TCP layer makes TCP congestion control ineffective. If a TCP flow is transmitted through a wired part without packet errors and a wireless link with bit errors, TCP reduces its congestion window in response to packet losses caused by bit errors on the wireless channel. Therefore, the throughput of TCP flows may be close to zero although the network is not congested at all.

Several options exist to optimize TCP performance over wireless networks. A major part of today's operational wireless networks employ an ARQ mechanism at the link layer for IP based data services (e.g. 802.11, UMTS acknowledged mode). Link layer retransmission has been shown to be efficient and scalable in many scenarios. The number of retransmissions attempts before a link layer frame is considered lost, however, is usually rather small in order to avoid high delay variation and buffering. Thus, even in the presence of link layer retransmission a residual packet loss probability exists from the perspective of the transport layer. Authors in [1] report an average residual packet loss probability of more than 6% for the case of WLAN in a standard industrial environment. Therefore, additional mechanisms are needed at the TCP layer to improve performance in the case of critical radio conditions even in the presence of link layer ARQ. In case of no link layer ARQ and FEC (e.g. UMTS transparent mode), TCP is obviously unprotected against packet loss due to bit errors, making a clear case for protection at the transport layer.

Several proposals have been worked out for transport layer mechanisms to help TCP distinguish between losses due to bit error and losses due to congestion. The indirect TCP approach [2] splits a TCP connection between a fixed and a mobile host into two separate connections and hides TCP from lossy link by using a protocol optimized for lossy links. The Berkeley SNOOP protocol [3] caches packets at the base station and performs local retransmissions over the lossy link. Note that both approaches require storage of extensive per micro-flow states and even data packet retransmissions at the base station make them scale badly in terms of link bandwidth and number of flows. Bakshi et al. [4] investigated the performance of TCP over wireless links varying the MTU size and proposes to send ICMP packets to inform the TCP source about the wireless link experiencing bit-errors.

Balan et al. [8] proposed the TCP header checksum option to make TCP detect whether a bit-error happened in the body of the TCP segment or in the header. In the first case, the congestion window is not reduced.

TCP Westwood [5] has the potential to improve the performance over lossy channels as it adapts the congestion window based on throughput variations rather than packet losses. TCP Westwood is a sender-only modification to TCP thus it scales and is easy to deploy. The mechanism of throughput based window adaptation, however, needs some initial time to become effective and TCP Westwood cannot avoid unnecessary window reductions in the presence of packet losses due to bit-error. Thus Westwood will not show significant improvements of TCP flows over lossy channels.

Explicit Loss Notification (ELN) [6] is a mechanism by which the reason for the loss of a packet can be communicated to the TCP sender. The base station monitors all TCP segments that arrive over the wireless

link but does not cache any TCP segments since it does not perform any retransmission. Rather, it keeps track of the segments that have been lost over the wireless link and sets the ELN bit in the corresponding ACKs. This original ELN proposal implies the drawbacks of requiring a list of the lost sequence numbers per microflow at the base station which is impossible in case of IPSEC encryption. Moreover, this proposal does not take ACK loss into account and provides poor signaling information for the TCP sender.

In [7] we have redefined the ELN technique keeping only the fundamental idea of it, informing the TCP source explicitly on packet losses on the wireless channel. We proposed a Reno based TCP congestion control algorithm that can improve the performance of TCP significantly if the source is able to inform the sender about the number of the corrupted TCP segments. However, the technique to gain loss information that makes the protocol work is only briefly described in that paper and implemented in the simulator in a course grained manner.

In this paper we give the details of gaining loss information from corrupted packets, taking the different configuration cases of wireless environments into account. We also introduce a novel enhanced congestion algorithm that exploits all information from the loss notifications in order to maximize its performance.

The rest of the paper is organized as follows. The ELN based TCP congestion control algorithm is discussed in section 2. In section 3 we describe the details of our loss recovery technique. Section 4 specifies the simulation environment and shows the results. Finally section 5 concludes the paper.

## 2. The TCP-ELN proposal

In order to improve the performance of TCP by collecting loss information, both the server and the client side of the TCP agent must be modified. Note that the approach for ELN taken in this paper does not necessarily require modification of network elements. Therefore, it can be implemented in an end-to-end manner. The TCP receiver must be able to receive loss information from lower layers and notify the sender, who needs to act properly in reaction to the reception of loss and congestion information.

### 2.1. Receiver side modifications

By using a loss recovery technique the TCP receiver may receive not only intact TCP segments but also loss notifications. When the TCP data receiver receives intact TCP segments it operates in compliance with standard TCP behavior i.e. an acknowledgement is generated. If a loss notification is received then it is stored. This loss information is sent to the TCP source embedded into the option field of the acknowledgement. In case of bursty bit errors several TCP segments can get damaged possibly producing a high amount of loss notifications before an ACK is generated. Thus these notifications *must be stored* at the receiver to avoid dropping of loss information.

No negative ACKs are used thus loss notifications can be sent to the source only piggybacked on ACKs. As a consequence the use of *delayed acknowledgements is feasible though not recommended* since it decreases the signaling speed.

The TCP receiver must take also ACK loss into account. Since loosing an ACK causes the loss of loss notifications the notifications *must be sent redundantly*. Depending on the implementation it must be defined how many notifications can be stored in an ACK and how many times a notification is resent. In our implementation we have implemented a SACK like solution with sending loss notifications three times.

During the lifetime of a TCP flow timeouts can occur at the sender. After the timeout the sender starts sending packets regardless of the ongoing TCP ACKs. In this case all loss notifications must be cleared by the receiver since these losses occurred before the timeout and sending them to the source can cause inconsistent TCP behavior. Thus the sender must be aware of sending timeout notifications and the receiver must be aware of *handling these signals*. How timeout signals are generated by the source can be found in the next subsection.

### 2.2. Sender side modifications

The flow control algorithm of TCP-ELN is based upon TCP Newreno's standard. If no loss notification is received by the source TCP-ELN behaves exactly the same way as Newreno. If the source receives loss notifications in the ACKs from the receiver then it stores the notifications and resends the lost segments. Until receiving the 3rd duplicate ACK TCP-ELN operates according to Slow Start or Congestion Avoidance algorithms. The 3rd duplicate ACK can be a sign of congestion loss or loss due to bit-error. To decide which recovery state the TCP source should enter the stored loss notifications are examined. If a notification of the lost segment indicated by the duplicate ACKs was stored previously then TCP-ELN enters the wireless recovery state. If no loss notification is stored for the segment indicated by the triple duplicate ack event then TCP-ELN treats the 3 consecutive duplicate ACKs as a sign of congestion. Thus Fast Retransmit and Fast Recovery are invoked.

In the wireless recovery state the source waits for a recovery ACK that acknowledges all segments that were lost on the wireless channel. During this recovery phase reported segments that are lost due to wireless bit-error are resent. Every time a duplicate ACK arrives the source increases the size of its congestion window with one segment. As a consequence, the TCP data flow and self clocking is kept going, except if the effective window size is limited by the maximum value of the send window. If a partial ACK arrives at the source then the source decides whether it should stay in the wire-

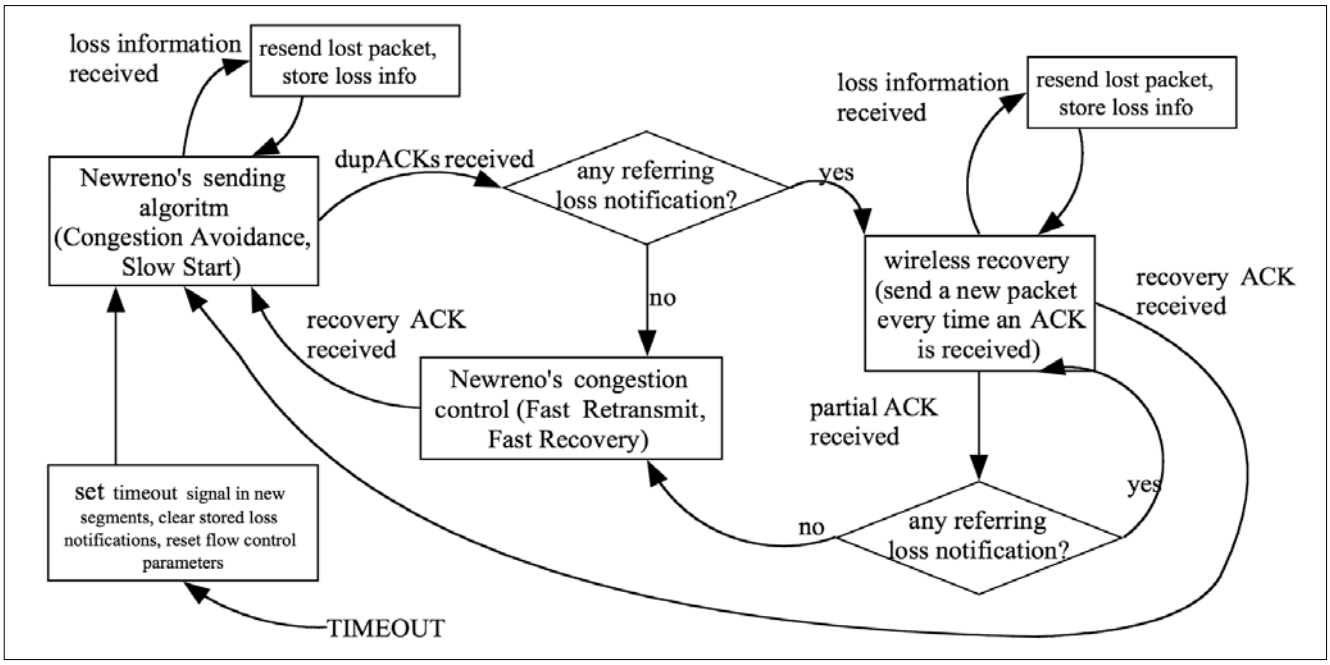



*Figure 1. The operation of the TCP-ELN sender*

less recovery state or enter the congestion recovery state. To make the decision the stored loss notifications are checked just like after receiving the 3rd duplicate ACK. If there is a referring loss notification then the source stays in the wireless recovery state otherwise it enters the congestion recovery state. If the recovery ACK is received then Slow Start or Congestion Avoidance, respectively, continues. On The operation of the TCP-ELN sender can be seen in *Figure 1*.

During the operation of the flow control algorithms timeouts can occur. In this case all flow control parameters are reset like in Newreno, the list of the loss notifications is cleared and a timeout signal is sent to the receiver. Since this timeout signal can be lost it must be sent in a redundant manner. The sender includes the number of timeouts since the start of the flow in every TCP segment's header option field. Thus even if packets get lost the timeout signal will reach the receiver. It will be received when the first intact TCP segment arrives.

## 3. Loss information recovery techniques

It is crucial for the operation of TCP-ELN to gain loss information from packets damaged on wireless channel. Three parameters the TCP sequence number, port number, and the source IP address of the lost segments must be collected by the wireless client in order to identify a single TCP flow.

There are two major methods to recover loss information; both of them influenced by the network topology and configuration. To demonstrate the recovery techniques the topology of *Figure 2* is used. This is a widely deployed scenario that contains a wireless link between the client and the access point.

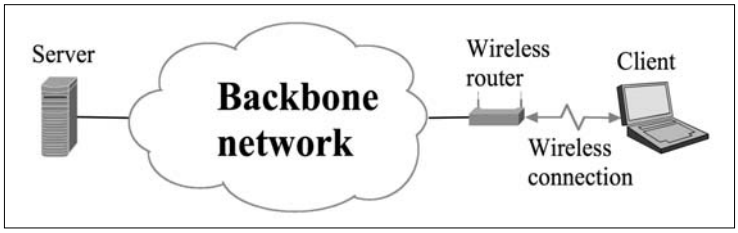
### 3.1. Cross layer communication

The first method is built upon the fact that MAC frames are not simply lost but received in a damaged manner by the destination node in the case of radio link error. Sensitive information can be found in the TCP/IP headers so extracting the intact headers from a damaged MAC frame solves the identification of a single TCP flow. To check whether the TCP header is damaged or not an additional header checksum is needed since TCP header has no dedicated protection [8].

This checksum can be inserted in the TCP header's option field. When the modified MAC layer of the client receives a corrupted frame instead of dropping it, it propagates the payload of the frame to the upper layers. If all the headers belonging to the different protocol levels are undamaged then the TCP segment reaches the TCP layer where loss information is extracted and sent to the source of the data flow. As the payload is much longer than the headers, bit errors are likely to occur in the payload leaving the headers intact. So this recovery method can be very efficient to collect loss information.

The advantage of the MAC modification method is that only the destination peer needs to be modified in order to collect loss information. All other network nodes remain untouched making this method absolutely

*Figure 2. Typical wireless access scenario*



end-to-end compliant. Since the MAC layer is modified only at the wireless client only the last wireless link can be monitored and loss information can be extracted only from the download flows.

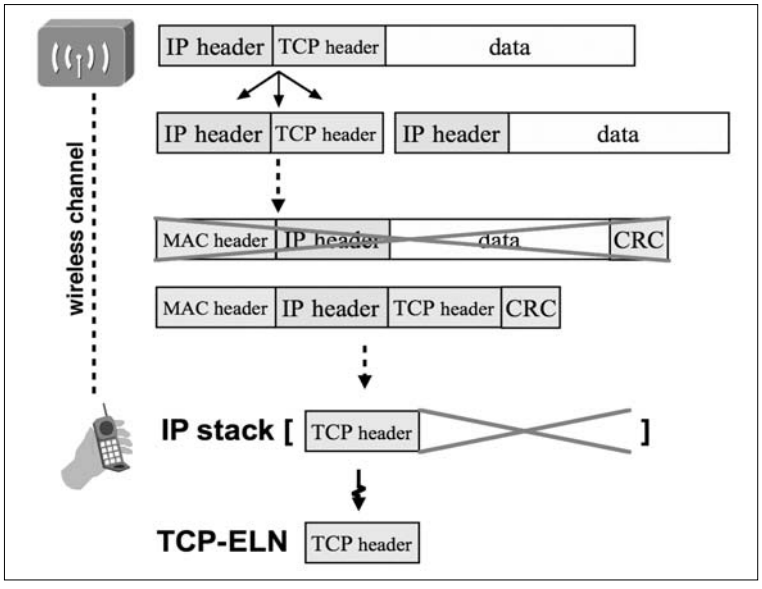
If modifying the MAC layer or recovering damaged packets at the MAC level is impossible, the ELN proposal may remain applicable by using the IP fragmentation method.

### 3.2. IP fragmentation

The fundamental idea of the IP fragmentation method is that the probability of loosing small packets on the wireless channel is smaller than the probability of loosing large packets. So if every TCP segment reaches the access point encapsulated in a single IP packet then fragmenting these IP packets at the access point into at least 2 packets where the first small part contains only the TCP header and the 2nd large part contains the payload can be used to collect loss information efficiently.

This method assumes that IP packets are not fragmented in the wired routers. The wireless router sends the IP packets embedded into MAC frames through the wireless channel. Since the MAC layer of the receiver is untouched, the damaged frames get dropped as usual. The IP layer tries to reassemble the received IP packets, but if any of them is lost then the original TCP segment cannot be reassembled. This means that at least one IP fragment was lost due to wireless bit error. But if the first IP packet with the TCP header included has arrived then the receiver's IP layer can partially reassemble the damaged TCP segment neglecting the missing parts. These missing IP packets contain the payload that is not used for loss information recovery. The damaged TCP segment is propagated then to the TCP layer where loss information can be extracted after verifying the TCP header checksum. The operation of this technique can be seen in *Figure 3*.

*Figure 3. The IP fragmentation technique*




The drawback of the IP fragmentation method is that it violates the end-to-end semantics since it needs the wireless router to be slightly modified. However using this technique can be very efficient since it is *scalable* in contrast to other non-end-to-end proposals like Snoop, the needed modifications at IP layer are in accordance with the *normal IP level operation* (fragmenting) and it is *independent* of the MAC layer. This is a trade-off between the possibility of MAC level modifications and the end-to-end schematics.

## 4. Case study

### 4.1. Simulation environment

To measure the performance of TCP-ELN a wide variety of simulations were run.

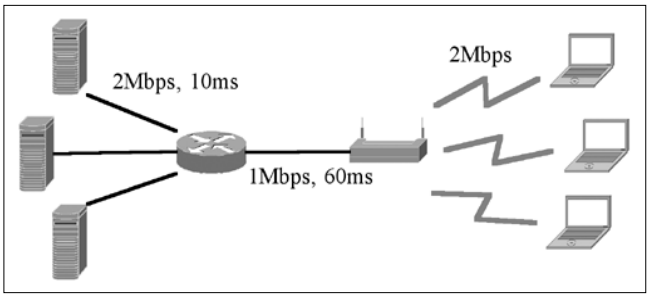



*Figure 4. Simulation topology*

The simulation topology illustrated in *Figure 4* consists of 3 servers, 3 wireless clients, a router and an access point. The bandwidth of the links between the servers and the router are 2 Mbps with 10 ms of delay. The router is connected to the access point with a 1 Mbps bandwidth link having a latency of 60 ms. The access point offers 2 Mbps for each client. All links are full-duplex. No specific MAC layer is applied on the wireless links, since the proposed techniques are general and do not depend on a specific wireless technology as strongly as on the topology. Due to the bandwidth setting congestion can occur only at the router in the wired part of the network.

Two different types of traffic: FTP and web traffic were used to examine the behavior of TCP-ELN. While FTP traffic is used to investigate the steady state behavior of the protocol web traffic is used to examine the dynamic behavior. In the first case one FTP session is started between each server-client pair. All servers used the same TCP version – TCP-ELN or TCP-Newreno. The throughput of the two TCP versions were measured and compared.

In the web traffic case each client simulates a web browsing user that connects to one of the servers. The parameters of the web traffic are in accordance with the SURGE model [9] that is based upon real network traces and widely used to generate realistic worklo-

ads. The download speed (page size/download time) of the web pages with the different TCP versions are measured and compared.

To simulate wireless channel errors two error models were used. Uniform and Markov error models are two different approaches of modeling the lossy link behavior. Both of them are used to simulate packet loss at the TCP level. The mean value of the uniform distribution is varied between 0 and 0.2.

The Markov model presented in [10] is applied in the simulation study, which integrates channel fading and radio link parameters to take into account the characteristics of wireless channels.

The Markov error model simulates how the radio links with different drop probabilities are experienced by users with different speed *(Table 1)*.

| Model number | User speed | Average error rate | Average error burst length |
|---|---|---|---|
| 1 | Pedestrian | 0.001 | 1.4913 |
| 2 | | 0.01 | 4.0701 |
| 3 | | 0.1 | 13.6708 |
| 4 | Middle | 0.001 | 1.0083 |
| 5 | | 0.01 | 1.0838 |
| 6 | | 0.1 | 1.8629 |
| 7 | Vehicular | 0.001 | 1.0024 |
| 8 | | 0.01 | 1.012 |

*Table 1.*
*The parameters of the Markov error model*

Since the efficiency of the loss recovery algorithms depends on how TCP segments are damaged, it is assumed that the probability of recovering a damaged segment is equal to the ratio of the payload size and the total packet size. In real environments this ratio is about 95% so in the simulations 5% of the damaged TCP segments cannot be recovered.

### 4.2. Results

Inspecting the throughput over FTP traffic TCP-ELN shows significant improvements in most cases, reaching as high as 400% depending on the particular error rate. As compared to Newreno, TCP-ELN is much better in the presence of error on the wireless link. In the error free case, i.e. without any packet loss on the wireless link, the performance of TCP-ELN is insignificantly lower (0.5%) than the performance of Newreno. This is due to the overhead, the larger header size of TCP-ELN segments. However this phenomenon also appears when using TCP-SACK.

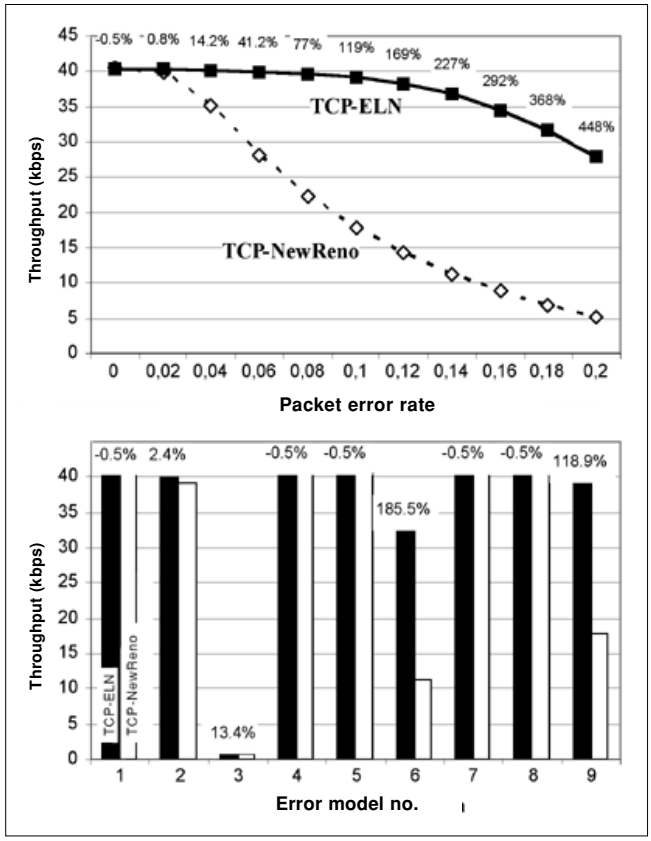
*Figure 5* shows that if the error rate is increased, the proposed flow control algorithm makes TCP-ELN throughput decline more gently than the rapid decrease observed for Newreno. For instance, while the throughput of TCP-ELN drops to 90% at the error rate of 14% Newreno's throughput drops down to 90% at the error rate of 4%. The results with the Markov error model show that in the high loss scenarios (0.1 mean PER) model 3, 6 and 9, the improvement depending on the burstiness can reach 185%. In the low loss scenarios (0.001, 0.01 mean PER) model 1, 4, 5, 7 and 8, an insignificant (0.5%) decrease can be observed, in model 2 the improvement is 2%.

Comparing the relative improvement measured over the uniform error with the ones measured for the Markov error model an interesting correlation can be recognized. As the speed parameter of the Markov error model increases the average length of error bursts decreases. Thus for the considered scenarios at higher speeds – as packet loss becomes sporadic – the Markov error model converges to the uniform error model.

In case of web traffic, less dramatic, but still significant improvements can be observed as shown in *Figure 6*. The improvement can reach 33% with the uniform error model at high error rates. When no packet loss occurs on the wireless channel then the performance decreases 0.1% due to the larger header size. Up to an error rate of 0.02 the performance of Newreno is almost identical to ELN performance. In case of the Markov model the improvement remains moderate 4% to 16% for the high error models (3, 6 and 9) and lower values for the low error models.

Comparing the web and FTP traffic results it can be seen that in the case of FTP traffic the improvement is significantly higher. The reason for this is that in the case of the used HTTP 1.1-like model for web traffic the majority of flows only has a few packets to send. When

*Figure 5.*
*The relative improvement and the throughput over FTP traffic with the Markov and uniform error model*



the TCP flows are very short Newreno's congestion window remains small. Thus ELN's flow control algorithm cannot improve the performance with avoiding window halving as much as in the case of longer flows and larger congestion windows.

## 5. Conclusions

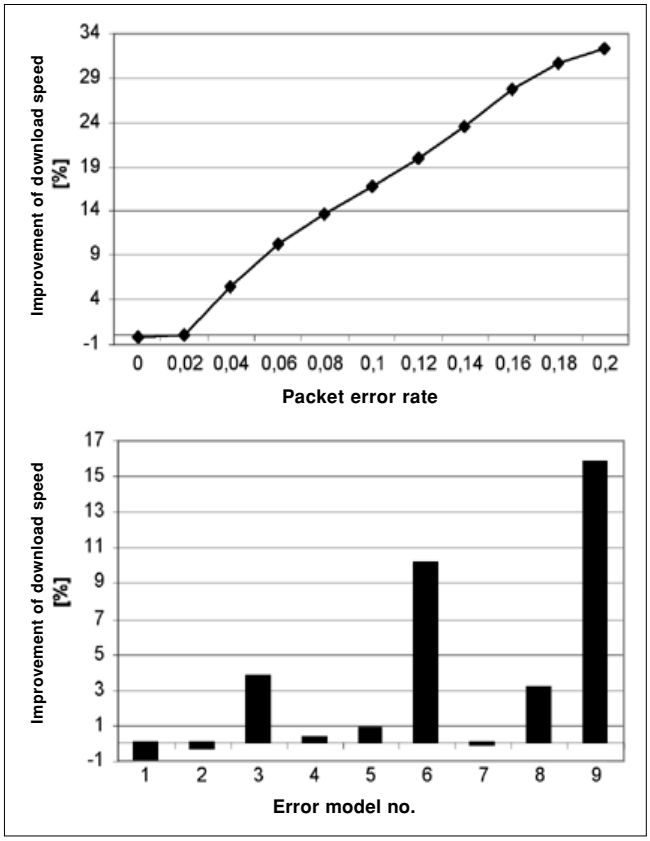
In this paper we proposed a new mechanism to improve the performance of TCP over wireless channels. Our solution is based on the idea of Explicit Loss Notification (ELN) which allows TCP to differentiate between packet losses due to congestion and packet losses due to radio channel errors. Thus unnecessary window reduction can be reduced substantially in the presence of radio channel errors. The resulting performance of the modified TCP protocol is greatly enhanced compared to TCP-Newreno.

We developed two alternatives to gain loss information in order to use ELN. The first solution is based on modifying the MAC layer; the second solution uses a special IP fragmentation technique. These methods can be used in a wide range of network environments depending on the topology and the setup of cross layer communication.

We developed a new TCP variant, TCP-ELN that integrates the loss recovery methods to improve its performance. The proposed receiver and sender side algorithms and processes were discussed in detail. Security issues concerning TCP-ELN were also considered.

Our simulation experiments have shown that TCP-ELN can improve the performance of TCP over the radio channel substantially in a wide range of environments. The protocol was tested over random and bursty erroneous radio links with two different types of traffic, static FTP bulk load and dynamic web traffic. Results proved that TCP-ELN produces better performance than TCP-Newreno in all the cases. The performance improvement can reach up to 400% when using FTP traffic and 30% when using an HTTP 1.1-like model for web traffic with high error probabilities.

*Figure 6.*
*The relative improvement of the download speed over web traffic with the uniform and the Markov error model*




### References


[1] A. Willig, M. Kubisch, C. Hoene and A. Wolisz, Measurements of a Wireless Link in an Industrial Environment Using an IEEE 802.11-Compliant Physical Layer. IEEE Transactions on Industrial Electronics, Vol. 49, No. 6, December 2002.

[2] A. Bakre and B. R. Badrinath, I-TCP: Indirect TCP for Mobile Hosts. In Proc. 15th International Conf. on Distributed Computing Systems (ICDCS), May 1995.

[3] H. Balakrishnan, V. N. Padmanabhan, S. Seshan, and R.H. Katz, A Comparison of Mechanisms for Improving TCP Performance over Wireless Links, IEEE/ACM Trans. on Networking, 5(6), December 1997.

[4] B. Bakshi, P. Krishna, N. H. Vaidya, D. K. Pradhan, Improving Performance of TCP over Wireless Networks, 17th Int. Conf. Distributed Computing Systems, Baltimore, May 1997.

[5] C. Casetti, M. Gerla, S. Mascolo, M. Y. Sanadidi, and R. Wang, TCP Westwood: Bandwidth Estimation for Enhanced Transport over Wireless Links, In Proc. ACM Mobicom 2001, pp 287-297, Rome, Italy, July 16-21 2001.

[6] Hari Balakrishnan and Randy Katz, Explicit Loss Notification and Wireless Web Performance, Proc. IEEE GLOBECOM Global Internet Conf., Sydney, Australia, November 1998.

[7] G. Buchholcz, A. Gricser, T. Ziegler, T. Van Do, Explicit Loss Notification to Improve TCP Performance over Wireless Networks, 6th IEEE High-Speed Networks and Multimedia Communications (HSNMC), 2003.

[8] R. K. Balan et. al, TCP Hack: TCP Header Checksum Option to Improve Performance over Lossy Links, IEEE Infocom 2001.

[9] P. Barford, M.E. Crovella, Generating Representative Web Workloads for Network and Server Performance Evaluation, in Proc. of Performance '98/ACM SIGMETRICS'98, pp. 151-160, Madison WI.

[10] A. Chockalingam, M. Zorzi, Rames R. Rao, Performance of TCP on Wireless Fading Links with Memory, ICC 1998.

# Network Model of the Processor System

VLADIMIR HOTTMAR

*The University of Zilina, Department of Telecommunications, Slovak Republic*
*hottmar@fel.utc.sk*

***Keywords: processor, queuing theory, service centre, closed network, dual port memory, data request***

*The paper deals with modelling of a double-processor system by closed service network. The article presents the queuing system as a method of modelling. It analyses the performance and quantifies time characteristics of the processor systems.*

## 1. Introduction

Questions connected with monitoring and valuation of processor systems (PS) are actual in many different levels of it's utilization. One of many possible processor system applications appears from assumption of dividing the final number of processed tasks among fractional processor systems, which activity is mutually and relatively independent, and every processor system $Ps_j$ of general system *M* process certain class of tasks $r_i$ of total set *R*, where $Ps_j \in M$ and $r_i \in R$, when i,j = 1, 2,...

Our task will be to understand a certain input set of data processed by given programme, and activities resulting from this processing will be the output set of data. Data transmission among processor systems is realised in the form of messages *S* that, on basis of routing and information content of message, is possible to consider as requests (answers) *N*. Requests are in every processor system saved into memory where they create a family of requirements. Systems processors solve the operation of requests. Lets define data transmission (independently of its information content) as requests coming (outgoing) into (from) PS and lets define processing of messages as operation of requests by processor system. Assign to the set of data the set of requests $S \rightarrow N$, where

$S = \{1, 2,...s\}$ a $N\{1, 2,...p\}$ and to the subset of processors ($P_i$) the set of service systems – service centres $\Sigma_i$. Then $\{P_i\} \rightarrow \{\Sigma_i\}$, where i = 1, 2,...

Computing memories are specific. Function of these memories is to save results and intermediate results of operations with data. As the processor always has free access to all data in the memory, it is needful to define the strategy of saving and selection of requests into (from) the memory. Then it is possible to transform operation memories to the family of requests, and to define the strategy of its service. Service strategies are worked out in detail in [1,2,3,4].

Assignment of the set of data to the set of requests, and the set of processors to the set of service centers allow to model every local processor system by stochastic model and to solve this model by queuing theory.

Lets apply queuing theory to a simple structure of a double-processor system with processors $P_1$ and $P_2$. Block structure of double-processor system and its model are shown in *Figure 1*. Processors perform their activity autonomously and the reciprocation of data through dual port memory. Reciprocation of data with the environment they can perform independently from each other. Data entry from environment (terminals) is realised only through input processor circuit $P_1$.

Processors $P_1$ and $P_2$ process tasks, and transmission of tasks between service centres are defined by probabilities of transition *pji*.

Terminals $T_1$, $T_2$,..., $T_n$ generate demand for service, which are serially stored in queue $L_1$. As the system doesn't distinguish between priority or non-priority requests, the requests for service will be selected from queues $L_1$, $L_2$ serially, as they are arriving. $L_1$, $L_2$ present the set of requests with FIFO service strategy. Service centres $\Sigma_1$, $\Sigma_2$ perform service of requests. If the request is serviced in service centre $\Sigma_1$ then it leaves from the service system with probability $p_{11}$, or with probability $p_{12}$ it requires the service from service centre $\Sigma_2$. If $\Sigma_2$ is free, the requests are served immediately, otherwise it takes stand in the queue $L_2$. After servicing has finished, the requests leave the service system with probability $p_{2T}$, or with probability $p_{12}$ they come to the queue $L_1$ and they again require the service from service centre $\Sigma_1$. After the service has finished in $\Sigma_1$, they leave from system with $p_{21}$ probability, or with $p_{12}$ probability they require another service from the service centre $\Sigma_2$, etc...

The system and its network model are shown in *Figure 1*.

## 2. Network model of double-processor system

Network model comes from the principle of **closed service network**. Characteristic feature of closed service network is the uniformity of total number of requests. According to this, the intensities of arrivals into particular service centres are non-constant.

The state of closed service network is at every instant of time defined as

$$K = \sum_{i=1}^{N} k_i \tag{1}$$

where

$N$ – is a total number of network nodes (Figure 1),
$K$ – is a total number of requests in network,
$k_i$ – is number of requests in $i$-th node.

We define the total number of network states as

$$Z = \binom{N + K - 1}{N - 1} \tag{2}$$

**Local balance condition** [1],[2] is the theoretical presumption for existence of closed service network solution by service centre. If every service centre in network fulfills the local balance condition, then even the whole network fulfil the balance condition. A solution is possible for service networks composed of local service centres of M/M/n/FIFO type [1],[2]. A precise solution of closed service networks was given by Gordon and G. F. Newell [1] and is defined as

$$p(k_1, k_2, ...k_N) = \frac{1}{G(K)} \prod_{i=1}^{N} \frac{x_i^{k_i}}{\beta(k_i)} \tag{3}$$

which is the probability that $k_i$ requests are in $i$-th service centre, where $x_i$ can be obtained from equation system solution

$$\mu_i x_i = \sum_{j=1}^{N} \mu_j x_j p_{ji} \tag{4}$$

where $i$ = 1, 2,...,$N$ and $G(K)$ is a **normalization constant**.

If every service centre has only one service device, then $\beta(k_i) = 1$ and (3) changes to the form

$$p(k_1, k_2, ...k_N) = \frac{1}{G(K)} \prod_{i=1}^{N} x_i^{k_i} \tag{5}$$

while, for the stationary state probability vectors, the following condition has to be fulfilled:

$$\sum_{S_K} p(k_1, k_2, ...k_N) = 1 \tag{6}$$

Closed service network is characterized by input parameters:

$N$ – number of service centres (nodes) of model service network
$K$ – total number of requests in the model
$\mu_i$ – service intensity of requests processing in $i$-th service centre, $i$ = 1, 2,...,$N$
$p_{ji}$ – transmission probability, that means – the request for service in $j$-th service centre will also require service in $i$-th service centre.

If we put $j$ = 1 and $i$ = 2,3,...,$N$ in equation system (4) then this will take the form

$$x_i \mu_i = \mu_1 x_1 p_{1i}, \tag{7}$$

or

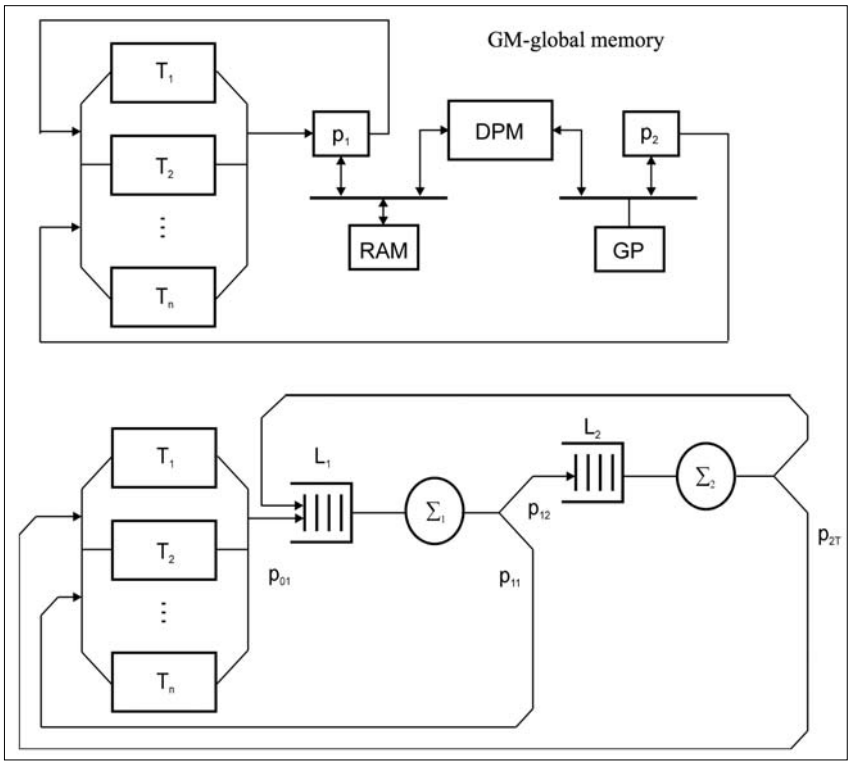
$$x_i = \frac{\mu_1}{\mu_i} x_1 p_{1i} \quad \text{for } i = 2, 3, ..., N \tag{8}$$

For $x_i$ we get an equation system of ($N$-1) equations which are independent. This equation system enables to express unknown $x_i$ through the parameter, for instance $x_1$.

If $x_1$=1, the normalization constant we define from equation

$$G(K) = \sum_{i=1}^{N} \prod_{i=1}^{N} x_i^{k_i} \tag{9}$$

*Figure 1. Block structure of computing system and its network model*




The basic indicator is utilization (loading) $\rho_1$, $\rho_2$ of service centres $\Sigma_1$ and $\Sigma_2$ [1,3,7]. By analysing (9) we arrive to the following conclusion. Exponents $k_i$ define possible number of requests in $i$-th service centre. Marginal attributes for $k_i$ are zero (condition, when $i$-th service centre is empty) and $K$ (all service centres are empty accept $i$-th). Consequently, two following alternatives can occur for $k_i$

$$(k_i = 0) \cup (k_i = K) \tag{10}$$

If we want to define $\rho_1$ then we consider only those states, in which $k_1$>0, that means, $k_1$=$K$ for $i$ = 2,3,...,$N$ can never happen at the some time.

As the minimum of requests in first service centre ($i$ =1) is $k_1$=1, then the maximum

$$k_{i_{\max}} = (K - 1) \tag{11}$$

for $i$ = 2,3,...,$N$ is divided into the rest of service centres [2],[9].

According to (9) we can write

$$\sum_{S_K}\prod_{i=2}^{N} x_i^{k_i} = G\left(k_{i_{\max}}\right) = G(K-1) \tag{12}$$

Utilization of *i*-th service centre is defined as

$$\rho_i = \sum_{S_K:k_i\rangle 0} p(k_1,k_2,...k_N) = \sum_{S_K:k_i\rangle 0} \frac{1}{G(K)} \prod_{i=2}^{N} x_i^{k_i} \tag{13}$$

Then the utilization of $\rho_1$ service centre $i=1$ in terms of (12) and (13)

$$\rho_1 = \sum_{S_K:k_1\rangle 0} p(k_1,k_2,...k_N) = \sum_{S_K:k_1\rangle 0} \frac{1}{G(K)} \prod_{i=2}^{N} x_i^{k_i} = \frac{G(K-1)}{G(K)} \tag{14}$$

We can interpret the meaning of (14) as follows.

Utilization of $\rho_1$ service centre $i=1$ is given by summary of state probabilities in which $k_1$ values larger than zero $k_1>0$ are gained. As an event, for $k_1=0$, surely occurs once at least, for $\rho_1$ we can write as

$$\rho_1 = \frac{G(K-1)}{G(K)} \tag{15}$$

Utilization of $\rho_1$ service centres $i=2,3,...,N$ we get from balance condition of input and output flow intensities of requests in steady-state regime in *i*-th service centre. Input flow intensity from service centre *i* is given by summary of service intensity $\mu_i$ and probability, that *i*-th service centre is occupied, which is just $\rho_i$.

Concerning (8) and (15) the following equation is valid

$$\rho_i = x_i.\rho_1 \quad or \quad \rho_i = x_i \frac{G(K-1)}{G(K)} \tag{16}$$

Intensities of $\lambda_i$ requests arriving in individual service centres are defined for steady state of input and output request flow equation of local service centre. For $\lambda_i$ will be valid:

$$\begin{aligned} \lambda_i &= \rho_i \mu_1 \quad for\ i=1 \\ \lambda_i &= \rho_1 \mu_1 p_{1i} \quad for\ i=2,3,...,N \end{aligned} \tag{17}$$

Average amount of requests $K_i$ in *i*-th service centre will be defined from the condition

$$K_i = \sum_{k_{ie}\rangle 0} k_{ie} p(k_1,k_2,...k_N) \qquad \text{for} \quad e=1,2,\ldots,K \tag{18}$$

where $k_{ie}$ is number of requests in *e*-th service centre and $p(k_1,k_2,...,k_N)$ is the already known vector of state line probabilities. We perform the summation through lines, in which $k_{ie}>0$, and according to (1), condition of requests uniformity in closed network has to be fulfilled. We define the average length of $L_i$ line in *i*-th service centre from deference of serviced $\rho_i$ and from average number of requests waiting for service, provided that number of service devices (processors) in *i*-th service centre is equal to one.

$$L_i = K_i - \rho_i \qquad \text{for}\ i=1,2,\ldots,N \tag{19}$$

Time characteristics of service centres – the average stopping time of $T_i$ requests in *i*-th service centre and average waiting period of $T_{wi}$ requests in $L_i$ line we can define from Little's law:

$$T_i = \frac{K_i}{\lambda_i} \tag{20}$$

and

$$T_{wi} = \frac{L_i}{\lambda_i} = T_i - \frac{1}{\mu_i} \tag{21}$$

## 3. Model analysis

We analyse the model of *Figure 1*. We assume a real technical environment and we set a number of connected terminals to 9, as an example ($T_1$ till $T_9$). Implementation of $\mu_i$ and $p_{ji}$ parameters will help us to define all performance parameters and time characteristics of the system. The number of closed network states for $K=9$ and $N=2$ define (2), according to which $Z=10$.

From system (4) we define the equations for parameters $x_i$ calculation, for two service centres, then

$$\begin{aligned} x_2\mu_2(p_{21}+p_{2T}) &= x_1\mu_1 p_{12} \\ x_1\mu_1(p_{12}+p_{11}) &= x_2\mu_2(p_{2T}+p_{21}) + x_1\mu_1 p_{11} \end{aligned} \tag{22}$$

for complete probabilities it is valid that

$$p_{12}+p_{11}=1, \quad p_{21}+p_{2T}=1 \tag{23}$$

By elimination of equations for $x_1=1$ we get

$$x_2 = \frac{\mu_1}{\mu_2} p_{12} \tag{24}$$

We see that parameter $x_2$ is dependent on ratio $\mu_1/\mu_2$ and on $p_{12}$ probability. These three variables crucially affect performance parameters of the system model. Let's analyse the model when the service intensity $\mu_1,\mu_2$ of both processors $P_1$, $P_2$ are the same, and requests for service in service centre $\Sigma_1$ demand the service in service centre $\Sigma_2$ with $p_{12}=1/2$ probability. According to (24), the parameter $x_2$ will obtain value $x_2=0,5$.

Substitution of $x_1$ and $x_2$ into (9) is normalization constant

$$G(K) = G(9) \doteq 1{,}998046$$

For performance parameters definition we use numerical figures from *Table 1*. Table shows in the first column the number of S network states, in second and third column we can see the distribution of requests $k_1$ and $k_2$ in service centres $i=1,2$.

*Table 1.*

| $S$ | $k_1$ | $k_2$ | $x_1^{k_1}x_2^{k_2}$ | $p(k_1,k_2)$ |
|---|---|---|---|---|
| 1 | 0 | 9 | 0,001953 | 0,000977 |
| 2 | 1 | 8 | 0,003906 | 0,001955 |
| 3 | 2 | 7 | 0,007812 | 0,00391 |
| 4 | 3 | 6 | 0,015625 | 0,00782 |
| 5 | 4 | 5 | 0,03125 | 0,01564 |
| 6 | 5 | 4 | 0,0625 | 0,03128 |
| 7 | 6 | 3 | 0,125 | 0,062561 |
| 8 | 7 | 2 | 0,25 | 0,125122 |
| 9 | 8 | 1 | 0,5 | 0,250244 |
| 10 | 9 | 0 | 1,0 | 0,500488 |
| $\Sigma$ | | | 1,998046 | 0,999997 |

The fourth column defines on the right side of the relation fractional conjunction of (9) and in the last column there are values of state line probabilities.

The value $G(K\text{-}1)$ is defined for $i$=1 from (12), or from Table 1 for $G(8)$=1,99414, then from (15) we define $\rho_1 \doteq$ 0,998, and from (16) we get by substitution of $\rho_1$ the value for $\rho_2 \doteq$ 0,499.

For the definition of $\lambda_1$, $\lambda_2$ we need to know the numerical value for service intensity $\mu_1$. Then by means of (17) the average number of requests in service centres will be $K_1$=8,00975 and $K_2$=0,990214, while the relation (1) has to be valid for total number of requests in network, that means $K$=8,999964.

The average length of lines $L_1$=7,01175 and $L_2$= 0,491214 requests waiting for service. We will define time characteristics of $T_1$, $T_2$ and $T_{W1}$, $T_{W2}$ from well known $\lambda_1$ and $\lambda_2$, and from (20) and (21).

Now let's analyze the model in the case when the service intensities are mutually deferent $\mu_1 \neq \mu_2$, and for probability of transition it will be again valid $p_{12}$=1/2. On the basis of (24) $x_2$=1, and $x_1$=1 we leave unchanged. After replacement $x_1$ and $x_2$ into (5) we obtain numerical data shown in *Table 2*.

*Table 2.*

| $S$ | $k_1$ | $k_2$ | $x_1^{k_1} x_2^{k_2}$ | $p(k_1,k_2)$ |
|---|---|---|---|---|
| 1 | 0 | 9 | 1 | 0,1 |
| 2 | 1 | 8 | 1 | 0,1 |
| 3 | 2 | 7 | 1 | 0,1 |
| 4 | 3 | 6 | 1 | 0,1 |
| 5 | 4 | 5 | 1 | 0,1 |
| 6 | 5 | 4 | 1 | 0,1 |
| 7 | 6 | 3 | 1 | 0,1 |
| 8 | 7 | 2 | 1 | 0,1 |
| 9 | 8 | 1 | 1 | 0,1 |
| 10 | 9 | 0 | 1 | 0,1 |
| $\sum$ | | | 10 | 1 |

Achievement parameters:

$$G(9) = 10,\ G(8) = 9,\ \rho_1 = \rho_2 = 0{,}9,\ K_1 = K_2 = 4{,}5,\ K = K = 9,\ L_1 = L_2 = 3{,}6,\ \lambda_1 = 0{,}9\ \mu_1,\ \lambda_2 = 0{,}45\ \mu_1$$

## 4. Discussion of the results

Our conclusions are as follows.

- Equal service intensities ($\mu_1 = \mu_2$) cause different coefficients of utilization of $\rho_1$ and $\rho_2$, while service centre $\Sigma_1$ (that is the processor $P_1$) works on saturation limit and creates a bottleneck in the system [7], the processor $P_2$ works with approximately 50% utilization.
- Coefficient $\rho \rightarrow 1$, which means that we can expect a breach of local balance in model (the steady state stops to be valid), what will be at the real processor $P_1$ expressed by not accepting the request for service and by memory conflicts.
- Unequal dividing of the requests in lines $L_1$, $L_2$ represent the increase of memory claims at real memories.
- Change in service intensity ratio ($\mu_1/\mu_2$=2) at $p_{12}$= 1/2 result in equal dividing of requests $K_1$,$K_2$ in service centres $\Sigma_1$ and $\Sigma_2$, utilization ratio $\rho_1$, $\rho_2$ gain equal values, and the processors $P_1$ and $P_2$ work with approximately 90% utilization.

**Acknowledgement**

The author gratefully acknowledges the support by the VEGA project No. 1/2043/05 "Systems for the interactive television cabling distribution ICATV"

**References**

[1] Rukovansky, I.:
Valuation of computer system efficiency, 1989., SNTL, pp.61–97.
[2] Neuschl, S. et al.:
Modelling and simulation, pp.304–354.
[3] Mitrani, I.:
Modelling of computer and communication systems, Cambridge University Press, 1987., pp.73–77.
[4] Zitek, F.:
Lost time, Academia Prague 1969, pp. 42–47.
[5] Kluvanek, P. and Brandalik, F.:
Operation analyse I. Bulk service theory, pp.40–42.
[6] Pesko, S. and Smiesko, J.:
Stochastic models of operation analyze, University of Zilina, 1999., pp.75–80.
[7] Hottmar, V.:
Performance model of network calculating systems ITKR, University of Zilina / EDIS – editorship ZU, 2005. ISBN 80-8070-338-8
[8] Hottmar, V.:
Digital television – text-book from international seminar, University of Zilina, May 1999., pp.66–90.
[9] Robertazzi, T. G.:
Computer Networks and Systems:
Queuing Theory and performance Evaluation, Springer-Verlag New York, 1990, pp.181–233.

# MAIPAN – Middleware
## for Application Interconnection in Personal Area Networks

MIKLÓS AURÉL RÓNAI, KRISTÓF FODOR, GERGELY BICZÓK, ZOLTÁN TURÁNYI, ANDRÁS VALKÓ

*Ericsson Hungary, Traffic Lab*
*{Miklos.Ronai, Kristof.Fodor, Gergely.Biczok, Zoltan.Turanyi, Andras.Valko}@ericsson.com*

***Keywords: pervasive/ubiqutious computing, access control, dynamic session management***

*This paper proposes the Middleware for Application Interconnection in Personal Area Networks (MAIPAN), a middleware that provides a uniform computing environment for creating dynamically changing personal area networks (PANs). The middleware hides the device configuration and physical scatteredness of the PAN and presents its capabilities as a single computer to the applications. The solution provides easy set-up of PAN-wide applications utilizing multiple devices and allows transparent redirection of ongoing data flows, whenever the configuration of the PAN changes. The proposed middleware interconnects services offered by applications running on different devices by creating virtual channels between the input and output outlets of the applications. Channels can be reconfigured when configuration or user needs change. In contrast to the approaches found in the literature, this paper presents a solution where session transfer, dynamic session management are tightly integrated with strong and intuitive access control security.*

## 1. Introduction

The ever-growing number of wireless terminals, such as smart phones, personal digital assistants (PDAs) and laptops, raises the need to set up, configure and reconfigure personal area networks (PANs) in an easy and ergonomic way.

This paper addresses the goal of creating a dynamically changeable, but uniform computing environment for PANs, which integrates wireless and wired, stationary and mobile devices that are connected to each other. The paper presents the Middleware for Application Interconnection in Personal Area Networks (MAIPAN) solution-which is situated below the applications and above the network layer-that hides the individual devices participating in the PAN and presents the capabilities of applications running on the devices as if they were located on a single computer. This provides a standard "PAN programming platform", which allows the easy set-up of personal area networks and dynamic connection and disconnection of applications running in the PAN. Application programmers using the uniform application programming interface (API) offered by the middleware can develop software without taking care of the various PAN configurations or PAN dynamics. They can assume certain capabilities, but disregard whether these capabilities are provided by one application running on one device or by a set of applications running on several devices. They only have to register the inputs and outputs of their applications at the middleware, and they do not have to take care of which kind of devices or applications will be connected to these outlets and will use their programs.

The presented middleware contains access control, flexible session management and transferable session control solutions. The middleware contains some intelligent functions, as well, which helps the user to control the PAN and improves human computer interaction (HCI). In theory all kind of solutions for service discovery, physical, link or networking layers can be used with MAIPAN.

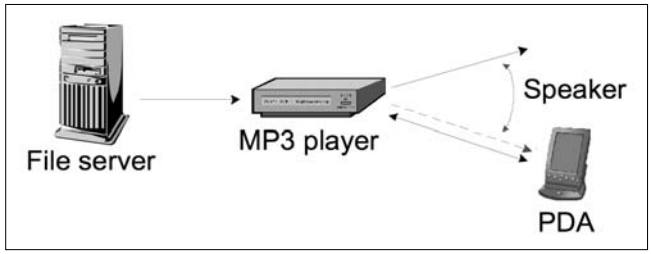



*Figure 1. Listening to music*

An example of a PAN application can be seen in *Figure 1.* Assume a MAIPAN-aware PDA user, who enters a MAIPAN-aware room where speakers are placed in the corners. The user decides to listen to music, but she neither has mp3 files nor an mp3-player. She turns on her PDA to check what kind of devices and services are available in the room. First, independently from MAIPAN the PDA's networking protocol establishes network connections to the devices located in the room, and the service discovery protocol (SDP) looks for services offered in the environment. After these steps she can configure a PAN with the MAIPAN dispatcher, which contains her PDA, the speakers in the corner, a fileserver and an mp3 player that are available through the network. MAIPAN establishes the necessary connections, that is, it connects the fileserver to the mp3 player's file input, the sound output of the mp3 player to the speakers and the control input of the mp3 player to the PDA. From now on she can select the mp3 file on the fileserver she wants to listen to, and instruct the

mp3 player via her PDA to play the music on the speakers. In the figure the arrows show the data flow between the devices.

The paper is organized as follows. Section 2 is about related work, in Section 3 the basic concepts and the middleware's architecture is introduced. Finally, Section 4 concludes the paper.

## 2. Related Work

Mark Weiser presented the concept of *ubiquitous computing* in the early 90s [1],[2]. Ten years after Weiser's publication, Satyanarayanan discussed the vision and challenges of ubiquitous computing in [3], where he introduced the expression *pervasive computing*, which is the end result of mixing distributed systems, mobile computing, smart spaces, invisibility, localized scalability and uneven conditioning [26].

Middleware are essential parts of pervasive and mobile computing environments. Mascolo et al. in [4] among others discussed why traditional middleware (such as CORBA [5]) is not well suited for mobile environments and how a mobile computing middleware should be designed. Another key component of ubiquitous computing is security. Chandrasiri et al. [6] proposes the concept of Personal Security Domains (PSDs)-where a PSD consists of a user's personal devices, and investigates security issues regarding PANs. Nowadays several projects are running in these research topics, some of them are presented in the followings and in our previous papers [26],[27].

The goal of the AURA project [7] is to provide each user with an invisible aura of computing and information services that persists regardless of location.

The project Gaia [8] designs a middleware infrastructure to enable active spaces in which data and tasks are always accessible and are mapped dynamically to convenient resources present at the current location of the user.

The Oxygen project [9] aims to develop very intelligent, user-friendly and easy-to-use mobile devices enabling users to communicate with the system naturally, using speech and gestures that describe their intent.

The Portolano project [10] focuses on highly adaptive user interfaces, an intelligent, data-centric network infrastructure and a new environment for developing distributed services. In the frame of this project the one.world architecture [11] is designed, which is a comprehensive framework for building pervasive applications. It includes services like service discovery, checkpointing, and migration in order to support programmers by creating applications for dynamically changing environments.

The extrovert-Gadgets (e-Gadgets) project [12] investigated architectures for the composition of ubiquitous computing applications and proposed the GAS-OS middleware to interconnect and control sensors and actuators with more intelligent devices.

The 2WEAR [13] project investigated the vision of a personal wearable system that can be dynamically composed out of different devices that are heterogeneous in terms of both form and function.

The Cortex project [14] addresses the emergence of a new class of applications that operate independently of human control.

The key objective of the EasyLiving [15] project is to create an intelligent home and work environment, which is based on the InConcert middleware solution.

The Speakeasy approach [16] focuses on the specification of minimal interfaces between devices using mobile agents and mobile code. MobiDesk [17] is a mobile virtual desktop computing hosting infrastructure that leverages improvements in network speed, cost, and ubiquity to address the complexity, cost, and mobility limitations of today's personal computing infrastructure.

Some concepts [18],[19],[27] are creating a big virtual device from the devices surrounding the users.

Similar to some of the above approaches, MAIPAN represents an entire personal area network as a single device to applications. In MAIPAN, users can easily reconfigure the PAN transparently to applications via a simple channel management mechanism. Application designs therefore do not have to take PAN configuration, its changes or even the notion of the PAN, into account. The basic ideas of MAIPAN were introduced in [20].

## 3. The MAIPAN Platform

### 3.1. Basic concepts and definitions

MAIPAN distinguishes among *devices, applications* and *services*.

The word "device" refers to the physical device and by "application" we refer to the software that offer the "services". For example, using these abstractions in case of a mouse we can say, that the mouse is a "device" where a "mouse application" is running, which offers a "mouse service". Or another example is the mp3 player: there is a device where the mp3 player application is running, which offers the mp3 player service. The distinction between application and service is necessary, because users want to use services, they do not want to care about the application offering the service.

MAIPAN is based on three concepts (see *Figure 2* – on the next page): *pins, channels* and *sessions*.

Applications offering the services have input and output outlets, which are called pins-borrowing the expression from the integrated circuit world. Pins are the connection points of the applications to the middleware, so the middleware sees the applications in the PAN as a set of input and output pins. A pin has a predefined type, which shows the type of data that the pin can emit or absorb, that is, the type of information the application can handle (e.g., mouse movements, key-

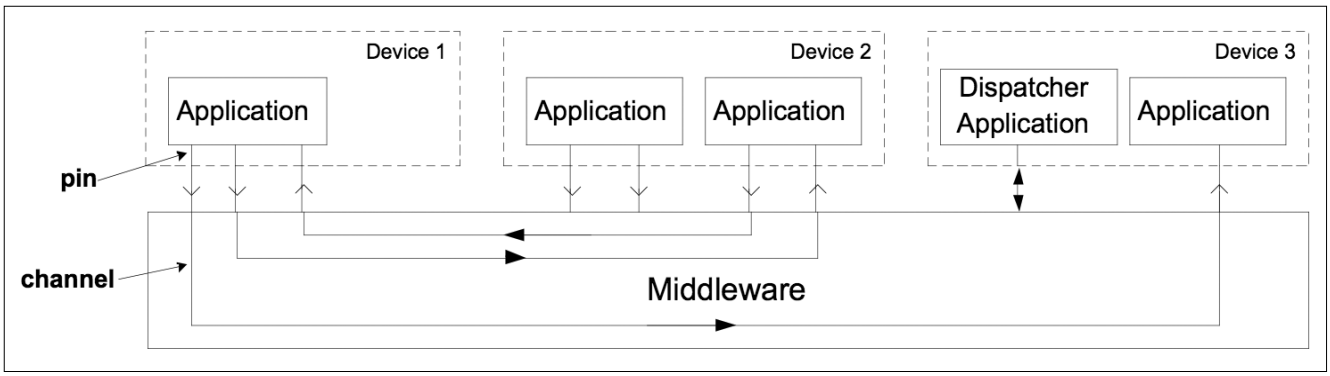



*Figure 2. MAIPAN session*

strokes). According to the needs new types can be defined any time. The dispatcher application is responsible to connect the pins with appropriate types.

To enable communication between pins, the middleware creates and reconfigures channels, which are point-to-point links that interconnect pins. The set of channels that are necessary to use a PAN service is called session.

For example, the channels between the entities that can be seen in the mp3 player scenario (Figure 1) compose the mp3 player session: there is a channel between the fileserver and the mp3 player, another channel between the mp3 player and the speakers, and a control channel between the mp3 player and the PDA. The PDA plays only the role of the *control entity*, it is the owner of the session, but it does not participate in it. Applications are neither aware of channels nor sessions, they only know about their pins.

### 3.2. Security and access control

Security and access control functions are defined and handled on device level. This means that access to services (i.e., access to application pins) are granted for devices, thus if a device gets the right to use a given service, then all applications running on this device will be able to access the service. If we assume that in the PAN there are small devices that offer one or two simple services (e.g., mouse, mp3 player), then in this case it is simpler to grant the access of services to devices instead to each application.

The dispatcher application running on a device, which plays the role of the control entity, can set up and reconfigure sessions. The control entity has to check and ask for the necessary access rights to enable the usage of a given service for the user. For instance, this main control entity can be a PDA, which has enough computing power to manage a PAN. All other devices participating in the PAN are called participants. In special cases, *participants* may delegate the access control rights to other devices, which will be referred to as *manager*.

In this paper we do not go into details regarding authentication and authorization, however, the solutions described in [6] can be used with MAIPAN to ensure secure communication between devices.

### 3.3. Transferring sessions

In the PAN at least one device playing the role of the controller entity is needed. In case this device disappears all concerned sessions will be automatically torn down. To keep up such sessions MAIPAN offers the possibility to transfer a running session from the current control entity to another one. For example, to make some music in a meeting room one of the users creates an mp3 playing session. This way the user's device becomes the control entity of the session. After a while, when the user wants to leave, she can transfer the session by telling to MAIPAN the identity of the new control entity, which can be for example another user's PDA.

### 3.4. Reconfiguring sessions

In case a participant disappears (e.g., the user leaves the room, or the device's battery is depleted), the concerned channels are automatically disconnected and the sessions have to be reconfigured. In the first step MAIPAN notifies the corresponding dispatcher(s) about the event. In the second step the dispatcher application(s) can decide which services to use instead of the disappeared ones. The dispatcher application can ask for user involvement, if there are multiple possibilities to replace the disappeared service(s), or it can decide on its own, if there is no or only one choice, or the user preferences are known. In the third step the dispatcher builds up the new channels or tears down the sessions concerned.

### 3.5. Architecture

Based on the concepts above we designed MAIPAN, whose architecture is depicted in *Figure 3.* Applications run over the middleware, while layers under the middleware provide end-to-end routing and data transmission functions. As it can be seen in the figure, the protocol stack is vertically divided into two parts. The aim of the *data plane* is to provide effective and secure data transport between applications. The *control plane* is responsible for managing pins, channels, sessions and for handling security and access control.

*Data plane*

The application sends data through a pin to the middleware, where the *channel connector layer* re-

directs the data to the corresponding channel. The *transport layer* creates packets and provides functions such as flow control, reordering, automatic re-transmission, quality of service, etc., if these functions are not supported in the layers below the middleware, but are necessary for the application. The *connection layer* adds information to the packet, which is needed for the delivery: address of the source and destination device and the identifier of the channel. Finally, the *cryptography* layer calculates a message integrity check (MIC) value and encrypts the packet if necessary.

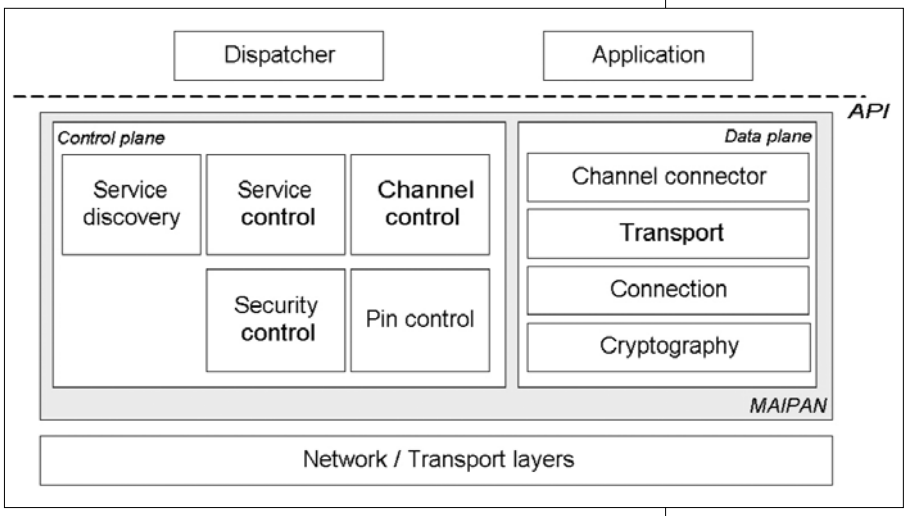



*Figure 3. MAIPAN's architecture*

*Control plane*

The control plane contains the control functions which are necessary to manage the PAN. The *service control* part registers local services offered by the applications, handles their access rights and communicates with the service discovery protocol. In theory, any kind of SDP can be attached to MAIPAN, such as SLP, UPnP or Salutation [21],[22]. The *channel control* creates and reconfigures sessions initiated by the dispatcher application. According to the needs of the dispatcher it asks the pin control parts of the participating devices to build up the channel between the pins. The *pin control* part instructs the channel connector layer to create a channel, activates the necessary transport functions for the given channel in the transport layer and sets the destination of the given channel in the connection layer. The *security control* part initiates and co-ordinates the authentication procedure between devices, manages the service access rights, and stores the necessary information for communication (e.g., security keys).

### 3.6. Implementation

An earlier version of MAIPAN is implemented in C on Linux [23]. We also created some applications (e.g., mp3 player, fileserver) and a dispatcher application in order to study the behavior of the middleware [24],[25]. The lessons learnt from these are also incorporated in the design of the currently presented version of the middleware.

## 4. Conclusion

This paper proposes MAIPAN, a middleware for application interconnection in personal area networks. The essence of the middleware is to create a PAN programming platform, whereby hardware and software resources are interconnected, and the scatteredness of the PAN is hidden from the services. Using the proposed architecture, developers creating distributed applications for PANs do not have to take care of PAN configuration or dynamics, furthermore, they can use the uniform application programming interface (API) offered by the system.

MAIPAN represents a novel approach in its secure access control mechanism and the use of a central control entity. MAIPAN access control ensures

1) seamless interworking of various devices of the same user,

2) protection of one user's devices from devices of another user,

3) still enabling controlled communication and lending between devices of different users.

MAIPAN manages device access and configuration via a convenient central control entity, the dispatcher. MAIPAN is also unique in enabling the change of the dispatcher role, that is, the session control rights can be transferred between devices, from the old dispatcher to a new one.

In the future, we plan to finalize the implementation of the current version of MAIPAN and to perform performance analysis of the middleware.

### References

[1] Mark Weiser:
"The Computer for the 21st Century",
Scientific American, September 1991.

[2] Mark Weiser:
"Some Computer Science Issues in
Ubiquitous Computing",
Communications of the ACM, July 1993.

[3] M. Satyanarayanan:
"Pervasive Computing: Vision and Challenges",
IEEE Personal Communications, August 2001.

[4] Cecilia Mascolo, Licia Capra, Wolfgang Emmerich:
"Middleware for Mobile Computing (A Survey)",
In Advanced Lectures in Networking.
Editors: E. Gregori, G. Anastasi, S. Basagni.
Springer, LNCS 2497. 2002.

[5] A. Pope:
"The Corba Reference Guide:
Understanding the Common Object Request
Broker Architecture",
Addison-Wesley, January 1998.

[6] Pubudu Chandrasiri, Ozgur Gurleyen, Yashar Shahabi, Christian Gehrmann, Annika Jonsson, Mats Naslund: "Personal Security Domains", Contribution to the 10th WWRF Meeting, New York, October 27-28., 2003.
[7] D. Garlan, D. P. Siewiorek, A. Smailagic, P. Steenkiste: "Aura: Toward Distraction-Free Pervasive Computing", IEEE Pervasive Computing, 2002., http://www-2.cs.cmu.edu/ aura/
[8] "Gaia Project: Active Spaces for Ubiquitous computing"; http://gaia.cs.uiuc.edu/index.html
[9] "MIT Project Oxygen", Online Documentation, http://oxygen.lcs.mit.edu/publications/oxygen.pdf
[10] M. Esler, J. Hightower, T. Anderson, G. Borriello: Next Century Challenges: "Data-Centric Networking for Invisible Computing: The Portolano Project at the University of Washington", Mobicom'99., http://portolano.cs.washington.edu/proposal/.
[11] Robert Grimm, Janet Davis, Eric Lemar, Adam MacBeth, Steven Swanson, Thomas Anderson, Brian Bershad, Gaetano Borriello, Steven Gribble, David Wetherall: "System support for pervasive applications", ACM Transactions on Computer Systems, 22(4), pp.421–486., November 2004.
[12] Achilles Kameas, Stephen Bellis, Irene Mavrommati, Kieran Delaney, Martin Colley, A. Pounds-Cornish: "An Architecture that Treats Everyday Objects as Communicating Tangible Components", in proceedings of the First IEEE Int. Conference on Pervasive Computing and Communications (PerCom '03), Fort Worth, Texas, USA, March 23-26., 2003, p.115.
[13] 2WEAR project: "A Runtime for Adaptive and Extensible Wireless Wearables"; http://2wear.ics.forth.gr
[14] "CORTEX Project: CO-operating Real-time sentient objects: architecture and EXperimental evaluation"; http://cortex.di.fc.ul.pt/index.htm.
[15] Barry Brumitt, Brian Meyers, John Krumm, Amanda Kern, Steven Shafer: "EasyLiving: Technologies for Intelligent Environments", in Proc. of Handheld and Ubiquitous Computing Symposium, (Bristol, England), 2000.
[16] W. K. Edwards, M.W. Newman, J. Sedivy, T. Smith: "Challenge: Recombinant Computing and the Speakeasy Approach", MobiCom'02, September 23-28., 2002, Atlanta, Georgia, USA.
[17] Ricardo Baratto, Shaya Potter, Gong Su, Jason Nieh: "MobiDesk: Mobile Virtual Desktop Computing", Proc. of the 10th Annual ACM International Conference on Mobile Computing and Networking (MobiCom 2004), Philadelphia, PA, September 26-October 1., 2004.
[18] Jonvik, T.E., Engelstad, P.E., Thanh, D.V.: "Building a Virtual Device on Personal Area Network", Proc. of 2003 International Conference on Software, Telecommunications and Computer Networks (SoftCOM'2003), Dubrovnik (Croatia) / Ancona, Venice (Italy), October 7-10., 2003.
[19] Jonvik, T.E., Engelstad, P.E., Thanh, D.V.: "Dynamic PAN Based Virtual Device", Proc. of 2nd IASTED International Conference on Communications, Internet and Information Technology (CIIT'2003), November 17-19., 2003.
[20] Miklós Aurél Rónai, Kristóf Fodor, Gergely Biczók, Zoltán Turányi, András Valkó: "MAIPAN: Middleware for Application Interconnection in Personal Area Networks", poster at Mobiquitous 2005 conference, San Diego, CA, USA, July 17-21., 2005.
[21] Rekesh John: "UPnP, Jini and Salutation – A look at some popular coordination frameworks for future networked devices", California Software Laboratories Inc., Technical Report, June 17., 1999.
[22] Feng Zhu, Matt Mutka, Lionel Ni: "Classification of Service Discovery in Pervasive Computing Environments" MSU-CSE-02-24, Michigan State University, EastLansing, 2002.
[23] Kristóf Fodor: "Implementation of a Protocol Stack for Personal Area Networks", Master's Thesis, Budapest University of Technology and Economics, June 2003.
[24] Fodor Kristóf, Kovács Balázs: "A Blown-up rendszer megvalósítása", HTE-BME Student conference, Budapest, Hungary, May 2003.
[25] Balázs Kovács: "Design and Implementation of Distributed Applications in Ad Hoc Network Environment", Master's Thesis, Budapest University of Technology and Economics, May 2003.
[26] Biczók Gergely, Fodor Kristóf, Kovács Balázs, Szabó Ágoston: "Pervasive computing – rejtett számítástechnika", Híradástechnika, Budapest, Hungary, March 2003.
[27] Biczók Gergely, Fodor Kristóf, Kovács Balázs, Szabó Ágoston: "Blown-up rendszer tervezése és megvalósítása", Students' National Scientific Conference, first prize, (első helyezett TDK/OTDK dolgozat), Győr, Hungary, November 2002 / April 2003.

# Multiview Video Presentation and Encoding

LÁSZLÓ LOIS, TAMÁS LUSTYIK, BÁLINT DARÓCZY,
*Budapest University of Technology and Economics, Department of Telecommunications*
*lois@hit.bme.hu*
TIBOR AGÓCS, TIBOR BALOGH
*Holografika Kft.*
*t.agocs@holografika.com*

***Keywords: multi-view video, video encoding, MPEG-4, 3-dimensional animation and visualization, 3-dimensional television***

*In this paper the multi-view video encoding and presentation is introduced. In the first part the current multi-view video displays and systems are described. After then we review the developed multi-view video encoding algorithms and the related computer graphics tools. Based the OpenGL system, we show a Depth Image-base Representation (DIBR) method to render an image from several existing reference pictures in a multi-view environment. We also present a new method to encode the depth image efficiently and build a whole multi-view encoding system where the images in the reference views are encoded by using MPEG-4 AVC and the other images are rendered by DIBR. We compare the distortion of this hybrid algorithm and a standard video encoding method in order to establish new multi-view video formats for Holografika's holographic display. Finally we give some further research objectives to complete the developed hybrid method.*

## 1. Introduction

The three-dimensional television system is likely to play an important role in the future broadcasting. Currently this research area is mostly related to the computer graphics and animation since the multi-view camera systems and displays have not entered into practical use, hence the capturing of the multi-view sequences are implemented by a computer graphic software.

The hybrid video encoding systems like MPEG-1 and MPEG-2 are based on the motion compensated DCT coding scheme. The block-based prediction scheme could be easily used for removing the redundancy between the adjacent views (disparity compensation), hence the conventional video encoding methods could be used for motion and disparity compensation. The multi-view profile of the MPEG-2 video encoding contains both motion and disparity compensation, and the several tools in MPEG-4 support to view an object from any viewpoint. In an MPEG-4 scene the virtual and natural video objects could be mixed since in the MPEG-2 multi-view profile uses only binocular representation.

Beside the motion and disparity compensation hybrid schemes the Depth Image-based Representation (DIBR) could be also used for multi-view video coding. This toolkit is also the part of the MPEG-4 video encoding tools.

This paper organized as follows. In the next section, the prevalent multi-view rendering devices are introduced. In Section 3 we describe the Depth Image-based Representation of the three-dimensional objects. In Section 4 we review the multi-view image and video encoding systems where both the motion and/or disparity compensation techniques and the methods based on Depth Image-based Representation are introduced. In Section 5 we show the developed multi-view encoding system which uses DIBR. In our experiments the reference views are encoded and the others are rendered. The MPEG-4 AVC is used to encode the color information by using motion compensation and the depth information in encoded by using a lossless codec. Finally we give some further research objectives to complete the developed hybrid method.

## 2. Three-dimensional display systems

There are several companies and universities offering 3D display solutions worldwide but all of them can be categorized according to the next groupings of basic principle. Many of those provide new opportunities for 3D presentation, but also present new challenges. This section surveys the capabilities and characteristics of different three-dimensional display technologies.

### 2.1. Volumetric Displays

Volumetric displays use some media positioned or moved in space where they project light beams and so light beams are scattered/reflected from that point of this media which is generally a semi transparent or diffuse surface.

*Moving screen*

One solution is when there is a moving screen and the different perspectives of the 3D object are projected on it. A well known solution (Actuality Systems) is a lightweight screen sheet that is rotated at very high speed in a protecting globe and the light beams from an array of LED-s or microdisplays (DMD from Texas Instruments) are projected onto it. By proper synchronization 3D objects can be seen in the globe.

*Double or Multi-layer screen*

Other commonly used technique in volumetric display technology when two or more LCD layers act as a

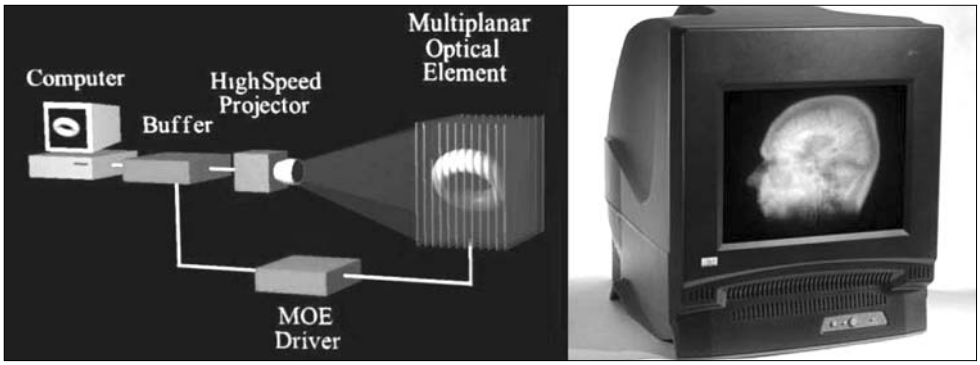



*Figure 1. A multi-layer based volumetric display solution*

projection screen, creating the vision of depth. Deep Video Imaging produced a 17" display which consist two LCD with the resolution of 1280x1024. The Depth Cube from LightSpace Technologies has 20 XGA (1024x 768) layers inside. The layers are LCD sheets that are transparent/opaque (diffuse) when switched on/off, and are acting as a projection screen positioned in 20 positions. Switching the 20 layer is synchronized to the projection and an adapting optics is keeping the focus.

In volumetric displays the portrayed objects appear transparent, since the light energy addressed to points in space cannot be absorbed by foreground pixels. Thus, practical applications seem to be limited to fields where the objects of interest are easily represented by wire frame models *(Figure 1)*.

### 2.2. Autostereoscopic Displays

Autostereoscopic displays provide 3D perception without the need for special glasses or other head-gear, the separation for the left/right eye could be implemented using various optical or lens raster techniques directly above the screen surface. Two basic technologies exist to make autostereoscopic displays: stereoscopic and multi-view displays.

#### 2.2.1. Stereoscopic Technology

It has long been known how to make a two-view autostereo display using parallax barrier, lenticular sheet or micropolarizer-based technology. These divide, into two sets, the horizontal resolution of the underlying, typically liquid crystal display device. One of the two visible images consists of every second column or row of pixels, the second image consists of the other columns or rows. The two images are captured or generated so that one is appropriate for the viewer's left eye and one appropriate for the right. The two displayed images are visible in multiple zones in space. If the viewer stands at the ideal distance and in the correct position he or she will perceive a stereoscopic image. The downside of this is that there is a 50% chance of the viewer being in the wrong position and seeing an incorrect, pseudoscopic image. These serious limitations necessitate the use of another autostereo solution. This is either to increase the number of views or to introduce head tracking.

*Passive Stereoscopic Technology*

This type of displays requires the viewer to be carefully positioned at a specific viewing angle, and with her head in a position within a certain range, otherwise the stereoscopic view will disappear. The information provided by these systems is only twice of the amount contained in a 2D image. Moreover, there are physiological side effects e.g., the contradiction between accommodation and focusing, that can produce discomfort.

*Tracking Stereoscopic Technology*

To overcome the aforementioned limitations, manufacturers of stereoscopic displays are developing head/eye-tracking systems capable of following the viewer's head/eye movement. Even if this solution cannot support multiple viewers, and there could be latency effects, it provides the viewer with parallax information and it is, therefore, a good solution for single user applications.

#### 2.2.2. Multi-view Technology

This display projects different images to multiple zones in space. The whole viewing space is divided into a finite number of horizontal windows. In each window only one image (view) of the scene is visible. The viewer's two eyes see a different image, and the images

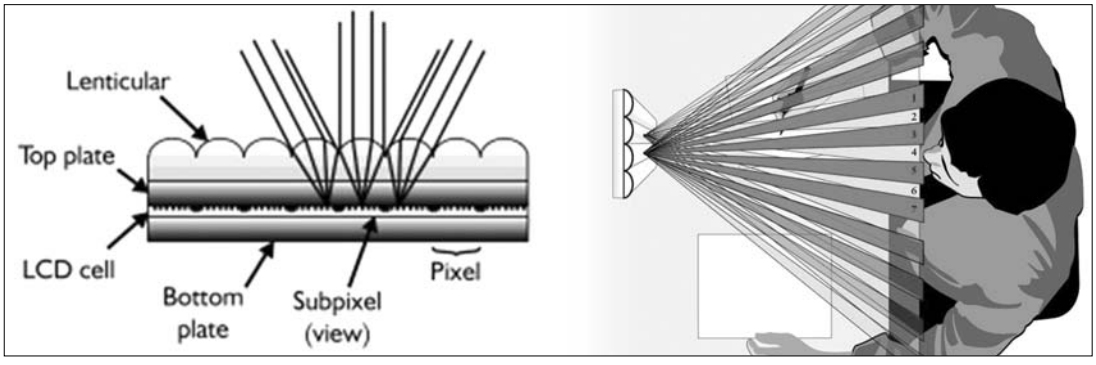



*Figure 2. Lenticular display*

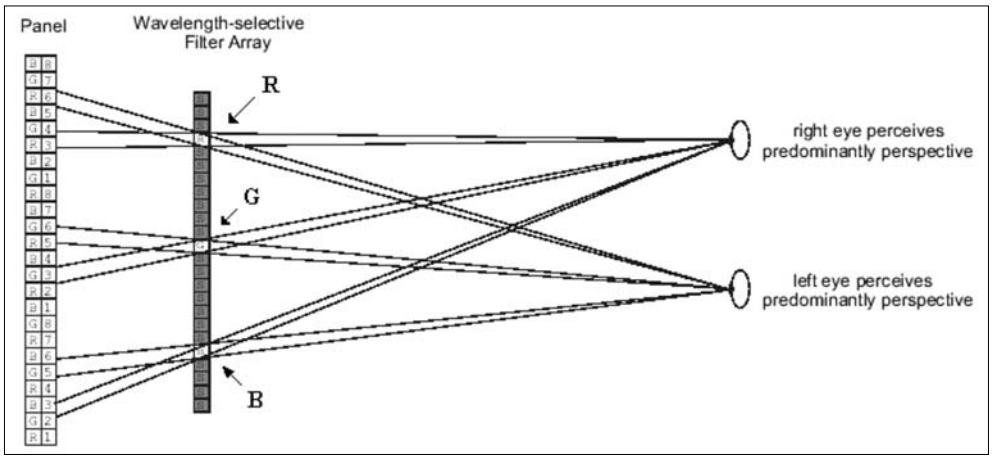



*Figure 3.*
*Wavelength selective filter based autostereoscopic display*

change when the viewer moves his head, causing "jumps" as the viewer moves from window to window. It does not require 3D eyeglasses and allows multiple simultaneous viewers, restricting them, however, to be within a limited viewing angle.

*Lenticular Displays*

Multi-view displays are often based on an optical mask, on a lenticular lens array. This is a sheet of cylindrical lenses placed on top of a high resolution LCD in such a way that the LCD image plane is located at the focal plane of the lenses. The effect of this arrangement is that different LCD pixels located at different positions underneath the lenticulars fill the lenses when viewed from different directions. Provided these pixels are loaded with suitable stereo information, a 3D stereo effect is obtained in which left and right eyes see different but matching information. Lenticular state of the art displays typically use 8-10 images *(Figure 2)*.

*Parallax Barrier Displays*

A parallax barrier, which comprises an array of slits spaced at a defined distance from a high resolution LCD, is one such a micro-optical component like the lenticular lens array. The parallax effect is created by this lattice of very thin vertical lines, causes each eye to view only light passing through alternate image columns.

*Displays with different optical filters and rasters*

These displays are based on wavelength dependent filters creating the necessary divided viewing space for the 3D vision. The wavelength-selective filter array is placed on a flat LCD panel, and a combination of several perspective views (state of the art displays provide eight views) is represented to the observer.

The images for wavelength-selective filter arrays contain different views which are combined in a regular pattern. The filter array itself is positioned in front of the display and radiates the light of the pixels from the combined image into different directions, depending on their wavelengths. As seen from the viewer position the different spectral components are blocked, filtered or transmitted. So a perception of different images in the viewing space is enabled *(Figure 3)*.

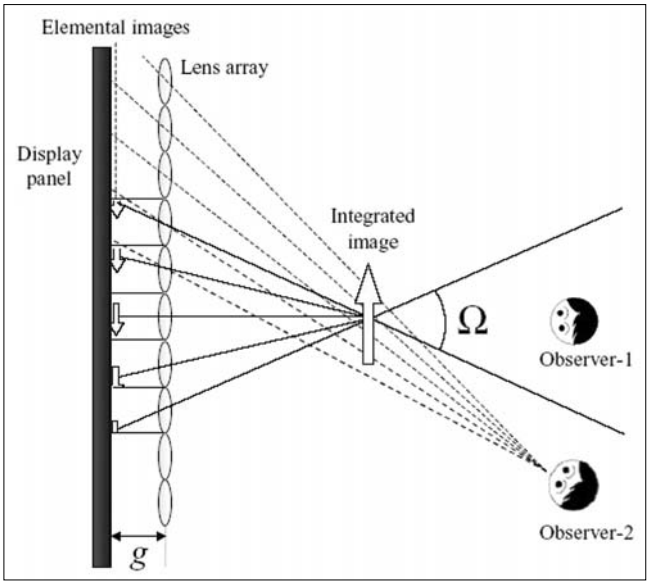
*Integral imaging*

Integral imaging (InIm) uses a lens array and a planar display panel. Each elementary lens constituting the lens array forms each corresponding elemental image based on its position relative to the object, and these elementary images displayed on the display panel are integrated at the original spatial position of the object forming a 3D image. A primary disadvantage of InIm is its narrow viewing angle. The viewing angle, the angle within which observers can see the complete image reconstructed by InIm, is limited due to the restriction of the area where each elemental image can be displayed. Generally, in the InIm system each elemental lens has its corresponding area on the display panel. To prevent image flipping, the elemental image that exceeds the corresponding area is discarded optically in direct pick up method or electrically in computer-generated integral imaging (CGII) method. Therefore, the number of the elemental images is limited and an observer outside the viewing zone cannot see the integrated image *(Figure 4)*.

### 2.3. Holographic Techniques

The holographic technology has the ability to store and reproduce the properties of light waves. Attempts have been made to use acousto-optic material. Further moving holograms have been created using optically addressed spatial light modulators. Pure hologram tech-

*Figure 4. Integral imaging displaying method*



nology utilizes 3D information to calculate a holographic pattern. This technology generates true 3D images by computer control of laser beams and the position of a system of mirrors. Compared to stereoscopic and multi-view technologies, the main advantage of a hologram is in the quality of the picture. The historical disadvantages are: the huge amount of information contained in the hologram which limits its use to mostly static 3D models; and the total incompatibility with existing displaying conventions.

*Holographic technology from Holografika*

Each point (voxel) of the holographic screen of Holo Vizio system emits light beams of different colour and intensity to the various directions (exactly how a point of a window does), in a controlled manner The light beams are generated through a patented specially arranged light modulation system and the holographic screen makes the necessary optical transformation to compose these beams into a perfect 3D view. The light beams cross each other in front of the screen or they propagate as if they were emitted from a common point behind the screen. With a proper software control of the light beams viewer or viewers see objects behind the screen or floating in the air in front of the screen. The system can be upgraded to large scale (wall-size holographic screens), resolution, brightness, etc. is not limited by principle *(Figure 5)*.

The main advantage of this approach is that, similarly to the pure holographic displays, it is able to provide all the depth cues and it is truly multi-user within a reasonably large field of view. This is a high-end solution compared to other technologies and fulfils all the requirements of real 3D displaying simultaneously, it creates all light beams that are present in a natural 3D view, that is the reason why one sees the same as in reality. The display is able to provide all the depth cues and is truly multi-user within a reasonably large field of view. This is qualitatively very different from other contemporary multi view technologies that force users into approximately fixed positions, because of the abrupt view-image changes that appear at the crossing of discrete viewing zones. By contrast, this holographic display provides continuous horizontal parallax with 0.8° angular resolution for the full 50° field of view, free of any transitions between view images. Reconstruction of the perfect 3D view makes this system a really working 3D display technology *(Figure 6)*.

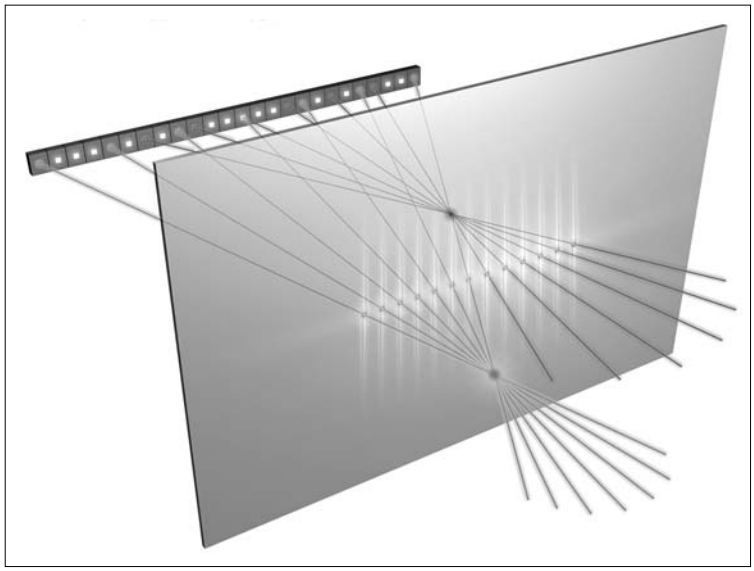
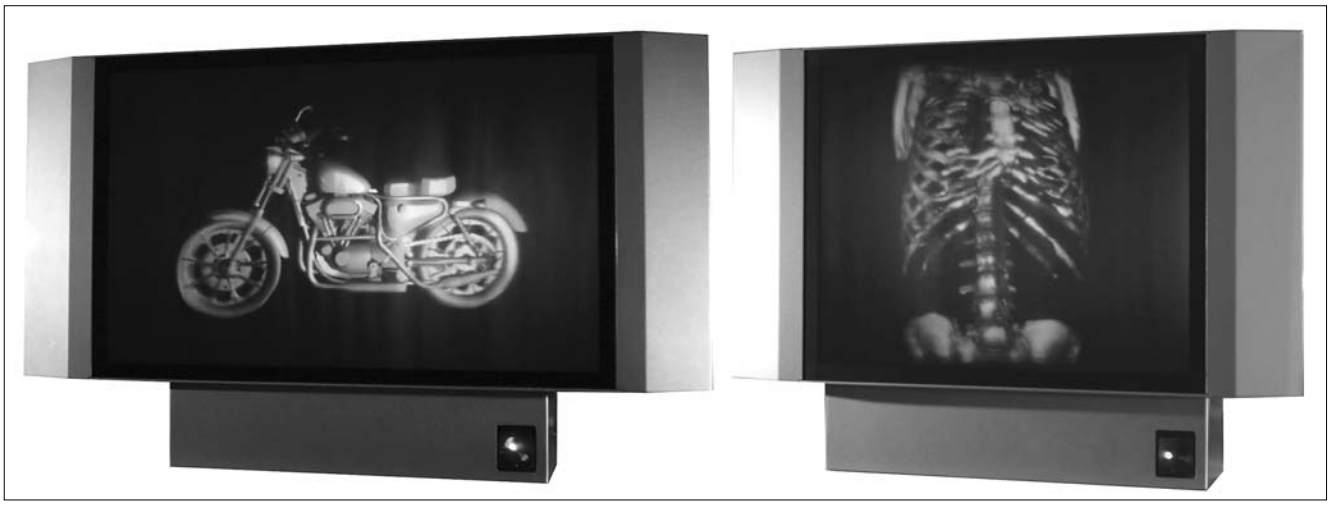


*Figure 5. HoloVizio system from Holografika*

## 3. Depth Image Based Representation of 3D objects

Image based rendering (IBR) techniques have been proposed as an efficient way of generating novel views of real and synthetic objects. The first step in the DIBR developed was the generalization of sprites, a technology already widespread in the 80's. The sprite is a planar projected image of an object or a part of a 3D scene, which is moveable by means of an affine transfor-

*Figure 6. HoloVizio in operation – CAD and medical application*



mation according to the camera. However, in the case of fast camera movement or changes in orientation, this method results in noticeable distortion. If we also store the depth value of the point of the sprites beside its colour (sprite with depth [1]), then we can handle the slight non-affine distortions of texture and shape depending on the camera parameters. The depth value of a pixel means the distance from the spectator.

Based on the sprites with depth values, an extended picture with depth information was defined (relief image) which contains the depth values in addition to the colour components.

There are some cameras on the market that can sense the distance between the point and the camera, but their accuracy is currently not acceptable, but in the near future it would possible to create relief images with cameras. Besides, there are several algorithms which are able to recognize the convergent points of different sequences and set the depth value of transparent points by triangulation.

When scanning the 3D scene, the screen position of points in an other view can be calculated with the help of the depth values. However, it is easy to see that there are several points which will be positioned outside the screen on the actual view or covered by an another point. This means that the rendered image will be incomplete. To fill the gaps in multi-view sequences, we can take another image from a different view, or for smaller gaps, apply an interpolation filter.

The missing pixels on a rendered image could be covered on the reference images by an another object. For these cases, the Layered Depth Image (LDI) was introduced in [1]. LDI contains potentially multiple depth pixels at each discrete location in the image. Instead of a 2D array of pixels with associated depth information, it stores a 2D array of layered depth pixels. A layered depth pixel stores a set of depth pixels along one line of sight sorted in front to back order. The front element in the layered depth pixel samples the first surface seen along that line of sight; the next pixel in the layered depth pixel samples the next surface seen along that line of sight, etc. When rendering from an LDI, the requested view can move away from the original LDI view and expose surfaces that were not visible in the first layer. The previously occluded regions may still be rendered from data stored in some later layer of a layered depth pixel. Using this description, we can calculate some of the missing points, but it is not possible to create such descriptions with a camera.

## 4. Multi-view video coding

In the case of multi-view video coding, a number of cameras are used for recording a given scene, which operate in a synchronized manner. Consequently, the pixels that are visible from more than one camera will be sampled at the same time. Owing to the spatial arrangement of the cameras, pixels representing the same point map to different coordinates along the views. Moreover, in case of a real illuminated surface, intensity and chrominance are likely to vary as well.

### 4.1. MPEG-2 multi-view profile

While MPEG-1 does not contain any recommendations aiming at the efficient coding of stereo images or other redundant sequences, MPEG-2 comes with the support for stereo coding. The MPEG-2 multi-view profile defined in 1996 as an extension to the MPEG-2 standard, introduced a new concept for temporal scaling and also made it possible to send various camera parameters within the standard bitstream [3].

This profile is based on that of MPEG-2 temporal scalability mode. The first, higher priority bitstreams codes video at a lower frame rate, and the intermediate frames can be coded in a second bitstream using the first bitstream reconstruction as prediction. The same principle is used in the multi-view profile, though in this case multiple refinement layers can be predicted from the view considered as the 'middle one'.

Here, of course, the term 'refinement layer' refers to a separate view along which the appropriate camera parameters are also transmitted in the bitstream. This way the prediction between layers becomes prediction between views *(Figure 7)*.

In MPEG-2 multi-view profile, coding gain is achieved only when the additional layer(s) could be encoded at a lower bit-rate than the sum of the bitrate of the separately coded original sequences. According to the observations, layers refining the side views cannot be compressed significantly better at the same quality, hence the bit-rate needed is proportional to the number of views.

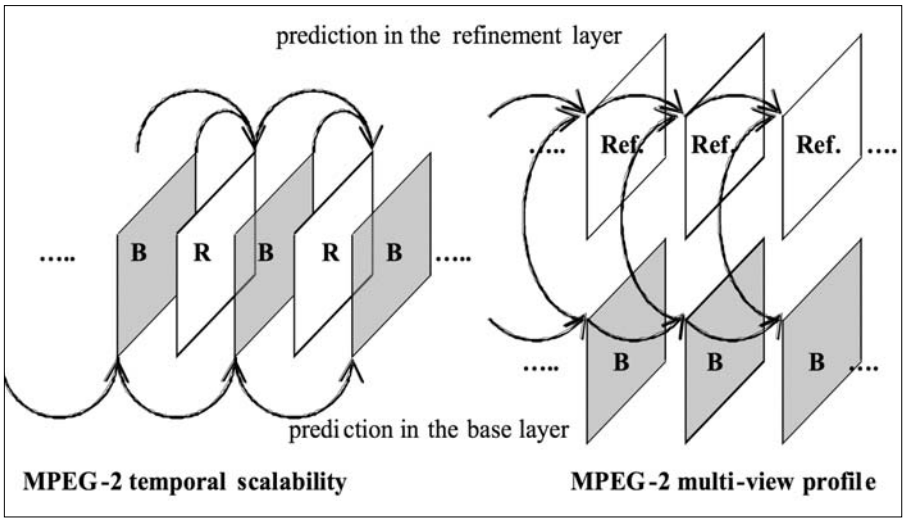



*Figure 7. Prediction method of MPEG-2 temporal scalability and multi-view profile*

#### 4.2. Implementing inter-view prediction in another MPEG encoder

The aforementioned principle can be applied in case of other codecs using inter-frame prediction. Choosing MPEG-4 AVC, the most efficient MPEG video coding tool at the moment, is beneficial since the use of AVC inherently means a significant increase in coding gain.

MPEG-4 AVC [4] is designed to encode rectangular objects, including natural camera sequences, too. In multi-view systems, one of the toughest problems concerns the proper calibration of cameras, since the more cameras we have, the more difficult it becomes to set the same level of light sensitivity. For this reason, in [5] block-based inter-frame motion compensation is complemented by an illumination compensation which models the effect of illumination by offsetting and scaling the intensity component. These two illumination parameters are also encoded beside the usual motion vector and motion segmentation information. Regarding the tests, this method resulted in 0.5-1.0 dB coding gain, and subjective evaluations also stated that the number of errors appearing in critical regions decreased significantly.

#### 4.3. MPEG-4 Part 16: Animation Framework eXtension (AFX)

MPEG-4 AFX [6] contains the tools for depth-image based representation to a large extent. Static and animated 3D objects of AFX have been developed mainly on the basis of VRML (Virtual Reality Modeling Language), consequently the tools are compatible with the BIFS node and the synthetic and audiovisual bitstreams as well [7-9]. There are two of the main data structures of AFX which are essential concerning DIBR: these are *SimpleTexture* and *PointTexture*.

The *SimpleTexture* data structure holds a 2D image, a depth image and the camera parameters (position, orientation, projection). The pixels of the depth image can be transmitted either pixel-by-pixel as a fourth coordinate beside the intensity and the chrominance signals, or as a separate grayscale bitstream. In most cases, the coding of depth images is lossless since the distortion of depth information along with changing the camera parameters can lead to the miscalculation of pixel positions.

However, a SimpleTexture structure is not enough for describing an object, apart from viewing it only from one direction and allowing minimal deviations of position and orientation from that of the middle camera. For the sake of more complex cases, the *PointTexture* data structure should be used for characterizing objects. Point Texture stores intersections of the object and various straight lines by means of assigning depth value and colour information to each point of intersection.

#### 4.4. Depth-image based multi-view image and video coding

A number of articles deal with depth-image based coding beyond the tools of MPEG-4. One of the most comprehensive achievement is described in [10], which introduces a new algorithm for the efficient coding of the MPEG-4 AFX *OctreeImage* representation and also gives an exhaustive and interesting overview of DIBR issues.

MPEG-4 AFX is not the only available tool for LDI coding. A solution enumerating a couple of alternatives can be found in [11]:

• Store the number of layers (namely the number of colour-depth pairs) for each pixel. This image can be encoded as a grayscale image with usually a few gradation levels. The JPEG-LS [12] algorithm has proven to be significantly more efficient than the Deflate (ZIP) algorithm regarding either compression ratio or elapsed time.

• The different colour component and depth layers are encoded separately. Since the larger are the layers, the smaller is the number of pixels, the following alternatives have been examined:
- MPEG-4 Shape Adaptive DCT
- Rectangular shape image completion, encoded by JPEG 2000
- MPEG-4 arbitrary shape codec
- VOW (Video Object Wavelet) codec that is able to encode arbitrary shaped images just the one above. According to the observations, VOW performed best at both the color and distance layers.

JPEG 2000 is also used for encoding a still picture in a multi-view manner [13]. In this research, the authors experienced that in case of 16 bit depth images the JPEG 2000 encoder produced such reconstructed picture quality at 0.3 bpp compression ratio that is just acceptable for DIBR. Important regions for ROI (Region of Interest) coding in JPEG 2000 are set by examining the depth image and the picture itself. Additionally, depth information is companded before encoding.

## 5. Experiments and results

In this section, we enumerate the effectiveness of the depth image based multi-view video encoding by implementing the DIBR rendering algorithm, the encoding of the depth and the color images and generating a natural-like synthetic multi-view video sequence. The target bit-rate was between 24 and 30 Mbps which fits into one DVB-T multiplex and the target number of cameras was 60 which corresponds the holographic display by Holografika.

The 60 cameras capture a virtual reality scene where 12 different cars turn in a simple 90-degree bend while there are trees and a high-rise blocks of flats in the background. The resolution was 512x320 for each camera. The car models of the Need For Speed 3 Hot Pursuit are used, this format could be easily imported and contains a car-specific texture-mapped polygon representation of 3-D objects (car body and several – usually four – wheels). Large amount of overlaps occurred on the rendered images and due to the transpa-

rent texture parts some image areas could be captured only by one camera. Only some viewpoint was compressed where the RGB image is encoded by using MPEG-4 AVC and the depth image is compressed by a lossless method. The missing views are rendered by using the decoded RGB and depth image.

In the first step we determine that 9 bit depth value in each pixel is capable of rendering the images at PSNR of 38 dB when the DIBR rendering uses the original images as references.

| Number of reference views | Camera angle |
|---|---|
| 13 | 0°, ±5°, ±10°, ±15°, ±20°, ±25°, ±30° |
| 11 | 0°, ±6°, ±12°, ±18°, ±24°, ±30° |
| 7 | 0°, ±10°, ±20°, ±30° |
| 5 | 0°, ±15°, ±30° |

*Table 1. The angle of camera of the reference views*

### 5.1. Image synthesis by using depth information and reference images encoded by MPEG-4 AVC

To develop a depth-based image rendering system we must define the proper parameters of each camera. Since we use OpenGL to produce the multi-view sequence the camera parameters could be extracted from the OpenGL transform matrices.

In the OpenGL, a point (*x,y,z*) in the virtual space is transformed to display coordinates (*u,v,d*) using homogeneous coordinates as follows:

1. Modelview transform:

$$[x', y', z', h'] = [x, y, z, 1] \cdot T_{view} \quad (1)$$

where $T_{view}$ denotes the modelview transform matrix which can be requested from the OpenGL system after defining the viewing transform.

2. Perspective transform:

$$[x'', y'', z'', h''] = [x', y', z', h'] \cdot T_{pers} \quad (2)$$

where $T_{pers}$ is the perspective transform matrix.

3. Display transform:

$$u = (x'' + 1) \cdot \frac{V_{sx}}{2} + V_x \qquad v = (y'' + 1) \cdot \frac{V_{sy}}{2} + V_y$$
$$z_d = \frac{Z_{max} - Z_{min}}{2} \cdot Z'' + \frac{Z_{max} + Z_{min}}{2} \quad (3)$$

where *u* and *v* denotes the horizontal and vertical pixel position on the rendered picture and $z_d$ denotes the depth information, $Z_{min}$ and $Z_{max}$ denote the minimal and maximal depth value, respectively, the width and height of the display is $v_{sx}$ and $v_{sy}$ and the top-left coordinate is ($v_{sx}$, $v_{sy}$).

This transform could be also represented using a 4x4 matrix, this matrix is simply called $T_3$- in this paper, hence

$$[u, v, z_d, \tilde{h}] = [x'', y'', z'', h''] \cdot T_3 \quad (4)$$

By using the above equations, the DIBR-based image reconstruction could be implemented as follows:

• Let $R_k(,)$ denote the reference image in the *k*-th view and $Zbuff_k(,)$ the corresponding depth image which contains the depth value of rendered pixels and there are $+\infty$ in the untouched pixel positions. To render the target image in the *n*-th view by using the reference image in the *k*-th view we use the z-buffer algorithm.

• For every $(u_k,v_k)$ on the :

– Get the depth information $z_k$ at the $(u_k,v_k)$ position on the depth image

$$z_k = Zbuff_k(u_k, v_k) \quad (5)$$

– Determine the original 3D coordinate (*x,y,z*) as follows (6):

$$[x, y, z, h] = [u_k, v_k, z_k, 1] \cdot (T_{view,k} \cdot T_{pers,k} \cdot T_{3,k})^{-1}$$

– Calculate the position and depth on the picture in the *n*-th view as follows (7):

$$[u_n, v_n, z_n, h_n] = [x, y, z, h] \cdot T_{view,n} \cdot T_{pers,n} \cdot T_{3,n}$$

After dividing the vector by scalar to keep the last component = 1 in the homogenous representation we get the position and the depth value on the target image. Both coordinates of the position must be quantized to get an integer value and the pixel of the target image at this integer coordinate position must be updated according to the z-buffer algorithm. By scanning all reference pictures the target image could be rendered but some pixels could be untouched.

The quantization of the position coordinated could cause a misplacement error by one pixel position, while the quantization of the depth value can enlarge the misplacement error by more pixel position.

The steps of the image synthesis are shown in *Figure 8.*

First the nearest reference view is used to render the target image, after that all reference views are used in ascending order of distance from the target image. Finally, the missing pixels are produced by an interpolating filter. The covered area is mainly covered by more pixels than the area contains since several pixel occurs on more reference images but the shading could be different due to the illumination. This case could be detected by comparing the shading, the position and depth value on the target picture, and if these values show significant similarity we assume that these pixels from the reference images correspond a same pixel in the 3D space hence we use the shading of the pixel form the most adjacent view.

Due to misplacement, illumination and coverage problems the subjective and objective quality of the rendered images are very different. While the objective quality in PSNR of the rendered image was very poor, in contrast the subjective tests showed significant degra-

*Figure 8/a.*
*Prediction from the nearest left reference image (PSNR = 24.10 dB)*

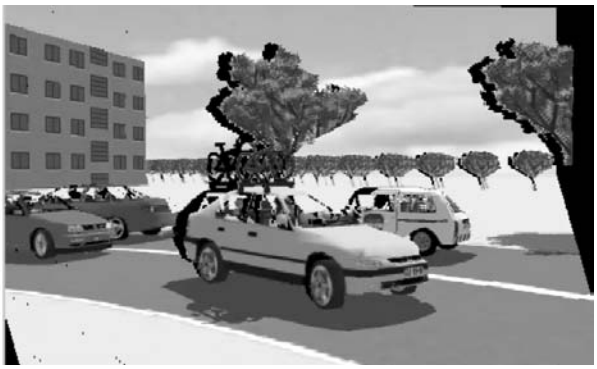


*Figure 8/b.*
*Prediction from the nearest right reference image (PSNR = 27.63 dB)*

*Figure 8/c.*
*Prediction from the nearest left and right reference image (PSNR = 33.45 dB).*
*There are uncovered pixels around the tree on left and the white car.*

*Figure 8/d.*
*Prediction by using all reference views (PSNR = 34.64 dB).*
*The missing pixels are almost covered.*

*Figure 8/e.*
*The result of the interpolation filtering of the missing pixels on Figure 8/d (PSNR = 36.25 dB)*

*Figure 8/f.*
*The original image*

dation only by the boundary of the objects. To find a suitable objective distance measure, we define two new distortion measures based on the well-known PSNR measure. The first measure called $PSNR_{covered}$ is used to handle the coverage problem, this measure is the average PSNR evaluated only on the covered pixels. To handle the misplacement error we define $PSNR_u$ where the horizontal coordinate denoted by $u$ in (7) is rounded to the nearest integer position where the squared error between the rendered pixel and the original pixel is smaller. The latter solution can not be applicable in a real system and used only for test purposes.

*Table 2* shows the results of these distance measures for different number of reference views where the reference views are compressed by using our H.264/AVC codec with quantization parameter of 32.

The results in Table 2 show that the best performance could be achieved by a proper rounding of the horizontal coordinate $u$ in (7) but this could not be implemented in a real system in this way. Furthermore, the uncovered pixels cause only slight quality degradation.

Finally, the larger number of reference views generate insignificant quality improvement while the encoding of the depth images of reference views are very

costly. Hence we suggest to use 5 or 7 reference views in the following experiments.

### 5.2. Lossless compression of depth images

The depth image pixels are quantized with 9 bits and these 9 bit symbols must be restored without error. The current lossless compression algorithms are designed for 8 bit alphabet hence one has to develop a new method or modify the current algorithms. In our experiments we modified the following well-known algorithms to support 9 bit images:

**Deflate (ZIP):**

Deflate compression is an LZ77 derivative used in zip, gzip, pkzip and related programs. In 1977 Abraham Lempel and Jacob Ziv presented their dictionary based scheme for text compression which outputs offset and lengths to the previous text seen and also outputs the next byte after the match.

**GIF87a (LZW):**

The GIF is an implementation of LZW (Lempel-Ziv-Welch algorithm) with some special code. LZW is a dictionary based scheme based on LZ78 which outputs bytes and codewords: pairs of offsets and lengths. LZW outputs always codewords, they refer to a dictionary that has 4096 entries.

**Burrows-Wheeler transform with WFC and RLE-BIT algorithm:**

This hybrid method is based on the implementation of Burrows-Wheeler transform by Jürgen Abel [14]. It is based on a permutation of the input sequence – the Burrows-Wheeler Transformation (BWT) –, which groups symbols with a similar context close together. It the used version, this permutation was followed by a Weighted Frequency Count (WFC) ranking algorithm and a final entropy coding stage using RLE-BIT algorithm and adaptive arithmetic coding.

**PNG (Portable Network Graphics):**

The image compression algorithm of PNG works on a byte basis. First, the pixels of the image are encoded as bytes. Optionally, the bytes are filtered by one of the 4 predictors to improve compression. The prediction error bytes are compressed to by using the deflate algorithm and the resulting symbols are Huffman encoded.

**JPEG-LS:**

This method is the predictive lossless part of the JPEG image coding standard. Similarly to the PNG, is uses 8 predictors and the prediction error in each pixel is Huffman encoded.

The JPEG-LS is only method that supports the 9-bit alphabet, hence the other methods are modified to support 9 bit input symbols. The result of the 9 bit variants are shown in *Table 3*. The performance of the 9 bit BWT algorithm is significantly better compared to the other methods hence we use this result to calculate the required bit-rate for encoding the multi-view video sequence.

### 5.3. The bandwidth of the encoded RGB and depth images

Based upon the latter results, *Table 4.* shows the required bandwidth of the encoded RGB and depth images of the reference views. The results show that at 20.74 Mbps the PSNR value of 26.7 dB can be achie-

| Number | Images rendered by DIBR | | | Every image | | |
|---|---|---|---|---|---|---|
| of ref. views | PSNR | $PSNR_{covered}$ | $PSNR_u$ | PSNR | $PSNR_{covered}$ | $PSNR_u$ |
| 5 | 24.988 | 24.625 | 28.029 | 25.726 | 25.394 | 28.518 |
| 7 | 25.589 | 25.282 | 28.831 | 26.555 | 26.283 | 29.425 |
| 11 | 25.783 | 25.527 | 29.082 | 27.269 | 27.059 | 29.973 |
| 13 | 25.776 | 25.535 | 29.073 | 27.535 | 27.345 | 30.129 |

*Table 2. The resulting values of the distance measures for different number of reference views*

| Lossless image compression method | Total size of 610 images [Bytes] | Average bit rate for one camera [Mbps] |
|---|---|---|
| LZW: 9 bit..13 bit | 19332712 | 6.34 |
| WinZIP (LZ77, 8 bit) | 12049986 | 3.95 |
| 9 bit LZ77 encoding | 10474861 | 3.43 |
| 9 bit PNG encoding (PNG prediction + LZ77) | 11782353 | 3.86 |
| JPEG-LS | 16188900 | 5.31 |
| 9 bit BWT + RLE-BIT + WFC | 8321324 | 2.73 |

*Table 3. Lossless encoding of the depth images in the test sequence by using the 9 bit variant of the compression algorithms*

| Number of reference views | AVC QP | RGB images [Mbit/s] | Depth images [Mbit/s] | Total bitrate [Mbit/s] | Average PSNR [dB] |
|---|---|---|---|---|---|
| **5** | **32** | 7.44 | 13.65 | 21.09 | 25.726 |
| **7** | **32** | 10.42 | 19.1 | 29.52 | 26.555 |
| **11** | **32** | 16.37 | 30.01 | 46.38 | 27.269 |
| **13** | **32** | 19.34 | 35.47 | 54.81 | 27.535 |
| **61** | **42** | 20.74 | none | 20.74 | 26.678 |

*Table 4. Comparison of total bit rate and average PSNR values for the hybrid DIBR-AVC schemes with several number of reference views. Note that the rendered images require no bandwidth in this scheme but their PSNR values are also calculated in the average PSNR values.*

ved by using MPEG-4 AVC with quantization parameter of 42. Almost the same quality could be obtained at 21 Mbit/s by using 5 reference views, where the RGB and depth images of the 5 views are encoded and the images in the remaining views are reconstructed by using DIBR.

The two most promising configurations are capable of transmitting the multiview sequence of 61 cameras on a single DVB-T channel. The single MPEG-4 AVC based scheme has a slightly better performance in quality, but the implementation of DIBR reconstruction is cheaper.

## 6. Conclusions and further work

In this paper, the two basic methods of multi-view video encoding were shown and compared. First we introduced the multi-view video representation and the devices capable of displaying the multi-view video sequences. We described the Depth Image-based Representation of the three-dimensional objects and reviewed the multi-view image and video encoding systems.

In Section 5 we showed the developed multi-view encoding system using DIBR. In our system the MPEG-4 AVC is used to encode the color pictures of reference views, and the corresponding depth information is encoded by a 9-bit lossless codec based on BWT. The other views are rendered by using the decoded depth and color image of the reference views. The developed hybrid method AVC could achieve the performance of the single-view MPEG-4 AVC scheme.

Further research will be performed on improvement of the performance of lossless compression of the depth images by utilizing the redundancy between the depth images. This redundancy can be by exploited since several depth value in a depth picture of a reference view can be predicted from an other reference view by using (7). As we have shown in the results the misplacement problem after (7) is also an important topic, this could be handled i.e. by using offset buffer on sub-pixel basis or an object warping technique could be also capable of solving this problem.

Another interesting aspect could be a stronger integration of the DIBR reconstruction and the hybrid motion compensated transform coding. In this topic the rendering error of the rendered images could be further compressed by a transform coding engine or the rendered images could be used as reference images in motion compensation.

**References**

[1] Jonathan Shade, Steven Gortler, Li-wei Hey, Richard Szeliski, “Layered Depth Images”, SIGGRAPH 98, Orlando Florida, July 19-24, 1998, Computer Graphics Proceedings, pp.231–242.

[2] M. Oliveira, G. Bishop, D. McAllister, “Relief textures mapping,” in Proc. SIGGRAPH, July 2000, pp.359–368.

[3] Jens-Rainer Ohm, “Stereo/Multiview Video Encoding Using the MPEG Family of Standards”

[4] ISO/IEC 14496-10:2003, Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding.

[5] Joaquin Lopez, Jae Hoon Kim, Antonio Ortega, George Chen, “Block-based Illumination Compensation and Search Techniques for Multiview Video Coding”, Picture Coding Symposium, San Francisco, CA, Dec. 2004.

[6] Information Technology – Coding of Audio-Visual Objects – Part 16: AFX – Animation Framework eXtension, ISO/IEC Standard JTC1/SC29/WG11 14 496-16:2003

[7] Information Technology – Coding of Audio-Visual Objects – Part 1: Systems, ISO/IEC Standard JTC1/SC29/WG11 14 496-1.

[8] Information Technology – Coding of Audio-Visual Objects – Part 2: Visual, ISO/IEC Standard JTC1/SC29/WG11 14 496-2.

[9] Information Technology – Coding of Audio-Visual Objects – Part 3: Audio, ISO/IEC Standard JTC1/SC29/WG11 14 496-3.

[10] Leonid Levkovich-Maslyuk, Alexey Ignatenko, Alexander Zhirkov, Anton Konushin, In Kyu Park, Mahnjin Han, Yuri Bayakovski, “Depth Image-Based Representation and Compression for Static and Animated 3-D Objects”, IEEE Transactions On Circuits and Systems for Video Technology, Vol. 14, No.7, July 2004, pp.1032–1045.

[11] Jiangang Duan, Jin Li, “Compression of the Layered Depth Image”, IEEE Transactions On Image Processing, Vol. 12, No.3, March 2003, pp.365–372.

[12] M. Weinberger, G. Seroussi, G. Sapiro, “The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS,” IEEE Trans. Image Processing, Vol. 9, Aug. 2000. pp.1309–1324.

[13] Ravi Krishnamurthy, Bing-Bing Chai, Hai Tao, Sriram Sethuraman, “Compression and Transmission of Depth Maps for Image-Based Rendering”.

[14] www.data-compression.info/JuergenAbel/Preprints/Preprint_After_BWT_Stages.pdf

# Evaluation of IPv6 Services in Mobile WiFi Environment

ZOLTÁN GÁL, ANDREA KARSAI, PÉTER OROSZ

*University of Debrecen, Service Center of Informatics*
*zgal@cis.unideb.hu, kandrea@fox.unideb.hu, oroszp@delfin.unideb.hu*

***Keywords: Internet2, IPv6, WiFi, L2 and L3 roaming, TCP Slow Start algorithm, TCP Windowing algorithm***

*IPv6 may play an important role in the introduction of mobile services in next generation networks. One of the key questions in this context is the impact of mobility on the TCPv6 and UDPv6 services. In this work, comparative measurements have been carried out in an outdoor WiFi test system, containing IEEE 802.11b access points and mobile clients, to understand the effects of the processes occurred during the roaming phase of the WiFi system on the IPv4 and IPv6 connections. One of the conclusions is that the TCP connections are significantly affected by the interaction between the relative speed of mobile clients to the APs and the execution of roaming, whilst it has minor effect to the UDP transfer. Furthermore, we demonstrated that the IPv6 protocol provides a higher quality in a mobile environment than its predecessor, the IPv4.*

## 1. Introduction

The appearance and spreading of the sophisticated mobile services over IPv6 is the most significant advantage of the Internet2. The effect of mobility on the TCPv6 and UDPv6 protocols reveals an exciting user and professional issue. In order to picture the qualitative answer with quantitative attributes, comparative measurements are needed.

The operability, availability of services during the physical movement of the node is a requirement that appears as one of the important demands from an advanced network with good reason. The development of wireless IP telephony, wireless laptops and PDAs show the way to this direction. The mobility feature of wireless LANs drives to a new result:

– *Innovative application development:*
  alarm and message sending. Appearance of the constantly online workflow systems.
– *Increase of efficiency and productivity:*
  the permanent network connectivity makes it possible to perform tasks from any place without delay.
– *Increase the authenticity of data:*
  data is available any time, any place.
– *Availability:*
  user can be virtually online at his home, on the street and in the workplace as well.

In this paper we intend to show a quantitative comparison of different applications of the IP terminals communicating over a mobile WiFi network. Furthermore we provide the explanation of the experienced phenomenon, and draw the conclusions about the expected directives of developments.

## 2. Mobile data transmission

As we already know, the IP versions 4 and 6 are able to provide mobile functions beyond the conventional fixed, wired network communication. The wireless data-link layer connections can be also capable of forwarding frames for the network layer. Thus for transport layer protocols the behavior of the lower layers will be considered more or less depending on the IP version.

The mobile function of IP means that the terminal moves physically from his place during the communica-

*Table 1.*

| Distinctive features | Mobile IPv4 | Mobile IPv6 |
|---|---|---|
| Special routing function (foreign agent) | Yes | No |
| Route optimization ability | Part of the protocol | Extension |
| Symmetrical connection between the MT and the router in the current location | No | Yes |
| Routing overhead bandwidth requirement | High | Low |
| Ability to disconnect from Layer 2 | No | Yes |
| "Tunnel soft" state handling required | Yes | No |
| "Dynamic home agent" address discovery | No | Yes |

*Table 2.*

| Network | Protocol | |
|---|---|---|
| | IPv4 / IPv6 | Mobile IPv4 / Mobile IPv6 |
| Wired | √ | √ |
| Wireless | √ | √ |

tion, therefore his network layer environment changes. At the new location an IP router with foreign agent function keeps the connection henceforward with the original home agent router by means of an IP tunnel [1]. Thus the IP terminal is able to communicate at the new place. The speed of interaction between the agent processes and the additional network load are important issues. The IP version 4 and 6 show different behaviors from this aspect as well. These features are listed in *Table 1*.

In the case of wireless data links there is a possibility for the mobile terminal to remain in the same broadcast domain, therefore the route of the forwarded frames changes, but the IP address of the terminal doesn't. This is the typical event of L2 roaming, when the terminal switches to another access point and only the data-link layer devices changes theirs CAM tables. This paper studies the effect of roaming occurred during cell changes generated by a terminal placed in a moving vehicle in a real outdoor WiFi environment for IP v4 and v6 connections *(Table 2)*.

## 3. Roaming mechanisms

In wireless LANs, nodes are able to virtually connect to the corporate network. The cell change (roaming) is a time consuming process during that the terminal re-associates with a new AP from the current AP. We are talking about data-link (L2) roaming when this process occurs between APs belonging to the same subnet *(Figure 1)*.

If the terminal connects to an AP in a different subnet, then network (L3) roaming will happen. Network roaming can occur only after a successful L2 roaming [2].

Changing the cell is based on the decision of the client whose task is to discover the possible APs, to evaluate their parameters, then to choose between the selected APs. The data-link cell change involves the followings:

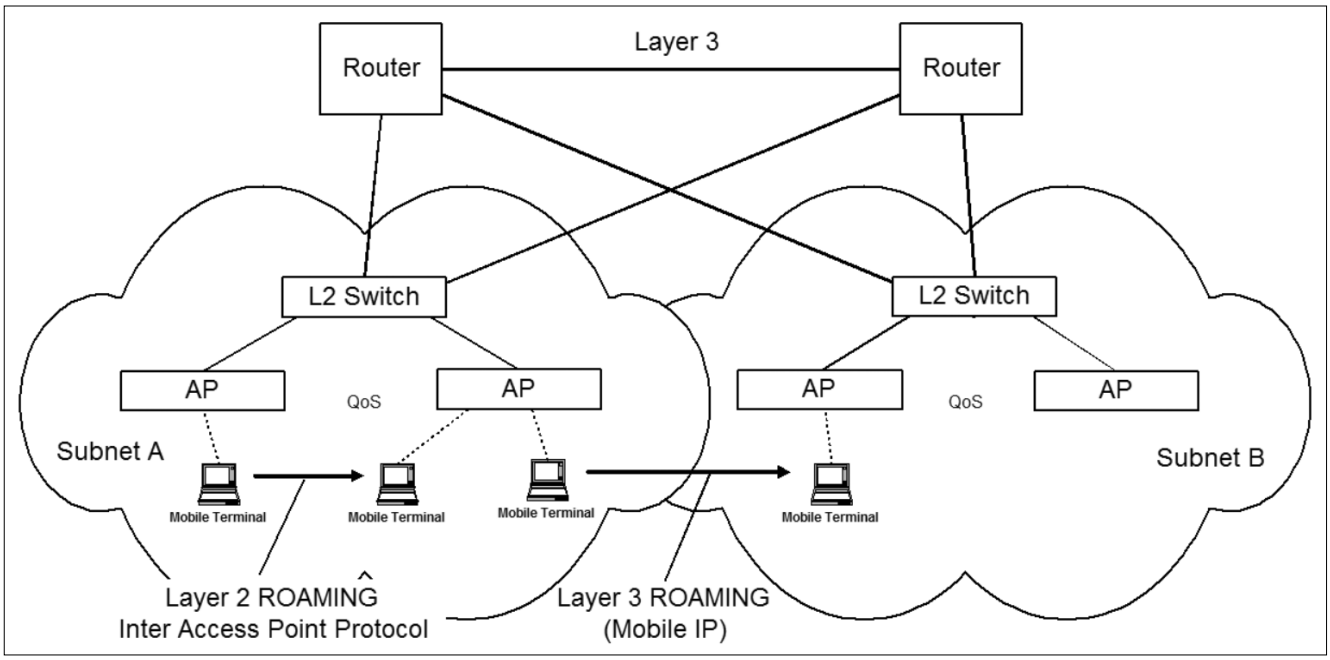
1) The terminal moves from cell "A" to cell "B". The access points belong to the same subnet, so it's an L2 roaming. As soon as the client gets out of cell "A", one of the parameters of the connection with $AP_A$ reaches the given threshold and generates the roaming process.

2) The client scans all of the IEEE 802.11 radio channels, and looks for an available AP. After finding the $AP_B$, authentication and association phases occur on the physical radio channel.

3) The $AP_B$ sends a zero content multicast message with a source MAC address identical to the address of the mobile terminal. Switches in the wired LAN update their CAM tables on the basis of this message. Therefore Ethernet frames addressed to the terminal reach $AP_B$ instead of $AP_A$.

4) $AP_B$ sends a multicast message with its own MAC address notifying all APs on the subnet that the terminal with the given physical address is associated to it. As the $AP_A$ gets this message, deletes the MAC address of the mobile terminal from its association table.

*Figure 1. The L2 and L3 roaming*



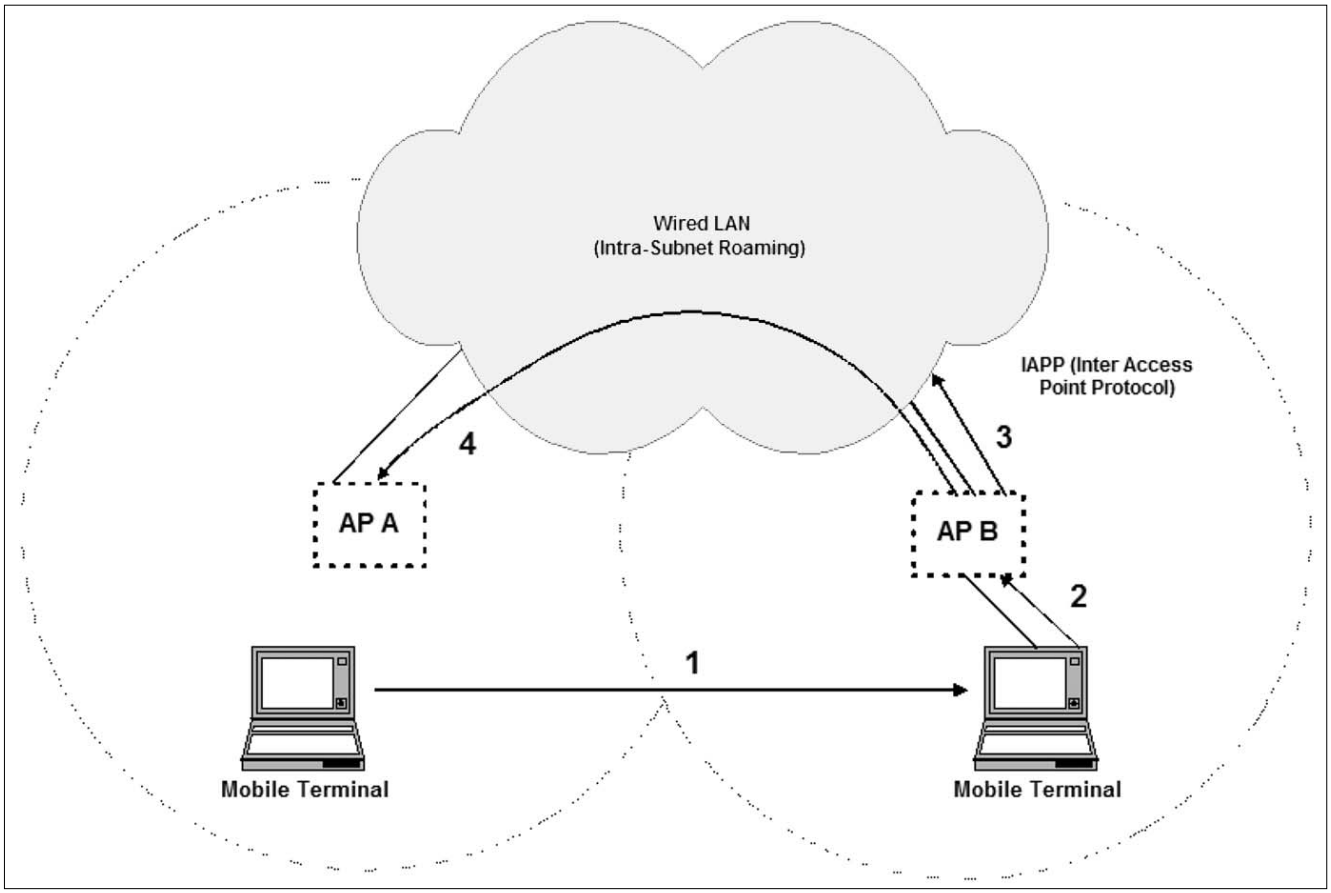



*Figure 2. Steps of the L2 roaming*

The roaming process is always initiated by the client, but no IEEE standard exists currently about this process [7]. Generally the roaming is triggered by the following events:

a) *Exceeding the Maximum Data Retry Count threshold:* when the client is uable to transmit the data after its preconfigured retry count, the station initiates a roaming process. This counter is set to 16 by default and is configured in the wireless adapter management software (e.g. Aironet Client Utility).

b) *Missing too many beacons:* all clients associated to an AP receive beacons periodically. APs send a beacon every 100 ms by default according to the beacon period configuration setting. A client learns about the AP's beacon interval from a beacon component. If the station doesn't receive eight consecutive beacons, the roaming event occurs and the roaming process is initiated. Even an idle client is able to detect the loss of radio link quality by monitoring the incoming beacons.

c) *Change of data rate:* Under normal conditions packets are transmitted at the AP's default rate. This rate is the highest that can be set as "enable" or "required" parameter on the AP. Every time a packet has to be retransmitted at a lower rate, the retransmit counter is increased by three. For packets transmitted successfully at the default rate, the retransmit counter is decreased by one, until it reaches zero. When this counter reaches 12, one of the following scenarios occurs: 1. If the client has not attempted roaming in the last 30 seconds then the roaming process occurs. 2. If the client has already attempted to roam in the last 30 seconds, the data rate for that client is set to the next lower rate. A client transmitting at a lower rate than the default one will increase the date rate to the next higher rate after a short time interval if transmissions are successful.

d) *Periodic Client Interval:* For Cisco Aironet v6.1, the rate and the signal intensity for the mobile terminal's search for a base station with better reception can be configured. With these settings, the client will look for a better base station when both of the following conditions have been met:

– The client has been associated to its current access point for at least 20 seconds. This restriction is needed to prevent a client switching between access points too rapidly. Valid values are from 5 to 255 seconds.
– The signal intensity is less than 50%. Valid values range from 0 to 75%. The periodic scan is a roaming event that causes the occurrence of the roam process.

e) *Initial Client Start-up:* When a client starts up it goes through the roaming process, to scan for and associate with the most appropriate access point.

Searching for a new AP is needed for the roaming process [3]. When a roaming event occurs the client station scans each 802.11 channel in order to determine the list of available APs and to select the best

one. On each channel the client station sends a probe, and waits for probe responses or beacons from access points on that channel. The probe responses and beacons received from access points are discarded unless they have matching Service Set Identifier (SSID) and encryption settings.

When the scan process is completed and the mobile client has a list of responding access points, it selects the access point to compare with the others. If the terminal is in its initial start up phase, then the new AP will be the first element of the list, when the terminal is in roaming phase then the new AP will remain the same if it answered to the test probe frames, in case of no answer the first element of the list will be the new AP.

The current access point is compared to each of the access points in the list. In order to be considered as a new current access point, each access point must meet all of the following conditions:

1) Signal intensity of the potential new AP is at least 20%. If signal is more than 20% weaker than the current one, signal intensity must be 50% or more.
2) If the potential new AP is in repeater mode, and it's distance from the backbone, in radio hops, is larger than the current AP, it must have 20% more signal intensity than the current base station.
3) The transmitter load of the roaming target AP can be at most 10 % higher than the current AP's load. The mobile client compares the access points that meet the base conditions with the current access point. If an accepted AP fulfills any of the following conditions, the terminal selects it as the new AP, henceforward the next AP of the list is compared to this new one: signal intensity is 20% higher than current, smaller hop distance to the backbone, at least four fewer clients associated to it than current access point, transmitter load is at least 20% less.

From 12.2.(11)JA IOS version, Cisco's "fast secure roaming" implementation has two additional features: more efficient 802.11 channel scanning during physical roaming, and an effective re-association mechanism applying more advanced encryption key management. The improved channel scanning results in faster L2 roam irrespective of the authentication method.

The key management speeds up the Cisco LEAP authentication process, therefore roam will be faster and safer. Both on Cisco terminals and base stations the channel scanning is enabled by default. Fast secure roaming is preceded by a channel scanning.

Before 12.2.(11)JA IOS version for a client station it took 37 ms to check a channel that took 431 ms total for 13 channels in case of European standards. For each channel the mobile client executes the following steps: when the radio hardware of the client tuned to the given WLAN channel, listens to avoid collision, then sends "probe" frames and waits for "probe response" or "beacon".

The channel scanning of fast secure roaming is more efficient: Re-associating clients now communicate information to the new access point such as the length of time since they lost association with the previous access point, channel number, and SSID. Using the information from client associations, an access point builds a list of neighbour access points and channels these access points were using. If the client reporting an neighbour access point was disassociated from its previous access point for more than 10 seconds its information is not added to the new access points list. Access points store a maximum list of 30 neighbour access points. This list expires over a one-day period. When a client associates to an access point, the associated access point sends the adjacent access point list to the client as a unicast packet.

When a client station needs to roam, it uses the list of neighbour APs it received from its current AP to reduce the number of channels it needs to scan. There are three roam types. Client uses them according to its activity.

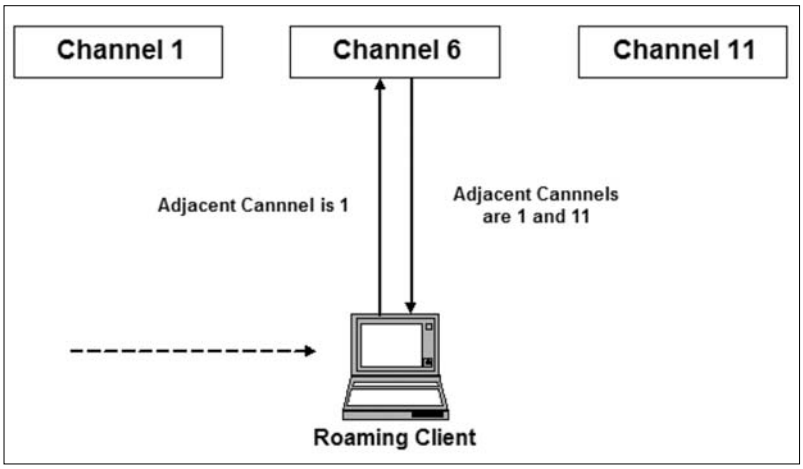
– *Normal Roam:* The client has not sent or received a unicast packet in the last 500 ms. The client does not use the neighbour access point list obtained from the previous access point. Instead it scans all channels valid for the operating regulatory domain.

– *Fast Roam:* The client has sent or received a unicast packet in the last 500 ms. The client scans the channels on which it has been told there is an adjacent access point. If no new access points are found after scanning the adjacent access point list, the client reverts to scanning all channels. The client limits its scan time to 75 ms if it is able to find at least one better AP.

– *Very Fast Roam:* the client has sent or received a unicast packet in the last 500 ms, and the client increases the load of the cell with a non-zero percentage. Same as Fast Roam except the scan is terminated as soon as a better AP is found.

Devices used in our test environment can operate in all of these three roaming modes.

*Figure 3. Channel management of the fast roaming*



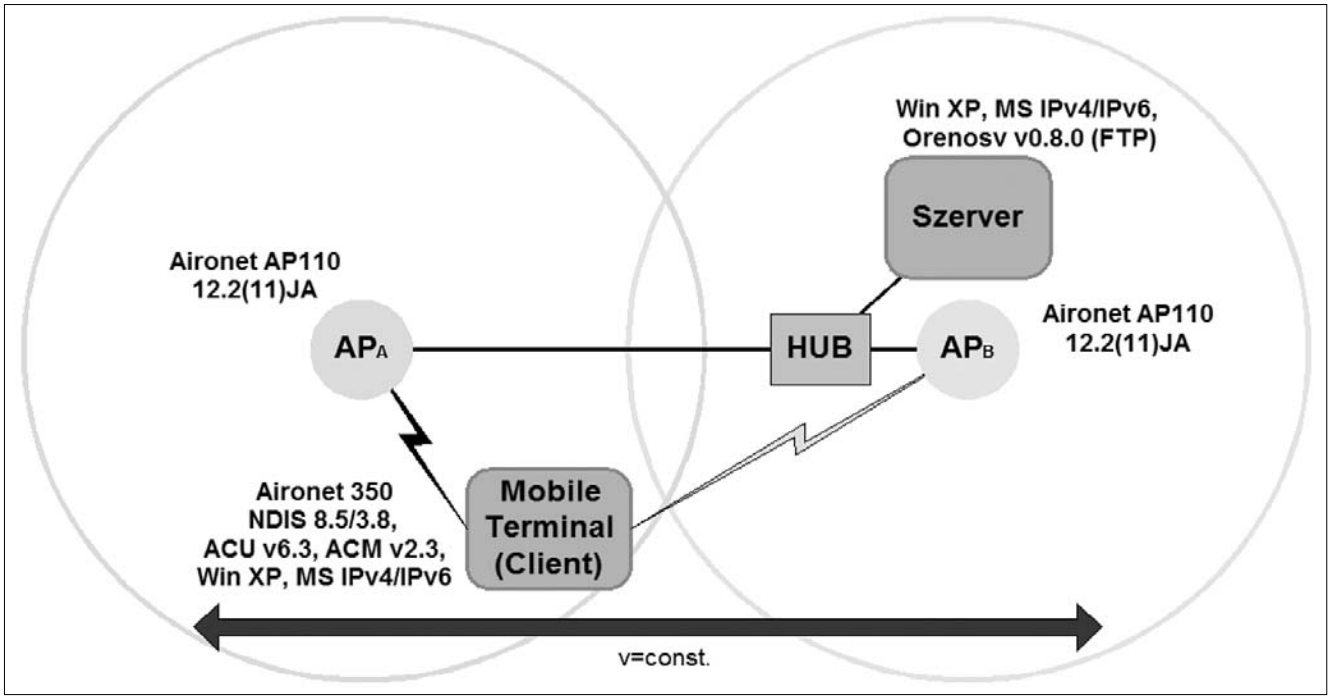


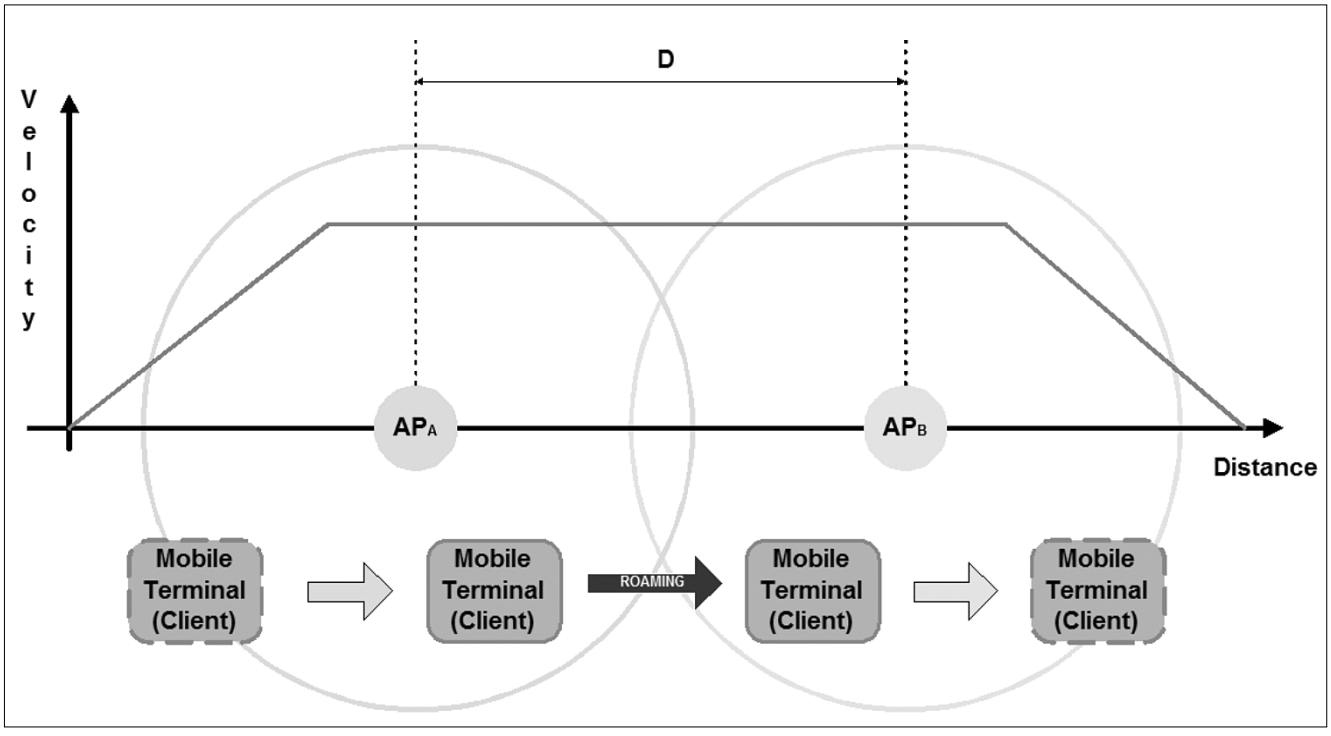

*Figure 4. The measurement environment*

## 4. Test environment

We set up an outdoor test WiFi system in order to study the behavior of the IPv4 and IPv6 protocols in mobile environment. The test is based on two APs operating according to the IEEE 802.11b standard, placed in 100 m distance from each other, linked with wired Ethernet connection, and a mobile terminal. As we know roaming is supported in the 11 Mbps WiFi standard. Whereas the speed of data transmission is strongly depend on the distance between the AP and the client station. During movement the client approaches to and diverges from the access point that causes the alteration of transmission speed in the data link layer.

In our measurements we fixed the transmission rate to 11 Mbps, this way the alteration of the radio signal strength forces the client station to initiate the roam process. We studied the behavior of the different transport layer protocols while the client station was moving on a vehicle (with constant speed during the roaming) paral-

*Figure 5. The roaming process*



| Parameters | | Independent measurements | |
|---|---|---|---|
| Access Point | Cisco Aironet AP1120 | L4 protocol | TCP (FTP), UDP (Spray) |
| Mobile Terminal | Cisco Aironet 350 series | L3 protocol | IPv4, IPv6 |
| AP IOS (1 mW) | 12.2(11)JA | TCP traffic | MT > Server (Up)<br>Server > MT (Down) |
| MT and server OS | Windows XP | UDP message [B] | 18, 1472, 31970 |
| MT Radio Firmware | Win/NDIS Driver 8.5/3.8, ACU v6.3, ACM v2.3 | Speed [Km/h] | 10, 30, 50 |
| FTP server (IPv4/IPv6) | Orenosv v0.8.0 | D(APA,APB) | 100 m |

*Table 3.*

lel to the line between the two access points. On the server side, the Ethereal snoop program was run, storing all of the L2's frames with unique timestamps for further analysis. During roaming the direction of data stream is important, because due to roaming, buffering is needed on the wired side (down) or on the wireless side (up) which significantly affects the operation of TCP connections.

We've selected three different sizes of ICMP message according to 64 bytes, 1500 bytes data link frame and 32 Kbytes in multiple frames. It's important for the minimal and maximal size of frame (MTU), and for the segmentation of IP packet. The speed of the vehicle was constant between the two APs, we were driving according the normal traffic rules applied for built-in areas (10 Km/h, 30 Km/h, 50 Km/h).

## 5. Measured parameters and evaluation

Using the frame sequence captured by Ethereal, we can interpret the signals of the roaming process, furthermore the traffic of the transport layer. So thus the roaming R[ms] and the traffic's drop out T[ms] became measurable.

FTP transmission of large files is realized based on the TCP's "Slow Start" and "Windowing" algorithms. The regulation of the window size is made necessary by the data rate alteration of the data link layer. The duration of the WiFi transmission's roaming phase strongly influences the efficiency of TCP.

UDP transport is much more adaptive by nature. We were sending 100 packets with spray ping which loaded the radio channel by 0.93, 21.82 and 100.0 percent depending on the packet size. Based on the measured parameters we can make the following observations *(see Figures 7-14)*:

• In case of ICMPv4 the time required for roaming is decreasing with the speed when frame size is under the default Ethernet MTU (1500 bytes), while with frames above the MTU it shows an increasing tendency. In case of segmentation it takes more time to reorder the packets. The ICMPv6 roaming time is decreasing with the speed at all frame size due to the medium sensing feature of the IPv6.

• Under the default Ethernet MTU frame size the loss of ICMPv4 traffic doesn't depend on the speed, whereas the delay is reduced by the segmentation. This virtual inconsistency is due to the persistent load of the radio channel. Strong fluctuation can be experienced with ICMP v6 in function of the speed, because IPv6 disconnects from the data link layer. The IPv4 doesn't perform this disconnection, so the ICMPv4 is less sensitive.

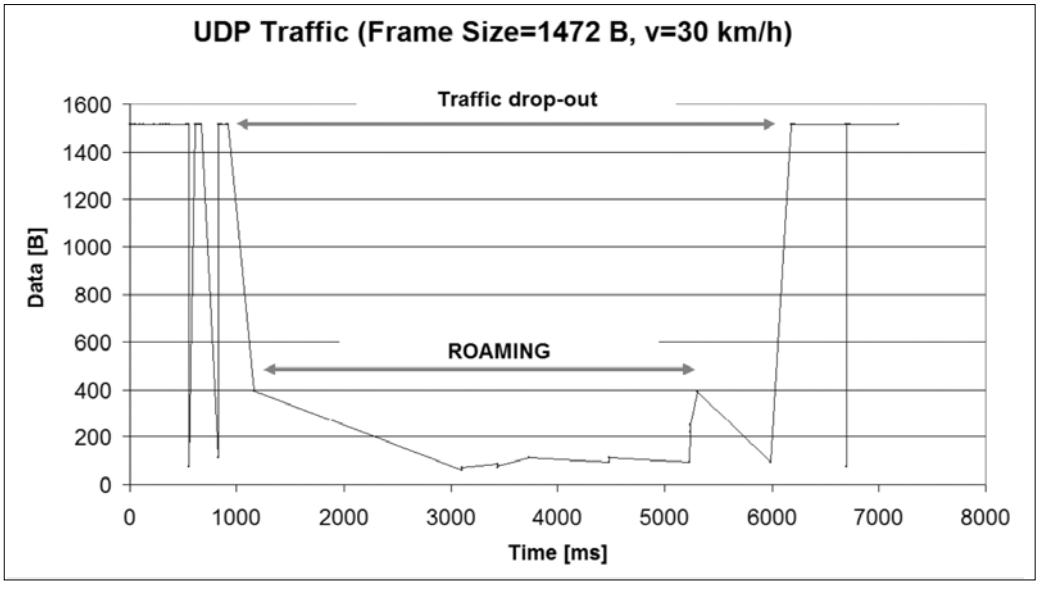



*Table 6.*
*Measured parameters (R[ms], T[ms])*

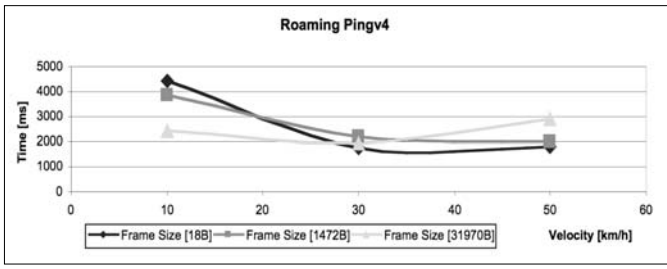



*Figure 7.*
*Effect of the frame size to the roaming (ICMPv4)*

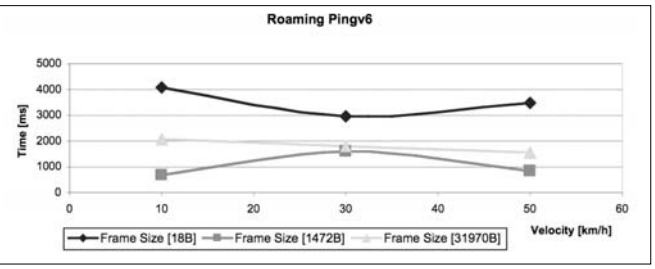



*Figure 8.*
*Effect of the frame size to the roaming (ICMPv6)*

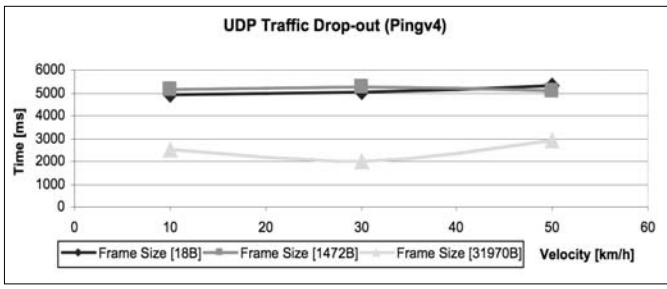



*Figure 9.*
*Effect of the roaming to the UPD v4*

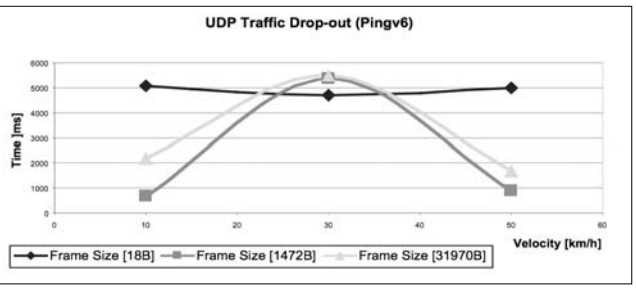



*Figure 10.*
*Effect of the roaming to the UPD v6*

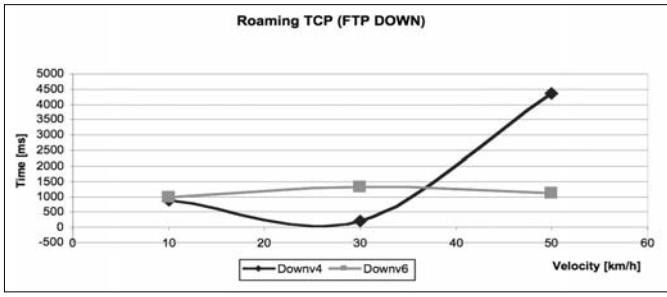



*Figure 11.*
*Effect of the roaming (TCP download traffic)*

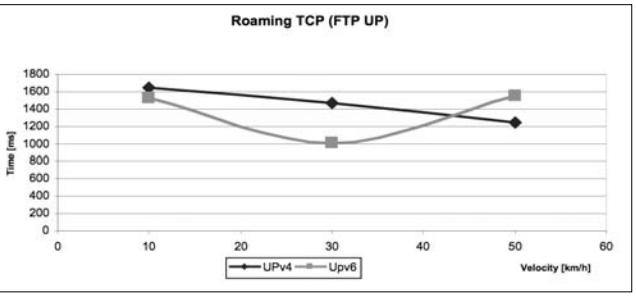



*Figure 12.*
*Effect of the roaming (TCP upload traffic)*

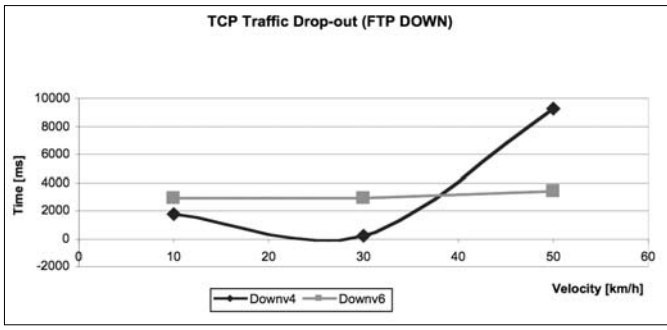



*Figure 13.*
*Effect of the roaming (TCP download traffic)*

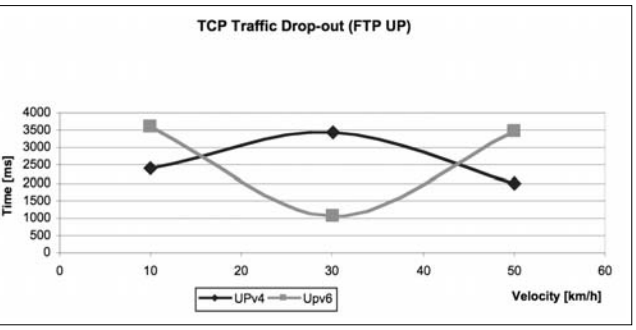



*Figure 14.*
*Effect of the roaming (TCP upload traffic)*

• The traffic dropout of ICMPv4 is 1.5 second longer, while the ICMPv6 dropout is 1 sec. longer than the roaming time. With smaller frame size the dropout depends less on speed, whereas at larger frames only IPv6 is sensitive to the speed.

• The dropout of TCPv6 is independent of the direction of data stream. In case of TCPv4 the download bears significantly more data loss than the upload. This is due to the quick window size alteration of the IPv4, therefore at download there's a high data loss because a great number of frames are sent to the previous AP.

• TCPv4 uses larger window size and slow dynamics, while TCPv6 applies smaller windows that are controlled quickly. Thus the IPv6 tolerates the roaming events of WiFi environment, resulting in the decrease of traffic loss.

• In case of download the loss of TCPv4 traffic strongly depends on the speed of the moving terminal station. It can produce a 9.2 seconds of dropout at speed of 50 Km/h that makes impossible to communicate from fast moving vehicle. In case of TCPv6 download this value doesn't depend on speed and can be

kept under 3.6 seconds. The dropout of TCPv4 upload traffic changes little with speed, while TCPv6 produces significant alterations.

## 6. Summary

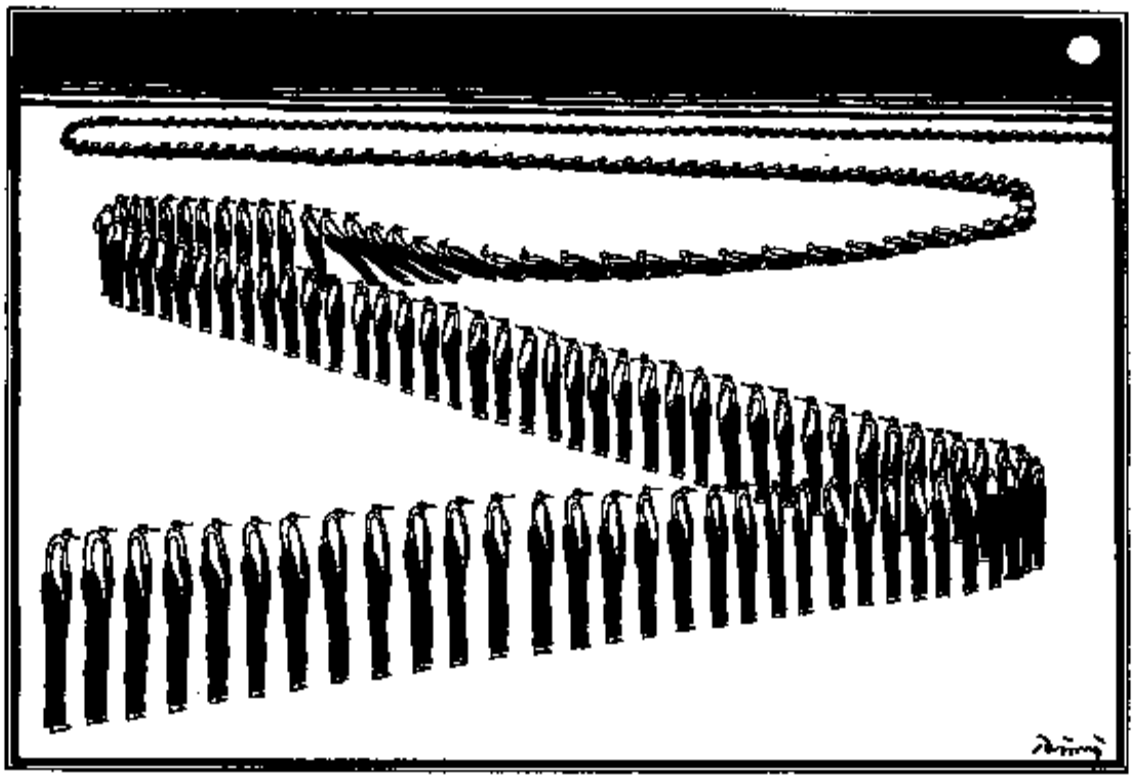
TCP connections are affected significantly, while UDP connections are affected less significantly by the interaction between the mobile station's relative speed and the roaming execution [5],[6]. The results gained statistically from the comparative measurements provide us with a practical view on the behavior of IPv4 and IPv6 protocols in mobile environment, furthermore we can answer the question if the performance of IPv6 over wireless data link layer is really higher compared its predecessor IPv4.

The conventional applications provide slower transmission over mobile links due to the best effort nature of IPv4, while IPv6 assures effective data transfer because of its quick adaptation to lower layers. Time sensitive applications (IP phone, video conference) suffer significant dropouts due to the limitations of IPv4's QoS in mobile WiFi environment therefore the provided quality is unacceptable. The quick adaptation of IPv6 decreases the interval of dropouts [4].

A possibility for further investigation may be the analysis of quality influencing factors for communication services of mobile stations that move outside of built-up areas, at higher speed (on highways, train, etc.). Now we can also clearly see the necessity of future development: speeding up the roaming phase and development of IPv6 specific applications.

**References**

[1] Microsoft TechNet, The Cable Guy – Sept. 2004, "Introduction to Mobile IPv6": http://www.microsoft.com/technet/community/columns/cableguy/cg0904.mspx
[2] Charles E. Perkins (Sun Microsystems), "Nomadicity: How mobility will affect the protocol stack", http://www.computer.org/internet/v2n1/nomad.htm
[3] Microsoft Corporation, "Understanding Mobile IPv6", http://www.microsoft.com/downloads/details.aspx?FamilyID=f85dd3f2-802b-4ea3-8148-6cde835c8921&displaylang=en
[4] Ye Tian, Kai Xu, Nirwan Ansari, "TCP in Wireless Environments: Problems and solutions" IEEE Radio Communications, March 2005.
[5] Zoltán Gál, Andrea Karsai, "Videokonferencia rendszerek minőségi garancia jellemzőinek elemzése", NetworkShop 2004 – konferenciakiadvány, Széchenyi István Egyetem, Győr, 2004. április 5-7.
[6] Zoltán Gál, György Terdik, "Multifractal Study of Wireless and Wireline Datanetworks", 8th International Conference on Advances in Communications and Control, Telecommunications/Signal Processing – Proceedings, Crete, Greece, 25-29. June 2001.
[7] Cisco Systems, Inc., "Cisco Fast Secure Roaming"



# World Telecommunications Congress 2006

***embracing for the first time WTC/ISS and ISSLS***
***"Emerging Telecom Opportunities"***

**Budapest Congress & World Trade Center, Hungary**
**30 April - 3 May 2006**

**The telecommunications network has undergone dramatic changes from the days of voice switches and copper loops. WTC2006 - World Telecommunications Congress brings together and preserves the traditions of two series of conferences: ISS (International Switching Symposium) and ISSLS (International Symposium on Services and Local AccesS) that for over three decades have tracked this evolution.**

**A significant number of outstanding representatives of the communications industry from different corners of the world will share opinions in plenary sessions about key issues of the day.**

**Over 90 papers thoroughly reviewed and selected by the International Technical Committee in 19 thematic and two poster sessions will focus and address the needs and opportunities of the 21st Century telecommunications industry. In four tracks the conference takes the chance to catch emerging telecom opportunities in the processes of transforming network technologies, services, business approaches and customers' experiences.**

**The Scientific Association for Infocommunications, Hungary (HTE) and the Association for Electrical, Electronic & Information Technologies, Germany (VDE/ITG), organizers of WTC2006 cordially invites you to participate and visit historic and beautiful Budapest, birthplace of Tivadar Puskás (1844-1893), intellectual father of the telephone switching office, business partner of Edison and John von Neumann, pioneer of computer technology.**

### Conference Tracks

• *Transforming the business of telecommunications:* about changes in the business environment affecting telecom companies, including regulation, new business models and competition, network and service convergence, and the management of networks and services.

• *Transforming the customer's experience:* about new developments in the Next Generation Networks and the Services that will emerge to radically change the way customers experience voice, video, data and mobility.

• *Transforming the network's technology:* about the latest developments in the underlying technology enabling higher broadband speeds, more mobility and portability options, and more efficient IP and optical networking.

• *Transforming the quality of service:* about the latest ideas on how to improve the overall quality of service, including how to engineer networks to ensure quality and how to measure the quality of the delivered services.

### Keynote Speakers

- Miklós Boda, *Secretary of State, National Office for Research and Technology, Hungary*
- Tadanobu Okada, *Associate Senior Vice President, NTT*
- Anton Schaaf, *Chief Technology Officer, Deutsche Telekom AG*
- Pradeep Sindhu, *Chief Technology Officer, Juniper*
- Vicente San Miguel, *Deputy Director General, Telefónica de España*
- Prof. Lajos Hanzo, *Chair, University of Southampton*
- Robert Cowie, *Chief Engineer, Openreach, BT*
- Peter Janeck, *Chief Technical Officer, T-Com Hungary*
- Tim Stone, *Senior Marketing Manager, Cisco Systems*
- John Cioffi, *Professor, Stanford University*
- Michael Chamberlain, *Director of Solutions, Microsoft*
- Rick Missault, *Vice President Marketing & Communications, Alcatel*
- Guido Roda, *Director of Network Service Engineering, FastWeb*
- Reza Jafari, *Member of the Board, ITU TELECOM*
- Herbert Mueller, *Chief Operation Officier, Slovak Telekom*
- Wolfgang Schmitz, *Senior Executive Vice President, Deutsche Telekom AG*
- Camille Mendler, *Director, Yankee Group*
- Oscar Gestblom, *Marketing Manager, Business Unit Systems, Ericsson AB*
- Dave B. Payne, *Manager, Broadband Architecture and Optical Networks, BT One IT*

**Conference Secretariat**

**HTE - Scientific Association for Infocommunications**
**Budapest, Kossuth Lajos tér 6-8, Hungary - 1055**
**Phone: +36 1 353 1027, Fax: +36 1 353 0451**
**e-mail: info@hte.hu, web: http://www.wtc2006.com**

# World Telecommunications Congress 2006

Budapest Congress & World Trade Center, Hungary • 30 April - 3 May 2006

| | Sunday, 30 April | Monday, 1 May | | | | Tuesday, 2 May | | | | Wednesday, 3 May | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 08:00 | | Registration | | | | Registration | | | | Registration | | |
| 08:30 | | Opening Addresses – to be confirmed<br>Miklós Boda, Secretary of State, Hungary<br>Tadanobu Okada (NTT)<br>Anton Schaaf (Deutsche Telekom AG)<br>Chair: Prof. Géza Gordos (BME) | | | | Plenary Session<br>Robert Cowie (Openreach) to be confirmed<br>Peter Janeck (T-Com Hungary)<br>Tim Stone (Cisco Systems)<br>John Cioffi (Stanford University)<br>Chair: Prof. Gyula Sallai (BME) | | | | Plenary Session<br>Reza Jafari (ITU TELECOM)<br>Herbert Mueller (Slovak Telecom)<br>Wolfgang Schmitz (T-Com Germany)<br>Chair: to be confirmed | | |
| | | | | | | | | | | Break | | |
| 10:30 | | Break | | | | Break | | | | | | |
| 10:50 | | Regulatory & Policy Issues (4 papers)<br>Carlo Mazzetti | QoS Engineering(4)<br>Prof. György Lajtha | IP Access & Networking (4)<br>Gábor Kolláth | | IP Infrastructure Developments (4 papers)<br>Frank Mellor | Dynamic Wireless Networks (4)<br>Chair to be confirmed | Quality of NGN Services (4)<br>Guido H. Petit | Poster Session 1<br>Volker Schanz | Convergent Networks & Services (5 papers)<br>Eugenio Guarene | Optical Core Networks (5)<br>Chair to be confirmed | NGN Infrastructure & Services (5)<br>Attila Koós |
| 12:20 | | Lunch | | | | Lunch | | | | Lunch | | |
| 13:45 | | Plenary Session<br>Sindhu (Juniper)<br>Vicente San Miguel (Telefonica)<br>Prof. Lajos Hanzo (University Of Southampton)<br>Chair: Cornelis Hoogendoorn (Siemens) | | | Pradeep | Plenary Session<br>Michael Chamberlain (Microsoft)<br>Rick Missault (Alcatel)<br>Guido Roda (FastWeb)<br>Chair: Michael A. Schwartz (Telcordia) | | | | New Business Models & Business Issues (4)<br>Michael A. Schwartz | Routing (4 papers)<br>Sándor Molnár | QoS Measurement & Monitoring (4)<br>Cornelis Hoogendoorn |
| 15:15 | | Break | | | | Break | | | | Best Paper and Presentation Awards<br>Closing Address: Leith H. Campbell | | |
| 15:35 | | Service Management (5 papers)<br>Chair to be confirmed | QoS Topics (5)<br>Leith H. Campbell | Access Networks: Technology & Modelling (5)<br>Sven Symalla | Multi-Services Networks & Applications (5)<br>Philippe Fouquart | Network Management (5 papers)<br>Tom Borowiecki | Broadband Access (5)<br>Stephen Hornung | Broadband Services (5)<br>David L. Waring | Poster Session 2<br>Peter Nagy | | | |
| 17:30 | Registration, BCC | | | | | | | | | | | |
| 18:00 | Welcome Reception<br>BCC | 19.00 -<br>Optional Social Event | | | | Gala Dinner<br>Patronship: Magyar Telekom | | | | | | |

**Transforming the business of telecommunications**
**Transforming the customer's experience**
**Transforming the networks' technology**
**Transforming the Quality of Services**

**Conference Secretariat**
HTE – Scientific Association for Infocommunications
H-1055 Budapest, Kossuth Lajos tér 6-8.
Phone: +36 1 353 1027 • Fax: +36 1 353 0451
e-mail: info@hte.hu • web: http://www.wtc2006.com