

Performance of the Neural network controlled ATM switch

V.M. Skulic, Z.R. Petrovic

Faculty of Electrical Engineering, University of Belgrade

Bulevar revolucije 73, 11000 Belgrade, Yugoslavia

Phone: +381 11 3370106, fax: +381 11 3248681, e-mail:

skule@etf.bg.ac.yu, epetrozo@etf.bg.ac.yu

Abstract

In this paper a new $N \times N$ space division ATM switch architecture based on banyan and neural network is presented. The switching function is performed by extended banyan network and the neural network controller is introduced to allow the recirculation of the misrouted cells. The paper analyses the switching performances, especially cell loss probability, and shows the advantages of the proposed switch.

Keywords

Banyan, HOL blocking, recirculation, neural network

1 INTRODUCTION

A number of switching architectures has been proposed to implement ATM standard. The solutions can be generally placed in three categories: shared-memory, shared-medium (bus) and space-division type. In this paper a new space

division switching architecture based on self-routing banyan network is analysed. The advantage of the banyan network is its capability to decentralise switching control. Hereby, big switches, with huge capacity, can be built. The deficiency of the banyan network is low throughput, due to the HOL blocking (Karol et al, 1987), which further decreases when the size of the switch grows.

Many architectures, such as Tandem banyan (Tobagi et al, 1991) and Rerouting network (Urushidani, 1991), which overrides HOL blocking by offering multiple paths from each input to each output and by replacing input with output buffering, have been proposed. In Section 2.1, we briefly describe the architecture and the performances of the Extended banyan network (Petrovic et al, 1997), which we choose as the core of the switch proposed in this paper.

There are other solutions using the windowing technique to eliminate HOL blocking. In that case, the cells in the input buffers are analysed by the centralised control unit to resolve the conflicts between them. In order to satisfy high speed requirements, the control unit may be realised as a neural network (Brown et al, 1990 and Park et al, 1994). In our solution, recirculation of the misrouted cells is introduced, and the main task of the neural network is to maintain cell sequencing. The architecture of the whole switch is described in Section 2.2. The neural network is the modification of the well known Hopfield network (Hopfield et al, 1985) and is given in Section 2.3.

The mathematical model of the switch is given in Section 3. According to this model, we calculate the throughput and the cell loss probability of the switch in Section 4. These results are compared with the results obtained by the computer simulation and good agreement is achieved.

2 ARCHITECTURE OF THE SWITCH

2.1 The extended banyan network

The switch is based on extended banyan network with m stages, $m > n = \log_2 N$, shown in Figure 1. The main feature of this arrangement is that any n consecutive stages build a banyan network. Therefore, a cell that fails to take its desired route due to cell contention at stage k can restart its routing from stage $k+1$. The cell may start and likewise finish routing at any stage, when it is directed toward output buffer. Therefore, a cell routing tag should have an additional field indicating how many stages the cell should pass from the present stage to its destination called RNS. The routing algorithm in every cascade is as follows: in the case of conflict the cell with lower RNS will be passed to its desired output of the switching element, and its RNS is decremented; the RNS of the cell that is misrouted is reset to n ; there is a logical filter at the output of the switching element directing cells with RNS=0 towards output buffer, and cells with RNS>0 to the next cascade.

An initial value of RNS at the switch input may be lower than n . Namely, if input address $s_{n-1} \dots s_1 s_0$ and desired output address $d_{n-1} \dots d_1 d_0$ are the same in h

consecutive bits $s_{h-1}...s_2s_1$ and $d_{h-1}...d_2d_1$ then RNS is set to $n-h$. The routing at stage k is performed according to the cells RNS's and two bits of the destination addresses, d_0 if $RNS=1$, or $d_{n-1-((k-1) \bmod (n-1))}$ if $RNS \neq 1$. The switching element performing the routing function is simple, as well as the logical filters at the output of the element.

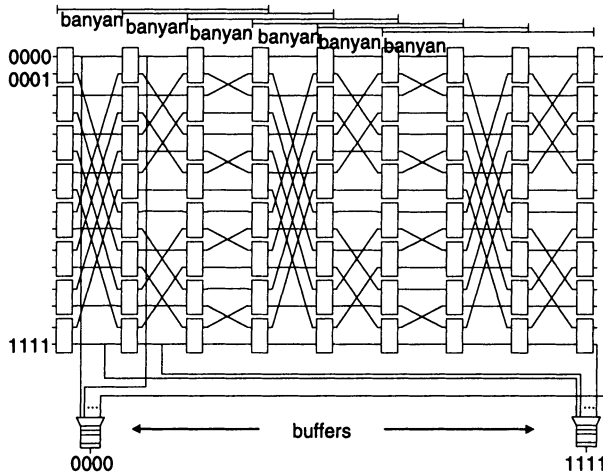


Figure 1 An extended banyan network with $N=16$ inputs and $m=9$ stages.

It is possible to calculate the cell loss probability as the function of the number of cascades in the banyan network (Petrovic et al, 1997), for uniform traffic (incoming traffic is Bernoulli process, independent for each input, with uniform distribution of requested output ports), if we neglect correlation between the traffics at different links of the same cascade.

We introduce ρ_0 as the offered traffic per input line, $\rho_0(r)$ as the offered traffic with $RNS=r$, and $\rho_k(r)$ output traffic intensity at stage k with $RNS=r$. So,

$$\rho_0(0) = 0; \quad \rho_0(1) = \left(\frac{1}{2}\right)^{n-1} \rho_0; \quad \rho_0(2 \leq r \leq n) = \left[\left(\frac{1}{2}\right)^{n-r} - \left(\frac{1}{2}\right)^{n-r+1} \right] \rho_0. \quad (1)$$

The relations between traffic intensities in the consecutive stages are derived in (Petrovic et al, 1997). They are:

$$\rho_k(0) = \rho_{k-1}(1) - \frac{1}{4} \rho_{k-1}^2(1) \quad (2)$$

$$\rho_k(r) = \rho_{k-1}(r+1) \left[1 - \frac{1}{4} \rho_{k-1}(r+1) - \frac{1}{2} \sum_{i=1}^r \rho_{k-1}(i) \right]. \quad (3)$$

$$\rho_k(n) = \sum_{i=1}^n \rho_{k-1}(i) - \sum_{i=0}^{n-1} \rho_k(i) \quad (4)$$

The above equations can be recursively solved to get the probability of cell misrouting, P_{mr} , as the function of the number of cascades m ,

$$P_{mr} = \frac{\sum_{i=1}^n \rho_m(i)}{\rho_0}. \quad (5)$$

The results are given in Figure 2 for different number of inputs and input traffic loads. These results are well proved by the computer simulation.

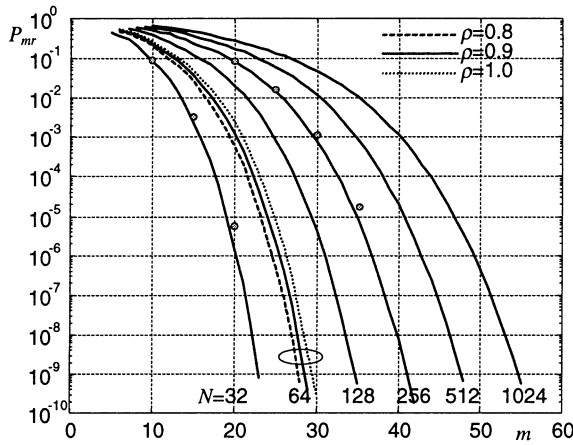


Figure 2 The probability of cell misrouting in the extended banyan network. The points are for computer simulations with $\rho=0.9$, and $N=32, 256$.

2.2 Proposed architecture

It is obvious that traffic intensity decreases from cascade to cascade, since it is proportional to the cell loss probability. We can see that we need 54 cascades for the probability of misrouting 10^{-8} , with input traffic load 0.9 and 1024 inputs. After 27 cascades, cell loss probability is already 0.1, which means that input traffic intensity decreases under 0.09. In the second half of the extended banyan traffic intensity is very low, and these cascades are not utilised well. So, we shorten banyan network and introduce recirculation of the misrouted cells (Figure 3).

The cells that didn't find the wanted output in banyan network, are sent in the recirculation buffers and they are waiting for the next time-slot (interval equal to the cell's duration) to enter banyan and try again. The cells that reached the desired outputs are directed towards concentrators at the inputs of the output buffers. Only m_k cells can pass through the concentrator in one time slot. The output buffer speed has to be m_k times the speed of the input buffer in order to avoid losses.

It is necessary to solve the possible congestion at the input of the banyan network, because it may happen that there are cells in the recirculation buffer and in the

input buffer at the same time. The answer is to give priority to the cell from the recirculation buffer in order to reduce the cell delay. The second problem is that there is possibility to loose cell sequence. The cells that enter on the same input and want to exit on the same output would be from the same connection, and so, they must be delivered in the same sequence they entered the switch. Neural network controller is used in order to accomplish cell sequence, because it is very fast.

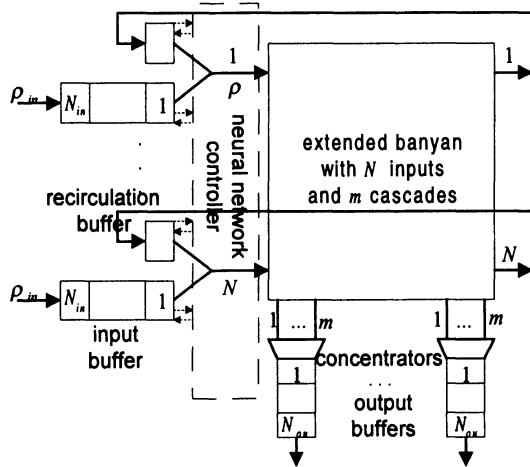


Figure 3 Block diagram of the proposed switch with recirculation of the cells.

2.3 Neural network controller

The task of the neural network controller is to choose as many cells as possible from N recirculation and N input buffers, but less than m_k , satisfying certain constrains:

1. If there is a cell in the recirculation buffer i , it enters the banyan network. If there is a cell in the input buffer i too, it has to wait for the next time slot.
2. If there is a cell in the recirculation buffer i_1 with destination j , and there is a cell on the head position of the input buffer i_2 with the same destination j , then the cell from recirculation buffer enters the switching network and the cell from input buffer waits for the next time slot, because the cell from recirculation buffer might be from the same input originally, and it must reach the output before the cell from the input buffer i_2 . If we pass them both, it is possible that the cell from the input buffer i_2 exits the banyan network before the cell from recirculation buffer i_1 .
3. If there are more than m_k cells, with the same destination j , that may enter the banyan network concerning 1. and 2., then only m_k cells are passed, and others have to wait in order to avoid losses on the concentrators at the input of output buffers. The priority is given to the cells from the buffers that are more occupied at that moment optimising the length of the buffers.

The neural network designed to fulfil these requirements is a modification of the continuous Hopfield neural network. It consists of $2N$ neurones, which are

associated with N input and N recirculation buffers. The block diagram of the neural network is presented in Figure 4, and the scheme of the neurone is given in (Brown et al, 1990).

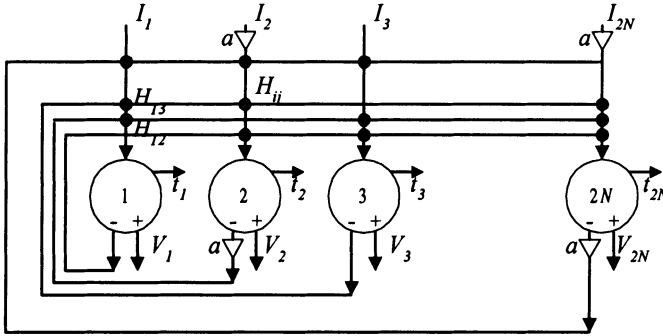


Figure 4 Block diagram of the neural network. Odd neurones are for input buffers, and even neurones are for recirculation buffers.

The output V of the neurone is the function of its input U :

$$V = \frac{1}{1 + \exp(-gU)} \quad (6)$$

where g is the gain of the neurone. The neurone is said to be OFF if $V=0$, and ON if $V=1$. Neurones are amplifiers with positive and negative output. The behaviour of the neurone is described by differential equation

$$\frac{dU_i}{dt} = a_i I_i - t_i - \sum_{j=1, j \neq i}^N a_j H_{i,j} V_j. \quad (7)$$

The external neurone input I_i is determined accordingly to the fact whether there is the cell in the head of the buffer i or not, that is whether its momentary length L_i is greater than 0

$$I_i = \begin{cases} m_k & , L_i > 0 \\ 0 & , L_i = 0. \end{cases} \quad (8)$$

The neurone threshold is set to t_i , which is calculated on the basis of buffer length L_i , in the way priority is given to the cells from more occupied buffers:

$$t_i = 0.75 - 0.5 \frac{L_i}{N_{in}}. \quad (9)$$

The parameter $H_{i,j}$ shows whether the connection between the output of the j th neurone and the input of the i th neurone is set up or not

$$H_{i,j} = \begin{cases} 1 & , w(i) = w(j) \\ 1 & , i = j - 1, \text{ odd } i \\ 0 & , \text{other} \end{cases} \quad (10)$$

where $w(i)$ stays for the wanted output of the cell from buffer i . The gain favouriting cells from the recirculation buffers is denoted by a_j , and it is applied to the neurones in connection with the recirculation buffers, only. It has been found that good results are obtained with $a_j = m_k$ and $g = 2$.

The neural network is very fast, because it works in parallel, and it is possible to reach steady state in the interval of the time slot. The states of the neural network controller neurones determine from which buffers the cells enter the banyan in this time slot; their neurones are ON. The design of the neural controller is verified by the extensive computer simulation.

3 MATHEMATICAL MODEL

In order to analyse performances of the proposed switch architecture, the original mathematical model is established according to the role of the neural network controller described in the previous section. The throughput T as the function of the number of cascades m and the cell loss probability P_{loss} as the function of m and the length of the input buffers N_{in} are found on the basis of this model.

The mathematical model is established for the Bernoulli traffic at the inputs of the switch, with the probability of the cell arrival in the time-slot ρ_{in} . Suppose that the intensity of the traffic at the input of the banyan is ρ_0 . This traffic is some kind of mixture of the traffic from the input and recirculation buffers and it is not of the Bernoulli type. Namely, the destinations are not uniformly distributed between the cells. But, we can see from Figure 2. that the probability of the misrouting isn't much sensitive if we change ρ_0 . It is expected that it's even less sensitive on small changes of the distribution of wanted destinations. In paper (Giacomazzi et al, 1996) it is validated for the similar extended banyan network. Hereby, we use the result (5) from Section 2.1, the probability of the misrouting in the banyan network, as the function of the traffic intensity on its input and the number of the cascades to find the traffic intensity of misrouted cells

$$\mu = \rho_0 \cdot P_{mr}(\rho_0, m) \quad (11)$$

Now, it is possible to find the probability p_{go} that a cell from the input buffer is not blocked due to the reasons given in Section 2.3. It is the probability that there is

no cell in the recirculation buffer at the same input of the banyan, and there is no cell in other $N-1$ recirculation buffers that want the same output. These events are independent because the cells from the input and the recirculation buffers are from different time-slots. Accordingly,

$$p_{go} = (1-\mu) \left(1 - \frac{\mu}{N}\right)^{N-1}. \quad (12)$$

Let q_L be the probability that there is L cells in the buffer at the end of the time-slot. The state of the buffer, the number of the cells in the buffer at the end of the time-slot, is modelled as the discrete birth and death process with the probability of birth ρ_{in} and the probability of death p_{go} . From this model, if $\rho_{in} < 1$, we get the probability that the buffer is empty

$$q_0 = \frac{1}{1-\rho_{in}} \frac{p_{go} - \rho_{in}}{p_{go} - \rho_{in} \left[\frac{(1-p_{go})\rho_{in}}{p_{go}(1-\rho_{in})} \right]^{N_{in}}}. \quad (13)$$

Now, we can find the traffic intensity on the input of the banyan network if we suppose the events that there is a cell in the recirculation buffer and in the same input buffer are mutually independent. It isn't possible that the cells from both buffers enter the banyan in the same slot, so

$$\rho_0 = \mu + [1 - q_0(1 - \rho_{in})]p_{go}. \quad (14)$$

It is possible to solve the equations (11)–(14) iteratively, starting with $\rho_0=1$. Upon the completion of this iterative process, using the birth and death model mentioned above, we can calculate the probability that there is no place in the input buffer for the incoming cell,

$$P_{loss} = q_0 \left[\frac{\rho_{in}(1-p_{go})}{(1-\rho_{in})p_{go}} \right]^{N_{in}} (1-\rho_{in}) \quad (15)$$

The throughput of the switch T is found by using a similar method. Generally, if the input traffic intensity is $\rho_{in}=1$, the input buffer saturates and the probability that it is empty is equal to zero, $q_0=0$, which simplifies (14). Equations (11)–(12) aren't changed. When we iteratively solve this changed set of equations, we gain

$$T = p_{go}'. \quad (16)$$

We assumed that the traffics in the inputs of the banyan are independent and with uniformly distributed destinations, that is not true because of the recirculation. But, if the number of the inputs N is not very small, and the intensity of the recirculated traffic is low, these assumptions are reasonable. The constrain 3, mentioned in Section 2.3 was not considered in the mathematical model. So, the model is valid if the probability of blocking the cell from the input buffer according to 3. is small compered to reasons named in 1. and 2. The fact that the cells from longer buffers have small priority is not examined in this model, too.

4 PERFORMANCES OF THE PROPOSED SOLUTION

Using the mathematical model derived in the previous section, we calculated the throughput of the switches of different sizes as the function of the number of cascades (Figure 5).

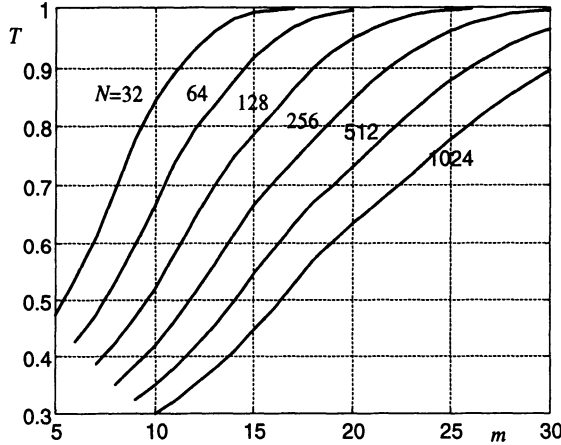


Figure 5 The throughput of the proposed ATM switch as the function of the number of cascades in the extended banyan network.

We can see that the switches have the throughput between 0.8 and 0.9 for m that is the half of the number of cascades needed for the probability of misrouting 10^{-8} in the extended banyan network (given in Figure 2). This is the benefit of the recirculation at the price of the slight throughput falling. But, the utilisation of the link at the output of an ATM multiplexer is hardly over 0.8–0.9 (Schwartz, 1996), and the throughput of 0.9 is sufficient. It is possible to find the cell loss probability using the mathematical model, and some of the results are presented in Figure 6. It could be pointed out that there is some value of m , the savings in the buffer length are not worth of further increasing, for the specified cell loss probability. We also analysed the behaviour of the switch with fixed number of cascades and different traffic intensities. As expected, one can find that the length of the input buffers increases rapidly as we approach the throughput of the switch.

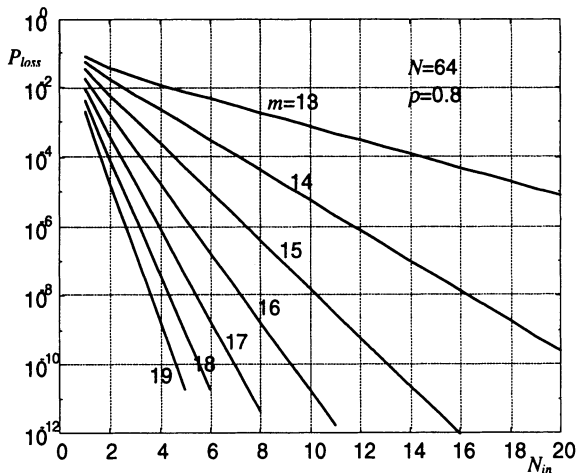


Figure 6 The probability of the cell loss as the function of the size of input buffers, for the switch with $N=64$ and $\rho=0.8$. The curves are for different number of cascades in the banyan network at the core of the switch.

The calculation of the output buffer lengths is given in (Petrovic et al, 1997). As we mentioned earlier, the results obtained using the mathematical model are verified using the computer simulations. The results are compared in Figure 7, and a good agreement is noticeable. The agreement is better if the ratio of the traffic intensity and the throughput is lower, and worse for smaller switch due to correlation between the processes in the links of the extended banyan network.

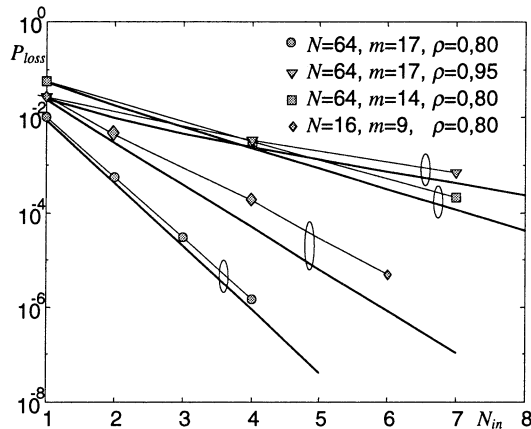


Figure 7 The comparison of the results obtained using the mathematical model (thick lines) and computer simulation (lines with marks) for the switch with 64 and 16 inputs, different numbers of cascades and different traffic intensities.

The extended banyan network stands the unbalancing of the traffic very well (Giacomazzi et al, 1996). We used the computer simulations to show that our

switch with recirculations keeps this property. The simulations were performed for unbalancing on the inputs and outputs. The input unbalancing means that the probability of cell arrival is not the same for all inputs. In the case of the output unbalancing the outputs that cell has to reach are not uniformly distributed. The detailed results of the simulations are in (Skulic et al, 1997) and some of them in Figure 8. The output unbalancing has small influence on the needed input buffer lengths. The effect of the input unbalancing is bigger, but we can control it well, if we put one or two cascades more in the banyan.

It is well known that the number of cells designated to the specific output is almost independent of N , if $N \geq 16$, and is the random variable with the Poisson distribution. The probability that 6 cells are designated to the same output is 3×10^{-4} , if the traffic intensity towards this output is 0.9. This probability is small enough to be neglected in (12). Hereby, we choose $m_k=6$, and this is exactly the required speed up of output buffers.

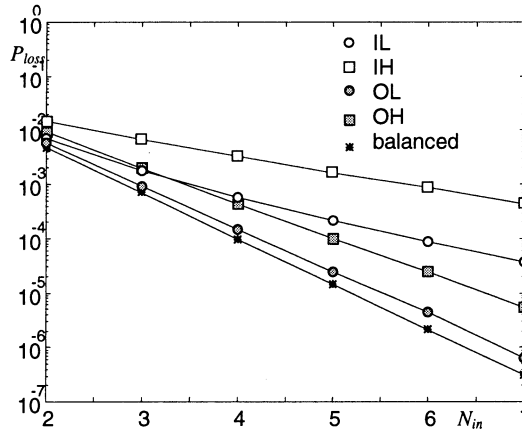


Figure 8 The results of the simulations for unbalanced traffic. The switch is with $N=64$, $m=14$, and the average traffic intensity 0.7. The ‘I’ stands for input and ‘O’ for output unbalancing. The linear decreasing of arrival rate, or the distribution of wanted outputs, over ports from 0.9 to 0.5 is denoted by ‘L’. The ‘H’ stands if some inputs, or outputs, are wanted with probability 0.9 and others are empty.

5 CONCLUSION

In this paper the new ATM switch architecture based on extended banyan network was presented and its performances were analysed. The core of the switch is banyan network with simple switching elements. The number of cascades in the banyan is almost halved by introducing recirculation of misrouted cells and neural network controller to preserve cell sequencing. This controller is a modification of the continuous Hopfield network and is much simpler than neural networks used in (Brown et al, 1990 and Park et al, 1994) to implement windowing technique.

The mathematical model was established to determine the performances of the switch. Computer simulations verified these results. It is shown that the proposed switch architecture with simple elements owns good performances. The small drop of the switch throughput is a disadvantage but it causes only the need for the input buffers that aren't too long. The switch retains good performances in the case of unbalanced traffic, especially output unbalanced.

6 REFERENCES

- Brown T.X. and Liu K.H. (1990) Neural Network Design of a Banyan Network Controller. *IEEE Journal on Selected Areas in Communications*, vol. 8 no. 8, 1428–38.
- Giacomazzi P. and Pattavina A. (1996) Performance Analysis of the ATM Shuffleout Switch Under Arbitrary Nonuniform Traffic Patterns, *IEEE Transactions on Communications*, vol. 44 no. 11.
- Hopfield J.J. and Tank D.W. (1985) Neural computation decisions in optimisation problems. *Biological Cybern.*, vol. 52, 141–52.
- Karol M.J, Hluchuj M.G. and Morgan S.P. (1987) Input vs. output queueing on space-division packet switch. *IEEE Transactions on Communications*, vol. COM-35, 1347–56.
- Tobagi F.A, Kwok T. and Chiussi F.M. (1991) Architecture, Performance, and Implementation of the Tandem Banyan Fast Packet Switch. *IEEE Journal on Selected Areas in Communications*, vol. 9 no.8, 1173–93.
- Urushidani S. (1991) Rerouting Network: A High-Performance Self-Routing Switch for B-ISDN. *IEEE Journal on Selected Areas in Communications*, vol. 9 no. 8, 1194–204.
- Park Y.K, Cherkassky V. and Lee G. (1994) Omega Network-Based ATM Switch with Neural Network-Controlled Bypass Queueing and Multiplexing. *IEEE Journal on Sel. Areas in Communications*, vol. 12 no. 9.
- Petrovic Z.R, Skulic V.M. and Cvetinovic M. (1997) Performance of the new ATM switch for B-ISDN. *J.C.Baltzer Telecommunication Systems*, vol.7, 379–90.
- Schwartz M. (1996) *Broadband Integrated Networks*. Prentice Hall, New Jersey.
- Skulic V.M. and Petrovic Z.R. (1997) Neural network controlled ATM switch, different types of traffic. *Proceedings of Telsiks '97, Nis, Yugoslavia*, 807–10

Skulic M. Vladimir received the B.S.E.E. degree from the Faculty of Electrical Engineering University of Belgrade, Yugoslavia in 1993. From 1994, he is with Department of Telecommunications. His research interests include switching systems and broadband networks.

Petrovic R. Zoran received Ph.D. degree in electrical engineering from the University of Belgrade on 1984. He is with Department of Telecommunications, Faculty of EE Belgrade, from 1973, where he is presently an Associate Professor. His research interest is in B-ISDN.