ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392 – 1215 2011. No. 2(108)
ELEKTRONIKA IR ELEKTROTECHNIKA

TELECOMMUNICATIONS ENGINEERING
T 180
TELEKOMUNIKACIJŲ INŽINERIJA

Adaptive Cross-layer Optimization Based on Markov Decision Process

Z. Velickovic
High Technical School Niš,
Aleksandra Medvedeva 20, 18000 Niš, Serbia, phone: +381 69 468 7155, e-mail: zoran.velickovic@vtsnis.edu.rs
M. Jevtovic
Yugoslav Engineering Academy,
Kneza Miloša 7, 11000 Beograd, Serbia, phone: +381 64 1319139, e-mail: vladmijev@ptt.rs

Introduction

Wireless wideband applications have been extremely fast developed in the last few years. The fundamental reason for this expansion lies in attractive multimedia services they offer to users. Depending on the offered service, the wireless communication systems should satisfy adequate Quality of Services (QoS). ITU-T has standardized in the specification G.1010 the key parameters which affect the quality of the service from the user's perspective [1]. According to G.1010 parameters which affect QoS are: *delay, delay variation* and *information loss.* Satisfying of the set QoS in wireless communication systems is not always an easy task. Difficulties in realization of QoS arise from the variable characteristics of communication media interference as a consequence of multi-user access medium, fading, mobility of users, dynamic change of the network structure and limits of resources for battery supply. Most of the mentioned difficulties have not been taken into consideration on the occasion of designing the stack of the standard communication protocols for wired networks. Hence the stack of communication protocols with excellent performances in wired environment suffers from degradation of performances in wireless environment. According to the concept of standard ISO/OSI stack, network resources are arranged in separate layers and not provided free communication between them. On the other hand, optimization of network resources in wireless environment demand jointly consideration of several network ISO/OSI layers. Jointly optimization of network parameters of several ISO/OSI layers according to [2] is called *cross-layer* (CL) design. Conditions under which optimization should be done relate most frequently to the available transmitting power of network nodes, allowed delay or demanded network throughput. Jointly consideration of PHY and MAC layers in order to satisfy QoS is presented in [3]. As an optimization criterion in wireless communication systems with many users the maximal mean value of MOS estimation of all the users can be set, which is realized through CL algorithms [4]. In [5] the optimization CL algorithm has been discussed about because it performs decomposition of the original optimization problem into several simpler ones by using of convex programming. In this paper, first of all Markov decision process (MDP) has been applied as tools for optimization of network parameters on many ISO/OSI layers [6]. If MDP is used, transition probability matrices depend on the system state and actions which "the software agent" can perform in every discrete time segment in which decisions are being made (decision epoch). The software agent can achieve rewards or costs depending on the performed action and the state of the system at that decision epoch. The aim of MDP is to found optimal policies for defining the series of actions which should be taken in every state of the system, to that utility functions could be maximized.

Further in this paper it has been defined the optimization problem of wireless communication system with one user for a proposed stochastic model of traffic. It has been developed as composite communication system model based on Markov chains of PHY and MAC-LLC layer. The solution of the optimization problem by using MDP is presented in a separate section. The obtained results and the analysis of the optimal policies structure in a simulated environment are presented in the final section.

Communication system model

The structural block diagram of the considered wireless communication system with one user is presented in Fig. 1. This communication system consists of a transmitter, which incorporates a buffer of the limited capacity and encoder-modulator and of a receiver. Loading of the buffer is done with packets from the higher ISO/OSI layers and is modeled by Poisson distribution. Transmission rate out of the buffer is adaptive, and it sustains the realization of optimization CL algorithms. The connection between the transmitter and the receiver is realized in Rayleigh fading channel, while information about the quality of the channel and the buffer state are being interchanged through the control channel as it has been presented in Fig. 1. The analysis of the work of the presented system has been considered through a series of successive time frames whose duration is T_f . Frame *i* comprises the time period of $\left[iT_f,(i+1)T_f\right]$ for which the number of incoming packets into the buffer A_i is specified. Only after the expiration of the time frame *i* to the existing state of the buffer, the number of the arrived packets A_i can be added.

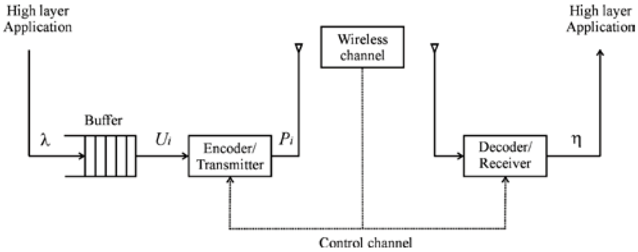

Fig. 1. Wireless communication system with one user and limited buffer capacity

The mean value of the number of packets coming into the buffer while the frame lasts, λ, is calculated as $\lambda = \mathrm{E}\{A_i\}$, where $\mathrm{E}\{\bullet\}$ is the operator of the mathematical expectation. It has been supposed that all the packets are of equal length and that the buffer is of limited capacity and that it can receive only *B* packets. Each the packet that arrive at the moment when the buffer is completely loaded will be rejected and regarded as lost. However, in the case when the buffer is not completely loaded, it will be complemented with the arrived packets until the full capacity has been reached (*B*), and all the rest packets are rejected and considered lost. Another source of the rejected packets in the considered system is the wireless communication channel with fading. Each error created during packets transmission will have as a consequence rejecting of that packet and it will be regarded as lost. The mean value of the number of packets η which will be delivered to the application on a higher layer of the receiver can be determined in the following way

$$\eta = \left(\lambda T_f\right)\cdot\left(1-P_o\right)\cdot\left(1-P_p\right), \tag{1}$$

where P_o is the buffer overflow and P_p is the packet error probability. Optimization problem can be defined as maximization problem in the given bellow way.

Prob_1. Maximize the long-term mean value of the throughput of the wireless communication system η engaging the proposed average transmission power $\overline{P}$ defined as

$$\lim_{T\to\infty}\ \sup\frac{1}{T}\mathrm{E}\left\{\sum_{i=0}^{T-1}P_i\right\}\le\overline{P}\ , \tag{2}$$

where P_i is the engaged transmission power of i-th frame.

Considering the fact that $P_p = f\left(P_b\right)$, for communication protocols of multimedia application which demand constant bit error rate P_b, it will be that the packet error probability P_p is also constant. The result of that is that maximization of the equation (1) can be realized only by minimization of the error probability issued of the buffer overflow P_o. In order to keep the bit error probability P_b constant, depending on code-modulation scheme (rate *u*) and fading level (*g*), it is necessary to engage certain level of transmission power $P = f\left(u,g,\overline{P}_b\right)$. On the base of these facts, maximization problem *Prob_1* can be redefined as minimization problem *Prob_2*:

$$\begin{cases} \underset{U_0,\dots,U_{T-1}}{\arg\min}\ \underset{T\to\infty}{\lim\sup}\frac{1}{T}\mathrm{E}\left\{\sum_{i=0}^{T-1}\left(L_o\left(B_i,U_i\right)\right)\right\}, \\ U_i\in\left\{0,1,2,\dots,B_i\right\},\ \forall i=0,1,\dots,T-1. \end{cases} \tag{3}$$

Parameters B_i and U_i represent the number of packets in the buffer, i.e. the number of packets for which the buffer is being emptied respectively during the frame *i*. Function L_o represents the number of rejected packets in node due to buffer overflow in function of buffer state (*b*) [3]

$$L_o\left(b,u\right)=\left(\lambda T_f\right)\cdot\left(1-\sum_{k=0}^{B-b+u-1}A\left(k,\lambda\right)\right)- \\ -\left(B-b-u\right)\cdot\left(1-\sum_{k=0}^{B-b+u}A\left(k,\lambda\right)\right). \tag{4}$$

Probability density function of arriving *k* packets with the mean value λ into the buffer of the communication node is modeled by Poisson distribution

$$A\left(k,\lambda\right)=\frac{1}{k!}\exp\left(-\lambda T_f\right)\cdot\left(\lambda T_f\right)^k . \tag{5}$$

Markov Decision Process

For the solution of optimization problem *Prob_2* in this paper MDP has been used. In order to describe a communication system by MDP, it is necessary to define five basic elements. These basic MDP elements consists of, for the considered communication system are:

- a set of discrete decision epoch: $t=\left\{0,1,\dots,T-1\right\}$ determined by expiration of the current frame,
- a set of state spaces of the communication system $S=\left\{s_0,s_1,\dots,s_S\right\}$,
- a set of allowed actions for each state of the communication system A_s,
- probability matrices for state transitions of the system P_s,
- a set of immediate costs *C* (or rewards *R*) associated to every state transition of the communication system.

With every state transition a reward is given depending on the actual state of the system, transition matrix and realized action. The aim of MDP is minimization of the mean value of rejected packets (4) or the expected discount cost sum (9) defined by the optimization problem. A set of discrete decision epoch consists of a series of moments determined by expiration of every frame T_f. State space of the system (S) is determined by states of the communication channel and the buffer in the communication node. Therefore, state of the communication system for frame i, s_i, is determined by two components: state of the buffer (b) and state of the communication channel (g). Set of actions (A_s) which can be performed for every state of the system is determined by adaptive possibilities of coding and choice of modulation techniques in the communication system. For the observed communication system the set of actions represents the number of packets through which the buffer in the communication node can be emptied during one frame. By fixing the symbol rate and adaptation of parameters of signal constellation in MQAM system the rate of emptying the communication buffer can be adapted. Transition probabilities of the communication system p_s in realization of the action u can be determined in the following way

$$p_S\left(s \mid s',u\right) = P_r\left\{ s_{i+1} = s' \mid s_i = s, U_i = u \right\} = p_G\left(g,g'\right) \cdot p_B\left(b,b',u\right), \quad (6)$$

where p_G and p_B are transition probabilities of the subsystem which describe the communication channel [7], that is, state of the communication buffer [8], respectively. The fifth element which describes MDP is the function of costs C [6], which are made with every transition of the system for the realized action u [9]

$$C\left(s(b,g),u\right) = P\left(u,g,\overline{P}_b\right) + \beta \cdot L_o(b,u). \quad (7)$$

The first addendum in the weight sum of costs (7) relates to the level of the engaged power (P), while the second addendum relates to the number of rejected packets (L_o). By the parameter β the mutual relation between the addenda is determined and the tradeoff between the set criteria is provided. Stationary optimal policy π^* should provide minimum of the function (8) along with meeting the set criterion (2), (3). Optimization problem *Prob_2* can be transformed into an equivalent discount costs problem *Prob_3* if (8) is used

$$\pi^* = \arg\min_{\pi} \left\{ \limsup_{T\to\infty} \frac{1}{T} \mathrm{E}\left[\sum_{i=0}^{T-1} \alpha^i C\left(S_i,U_i\right) \right] \right\}, \; 0 < \alpha < 1 . \quad (8)$$

Solution of the problem *Prob_3* can be presented by following notation $U_i = \pi^*\left(S_i\right)$, where S_i represents state of the communication system and α^i dicount factor of i-th frame. Therefore, optimal policies relates to every frame i on the base of the state of the communication system S_i and is realized with emptying the transmitting buffer in optimal number of packets U_i. By solving the optimization problem (8) for individual values of parameter β, *pareto-optimal* values L_o and $\overline{P}$ are determined. For comparison of results of the applied policies a *value function* $V_\pi(s)$ is formed and it determines the global revards R (*revards R= -costs C*) when applied policy π for the starting state $s \in S$

$$V_\pi(s) = \mathrm{E}\left\{ \sum_{i=0}^{T-1} \alpha^i \, R(s,u) \mid s_0 = s \right\} . \quad (9)$$

The value function $V_\pi(s)$ makes it possible for the results to the communication system to be compared. So, if it is: $V_\pi(s) \geq V_{\pi'}(s), \; \forall s \in S$, then on the base of the value functions of the applied policies it can be concluded that the greater global revards is realized when policies π is applied in relation to the policies π'. The optimal policy π^* is assigned to the greatest value of the value function $V^* = V_{\pi^*}$. Recursive Bellman's optimality equations allow determination of the optimal policy [6]:

$$V^*(s) = \max_{u\in A_S} \left(R(s,u) + \alpha \sum_{s'\in S} p_S\left(s' \mid s,u\right) \cdot V^*\left(s'\right) \right), \quad (10)$$

$$\pi^*(s) = \arg\max_{u\in A_S} \left(R(s,u) + \alpha \sum_{s'\in S} p_S\left(s' \mid s,u\right) \cdot V^*\left(s'\right) \right), \forall s \in S . \quad (11)$$

Simulation results

This section presents the obtained numerical results when optimization MDP algorithm presented in the previous section is applied. The results are obtained for the following simulation parameters: $\lambda = 1000$ packets/s, $L = 100$ bits in a packet, capacity of the buffer $B = 15$ packets, $P_b = 10^{-6}$, channel bandwith $W = 100$ kHz, noise power density $N_o = 2 \cdot 10^{-5}$ W/Hz, duration of the symbol is fixed to $T_s = 1/W$ and the frame duration is determined with 100 symbols. In MQAM systems this parameter arrangement allows for the buffer to be emptied with u packets/frame with signal constellation $M = 2^u$. The channel with the time correlated Rayleigh fading is modeled with eight states as in [7, Table I]. For different values of the weight parameter β, the solution of the optimization problem is the vector of the rates U_i in which the communication buffer should be emptied for every state of the communication system. On the base of the value of the vector U_i the number of rejected packets L_o is determined and the value of the engaged power $\overline{P}$. By averaging the obtained values for the engaged power and the number of rejected packets for all frames, pareto-optimal values have been obtained. In Fig. 2 the structure of the optimal policies for presupposed communication parameters is presented. The state of the communication system is presented by the quadrate surface in the plane *packets in buffer - states of fading channel* (Fig. 2). Optimal number of packets by which the network node should be emptied is presented as a parallelepiped of certain height - *transmitted packets/frame* whose base is the surface determined by the state of the communication system. It is evidently that the structure of the optimal policies in the fading channel does not correspond to the "*water-filling*" structure. In "water-filling" structure it is necessary to send great number of packets with better quality of the channel. In channels with time correlated fading with the increase of the communication channel quality does not necessarily come to the increase of the rate of sending the packets out of the buffer. The structure of the optimal policies depends on the model of the communication channel, i.e. transition matrix of the communication channel.

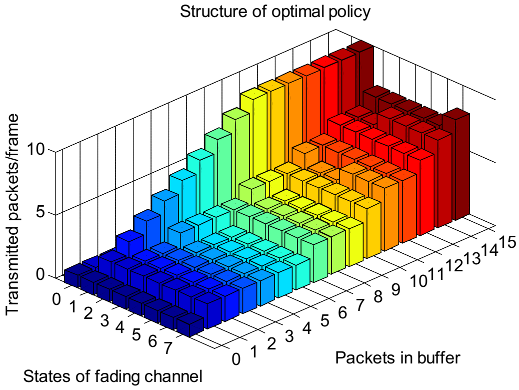

Fig. 2. Structure of the optimal policy obtained by MDP algorithm presented in previous sections for $\beta = 10^7$

Conclusions

Telecommunication systems with adaptive modulation and coding provide satisfying QoS of wireles multimedia applications. The set QoS is reached by designing the optimal protocols adapted to the dynamic changes of wireless channels. Cross-layer design is a new approach in optimization of parameters of communication networks which should provide access to parameters of several OSI layers. As basic tools in optimization procedures in this paper Markov decision process - MDP has been used. By the composite model of the communication system based on Markov chains PHY and MAC-LLC layer, application of MDP has been made possible. In order to realize the optimal policies information from two layers (PHY, MAC-LLC) has been used, while realization was done only by adjusting parameters on one layer (PHY). The obtained results in simulated environment of MATLAB show that the application of cross-layer design in modern wireless multimedia applications is justified.

References

1. **ITU–T** Recommendation G. 1010: End–user multimedia QoS categories – ITU–T, 2001.
2. **Srivastana V., Motani M.** Cross–layer design: a survey and the road ahead // IEEE Communication Mag, 2005. – Vol. 43. – No. 12. – P. 112–119.
3. **Wang X., Liu Q., Giannakis G.** Analyzing and Optimizing Adaptive Modulation Coding Jointly With ARQ for QoS–Guaranteed Traffic // IEEE Trans. Vehicular Techn. – March, 2007. – Vol. 56. – No. 2. – P. 710–720.
4. **Khan S., Duhovnikov S., Steinbach E., Keller W.** MOS–Based Multi–user application Cross–Layer Optimization for Mobile Multimedia Communication // Advances in Multimedia, Hindawi Publications, 2007.
5. **Palomar D., Chaing M.** A Tutorial on Decomposition Methods for Network Utility Maximization // IEEE J. on Selected areas in Communications, 2006. – Vol. 24. – No. 8. – P. 1439–1451.
6. **Puterman M.** Markov Decision Processes – Discrete Stochastic Dynamic Programming. – John Wiley & Sons, Inc., 2005. – 649 p.
7. **Wang H., Moayeri N.** Finite–state Markov channel – a useful model for radio communication channels // IEEE Trans. Veh. Technol., 1995. – Vol. 44. – No. 1. – P. 163–171.
8. **Ursulean R., Lazar A.–M.** Detrended Cross–Correlation Analysis of Biometric Signals used in a new Authentication Method // Electronics and Electrical Engineering. – Kaunas: Technologija, 2009. – No. 1(89). – P. 55–58.
9. **Burokas T., Staras S.** Properties of the Retard System Models Based on the Complex Cross Section Multiconductor Lines // Electronics and Electrical Engineering. – Kaunas: Technologija, 2008. – No. 4(84). – P. 3–8.

Received 2010 04 16

Z. Velickovic, M. Jevtovic. Adaptive Cross-layer Optimization Based on Markov Decision Process // Electronics and Electrical Engineering. – Kaunas: Technologija, 2011. – No. 2(108). – P. 39–42.

In order to satisfy QoS demands of wireless multimedia application it is necessary to make an optimization on several ISO-OSI layers in the protocol stack. In this paper an optimization cross-layer algorithm has been applied based on Markov decision process (MDP). The wireless communication system with one user has been optimized by the transmitting policies in order to maximize the throughput along with the optimization of the average value of the engaged power, satisfying the demanded BER and the average value of rejected packets. Simulation results show that the application of cross-layer design based on MDP is justified. Ill. 2, bibl. 9 (in English; abstracts in English and Lithuanian).

Z. Velickovic, M. Jevtovic. Markovo sprendimų metodo taikymas adaptyviesiems tarpusavio lygiams optimizuoti // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2011. – Nr. 2(108). – P. 39–42.

Būtina optimizuoti keletą ISO-OSI lygių siekiant per bevielio ryšio prieigą racionaliai išnaudoti aktyvinės įrangos resursus su įjungta paslaugos kokybės funkcija. Tarpusavio lygių optimizavimas pasiekiamas taikant Markovo sprendimų metodą. Bevielio ryšio sistema buvo optimizuota su vienu bevielio ryšio naudotoju. Atlikti tyrimai patvirtino Markovo sprendimų metodo taikymą tarpusavio lygiams optimizuoti. Il. 2, bibl. 9 (anglų kalba; santraukos anglų ir lietuvių k.).