ISSN 1392 – 124X INFORMACINĖS TECHNOLOGIJOS IR VALDYMAS, 2004, Nr. 3(32)

AUTOMATED MODELING AND PRACTICAL APPLICATIONS OF SYSTEMS DESCRIBED BY THE AGGREGATES

A. Zvironiene

Department of Applied Mathematics, Kaunas University of Technology
Studentu 50-323, 51368 Kaunas, Lithuania

A. Vrubliauskas

Department of Computer Science, Kaunas University of Technology
Studentu 50-208, 51368 Kaunas, Lithuania

Z. Navickas

Department of Applied Mathematics, Kaunas University of Technology
Studentu 50-326, 51368 Kaunas, Lithuania

Abstract. In this paper we present a concept of the aggregate, which is based on the "cellular" principle. Besides, some concrete application cases (in Ethernet network) and potentialities of the aggregate are presented.

Key words: the aggregate, the distribution module, the operation module, Ethernet network.

1. Indroduction

Recently, the investigations in the area of specification and simulation of complex systems are being performed by many scientists world-wide. The problems such as quality of a system and estimation of the system characteristics are most relevant. Methods and software tools are designed to help in creating analytical and simulation models of a system that are based on hybrid aggregates where system states change continuously and discretely [1], [2].

The scientists at the department of Business Informatics in Kaunas University of Technology, led by professor H.Pranevicius [3], [4], [5], are doing research in the area of mathematical modeling of systems. A piece-linear aggregate (PLA) formalism is used while designing the simulation models of systems found in various domains (telecommunication networks, transport, business etc.).

The aggregates are obtained by defining the finite automaton (Mealy automaton) [6]. Formally, it is a 5-tuple $A=(Z,X,Y,f,g)$. Here *Z, X, Y* – are the finite sets of states, inputs and outputs, respectively, f and g – the functions of the transitions and outputs, respectively, where $f: Z \times X \to Z$ and $g: Z \times X \to Y$, and g is the surjection. The Mealy automaton has a specific property that even with many states and transitions in the automaton the last transitions made depend on the last inputs.

The aggregate is designed having a precondition that the states of the aggregate are changing continuously. The classical concept of the aggregate is used [7], [8]: the aggregate is an object defined by the set of states *Z*, the set of input data *X* and the set of output data *Y*. The aggregate is operating during time $t \in T$. The states $z \in Z$, the input data $x \in X$ and the output data $y \in Y$ are the functions of time. Both, the operator *H*, which forms the transitions, and the operator *G*, which forms the outputs, have to be known. Differently from the Mealy automaton the input and output data of the aggregate are not synchronized with the change of the state.

In many complex systems, the size and complexity of tasks grows very fast. Different technologies are used to design more reliable, stable and error-prone systems. Formal analysis tools (tools that employ mathematical analysis to achieve provably secure systems) are becoming increasingly popular. A component-based approach to building formal analysis tools, where tasks can be decomposed into several subtasks, is widely used. Following the component-based approach, in this paper we propose a model of the aggregate, which is based on the "cellular" principle. In proposed model, different components are responsible for performing individual subtasks whose results are collected and presented to the user or coherent components. The main advantage of the aggregate model presented in this paper is the ability to model and simulate many complex systems from various domains.

In order to show the application of the proposed model, a case study of the Ethernet network is presented. In the study we present the aggregate capable of simulating and estimating the characteristics of the Ethernet network.

2. Model of the aggregate

One way to describe the aggregate as an object is the method of controlling sequences (I.N. Kovalenko and A.A. Borovkov [7], [8]), but it is rather difficult to automate it and it requires lots of time resources.

In this paper we propose a model of the aggregate (Figure 1), which consists of the input data generator, the data distribution module, the n-levels of the operation modules and the module of results.

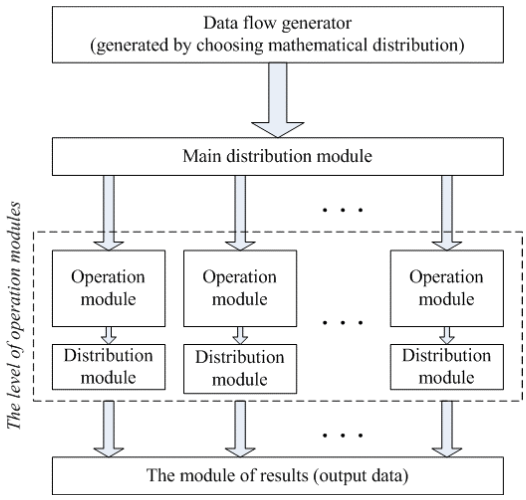

Figure 1. The model of the aggregate

The aggregate can be seen as a service, and we will use the concepts (such as input data flow, served packets, output data flow etc.) that are used when describing services.

To model the situation when at particular time moments a data flow enters the system, the incoming data generator was implemented. It can generate various data flows with various parameters according to the exponential, deterministic, uniform and other distributions or a combination of them.

In order to model how much time does it takes to process the data in the service, the exponential, uniform, deterministic or a combination of distributions can be used. Different distributions can be selected at different levels. Since the capacity of a buffer is unlimited, i.e. the system works without losses, the analytical methods for systems with series were used in the aggregate [9].

The operation of distribution module depends on the probabilities p_i . The probabilities can be selected by the user.

Behavior of the aggregate. The arrived data packets fall into the main distribution module, from which they randomly (by selected probabilities) are directed to the operation modules. In the operation module, the k-th data packet for some time a_{kl} is served in the l-th level. If the data packet arrives at the operation module and finds it busy, then the data packet is placed into the queue. After the data packet has been processed it is sent to the distribution module coherent to it, which again directs the data packet further to the next level or directs it to the results module. The results can be presented as tables or graphs.

A formal description of the aggregate model (Figure 1) is presented in [10]. A user interface for the aggregate is presented in Appendix A.

The analytical methods used in the implementation of the aggregate allow us to analyze various characteristics and processes that occur in complex telecommunication networks.

3. Applications of the aggregate

3.1. The aggregate for the estimation of the Ethernet network characteristics

Every telecommunication system has typical elements of service: the flow of input data, a service and the flow of output data (Figure 2.) [11].

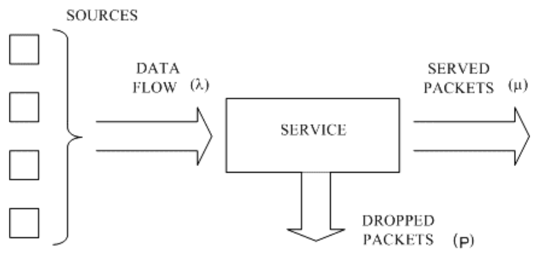

Figure 2. The general structure of a service

Usually, the input data flow is created by many users in the telecommunication networks. The nature of the data flow depends on the number and behavior of the sources, and service behavior. The random flows of input data occur mostly. Such data can be described by using various probabilistic distributions.

Data packets can be served differently: in the system with losses, if there's no possibility to serve a data packet, then it is dropped; in the system without losses, if there's no possibility to serve a data packet, then it waits for service in a queue.

The important parameter of a service is the servicing time a_i of data packet i, which can be deterministic or random, or in some cases distributed by the exponential distribution.

Because of a variety of telecommunication systems and networks, different mathematical models describing these systems, are needed. However, it is not always possible to create an adequate stochastic mathematical model for a very complex system. Such systems sometimes are analyzed piece by piece and then the conclusions are made about the complete system.

As the real object for the investigation the Ethernet network was taken [12], [13].

The simulation of the Ethernet network is performed as presented in Figure 3.

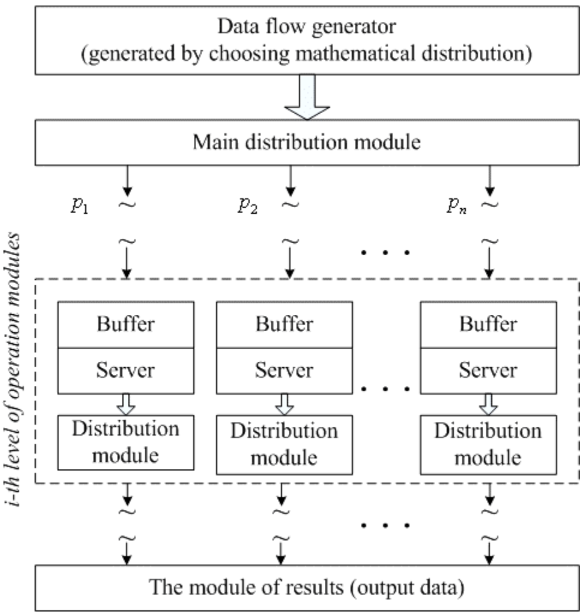

Figure 3. Ethernet network model

Ethernet network traffic is modelled as a data flow that consists of a stream of packets. A packet is a block of user data together with administration information attached, i.e. the time spent for servicing this packet and a reference time when this packet entered the aggregate.

The buffer and the server can be seen as the operation block. These blocks can be connected in parallel or serially, like those shown in Figure 4.

Some of the operation modules can be switched off by setting particular probabilities $p_i = 0$. As a result the workload on the other operation modules will be increased.

The aggregate presented in Figure 3 allows investigating the network load in the Ethernet networks and seeing how much time a data packet traverses the network, until it reaches the destination. These characteristics can be very useful to the designers of the telecommunication networks who seek to guarantee the high quality and optimal transmission of information in the network.

Other important network characteristics can be calculated: the time spent for servicing fixed number of packets, traffic load in Erlangs, channel utilization, the time a packet waits in the channel for servicing (packet delay), the maximum packet delay in the channels etc. Moreover, it is also possible to track an individual packet and analyse its characteristics.

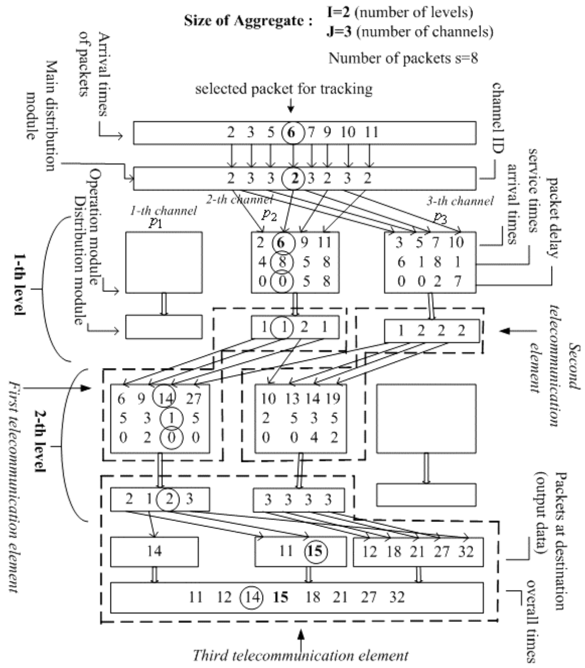

Figure 4. The coupling scheme of aggregate elements

The principle of operation of the aggregate containing two levels and three channels is illustrated in Figure 4.

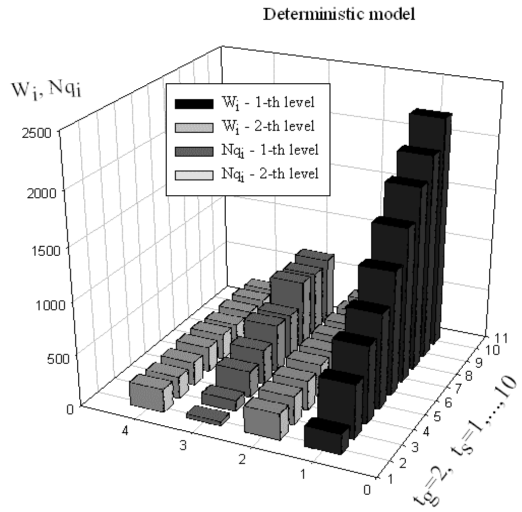

Figure 5. Simulation results, $t_g = 2$ and $t_s = \overline{1,10}$

The results of simulation using the aggregate model, which was designed for calculating Ethernet network characteristics, are presented in Figures 5, 6 and 7. Two important characteristics (W_i – delays of packets, Nq_i – the number of packets in the queues) were calculated. During the simulation, the data flow and servicing times were generated using the deterministic (simulation results presented in Figure 5) and uniform (simulation results presented in Figure 6) distributions. To generate the data flow, a deterministic distribution parameter t_g and the uniform distribution parameters A_g and B_g were selected as constants. To analyse the network having different servicing times, the deterministic distribution parameter t_s and the uniform distribution parameters A_s and B_s, were different for each simulation case.

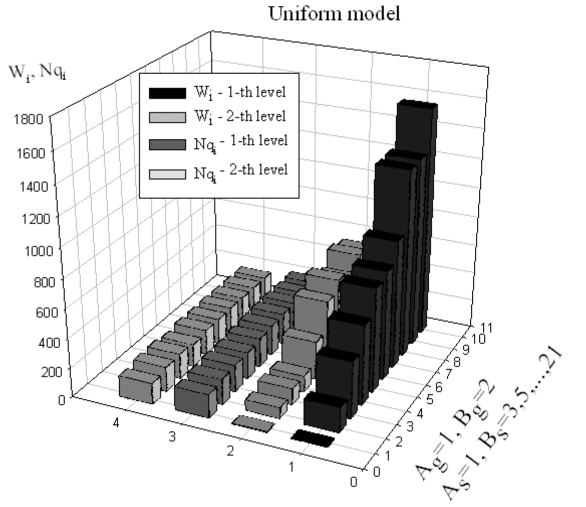

Figure 6. Simulation results, $A_g = 1, B_g = 2$ and $A_s = 1, B_s = \overline{1,10}$

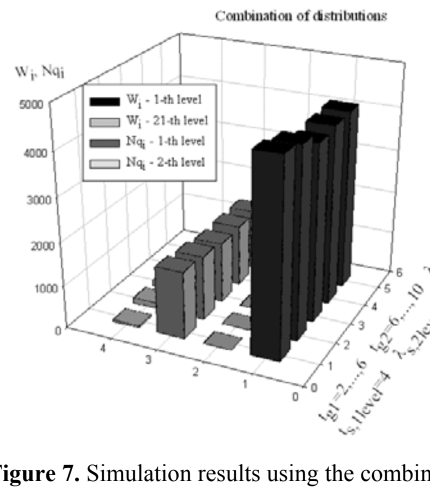

Figure 7. Simulation results using the combination of distributions

To create more advanced models, a combination of different probabilistic distributions can be used (simulation results presented in Figure 7). In this case, to generate the data flow a combination of two deterministic distributions (parameters $t_{g1} = \overline{2,6}$ and $t_{g2} = \overline{6,10}$) and exponential distribution (constant parameter $\lambda_g = 4$) were used. To analyse the network having servicing times generated using different distributions at different levels, a deterministic distribution at level 1 (parameter $t_s = 4$) and exponential distribution at level 2 (parameter $\lambda_s = 2$) were selected.

As can be seen from the simulation results (Figure 5, 6 and 7), there could be a problem at the first level, because of the high packet delay times. Since in real networks the queues are limited, there is a high risk that some packets may not arrive at their destination (arriving packets might be dropped by a service), causing the degradation of network performance.

3.2. Choosing between exponential and Weibull distributions

In some cases, telecommunication networks are modeled only by using exponential distribution [9, 11], which means that only exponential time modeling is expected. However, we propose to use a Weibull distribution (named after Wallodi Weibull), which is continuous probability distribution with the probability density function where $s > 0$ is the *shape parameter* and $\lambda > 0$ is the *scale parameter* of the distribution: $1 - e^{-(\lambda t)^s}$. A formal description is presented in Appendix B. The exponential distribution (when $s = 1$) and Raleigh distribution (when $s = 2$) are two special cases of the Weibull distribution. Many statistical analyses, particularly in the field of reliability, are based on the assumption that the data follow a Weibull distribution. Weibull distributions are often used to model the time until a given technical device fails. If the failure rate of the device decreases over time, one chooses $s < 1$ (resulting in a decreasing density *f*). If the failure rate of the device is constant over time, one chooses $s = 1$, again resulting in a decreasing function *f*. If the failure rate of the device increases over time, one chooses $s > 1$ and obtains a density *f* which increases towards a maximum and then decreases forever.

The results of Ethernet network simulation using four different cases of Weibull distribution (with parameter $s = 0.5, 1, 2, 5$) are presented in Figure 8.

As can be seen in Figure 8, the dynamics of the system characteristics, i.e. the waiting times W_i and the number of packets Nq_i waiting in a queue are decreasing if the value of parameter s is increased. One of the advantages of using the Weibull distribution is that by leaving the exponential nature of the distribution it is relatively easy to change the dynamics (data flow intensity, servicing times etc.) of the modeled system.

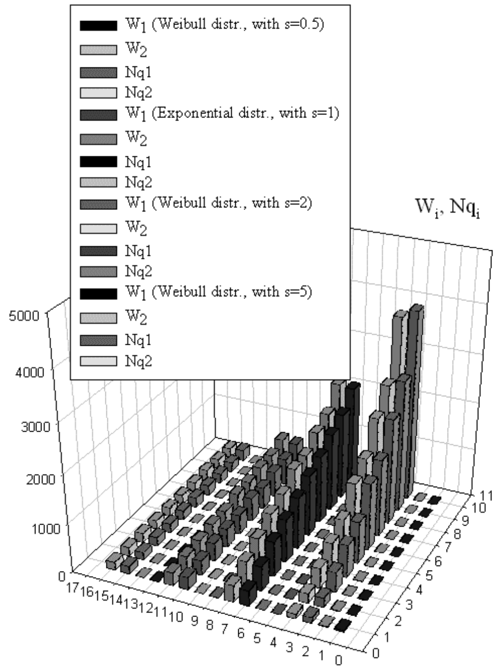

Figure 8. Simulation results using four different cases of Weibull distribution

4. Conclusion

Following the component-based approach, we proposed a model of the aggregate, which is based on the “cellular” principle. In the model, different components are responsible for performing individual sub-tasks whose results are collected and presented to the user or coherent components. The main advantage of the aggregate model presented in this paper is the ability to model and simulate many complex systems from various domains.

A case study of the Ethernet network has been made, showing the use of the aggregate model presented in this paper. As the case study showed, it is relatively easy to use and combine different statistical distributions available in the model. Moreover, a Weibull distribution has been presented as an alternative for the exponential distribution, which is widely used when modeling telecommunication networks. One of the advantages of using the Weibull distribution is that by leaving the exponential nature of the distribution it is relatively easy to change the dynamics of the modeled system.

To support a wider range of the systems to be modelled, in the future we plan to include many more options to the modelling system and implement alternative probability distributions (Parett etc.).

Appendix A. User Interface

Size of aggregate:

Levels *I*= ☐
Channels *J*= ☐

○ Select a packet to track : p_n= ☐

Data flow generation strategy:

○ By time: *Time*= ☐
○ By number of packets: *s*= ☐

Method of data flow generation:

○ - deterministic with parameter *t*
○ - uniform with parameters *A* and *B*
○ - exponential with parameter *LIAMBDA*
○ - combined: det+unif+exp with parameters *t, A, B, LIAMBDA*
○ - combined: det+unif with parameters *t, A, B*
○ - combined: det+exp with parameters *t, LIAMBDA*
○ - combined: unif+exp with parameters *A, B, LIAMBDA*
○ - combined: deter1+deter2 with t_1 and t_2
○ - combined: unif1+unif2 with A_1 and B_1 , A_2 and B_2
○ - combined: exp1+exp2 with $LIAMBDA_1$ and $LIAMBDA_2$
○ - Weibull with parameter λ and *s*

Parameter values:

t= ☐
A= ☐ *B*= ☐
LIAMBDA= ☐
t_1= ☐ t_2= ☐
A_1= ☐ B_1= ☐ A_2= ☐ B_2= ☐
$LIAMBDA_1$= ☐ $LIAMBDA_2$= ☐
λ= ☐ s= ☐

Distribution strategy:

○ - distribution module distributes packets by the same probabilities
○ - distribution module distributes packets from 1 to *J sequentially*
○ - by probabilities Q_p= ☐

The strategy of servicing times generation :

○ - servicing times for all operation modules generated by the same distribution:

○ ~ deterministic with parameter *aa*= ☐
○ ~ uniform with parameters *aA*= ☐ and *aB*= ☐
○ ~ exponential with parameter *MIU*= ☐
○ ~ combined: deter+unif+exp, with probabilities *p*= ☐
○ ~ combined: deter+unif, with probabilities *p*= ☐
○ ~ combined: deter+exp, with probabilities *p*= ☐
○ ~ combined: unif+exp, with probabilities *p*= ☐
○ ~ combined: deter1+deter2 with tt_1= ☐ and tt_2= ☐
○ ~ combined: unif1+unif2 with aA_1= ☐ and aB_1= ☐
aA_2= ☐ and aB_2= ☐
○ ~ combined: exp1+exp2 with Miu_1= ☐ and Miu_2= ☐
○ ~ Weibull with λ= ☐ and *s*= ☐

○ - servicing times of operation modules are different at several levels

var_a= ____ 1 level
____ 2
...

○ - servicing times of operation modules generated by exp distribution , but with different parameters MIU in each level

MIU = ____ 1 level
____ 2
...

Select which characteristics will be calculated :

- ☐ - Z – Mean value of an inter -arrival time
- ☐ - W – packet delays
- ☐ - P{W_J>0} - Probability, that a packet in J -th channel will wait in a queue
- ☐ - $W_J$max - maximum packet delay in channel
- ☐ - Ts – packet lifetime
- ☐ - Tserv – mean servicing time of packets in channels
- ☐ - T – system operation time
- ☐ - Erl – traffic load in Erlangs
- ☐ - N_J – the number of packets in each channel
- ☐ - Roch - channel utilization
- ☐ - LIAMBDAin - mean load of arriving data flow
- ☐ - t_{in} – packet arrival times
- ☐ - t_{out} – outgoing times of served packets
- ☐ - show results only for tracked packet

Appendix B. Weibull Distribution

The random value ξ is distributed by Weibull distribution:

$$M\xi^k = \frac{1}{\lambda^k}\Gamma\left(\frac{k+s}{s}\right).$$

Then

$$M\xi = \frac{1}{\lambda}\Gamma\left(\frac{1+s}{s}\right),$$

and

$$D\xi = \frac{1}{\lambda^2}\left[\Gamma\left(\frac{2+s}{s}\right) - \left(\Gamma\left(\frac{1+s}{s}\right)\right)^2\right].$$

However, in telecommunications, the intensity $\tau = \frac{1}{M\xi}$ is used. Then the intensity τ can be expressed through the distribution parameter λ:

$$\tau = \lambda \cdot \Gamma^{-1}\left(\frac{1+s}{s}\right).$$

Note that when $s = 1$ we have the exponential distribution with parameter $\tau = \lambda$.

References

[1] **T. Krilavičius.** Aggregate specifications using coalgebras and monoid actions. *http:// www3.vdu.lt/~ivus01/pdf/sm_krilavicius.pdf*

[2] **S. Bohacek, B. Rozovskii.** Models and Techniques for Network Tomography. *Proceedings of the* 2001 *IEEE. Workshop on Information Assurance and Security* T1C2 1030 *United States Military Academy, West Point, NY*, 5-6 *June* 2001. *http://www.itoc.usma.edu/Workshop/2001/Authors/Submitted_Abstracts/paperT1C2(44).pdf*

[3] **K. Wang, H. Pranevicius.** Applications of AI to Production Engineering. *Nordic-Baltic gummer school'*97, *Lecture Notes, KTU, Technologija*, 1997, 269-322.

[4] **H. Pranevičius, J.A.G. Knight, I. Pranevičienė.** Formal specification and analysis of distributed systems using aggregate approach. *Informacinės technologijos ir valdymas, Nr.*1(22), *KTU, Technologija,* 2002, 31-38.

[5] **H. Pranevičius.** Kompiuterinių tinklų protokolų formalusis specifikavimas ir analizė: agregatinis metodas, *Monografija, Kaunas, Technologija,* 2003, 19-32.

[6] **V. Brauer.** The automate theory. *Moscow,* 1987, (*in Russian*) .

[7] **И.Н. Коваленко, Н.П. Бусленко.** Лекции по теории сложных систем. *Москва,* 1973, 99-421.

[8] **А.А. Боровков.** Вероятностные процессы в теории массового обслуживания. *Москва: Наука (Главная редакция физико – математической литературы),* 1972, 5-367.

[9] **Z. Navickas, R. Rindzevičius, M. Šmočiukas.** Daugiakanalių sistemų su eilėmis, esant skirtingai paraiškų paskirstymo į laisvus kanalus tvarkai, analiziniai modeliai, *Elektronika ir elektrotechnika, Nr.*1(36), 2002, 15-19.

[10] **A. Žvironienė.** Automatic aggregate modelling, *Liet. matem. rink., T.*42*, spec.no., Vilnius*, 2002, 636–640.

[11] **R. Rindzevičius.** Informacijos srautų analizės ir jų aptarnavimo efektyvumo įvertinimo matematiniai ir imitaciniai modeliai, *Monografija*, *Kaunas, Technologija,* 2003, 137.

[12] **M. Tatipaluma, B. Khasnabish (Editors).** Multimedia Communications Networks Technologies and services. *Boston-London, Artech house,* 1998, 39-102.

[13] **Ng Chee Hock.** Queueing Modelling Fundamentals, *England,* 2001.

DOI: 10.5755/j01.itc.32.3.11864