EVALUATION OF EXPERIMENTAL DATA FROM MONITORING AND SIMULATION OF NETWORK COMMUNICATION PARAMETERS

Radi Romansky *

Department of Electronics and Electro-energy, College of Energy and Electronics, Technical University of Sofia
Bulgaria

* Corresponding Author, e-mail: rrom@tu-sofia.bg

Abstract: The main goal of the article is to investigate selected communication parameters at the access to remote resources in the digital spice. The goal is to analyse the possible losses in maintaining communication performance which is made based on preliminary mathematical formalization and organization of two types experiments – program monitoring and simulation. Obtained results are processed and assessments are summarized.

Key words: network communication, formalization, program monitoring, simulation, statistical analysis, assessments.

1. INTRODUCTION

The contemporary digital age requires sufficient reliable and legal access to information resources [1] and secure information services in network communications [2, 3]. This places new demands on the infrastructure (for example based on quantum computing [4]), the organization of the allocated resources, the protection of personal data upon access [5, 6], ensuring adequate information security, etc. It is obvious that information security is an important component of information services in the digital space, which is characterized by mass digitization of resources and objects. In this direction, a modern technology for digitalization for security purposes and proposes an effective adaptive method for digital transformation is presented in [7]. In addition, the information security risk when using cloud computing in different cases is discussed in [8, 9]. On the other hand, article [10] clearly emphasizes that information security policies are directly dependent on the actions of users to protect information resources, and when adequate security measures are not implemented, the probability of data breaches is significant.

The information service in the digital space is directly related to the organization of network communications [11], as the times for data transfer between the individual nodes (workstations) depend not only on the time of the actual transmission on the communication lines, but also on the additional influence of external factors causing losses of communication time. This leads to lower productivity of the transmission medium as a whole. This determines the importance of access to remote resources and the study of parameters of communication processes, which is important in increasing performance by determining "end-to-end" delay known as "latency" [12].

The purpose of the article is to present a study of communication parameters and increase the informativeness of the obtained estimates by joint validation of experimental data. The goal is to analyse the possible losses in maintaining communication performance. For this purpose, the formalization of the investigated object was performed and a two-stage study was planned, including conducting program monitoring in a real working environment, followed by simulation modelling under similar conditions. Such an approach allows comparability of experimental results and ensures the adequacy of the estimates obtained.

The structure of the article includes an analysis of related work on the topic (section 2) and initial discussion of the task statement (section 3), descriptions of the applied approach to mathematical formalization (section 4), as the results of program monitoring and simulation are presented in section 5, and section 6 summarizes the obtained experimental data.

2. RELATED WORKS

The contemporary development of network communications is based on new and promising technologies, but as noted in [11], despite the rich theoretical and practical experience of research, new approaches are needed to solve the problems of optimal management of large corporate computer networks. In support of this statement, the article analyses the existing routing and switching algorithms in order to determine good directions for improving the efficiency of routing, which is directly related to improving communication indicators and interaction performance. Problems related to the efficiency of network communications and management are also discussed in [2], where it is stated that in traditional networks this is done mainly manually via the command line, but in large networks this is inefficient. To overcome this shortcoming, the article proposes a network management approach using the Ansible playbook by dynamically configuring the multitude of devices and sending archive files to the Version Control Git.

The analysis of the productivity in network communications is the main task in the design and is a condition for proper implementation and maintenance of the processes and as stated in [12] is a guarantee for the effectiveness of the developed project. The article proposes the use of a research framework based on Finite State Machine and Mathematical Analytical method to calculate the expected average delay in the network topology, as well as the use of a network simulator. In addition, research using MATLAB and Cisco Packet Tracer has been conducted to confirm the effectiveness of the proposed framework. The importance of network efficiency, especially in wireless network communications, is also discussed in [13], identifying two important factors – data transfer rate and spectral efficiency. Because data transmission channels are subject to various damages that degrade network performance, the article proposes an approach to channel equalization through the use of machine learning in comparative error rate analysis.

Different approaches can be applied in the study of network communications and process efficiency, which depends on the goal and the object. For example, in [14] the theory of Finite State Machine is applied for investigation of software and functionality of the object. An analytical model of a finite automata for testing the processes is defined, its states are determined and a complete algorithmization of the research procedure is performed. In another direction is the study presented in [15], conducted based on a developed framework with the application of monitoring in assessing the quality of open government data. For the purposes of the research, an application was developed using a combination of indicators, based on the performed measurements. In [16] the situation with remote servers, which are divided into separate groups with equal capacity, is discussed. For each group there is a defined main server that stores up-to-date control information for the others. The task is to minimize the total time for searching and processing information throughout the system. To solve the problem, a statistical model has been developed examining the possibilities for a significant reduction in communication time when searching for certain data.

Smart technologies and Internet of Things (IoT) applications add new requirements when study the effectiveness of connections. In [17] attention is paid to Edge Mining and Cloud Mining as related to IoT scenarios and executed in accordance with the principles of Cloud or Edge calculations. To meet basic requirements such as bandwidth, energy saving, ensuring confidentiality and security, the article proposes a methodology based on simulation and intelligent monitoring and allowing to ensure the necessary efficiency of data mining processes. The combination of simulation and monitoring is an approach used in [18], confirming its effectiveness in studying the parameters of wireless sensor network (delay packet transfer, power consumption, network behaviour) with the main goal to facilitate the implementation of the real network by analysis of leading parameters. Simulation is performed on the firmware of the functionality (software) of the network nodes, which provide a specific application with the communication infrastructure, whose measured data are compared with the results of the simulation.

3. INITIAL STATEMENT OF THE TASK

The investigation of the interactions between distributed objects in a network environment [19] and the information and communication processes taking place in it is an important moment in the optimization of the distributed information service [20]. Possible alternatives in the organization of research experiments are related to the application of methods for modelling (determined, probabilistic, simulation, empirical, etc.) or for measurement (hardware or software monitoring). When choosing a research approach and specifying the main tasks, both the architectural features of the distributed information environment and factors such as the characteristics of network processes, the expected estimates of system performance, the possibilities of algorithmic provision, etc. should be taken into account.

Network traffic is very diverse and stochastic in nature and is divided into two categories: functional (transactional) traffic and background. There are two main approaches to its monitoring:

1. Monitoring the details of each, or at least many of the frames passed through the network environment – allows viewing of each frame passed through the network, or filtered by a special criterion frames. This method is especially suitable for dealing with network problems, as the content of the frames is extremely important in this process.

2. Monitoring of statistically processed results for the load of network segments and for the distribution of network protocols and services by segments. The detailed content of each frame is not available and in many cases is not buffered during the operation of the network analyser. This method is suitable for long-term monitoring of network segment parameters in order to perform optimization.

Some features of network traffic related to research and evaluation of information services can be summarized as follows:

- It is possible that in certain network segments with a significant load there is a large volume of network traffic, which can lead to research problems (for example, delays in the real environment, loss of real staff, etc.).
- The conduct of a study of network traffic, including surveillance, must be based on a predefined goal and strategy, in order to ensure the necessary effectiveness of the research tools.
- Due to the existence of different types of traffic (some of which are less important for the overall productivity of information services), it is necessary to focus specific research on appropriate types to ensure the adequacy of the estimates and conclusions.
- In order to optimize the experiments and present the estimates, it is necessary to choose the appropriate research method for the specific task, which will reduce possible "distortions" of the estimates. In this sense, the means used to illustrate the accumulated experimental results (sample of registrations) and the estimates formed on the sample are also of great importance.

4. PRELIMINARY MATHEMATICAL FORMALIZATION

For the general organization of the experiments and the research as a whole, a preliminary mathematical formalization of the main objects and processes has been performed. Two basic concepts related to data transfer between two workstations are defined: communication distance (*d*) and packet size (*s*). For each pair of workstations $<WS_A, WS_B>$ these parameters have individual dimensions because they take specific values for which the following formalization can be given.

- ***Communication distance.***

$$\forall < WS_A, WS_B > \Rightarrow \exists d_{AB} = d(WS_A, WS_B) \in D$$

$$D = [d_{\min}, d_{\max}] = \bigcup_{i=1}^{n} D_i$$

$$D_0 = d_{\min}; \quad d_{\min} \le D_i = D_{i-1} + \Delta D \le d_{\max}$$

Separate groups of distances (communication distances) D_i (i=1÷n) for which the experiments are planned are defined. Each individual distance falls into one of the group values, i.e. $d_{AB} \in D_i \in D = \{D_0, D_1, ..., D_n\}$. The size of set D is:

$$n = \left\lceil \frac{d_{\max} - d_{\min}}{\Delta D} \right\rceil$$

• ***Packet size.***

When planning the experiments, it is assumed that the packets have a size limited in the range of minimum and maximum value, i.e. $s_j \in S \equiv [s_{\min}, s_{\max}]$. To unify the planning, a step of change ΔS is introduced, which defines basic values:

$$S_0 = s_{min} \quad \text{and} \quad S_i = S_{i-1} + \Delta S \leq s_{min}$$

This determines a set $S = \{S_i\,;\, i=0 \div m\}$ with power $m = \left\lceil \frac{s_{\max} - s_{\min}}{\Delta S} \right\rceil$.

When planning experiments, the parameters d and s are controllable factors. The main observable factor is the communication time (CT). For its evaluation it is necessary to obtain generalized evaluations based on individual values under different communication conditions.

Let t_{ij} be the *individual time for transporting* a packet of size S_j between two workstations at a distance D_i, i.e.

$$t_{ij} = t(D_i, S_j) \quad ; i = 1 \div n; j = 1 \div m \cdot$$

The communication time depends on the size of the transmitted packets, which allows to define the *average communication time* for a given distance in the network environment:

$$t_i = t(D_i) = \frac{1}{m} \sum_{j=1}^{m} t_{ij}$$

After summarizing for the whole network environment, an expression for estimating the *total communication time* is defined on the set D:

$$CT = \frac{1}{n} \sum_{i=1}^{n} t_i = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{1}{m} \sum_{j=1}^{m} t_{ij} \right] = \frac{1}{n.m} \sum_{i=1}^{n} \sum_{j=1}^{m} t_{ij} \quad .$$

This formalization was used as a basis for research of network environment of type Ethernet 100BaseT. The maintained network traffic is generally self-similar, i.e. it has the same stochastic characteristics at different values for the controllable parameters. This allows the results obtained from several individual experiments to be summarized in a global assessment of observable factors. The experiments organization is made based on the formalization.

5. ORGANIZATION OF EXPERIMENTS

5.1. Program monitoring

Specific values for the controllable factors presented below are defined for the purpose of the program monitoring organization:

• range of transport distances in the environment $D = [0, 3000]$ with dimension [m] at step $\Delta D = 500$m and power of the set of distances $\|D\| = n = 6$;

• size range of the transferred packets $S = [0, 10240]$ with dimension [bit] at step $\Delta S = 512$b, which determines the power of the set of group sections $\|S\| = m = 20$.

According to the introduced formal rules, the initial values for both sets are zero and are not informative. This allows further planning of the experiment to use pairs of controllable factors $<D_i, S_j>$, applying a complete factor plan over the vectors $<D_1,\ldots, D_6> = <500, 1000, 1500, 2000, 2500, 3000>$ and $<S_1,\ldots, S_{20}> = <512, 1024, \ldots, 9728, 10240>$. Software products HoverIPO and Network Probe 1.1 were used to perform the measurement. A summary of the obtained experimental results is shown in Fig. 1.

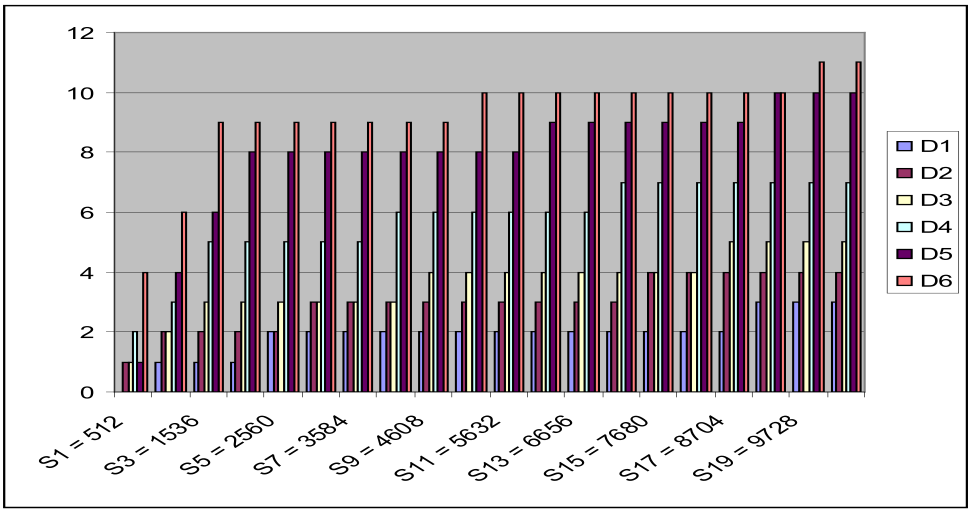

Figure 1. Experimental monitoring results for D_i [m] at S_j [b]

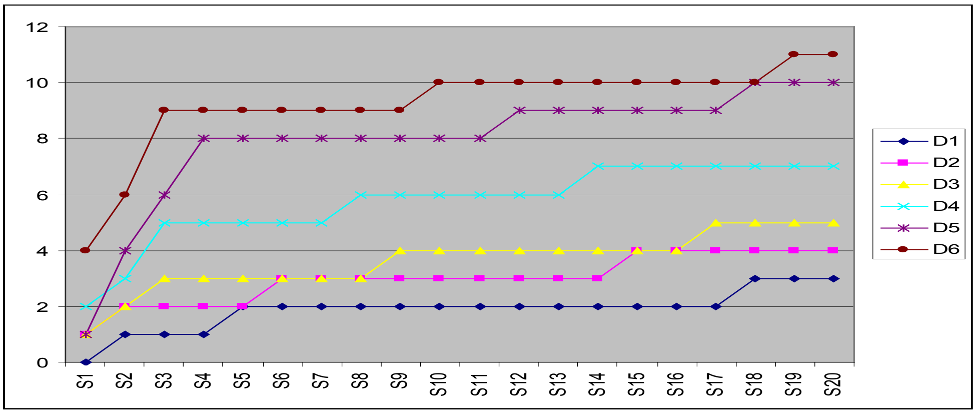

Figure 2. Experimental results for t_{ij} [s] for selected distances

The graphical interpretation of Fig. 2 shows for each distance Di the change of the transport time with increasing packet size, as well as the change with increasing with the selected step ΔD of the distance. These results give an idea of the average communication time for each individual distance in the network environment. Based on the obtained experimental results, the estimates for both times t_i and CT were calculated. The estimates for the average communication time for the individual distances are summarized in Fig. 3, as the calculated estimate for the total communication time is CT = 5.167 [s].

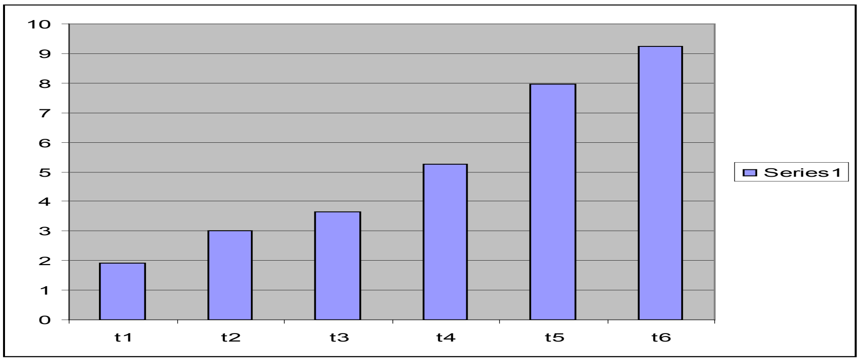

Figure 3. Estimates of average communication times for distances in [*s*]

5.2. Simulation

When organizing the simulation modelling, parameters and values similar to the measurement are selected. The aim is to run the simulation under conditions similar to monitoring so that the results can be compared. This achieves higher reliability of estimates, as well as verification of experimental results.

To develop the simulation model, a mathematical model based on queuing theory has been defined, which includes two parts – Model of the system and Model of the input flow (determined workload). The first model was developed as a service system with feedback and multiple access to incoming service requests. The system has no service priorities, i.e. applications are treated as equal. It is accepted that each workstation accesses the transmission medium on the basis of one or more attempts to occupy it. Attempts to access k are limited to the LNA value, and if it is exceeded, a failed access is registered. Network self-similar traffic is defined as multiple stochastic flow generated by the subscribers Z_j in the network environment. Each request is routed to the input buffer of the system model, from where it is routed to an access analysis environment that checks for each request the fulfilment of the condition $k_j \leq LNA$. An illustration of the general idea for creating the simulation model is shown in Fig. 4. The program implementation was performed in GPSS World for Windows environment, and two models were defined – control (for setting the model parameters and analysis of model sensitivity) and working (for conducting simulation experiments and accumulation of statistical estimates).

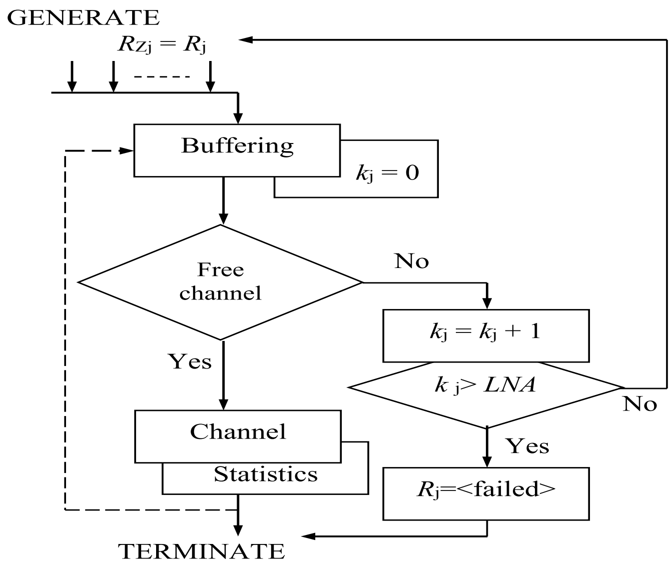

Figure 4. Basic principle of the simulation experiments organization

When conducting the simulation study, the step was reduced twice (256b), which leads to an increase of *m* to 40. The aim is to achieve higher detail of the study by increasing the registrations. Summary estimates of the simulation results are presented in the following two figures. Fig. 5 illustrates the development of the transfer time with increasing packet size for each of the studied distances, and Fig. 6 shows the influence of the increasing distance between workstations when transporting selected fixed length packets.

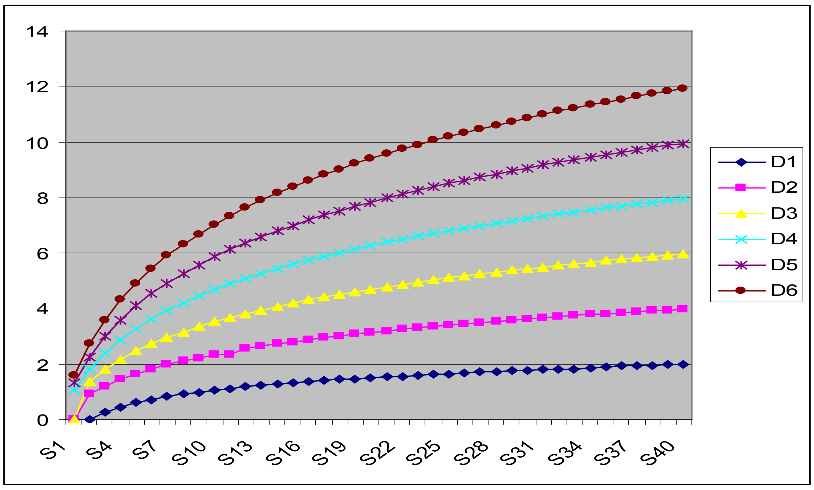

Figure 5. Simulation results for parameter t(S)

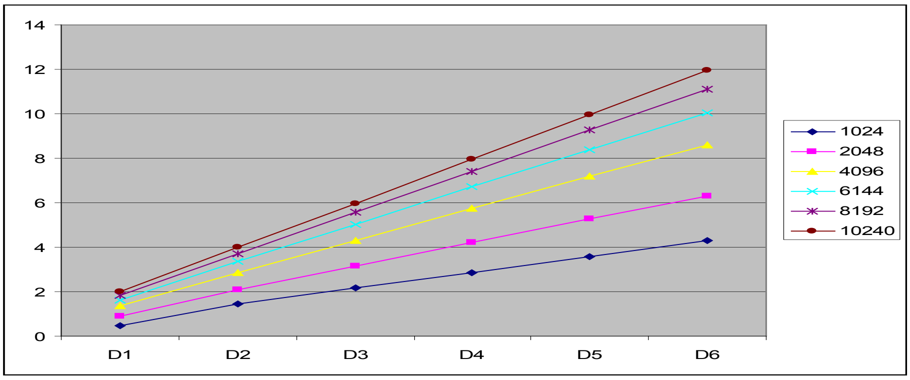

Figure 6. Experiment al results from simulation of parameter t(D) at fixed packets size

6. EVALUATION OF EXPERIMENTAL DATA AND ASSESSMENTS

The conducted comparative analysis is organized over samples of the two methods with the same values for the controlled factors. The larger number of registrations in the simulation study has its advantages in the formation of statistical estimates from the model study and does not interfere with the comparison with the measured data.

Figure 7 compares the estimates for selected sets obtained independently by monitoring (Sm) and by simulation (Ss). The results are presented for all studied distances with varying packet size, covering the range $S = [0, 10240]$.

The results of both types of research give close estimates of the time parameters in the organization of communications in the network environment. This can be seen from Fig. 8, which shows graphic dependences for three type groups of packets with sizes of 1024, 6144 and 10240 bits, respectively. It can be seen that the results of the monitoring (whit index m) and the simulation (whit index s) are close enough, which allows to confirm the summarized results for the communication time given in Table 1.

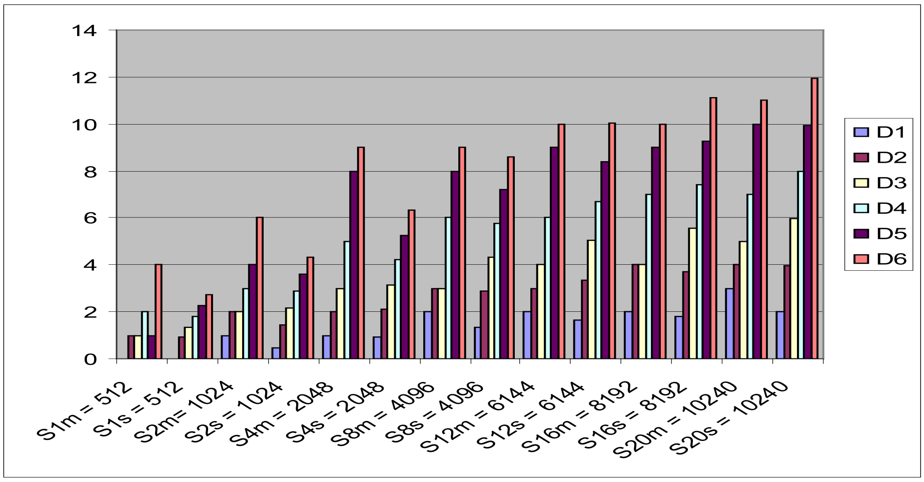

Figure 7. Comparison of monitoring and simulation estimates

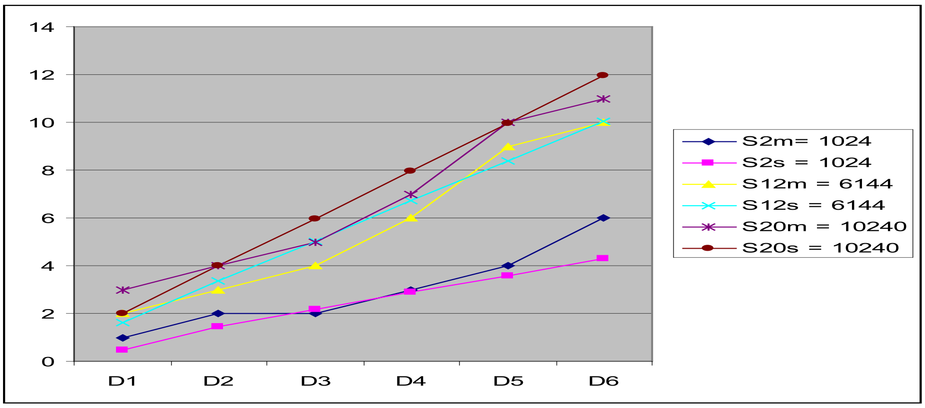

Figure 8. Deviation of estimates under the same experimental conditions

Table 1.

Average communication time t_i*:*					
t_1	t_2	t_3	t_4	t_5	t_6
1,90	3,00	3,65	5,25	7,95	9,25
1,366	2,905	4,361	5,841	7,301	8,761
Total communication time CT:					
Monitoring ➔ 5,167			Simulation ➔ 5,089132		

7. CONCLUSION

Policies for the management of contemporary communication systems, especially for 5G networks, need to provide advanced functionalities in order to adequately respond to the increasing demands on them, particularly in the so-called edge computing [21]. One of the possible approaches is to conduct an adequate study of network traffic and the behaviour of communication parameters, in accordance with the type of traffic and specific application [22].

The article discusses some features of network traffic and communications between distributed objects, and proposes am approach for combined investigation of communication parameters by using monitoring and simulation. Conducted measuring in a network environment allowed the accumulation of registrations and determination of the range of controllable factors used in the organization of simulation modelling.

The applied approach for combined research through a complete factor plan at the same values for the controllable parameters ensures the adequacy of the estimates. Preliminary formalization allows adequate structuring of the studied objects and limiting the influence of external factors. Thus, the results of this study can be successfully used in optimizing access and managing communications in a network environment.

REFERENCES

[1] Pardayev, S.M., S. K. Sindarov, S.K. Legal regulation of access to information resources of e-government interactive service. *Mental Enlightenment Scientific-Methodological Journal,* No. 1, 2022, pp. 288-299.

[2] Bwalya, B., Zimba, A. An SDN approach to mitigating network management challenges in traditional networks. *International Journal on Information Technologies and Security*, vol. 13, No. 4, 2021, pp. 3-14.

[3] Zimba, A., Mulenga, A. A Dive into the deep: Demystifying WannaCry crypto ransomware network attacks via digital forensics. *International Journal on Information Technologies and Security*, Vol. 10, No. 2, 2018, pp. 57-68.

[4] Avesani, M., Foletto, G., Padovan, M., Calderaro, L., Agnesi, C., Bazzani, E., Berra, F. et al. Deployment-ready quantum key distribution over a classical network infrastructure in Padua. *Journal of Lightwave Technology,* No. 6 (vol. 40), 2022, pp. 1658-1663.

[5] Romansky, R. Digital age and personal data protection, LAP LAMBER Acad. Publ., January 2022, 124 p.

[6] Tzolov, T. Data model in the context of the General Data Protection Regulation. *International Journal on Information Technologies and Security*, Vol. 9, No. 4, 2017, pp. 113-122.

[7] Ivanova, Y. Adaptive digitalization methods and digital transformation trends for security. *International Journal on Information Technologies and Security*, Vol. 12, No. 3, 2020, pp.51-62.

[8] Tsaregorodtsev, A.V., Kravets, O.Ja. Choporov, O.N., Zelenina, A.N. Information security risk estimation for cloud infrastructure. *International Journal on Information Technologies and Security*, Vol. 10, No. 4, 2018, pp. 67-76

[9] Zaslavskaya, O.Y., Zaslavskiy, A.A. Bolnokin, V.E., Kravets, O.Ya. Features of ensuring information security when using cloud technologies in educational institutions. *International Journal on Information Technologies and Security*, Vol. 10, No. 3, 2018, pp. 93-102.

[10] McLeod, Al. Dolezel, D. Information security policy non-compliance: Can capitulation theory explain user behaviors?. *Computers & Security*, Vol. 112, 2022, art. 102526. https://doi.org/10.1016/j.cose.2021.102526

[11] Kravets, O.Ja., Atlasov, I.V., Aksenov, I.A., Molchan, A.S., Frantsisko, O.Yu., Rahman, P.A. Increasing efficiency of routing in transient modes of computer network operation, *International Journal on Information Technologies and Security*, Vol. 13, No. 2, 2021, pp. 3-14.

[12] Alsheikhy, A. Estimating End-to-End delay on a networking environment using developed framework. *International Journal on Information Technologies and Security*, Vol. 14, No. 1, 2022, pp. 3-16.

[13] Hassan, Sh., Tariq, N., Ali Naqvi, R., Ur Rehman, A., Kaabar, M. K. A. Performance evaluation of machine learning-based channel equalization techniques: New trends and challenges, *Journal of Sensors (Special issue: "Physical Layer Security for Sensor-enabled Heterogeneous Networks")*, Vol. 2022, Article ID 2053086, 2022, 14p. https://doi.org/10.1155/2022/2053086

[14] Zozulya, M.M., Kravets O.Ja., Atlasov, I.V., Aksenov, I.A., Bozhko, L.M., Rahman, P.A. Algorithmization of the software testing system based on finite automata. *International Journal on Information Technologies and Security*, Vol. 14, No. 1, 2022, pp. 77-86.

[15] Raca, V., Velinov, G., Cico, B., Kon-Popovska, M. Development framework of monitoring and evaluation of open government data quality, *International Journal on Information Technologies and Security*, Vol. 13, No. 2, 2021, pp. 87-100.

[16] Atlasov, I. V., Bolnokin, V.E., Kravets, O.Ja, Mutin, D.I., Nurutdinov, G.N. Statistical models for minimizing the number of search queries. *International Journal on Information Technologies and Security*, Vol. 12, No. 3, 2020, pp. 3-12.

[17] Savaglio, C., Fortino, G. A simulation-driven methodology for IoT data mining based on edge computing. *ACM Transaction on Internet Technology*, Vol. 21, No. 2, art. 30, June 2021, pp.1-22. https://doi.org/10.1145/3402444.

[18] Särestöniemi, M., Pomalaza-Ráez, C., Kissi, C., Iinatti, J. Simulation and measurement data-based study on fat as propagation medium in WBAN abdominal implant communication systems. *IEEE Access*, Vol. 9, 2021, pp. 46240-46259, doi: 10.1109/ACCESS.2021.3068116.

[19] Goryachko, V.V., Choporov, O,N., Preobrazhenskiy, A.P., Kravets, O.Ja. The use of intellectualization management decision-making in the interaction of territorially connected systems. *International Journal on Information Technologies and Security*, Vol. 12, No. 1, 2020, pp. 87-98.

[20] Nedyalkov, I., Stefanov, A., Georgiev, G. Studying and characterization of the data flows in an IP-based network. *International Journal on Information Technologies and Security*, Vol. 11, No. 1, 2019, pp. 3-12.

[21] Atanasov, I., Pencheva, E., Nametkov, A., Trifonov, V. On functionality of policy control at the network edge. *International Journal on Information Technologies and Security*, vol. 11, No. 3, 2019, pp. 3-24.

[22] Shi Dong, Y. Xia. Network traffic identification in packet sampling environment. *Digital Communications and Networks*, 2022 (in press). https://doi.org/10.1016/j.dcan.2022.02.003

Information about the author:

Radi Romansky is a full professor at Technical University of Sofia, Doctor (Dr) in Computer Engineering and Doctor of Science (D.Sc.) in Informatics and Computer Science; Full member of European Network of Excellence on High Performance and Embedded Architectures and Compilation (HiPEAC). He has over 215 scientific publications and over 25 books. Areas of scientific interests: ICT, informatics, computer architectures, computer modelling, privacy and data protection, etc.

Manuscript received on 22 March 2022