III. ENGINEERING SCIENCES AND TECHNOLOGIES
III. ŞTIINŢE INGINEREŞTI ŞI TEHNOLOGII

UDC: 004.73

QUEUEING MODELS OF POLLING TYPE AND THEIR APPLICATIONS
MODELE POLLING DE AŞTEPTARE ŞI UTILIZAREA LOR

MISHKOY Gheorghe, academician, PhD, University Professor,
Free International University of Moldova, Chisinau
MITEV Lilia, PhD,
Free International University of Moldova, Chisinau

MIŞCOI Gheorghe, academician, doctor habilitat, profesor universitar,
Universitatea Liberă Internațională din Moldova, Chișinău
MITEV Lilia, doctor,
Universitatea Liberă Internațională din Moldova, Chișinău

Annotation: *Polling model is a multi-queue system with a single server which visits the queues according to a polling order and serves the customers from these queues. Furthermore, polling models are applicable in situations in which several users compete for access to a common resource which is available to only one at a time, such as communication systems, traffic and transportation systems, production systems, etc.*

Annotare: *Modelul polling este un sistem multi-queue, cu un singur server care verifică așteptările în funcție de o comandă de sondare și servește clienților aflați în așteptare. Mai mult, modelele polling sunt aplicabile în situațiile în care mai mulți utilizatori concurează pentru accesul la o resursă comună disponibilă doar la un moment dat, cum ar fi sistemele de comunicații, sistemele de trafic și de transport, sistemele de producție etc.*

Keywords: *polling model, queueing models, communication systems, traffic systems, transportation systems, production systems.*

Cuvinte-cheie: *model polling, șir de așteptare, sisteme de comunicații, sisteme de trafic, sisteme de transport, sisteme de producție.*

Introduction

Systematization and generalization of the theoretical results obtained in the study of Polling models by the year 1985 are summarized in the monograph of H. Takagi [1]. Further, the development of theoretical results in this direction published before 1995 are reflected in the monograph S. Borst, and the published papers during the years 1996-2006 are presented in the work [2].

Generalization and systematization of models and methods for studying stochastic systems with cyclic sounding (Polling system) and their use for the design of wireless broadband networks is dedicated monograph by Vishnevski and Semenova [2]. In this paper also are considered new models that describe the functioning of broadband wireless networks running under Wi-Fi and Wi-MAX protocols with a centralized control mechanism. The analysis of priority models with non-zero switching of semi-Markov type of priority queues, called generalized priority models is set forth in the monograph [3]. In this paper are presented new priority discipline, arising as a result of

formalization of lost time at switching, new analytical methods are developed and numerical algorithms for determination the characteristics of generalized systems are elaborated. It is showed the connection and continuity results for classical systems and generalized systems.

In the works of G. P. Klimov and G. K. Mishkoy [13], M. I. Volkovinskii and A. N. Kabalevskii a special case of polling systems, priority queuing systems with Poisson input flow and orientation time with a special form is studied. These monographs extend the results obtained in that period and are presented in the paper B. V. Gnedenko et al. [11].

One of the attempts, in a paper by S. Alfa [1], is to present queueing models in a clear aspect for the reader, on how to use available mathematical tools to assist in analyzing problems associated with queueing models. For this purpose, the content of this book is generally presented in a form that makes the applications to be easily seen and also because new developments and results from the area to be more accessible for the pragmatic user of the queueing models.

It is well known that queues with a single server give some very good prospects even to complex queues with more servers if they are properly approximated. So, waiting queues with a single server are very important in Queueing Theory and are much spread. For this reason, this paper focuses on queues with a single server. Waiting discrete-time networks are described in the monograph of H. Daduna.

An important area where Queueing Theory is applied very often is telecommunications sphere. Telecommunications systems are considered today as discrete time [2], as is generally they are based on discrete technology; the time is divided on intervals and the system was transferred from analogue technology (continuous time) to the discrete. Therefore, the discrete time of queueing models requires special attention in the queueing field and telecommunications.

Materials and methods of research

The main purpose of the study of Polling systems is to determine the probabilistic characteristics of system, such as busy period, queue length, workload, etc. But not always analytical formulas can be used directly to determine these characteristics, such great importance is given to development of new numerical methods and algorithms perform under these methods. For the analysis of Polling systems are proposed several methods, further it will be described briefly some of them.

Averages method. One of the methods proposed in scientific literature for the research of Polling systems is averages method. This method is widely described in paper by Winands et al and is designed to calculate the average length of queues from systems in an arbitrary time, for which can be obtained the averages durations of visiting the queues, in particular for cyclic survey system M/G/1 and exhaustive or closed service. Based on the average visit time of queue and the remaining average value is calculated the average number of requests in queues of the system as a solution of a system of linear equations. We will mention that averages method can be extended for the following Polling systems: Poisson grouped flow systems, periodically survey systems, discrete-time systems, also the application of this method to approximate analysis of other Polling models [7].

Method of catastrophes. A research method of Polling systems with a rich history of success in obtaining of new results in Queueing Theory, is the method of ”catastrophes” or method of introducing an additional random event [12]. The essence of the method of “catastrophes” consists in that by introducing an additional event (”catastrophe”) it is succeeds to assigned a clear probabilistic sense to Laplace and Laplace-Stieltjes transforms, after the evolution of the queueing system is analyzed and these probabilities are determined, this allows us to avoid some complicated structures.

Method of generating functions [12] is an effective method in researching problems from Queueing Theory. The main idea of this method consists in the assignment of probabilistic sense to generating function and this is achieved by coloring (marking, painting) procedure of requests that entry into the serving system. Thus, the abstract mathematical structure defined as generating function, due to its probabilistic sense becomes more comfortable and understandable in applied problems. Furthermore, due to this method, it is often possible to obtain analytical expressions for the generating function based on its probabilistic meaning. And this is possible without knowing the distribution function of the random variable. Just as it happens with the numerical values of the random variable like mean, variance, etc.

Method branched processes. Another method for the research of Polling systems based on theory branched processes is proposed in the work by van der R.D. Mei. [7].

This method applies to systems that operate in the conditions of high loads and enables the obtaining of approximate expressions for Laplace-Stieltjes transforms of queue length distribution and waiting times for broad classes of Polling systems whose behavior can be described by branching processes.

Polling systems analysis using **theory of decomposition of semi-regenerated processes** is detailed described in paper by V. Rycov and Gh. Mishkoy.

The survey of queues which is considered themselves periodically, it is made according to the given table survey. The Laplace transform of generating function of requests' number in waiting queues of the system is determined for different service disciplines (exhaustive, closed and limited).

Conditions for the existence of a steady state. In the work Z. Saffer is consider the system with periodically poll of the queues with input flows of BMAP type (grouped Markov flow of requests). For the system are obtained necessary and sufficient conditions for the existence of steady state and it is indicated the order of queues survey, for which the system does not have the working stationary regime.

After analyzing the specialty literature, especially of recent publications, we can see that the majority of probabilistic performance characteristics of Polling models and generalized priority systems are obtained in terms of functional equations, generating functions, Laplace and Laplace-Stieltjes transforms, etc. Thus, to determine the characteristics of these systems and such purpose as modeling of their evolution for analysis of practical problems requires the development of various methods, numerical algorithms and modelings. The purpose of these numerical methods and of developed numerical algorithms in their basis lies in resolving of those mathematical structures that do not have exact analytical solution.

Technological advances in various fields of contemporary practice lead to the formulation and study of new Polling models with characteristics that are not studied within the classics Polling models, for which are obtained and known some results today, these premises lead and will lead to a diversity of scientific challenges in the coming years. Among the remarkable and well known scientific work in this field are works of researchers from Republic of Moldova. Thus, we mention that Gh. Mishkoy and A. Bejan, proposed an improved algorithm for solving classical Kendall equation that can be used in multidimensional algorithms. Also, Gh. Mishkoy and O. Benderschi [3] have elaborated new methods, techniques and numerical algorithms of evaluation for coefficient traffic and others numerical characteristics for generalized priorities models. Gh. Mishkoy and D. Bejenari [5], have developed matrix methods and numerical algorithms of determining busy period for Polling models with semi-Markov switching and for generalized priority queueing models.

Results and considerations

Applications of Polling models. Polling systems find widespread applications in public health systems, in air and rail transportation as well as in communication systems. For this reason, studying them, which dates back to the late 1950s, it represents an enough important factor to economy of our country.

Polling model is a system with multiple queues with a single server which visits the queues according to the Polling table and serves customers from these queues. In addition, the Polling models have applications in situations where many users are competing for access to a shared resource which is available at a given time, such as communication systems, traffic and transportation systems, manufacturing systems, etc. Queueing systems are of interest, generally for two categories:

1) system users (customers);

2) service providers.

Customers want to use the system and to minimize the time of stay in the system, spending on services, the waiting time of serving start, delays, etc. On the other hand, the service provider wants to minimize the cost of providing service to customers, while ensuring that customers are "reasonably" satisfied. For example, the service providers do not want to assign too many servers, they do not want to provide a great length to free buffer space, which is often unoccupied, etc., and they do it without knowing exactly workload that will be sent by customers, because this component is stochastic. Sometimes, the used rule to operate with the system is a major factor in achieving the proper aims. For example, just by changing the operating rule from the first come-first served on a priority system it will change how the system is perceived by customers. So there is a need to understand how to operate with a queueing system while achieving at the same time all of these conflict goals caused both of random arrival of entities and the random duration of arrivals. The question arises whether in general it can be studied real systems where this random double phenomenon occurs. The answer is affirmative and it is well known. The research process consists of two stages: the first stage consists of elaborating mathematical model of the real system; the second stage is the research of developed model, to establish certain regularities, determine certain characteristics, etc.

Polling models find a wide range of applications in communication systems, where resources (e.g., bandwidth, CPU capacity) are divided among different users. Works that mentions applications of Polling models until the early 1990s, we can refer to Grillo's surveys, Levy and Sidi and Takagi [9]. For generalization, the main applications mentioned in this work are described below and completed with applications of Polling models in latest fields of communication systems.

Time-sharing computer systems. Classic applications of Polling models represent timesharing systems, these consist of a number of terminals connected via multi-drop lines to a central computer. The transfer of data from terminals to computer and vice versa is controlled by a Polling system where the computer interrogates terminals, requesting their details, a terminal at a given time.

In such applications of Polling models, the server is the central computer, queues represents the terminals and customers represents the data. In communication networks, in case that several terminals transmit or receive data from the environment at the same time, packet conflictss may occur and interference problems. For these reasons, many protocols such as medium access control (MAC) have been proposed for various network technologies, in most cases this conducts to the Polling systems.

Token-ring networks. W. Bux takes Polling models to analyze the performance schemes of crossing token in local area networks (LANs), where a token is the transmitting right and it is

transmitted through different users. In these situations, token passing scheme is habitually configured in a ring or bus topology. A token ring network can be described as a set of stations connected to a collective transmission medium in a ring topology. All messages traveling on a fixed route from a station to another station around a loop. A token can obtain one of the situations: busy or free. A station with data transmission reads the free token and modifies the state occupied before retransmits it. The busy token is then embedded as part of the header data which is transmitted from the ring to station. Therefore, other stations from the ring can view and read the header, remark the busy token and refrain from the transmitting. When the token returns to the station that it has changed into busy status and station decided to change the transmission right to another station, it modifies the status of the token in free. Token-ring network permits packet forwarding in a way without conflict. However, packet forwarding may still fail because of faults and distortions of the ring itself. In the scheme of Selective Repeat ARQ (SR-ARQ), a station which gets a negative acknowledgment send a wrong message to the station that transmitted information to indicate it that the message should be sent again. To evaluate the performance of SR-ARQ schemes, Levy and Sidi uses a Polling model, where each station consists of two queues, one is used for messages which need to be sent and one for negative acknowledgment which must be sent back in case that erroneous messages are get [8].

Token-bus networks. The token bus network is formed from a set of stations interconnected in a bus topology. The main idea of this technique is to combine the attractive characteristics of the bus topology of a medium access protocol without conflict. In a token-bus network, a logical ring is formed since the token is sent. Since the bus topology does not require any sequential ordering stations, logical ring is defined by a sequence of station addresses. The difference between a token ring and a token bus, from the viewpoint of modeling, is that in token ring network, the server visits queues in a cyclic order, while in the token bus model the server travels along the queues in a regular non-cyclic order, which can be modeled using Polling table.

Slotted-ring networks. Slotted-ring network represents another class of communication systems for networks with a ring topology network. This network function as follows, one or more slots move along the station. If there is a package ready to send to a station and a free slot comes, then the packet is placed in the slot with the address of destination station. That slot is then examined by each of the other stations in turn, until it is recognized by destination station. There are two ways to release the slot: either it is emptied by the source station or by the destination station. We make reference to the work of W. Bux, which describes a Polling system with source release and the work of van Arem which refers to a Polling system of a slotted-ring protocol, issuing destination.

Distributed Queue Dual Bus networks (DQDBN). The DQDB protocol is a protocol for multiple access communication networks, which consists of two unidirectional buses that carry information in opposite ways. The used Polling model for studying medium access of the mechanism with a station in the DQDB network is presented in the work by Bisdikian.

Optical networks. In the field of Ethernet Passive Optical Networks (EPONs) also Polling systems find applications. In such a network, packets from various Optical Network Units split the channel capacity. The OLT (Optical Line Terminal) is in local office interlocking the access network to the Internet system. The OLT allots the bandwidth to the Optical Network Units (ONUs), which are located at the client premises, providing interfaces between the OLT and the end user network to send different information. Simple time division multiple access (TDMA) schemes grounded on fixed allocation of a timeslot suffer from the lack of multiplexing statistics, making inefficient use of disposable bandwidth, which increases the necessity for dynamic bandwidth allocation (DBA) schemes. A dynamic system that reduces the size of a time slot when no data to

be transmitted would permit excess of bandwidth to be used by other ONUs. Kramer et all, propose an OLT based dithered Polling model similar to hub polling to support dynamic bandwidth allocation.

Bluetooth. Bluetooth is a wireless technology standard for the exchange of data between different mobile apparatus such as mobile phones, notebooks and headphones. These apparatus constitutes small networks known as Wireless Personal Area Networks (WPANs). Generally, the topology of Bluetooth network is named piconet, and Miorandi et all, observed that the structure of a piconet that has X slaves, can be represented using a Polling system, which also consists from $2X$ queues. Approximations of average delay are determined for 1-limited service, dependent and exhaustive disciplines [2].

Mobile networks. In mobile networks area the Polling models also can find applications. In such a scheme, the users compete for access because the shared radio resources are limited. In such situations, the base station is usually responsible for the allocation of time slots to various users in a certain way. In this context, the server is the right to transmit information and on the other hand, the clients are data packets that should be sent. Analogical examples of Polling models can be found in the context of Code Division Multiple Access (CDMA) based on High Speed Packet Access (HSPA), where the base station controller provides access to the medium in the base of per-timeslot. An implementation example is a simple Round-Robin (RR) scheduling, in which the medium access is transmitted through the terminals, regardless of the signal quality (see paper by van den J. Berg). This leads immediately to Polling systems where the service type is limited and the server serves the queues in a cyclic order. Another example can be the Weighted Round-Robin (WRR), which can be modeled as a Polling system where the server crosses the queues in periodic route. To increase the effectiveness of medium access HSPA networks, were proposed programming mechanism with open channels highly sophisticated, based on instantaneous measurement Signal to Noise Ratios (SINRs) at every of terminals. Due to aleatory in channel terms, and therefore the order in which the stations are allowed to the medium, definitely, conduct to modeling of Polling systems with server routing of random or Markov type [8].

Ferry-based Wireless LANs (FWLANs). As well, Polling systems have applications in the designing of message ferry routes in Ferry based LANs. In such FWLANs a number of isolated nodes are scattered over some geographical field where communication between a node and the outside world, or communication between nodes is possible via a ferry message. The ferry follows a path of predetermined cycle, collecting messages from and delivering messages to nodes whenever it is in the neighborhood of the node [6]. Considering the route of ferry that is of cyclic type, one can design an optimal route using Polling models. Results regarding FWLANs are presented on the work by V. Kavitha and E. Altman.

Mobile adhoc networks (MANETs). In the designing of mobile ad-hoc networks also find place Polling systems, which consist of both fixed wireless terminals and mobile. A typical characteristic of this type of networks is that wireless apparatus create their own wireless network in a distribution. Mobile phone users can change the location and therefore changes the communication connections in the network. An summary of applications of Polling models in MANETs is given in paper by de R. Haan.

Networks on chips (NOCs). Another interesting application field of Polling models is networks based on chips, that have appeared as a solution to the inefficiency caused by traditional bus links [15]. We will make reference to the paper of P. Beekhuizen, which is an overview of the applicability of Polling models for networks based on chips, and to the work by van B. Houdt, where it is presented an effectively numerical algorithm.

Conclusions

The extension of applications of the results for Polling models, which are known today, is limited by various factors such as: the study of models of certain distribution functions, although in the majority of the real processes they involve various situations, thus at an adequate modeling of a real process is necessary a wider range of distribution functions; the need of new methods for theoretical analysis of the queues, so the development of flexible and adequate methods would increase the applicability of theoretical results of Polling models; the assumption of having one server that serves several queues is also a major limitation to many real applications. As it was mentioned above, there are only few results for Polling models that are composed of multiple servers.

In the field of mobile and wireless networks, a future perspective is to increase bandwidth, thus to enhance the robustness of the quality of service provided to the user, we must make use of multiple antennas, allowing users to use simultaneous multiple networks (e.g., MIMO). In such medium, packet flows can be divided into multiple parallel networks, where each of them can be modeled as a Polling model, where the route package may depend on the current status of these networks. Studying and getting new Polling models results will continue to present a challenging research area for the coming years.

Bibliography

1. Alfa S. A. Queueing Theory for Telecommunications. London: Springer, 2010. 238p.
2. Bejenari D., Mitev L. Formule aproximative pentru sisteme Polling cu timp discret. In: Revista Ştiinţifică Studii Economice, 2012, vol. VI, nr. 3-4, p. 326-331.
3. Benderschi O. Heavy traffic analysis in queueing systems. În: Studia Universitatis, Seria Științe Exacte și Economice. Chișinău: USM, 2008, nr. 8(18), p. 38-48.
4. Benini L., De Micheli G. Networks on chips: a new SoC paradigm. In: Computer, 2002, vol. 35(1), p. 70-78.
5. Mishkoy Gh., Krieger Udo R., Bejenari D. Matrix algorithm for Polling models with PH distribution. În: Buletinul Academiei de Ştiinţe a Republicii Moldova, Matematica. Chişinău: IMI, 2012, nr.1, p. 70-80.
6. R.D. van der Mei, Hariharan R., Reeser P. K. Web server performance modeling. In: Telecommunication Systems, 2001, vol. 16, p. 361-378.
7. R.D. van der Mei. Winands E. Heavy traffic analysis of polling models by mean value analysis. In: Performance Evaluation, 2008, vol. 65, nr. 6-7, p. 400-416.
8. Sevcik K. C., Johnson M. J. Cycle time properties of the FDDI token ring protocol. In: IEEE Transactions on Software Engineering, 1987, vol. 13, p. 376-385.
9. Takagi H. Analysis and application of polling models. In: Performance Evaluation: Origins and Directions of Lecture Notes in Computer Science. Berlin: Springer, 2000, vol. 1769, p. 424-442.
10. Вишневский В.М., Семенова О.В. Математические методы исследования систем поллинга. В: Автоматика и телемеханика, 2006, № 2, с. 3-56.
11. Гнеденко Б.В. и др. Приоритетные системы обслуживания. Москва: Изд-во Московского Университета, 1973. 447с.
12. Климов Г. П. Стохастические системы обслуживания. Второе издание, Москва, 2011. 242с.
13. Климов Г.П., Мишкой Г.К. Приоритетные системы обслуживания с ориентацией. Москва: Изд-во Московского Университета, 1979. 128с.
14. Мишкой Г.К. Обобщенные приоритетные системы. Кишинев: Из-во. А. Н. Молдовы, 2009. 200с.