MATTEX 2018, **PLENARY LECTURES**
CONFERENCE PROCEEDING, v. 1, pp. 3 - 11

A NEW NETWORK SECURITY METHOD BASED ON THE STEGANOGRAPHIC DISPERSION

JĘDRZEJ BIENIASZ, KRZYSZTOF SZCZYPIORSKI
INSTITUTE OF TELECOMMUNICATIONS,
WARSAW UNIVERSITY OF TECHNOLOGY, POLAND

ABSTRACT: *The fog computing has emerged as the extension of cloud computing to the network edge. The idea could be considered as a promising mechanism for cybersecurity by assuming that higher uncertainty of information from perspective of adversaries would improve the security of networks and data. This approach is recognized as cyberfog security in which data, split into fragments, is dispersed across multiple end-user devices. Even if some of them would be compromised, the adversary could not decode information and the availability of data would not be affected. This paper considers applying two steganographic proposals (StegHash and SocialStegDisc) for a new distributed communication system by fulfilling assumptions of cyberfog security approach. The initial design of such system is proposed. Features and limitations were analyzed to prepare recommendations for further development and research.*

KEYWORDS: *fog computing, information hiding, steganography, distributed filesystems, StegHash, SocialStegDisc, cyber-physical systems*

1 Introduction

Cloud computing (centralized model) is extensively developed and widely applied in last years. Parallelly, Internet of Things with deployment of smart and interconnected devices, for example mobile phones, wearables, sensors, power grids etc. could overgrow 50 billion units by 2020 [1]. It triggers defining a new paradigm of distributed computing called fog computing [2] that complement the centralized cloud computing. National Institute of Standards and Technology (NIST) went further with unification of the description of computing systems. They defined a whole new framework called cyber-physical systems (CPS) [3] and IoT could be considered as the implementation of CPS. In [4] authors consider the fog computing as a network and data security mechanism. Cyberfog security approach means that higher dispersion of data supports greater attack resiliency. An adversary, which compromises the system, could take only a part of the system and it would be useless.

The authors noticed a possible connection of their current research with this topic. Recently, they revisited the idea of text steganography [5] in combination with social networks. The original idea of StegHash [6] method is to use multimedia objects as carriers of hidden information and to disperse them across open social networks (OSNs). A logical connection between them is established by means of a mechanism of hashtags as the text markers in the form of #<tag>, commonly applied in OSNs. In addition, the author of StegHash proposes an option of applying the StegHash technique to create a steganographic filesystem analogous to the existing classic filesystems such as FAT (File Allocation Table) or NTFS (New Technology File System). Research on that option was furtherly conducted in [7] to follow the pattern of researching steganography by development of applications preceded by the appearance of new steganographic technique. SocialStegDisc proposed the application of basic concepts of classic filesystems, such as create-read-update-delete operations or defragmentation process to the original idea of StegHash. Furthermore, time-memory trade-offs were proposed by the design.

This paper proposes to apply methods [6][7] directly to design a new kind of distributed communication system for cyberfog security approach. The design utilizes the indexing scheme introduced by StegHash [6] and provides the basic operations like SocialStegDisc [7]. Furthermore, the concept expands the ideas beyond the original ones by using end-user devices' memory as a carrier for objects with hidden data. The methods of routing and finding next parts physically in such peer-to-peer network is addressed. The security is majorly ensured by the character of logical connection between the distributed parts of data. The system could only be properly compromised when the adversary takes over the secret generation function. Without it, the captured parts, e.g. by compromising one or part of devices would be unusable what fulfill the basic cyberfog cybersecurity approach requirement.

2 Review of literature

Cyberfog security approach was introduced by Kott et al. in [4]. They proposed to use the fog computing [2] as a method for mitigating a cyber adversary. It assumes that presenting adversaries with uncertainty could provide greater attack resiliency. A direct realization of this approach is splitting data into numerous fragments and dispersing them across multiple end-user devices. Even if the system could be partly compromised, captured information would be useless for adversary, while still being useful to us. Authors recognized numerous benefits, but also formidable challenges with respect to data and network management complexity, bandwidth, storage, battery-power demands, data-reassembly latency and intermittent connectivity. They see that the network might need to manage a complex tradeoff between availability and confidentiality in real time depending on users' tasks and circumstances.

Two models of communication: high-entropy and low-entropy was presented by Beato et. al in [8]. In high entropy model the steganogram is transported by a single multimedia object such as pictures, video or music. This is considered as a classic method of steganographic communication. This model is characterized by high steganographic throughput, but the channel is easily detectable. Low-entropy model utilizes text data (e.g., status update, group text message) to carry secret information. To determine the steganogram location a pre-shared secret is used to decode the actual message. This could be used mainly for signaling due to the low steganographic throughput.

It should be noticed another low-entropy steganographic method proposed by Castiglione et al. in [9]. They took an advantage of the feature of inserting tags in images. The proposed stealth communication channel requires the uploading of multiple images and to tag multiple users.

3 Concept of the system

The proposed method of StegHash [6] is based on the use of hashtags on various open social networks to connect multimedia files, like images, movies or music, with embedded hidden data. For every set of hashtags containing n elements there is the factorial of n permutations, which are individual indexes of each message. With a secret value (password) and a secret transition generator (function), the link between these indexes could be established and then explored as a chain from one message to another, with each containing hidden content. The example of the chain is presented in the Figure 1. It establishes a new paradigm of unlimited data space, but limited address space. SocialStegDisc [7] is an application of StegHash which introduced filesystem construction on the top of StegHash chaining mechanism. Furthermore, time-memory trade-offs were proposed. Increasing the number n of hashtags is followed by the increasing volume of the dictionary proportionately to n! what for higher n would be unacceptable. Proposed mechanism of the linked list were taken directly from filesystem theory to introduce basic operations - read, write, update and delete – without comprising the level of security. To sum up main assumptions:

- An unique permutation of the set of hashtags is an index on data fragment;
- The set of *n* hashtags means *n*! of unique indexes to generate;
- There could be more than one service, where data is uploaded, so the index system for services need to be designed. For StegHash and SocialStegDisc was to
- use each of *n* hashtags on the last position as a key for one of *n* services.
- The used pseudorandom generation function should produce the same chain of indexes with the same seed;

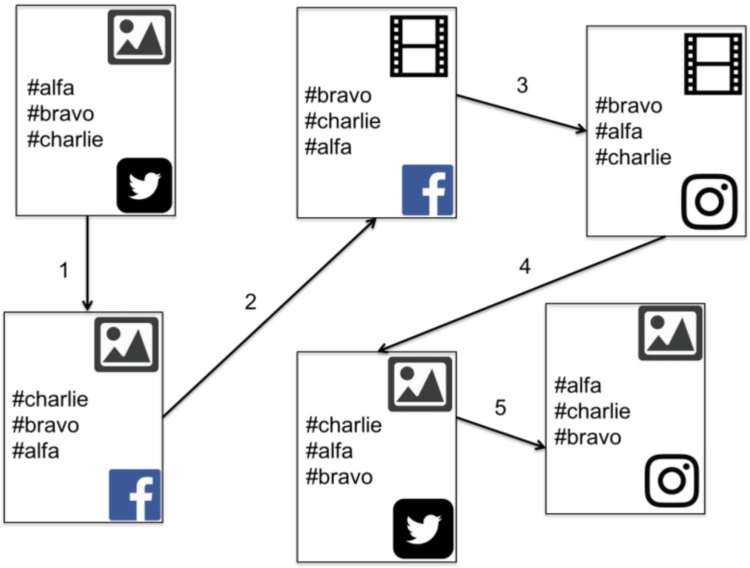

Figure 1 Example of StegHash method [6]

We need to identify how well-defined and examined concepts of StegHash and SocialStegDisc fit into a domain of fog of the devices. Firstly, there are no public and open internet services like open social networks. Instead, everything will operate between a defined set of end users' devices. The next main difference is a size of a problem. Earlier, the *n* was from a few to a dozen or so. Now, this number increases many times, so indexing the services for uploading data by one of *n* hashtags in an index is impossible. So, we propose to apply this mechanism in another setting:

- *n* devices would create a *memory sector* of size *n*. For such memory sector the mechanism of addressing the carrier of data (end-user device) by one of *n* hashtags is still applicable inside this memory sector.
- System needs a new layer for the communication on the level of the memory sectors;

Basing on this, there would be many memory sectors available in the system. Every device is locally associated with a particular memory sector. It also means that every device have its own instance of StegHash/SocialStegDisc as a memory controller to serve all the basic filesystem operations inside the memory sector. It should be noticed that the design could be scaled to support multiple membership of devices in memory sectors.

From the user perspective, sending data between devices or generally saving the data in the system is realized by using associated StegHash/SocialStegDisc controller to load data to the local memory sector by dispersing fragments between *n* other devices. Every device needs to have a synchronized copy of an allocation table of the sector. Figure 2 presents the conceptual model of applying StegHash/SocialStegDisc for the fog architecture.

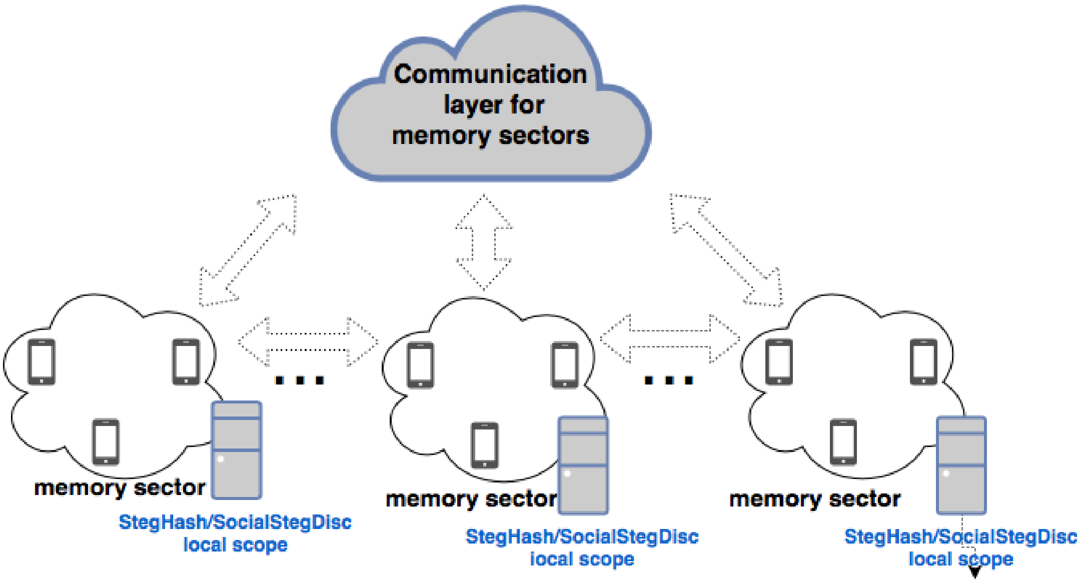

Figure 2 Applying StegHash/SocialStegDisc into the fog architecture

To complete the design of the system we needed to design a communication scheme between the memory sectors. We evaluated two ideas which are presented in this paper.

The first idea is to use a device as the memory sector's gateway. In this scheme when a device X saves the data inside its memory sector, it would inform other devices that starting from the index I, there are B bytes to read and the device X is a gateway to read. This kind of scheme is combining the classic methods of routing (networking) and the file allocating methods (filesystem theory). Other devices would have a view of allocated data and they would know where to send read requests to download data. This approach requires a synchronization of allocated addresses across the memory sectors, but the state of generation function is needed to be synchronized only across the memory sector.

Another idea is to read directly from the memory sector. This is about not using a one of memory sector's devices as a gateway. Instead, the routing information for uploaded data requires including all n devices keys and the coding scheme of how to choose the next device for reading data. With this information, the device could read a requested file directly from the memory sector. There are two options of sharing the routing records:

- *Proactive:* sharing the routing records with every single upload to the system and caching it by every device;
- *Reactive:* sharing only metadata of the file and gateway with every single upload to the system and caching it by every device. When the data is wanted to be retrieved, at first the request of memory scheme is issued to the gateway. The gateway responses back with all n devices keys and coding scheme. Next, the device could read a requested file directly from the memory sector.

The layer of network communication and reachability between devices is needed to be considered by design. This consists of two functions:

- *Streaming data end-to-end*. Providing communication by TCP/IP transport layer protocol – TCP or UDP – is an obvious choice. Other protocols from upper levels could be also used.

- *Reachability between devices.* Every device needs to know the network addresses of other devices in network. It could be achieved through static routing configuration or through dynamic routing protocols.

To serve the described communication layer we propose to use the TrustMAS platform [10]. This system would provide all required features, but in a steganographic manner. Every device would have use a StegAgent instance introduced by TrustMAS design to support covert streaming of system data. It would serve as a network communication endpoint for StegHash/SocialStegDisc memory controller. Figure 3 presents the scheme of the layer architecture of the platform.

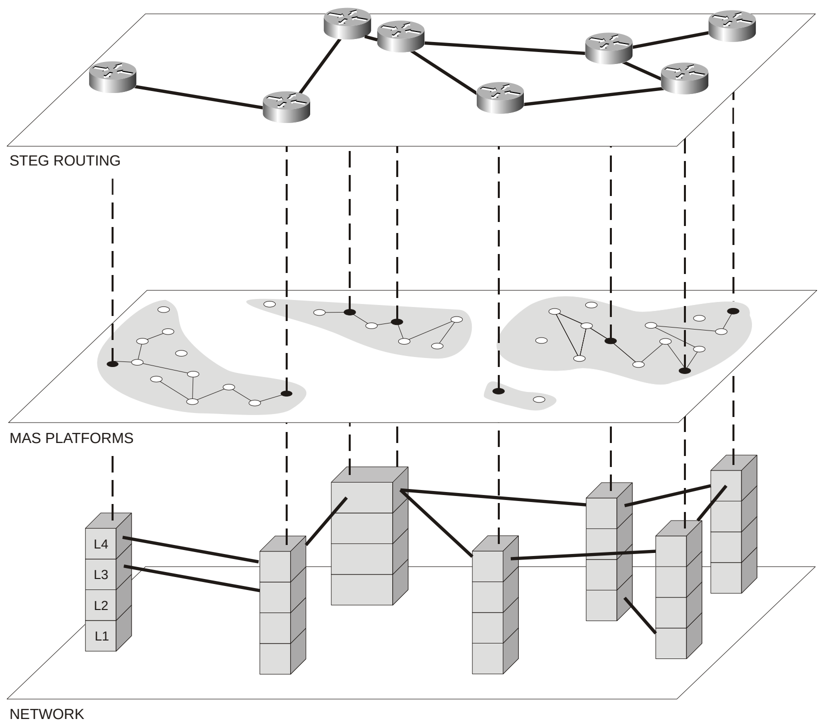

Figure 3 Three-layer architecture of TrustMAS [10]

Until this paragraph we introduced a new area of application for StegHash [6] and SocialStegDisc [7] which is a distributed communication system between end-user devices. We defined the memory sectors and how to manage them by StegHash/SocialStegDisc controller, the communication layer between them and we chose a platform for device-to-device communication – TrustMAS [10]. Figure 4 provides the conceptual scheme of components inside a particular end-user device and interaction with user, physical memory and another device. On Figure 5. a view of the system by a stack of operational layers is presented. Every layer provides the set of operations logically connected:

• *Application layer* – an application which uses the system from the perspective of user;

• *Filesystem/messaging service* – provides the interface for application layer with read-write-delete operations inside device and with managing point-to-point connection sessions; it passes commands to lower layers;

• *StegHash/SocialStegDisc* controller – the implementation of StegHash [6] indexing method and SocialStegDisc [7] memory management;

• *Memory sector and its controller* – defining set of operations on the memory sector as a group of n devices. It provides the view of the memory sector for StegHash/SocialStegDisc controller and manages the state of the memory sector. Furthermore, they cooperatively manage the scheme of allocation of the memory over devices, memory sectors and the whole storage system.

• *Physical memory* – a non-volatile memory of the device, where data is stored;

• *StegAgent* – treated as a network service introduced by TrustMAS [10]. It offers a network communication and reachability of devices.

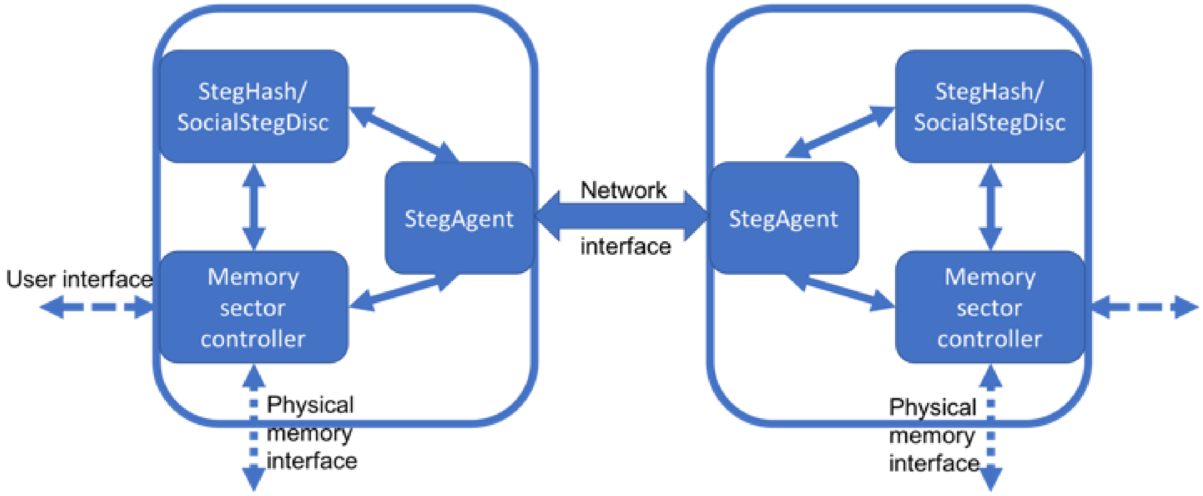

Figure 4 Components of the system and their interaction scheme

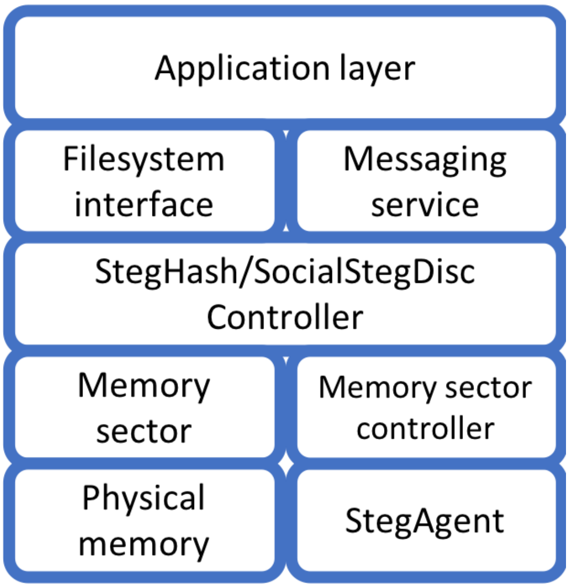

Figure 5 The layer view of the system

4 Discussion

The concept expands the idea beyond the original ones [6][7] by:

- Using end-user devices' memory as a carrier for objects with hidden data;
- Using any type of object as a carrier of hidden data;

- Data could be also written directly into end-user device memory with embedding hiddenly into a carrier;

The design introduced the concept of the memory sector created from n devices. For such memory sector the mechanism of addressing the carrier of data (end-user device) by one of n hashtags is still applicable inside this memory sector. This idea was needed to solve the incompatibility between number of a few services in StegHash/SocialStegDisc [6][7] and hundreds of devices in the fog architecture. The trade-off is a need of a new layer for the communication on the level of the memory sectors and for managing the allocation procedures over them. To support these operations, a new type of a distributed filesystem would be explored, but this is planned for the future extension on that topic by authors. The security of the design is provided from two perspectives:

• The fog of devices and dispersion of data between them supported by a type of a logical chain created by StegHash [6] indexing concept.

• Security of a data transmissions provided by the StegAgent of TrustMAS [10] platform;

For the fog architecture, security is majorly ensured by the distribution of data and by the character of logical connection between the parts. The system could only be compromised when the adversary takes over the generation function. The adversary could sniff the network or compromise one or a part of devices, but the captured parts of data are unusable without knowledge about logical connection between them. It could look like a chaotic set of bytes. The authorized operators of the system could still retrieve data properly as they have knowledge of the correct connection chain among parts of data. Only compromising the generation function module or taking over a unit of it is the strongest threat, so any mechanism for triggering the automative destruction would be taken under consideration.

Another level of security is provided by application of parts of TrustMAS [10] platform. For the design proposed in this paper, we would apply a trust management system for communicating agents and agents communicating through covert channels. Main trust model proposed for TrustMAS is based on following the behavior of agents. If they realize an expected scenario and the protocol of communication is correctly executed that means that agents are trusted. TrustMAS utilizes cross-layer steganography what means a capability of using many different types of steganography like network steganography in every TCP/IP layer or application layer steganography through images, videos or text hiding. Furthermore, between two StegAgents a communication path could be created from links with other methods of steganography used by every of them. It has an advantage of adapting the exchanging hidden data what is harder to uncover.

There is one more feature of TrustMAS [10] which is an anonymous technique based on the random-walk algorithm. The message is forwarded by an agent to next in a probabilistic manner, so any other agent cannot conclude about originator of the message. We see the possibility of including such mechanism in the design of the system, but we omit it now to address it in further research papers. We would consider other anonymity algorithms.

It should be mentioned following Kott et al. in [4] that in a cyberfog security approach, the network might need to manage a complex tradeoff between availability and confidentiality in real time depending on users' tasks and circumstances. Going through the proposed design we see that the every next mechanism of security means adding a greater level of complexity. In Table 1 we summed up if the mentioned mechanisms and algorithms affect one of the aspects: memory, time and reliability. We marked the positive effect on the aspect by "+" and the negative effect by "-". "+/-" means that there is no evidence or cause to mark them positive nor negative.

Table 1 Impact of the design on the memory, time and the reliability of the system

Design components	Aspect			
	Memory	*Time*	*Reliability*	*Security*
StegHash indexing [6]	+/-	-	+/-	+
SocialStegDisc operations and improvements [7]	+	-	-	+
Memory sector	-	-	+/-	+/-
StegAgent [10] covert communication	-	-	-	+
StegAgent [10] steganographic routing	+/-	-	-	+
Storage system	+/-	-	+/-	++
Messaging service	+/-	-	+/-	++

It is noticeable that all components impose a time penalty on the system. It is caused majorly by the fact that many mechanisms need to be executed in the order and time of computation is obviously higher.

In this paper we focused on the basic aspects such as memory, time and reliability. In the future work, we will evaluate executing the design in the limited environment of Internet of Things' devices, where the consumed energy, available memory and computing power. They are decisive features impacting the successfulness of the implementation of IoT solutions.

5 Summary

In this work a new concept of the communicating system realizing the idea of cyberfog security [4] was presented. The design combines and adapts a few components such as StegHash [6] for indexing data, SocialStegDisc [7] for filesystem operations and TrustMAS [10] for device-to-device data transmission. As they are adapted for the environment of the fog of devices, they establish a new kind of a secure communication platform. Security is characterized by the fact that the partly compromising of the system does not interfere the operations, whereas captured samples are useless for the adversary.

In the future work we would deepen the design of the system by considering more mechanisms for operations of the platform and for achieving higher level of security. A proof-of-concept implementation would be prepared to deliver results from the real working example of the system.

REFERENCES:

[1] D. Evans, "The Internet of Things. How the Next Evolution of the Internet Is Changing Everything". Cisco Internet Business Solutions Group (IBSG). 1. 1-11.

[2] M. Chiang and T. Zhang, "Fog and IoT: An Overview of Research Opportunities," in IEEE Internet of Things Journal, vol. 3, no. 6, pp. 854-864, Dec. 2016.

[3] NIST SP 1500-201, Edward R. Griffor, Christopher Greer, David A. Wollman, Martin J. Burns (June 2017), Framework for Cyber-Physical Systems: Volume 1, Overview, https://dx.doi.org/10.6028/NIST.SP.1500-201

[4] A. Kott, A. Swami and B. J. West, "The Fog of War in Cyberspace," in Computer, vol. 49, no. 11, pp. 84-87, Nov. 2016.

[5] M. Chapman, G. Davida, and M. Rennhard, "A Practical and Effective Approach to Large-Scale Automated Linguistic Steganography", Proceedings of the Information Security Conference, October 2001, pp. 156-165.

[6] K. Szczypiorski. 2016. “StegHash: New Method for Information Hiding in Open Social Networks”. IJET International Journal of Electronics and Telecommunication, 62 (4): 347–352.

[7] J. Bieniasz, K. Szczypioski: “SocialStegDisc: Application of steganography in social networks to create a file system”, In Proc. of 3rd International Conference on Frontiers of Signal Processing (ICFSP 2017), Paris, France, 6-8 September 2017

[8] F. Beato, E. De Cristofaro and K. B. Rasmussen, "Undetectable communication: The Online Social Networks case," Privacy, Security and Trust (PST), 2014 Twelfth Annual International Conference on, Toronto, ON, 2014, pp. 19-26.

[9] A. Castiglione, B. D'Alessio and A. De Santis, "Steganography and Secure Communication on Online Social Networks and Online Photo Sharing," Broadband and Wireless Computing, Communication and Applications (BWCCA), 2011 International Conference on, Barcelona, 2011, pp. 363-368.

[10] K. Szczypiorski, I. Margasiński, W. Mazurczyk, K. Cabaj, and P. Radziszewski, “TrustMAS: Trusted Communication Platform for Multi-Agent Systems”, In Proceedings of the OTM 2008 Confederated International Conferences, CoopIS, DOA, GADA, IS, and ODBASE 2008. Part II on On the Move to Meaningful Internet Systems (OTM '08), Robert Meersman and Zahir Tari (Eds.). Springer-Verlag, Berlin, Heidelberg, pp. 1019-1035.

Krzysztof Szczypiorski
Institute of Telecommunications,
Warsaw University of Technology, Poland
E-mails: K.Szczypiorski@tele.pw.edu.pl

Jędrzej Bieniasz
Institute of Telecommunications,
Warsaw University of Technology, Poland
E-mails: J.Bieniasz@tele.pw.edu.pl

QUEUEING SYSTEMS WITH LIMITED ACCESS TO SERVICE STATION

WOJCIECH M. KEMPA

SILESIAN UNIVERSITY OF TECHNOLOGY,
FACULTY OF APPLIED MATHEMATICS,
INSTITUTE OF MATHEMATICS, GLIWICE, POLAND

ABSTRACT: *Queueing models with different-type limitations in the access to the service station are discussed. Such systems are used in modelling of different-type phenomena occurring in telecommunication and computer networks, production management, transport and logistics. Exemplary models are described and characterized by analytical results. In particular, results for queueing systems with single and multiple server vacation policies and the threshold-type N-policy are given. Moreover, the Active Queue Management (AQM) is described as the set of techniques for reducing the risk of buffer overflows.*

KEYWORDS: *Active Queue Management (AQM), N-policy, queueing system, single/multiple vacation policy.*

1 Introduction

As it seems, queueing theory is nowadays one of the most rapidly developing scientific disciplines of the applied probability area. Fast technological progress in the design of computer and telecommunication networks, in particular relating to different aspects of wireless communication, causes that the range of problems and phenomena which can be investigated using the methods of queueing theory is essentially increasing. As one can observe, many analytical (theoretical) results in queueing theory are obtained, as one can say, "on the occassion" in the process of finding solutions of many practical problems under different-type grants executed by consortia including scientific and industrial individuals. Indeed, queueing theory is particularly applicable in the process of development and performance evaluation of packet-oriented computer networks. In this case the operation of a node of the network (like e.g. IP router) can be modelled by using an appropriate queueing system. Indeed, Leonard Kleinrock, who is the author of classical handbooks in queueing theory, can be deservedly called "a father of the Internet". In his laboratory, in Boelter Hall at University of California Los Angeles, one of two first nodes of the ARPANET network was built (this network can be named the "ancestor" of todays Internet). Obviously, applications of queueing theory reach much deeper and concern the production management and organization, transport problems, logistics etc.

Queueing theory, which was primarily considered as a branch of applied probability, now is in fact an independent scientific discipline. The result of its dynamic development and increasing applicability in solving different-type real-life problems is a huge number of textbooks, monographs, journal articles and scientific conferences devoted to theoretical aspects and practical applications of queueing theory. Among many items, let us mention about classical theoretical works by Bocharov et al. [2], Borovkov [3], Cohen [7], Kleinrock [19], Takács [24] or Takagi [25].

2 Models with service limitations

Queueing models with different-type limitations in the access to the service station are being intensively studied currently. Indeed, in practice one can very often meet a situation in which the service process for some reason is suspended or delayed. The limitation in free access to the service can be associated with the temporary deactivation of the service station (server vacation), during which the service process is completely blocked, or may be connected with the imposibility for a customer (a job, a packet etc.) to join the waiting room (a buffer) due to its finite capacity or some other mechanisms "filtering" the input flow. A detailed discussion on modelling of different-type practical issues leading to queues with limited access to the service station can be found e.g. in an interesting survey by Doshi, 1986 [8].

Indeed, the following practical situations and motivations can be mentioned here:

- machine failures – they may occur randomly, independently on the status of the queue and the number of customers present. Repair times can be considered as server vacations, during which the processing is completely blocked;
- maintenances – every time when the machine processing jobs becomes idle, it undergoes preventive maintenance of random (usually) duration. For example, processors in computer and communication systems execute testing and maintenance besides executing their primary tasks (like processing jobs, receiving and transmitting data etc.). The maintenance and testing is needed to preserve high quality of service and reliability of the system mainly.
- energy saving – a typical problem e.g. in production lines and wireless network communication. In this case each busy period of the appropriate queueing system may be considered as an active period (active mode) e.g. in the functioning of the wireless network node (e.g. wireless sensor network) while, similarly, each idle time can be treated as a power saving period (sleep mode).

In the literature different-types mechanisms of limitation of the access to the service station are proposed. The most important solutions are the following ones:

- single/multiple vacation policy (SV/MV) – in the case of SV policy, every time when the system becomes idle, the service station takes on a (single) vacation (usually of random duration) during which the processing of jobs is suspended. After closing the vacation, if there are jobs waiting for service, the processing restarts immediately; otherwise, the server waits in readiness for the first arrival. In the MV policy successive independent vacations are taken on as far as at least one job waiting for service is detected at the completion epoch of one of them;
- N-policy – after the idle time the processing (a new busy period) is being initialized simultaneously with arriving of the Nth job;
- T-policy – in this case the server is turned on only for a time of a fixed length T every time when at least one job waits for service (this solution can be used in modelling of TDM (Time Division Multiplexing));
- setup/closedown times – may occur if the server is switched off being idle: after the last service the service station needs some time to be deactivated safely; similarly, the first processing in a new busy period is preceded by a setup time during which the server becomes

full ready for job processing. During setup and closedown times the service process is completely blocked;

- server breakdowns (failures) – occur randomly if the service station is busy with processing. After the breakdown occurrence the server needs some time (repair period) to restart the service process;
- gated service discipline – in this mechanism a single processing period is divided into two subperiods. During the first one the arriving jobs are allowed to join the server; at the beginning of the second service subperiod, the service "gate" is being closed up and all incoming jobs must accumulate in the buffer and wait for the opening of the "gate". The gated service discipline is usually used in modelling of the operation of transportation networks (buses, trains, ferries etc.). However, such a phenomenon may also occur in parallel computing and the Internet traffic.

Evidently, due to vast literature which is still growing, it is impossible to present and discuss main results for each of the above mentioned queueing models. In the next sections we present results for some chosen models, both for transient and stationary stochastic characteristics.

3 Single and multiple vacation policies

In these two queueing models the limitation in the access to the service station is connected with the so called "server vacation". More precisely speaking, every time when the server becomes idle it initializes a vacation period during which the processing of accumulated and incoming jobs is completely blocked. In the single vacation policy the service station takes exactly one vacation (usually of random duration) at the end of each busy period. If, at the end of the vacation, the system is still empty, the server waits for the first arrival and the processing starts immediately. Otherwise, the completion epoch of the vacation coincides with the start of the service process. In the multiple vacation policy the service station keeps on taking vacations until, returning from a vacation, at least one customer is accumulated in the buffer queue and waits for service.

Miller in 1964 [23] was the first who investigated a queueing model of the $M/G/1$ type in that the service station is unavailable during some random time (referred to as vacation). Next Levy and Yechiali (1975, [21]) introduced several generalized models of the classical $M/G/1$ queue with temporarily unavailable server.

Queueing models with different-type vacation policies are nowadays intensively studied due to their applications e.g. in modelling of different phenomena occurring in telecommunications and computer networks. In particular, vacation queueing models are used in modelling and perfomance evaluation of the energy saving periods (sleep modes) occuring in wireless telecommunication (like LTE or WiMAX 802.16e). In a wireless network users need the continuous availability of a dedicated wideband data channel. To provide such an undisturbed networking it is necessary to exchange control packets frequently, even in the case of no data to be exchange using the network. As a consequence a large amount of energy is consumed to control such a high-speed connection (see e.g. Mancuso and Alouf, 2012 [22] for discussion on this topic). The power saving can be achieved, however, by using the so called sleep mode operations.

Let us take into consideration an infinite-buffer $M/G/1$ type model in which customers arrive according to a Poisson process with rate λ and are being served individually with a CDF $F(\cdot)$ of the service time with LST $f(\cdot)$. Moreover, let (V_n), $n \geq 1$, be a sequence of independent positive random variables with a common CDF $G(\cdot)$, with LST $g(\cdot)$, where V_n corresponds to the nth server vacation. We assume that at time $t = 0$ a customer enters the empty system with the server being on vacation, and that V_1 denotes the residual duration of the first vacation. Besides, assume that the sequence (V_n) is independent on the arrival and service processes.

Investigate the models with single and multiple vacations simultaneously. Let us take into consideration only moments of service and vacation completions (we treat them as one sequence). Let $(X_n^\star, I_n^\star)$ denote the state of the system at the nth such moment, where $X_n^\star$ stands for the number of customers present and

$$I_n^\star = \begin{cases} 0 & \text{if the } n\text{th moment is a vacation termination instant,} \\ 1 & \text{if the } n\text{th moment is a service completion instant.} \end{cases} \tag{1}$$

Let us note that for the system operating under single vacation policy we have (see [8])

$$(X_{n+1}^\star, I_{n+1}^\star) = \begin{cases} (X_n^\star + \eta_S - 1, 1) & \text{if} \quad X_n^\star \geq 1, \\ (\eta_S, 1) & \text{if} \quad (X_n^\star, I_n^\star) = (0,0), \\ (\eta_V, 0) & \text{if} \quad (X_n^\star, I_n^\star) = (0,1), \end{cases} \tag{2}$$

where η_S and η_V stand, respectively, for the number of customers arriving during one service time and during one vacation.

For the model with multiple vacation policy we obtain, similarly,

$$(X_{n+1}^\star, I_{n+1}^\star) = \begin{cases} (X_n^\star + \eta_S - 1, 1) & \text{if} \quad X_n^\star \geq 1, \\ (\eta_V, 0) & \text{if} \quad X_n^\star = 0. \end{cases} \tag{3}$$

Let $X^\star$ and $I^\star$ be the limitting random variables for $X_n^\star$ and $I_n^\star$ as $n \to \infty$, so $(X_n^\star, I_n^\star) \xrightarrow{d} (X^\star, I^\star)$. As it was noted in [8], the number of customers $\widehat{X}$ "seen" by an arbitrary service completion is just the random variable $X^\star$ conditioned by $I^\star = 1$. Using this conditioning it can be proved (see [8] or [21]) that for the queueing system with single server vacations the following representation is true:

$$\begin{aligned} \widehat{P}^{SV}(z) &\overset{def}{=} \mathbf{E}\{z^{\widehat{X}}\} = \mathbf{E}\{z^{X^\star} \mid I^\star = 1\} = \frac{(1-\rho)(1-z)f(\lambda - \lambda z)}{f(\lambda - \lambda z) - z} \cdot \frac{1 - g(\lambda - \lambda z) + (1-z)f(\lambda)}{[f(\lambda) + \lambda \mathbf{E}(V)](1-z)} \\ &= \widehat{P}(z) \cdot \frac{1 - g(\lambda - \lambda z) + (1-z)f(\lambda)}{[f(\lambda) + \lambda \mathbf{E}(V)](1-z)}, \end{aligned} \tag{4}$$

where $\rho = \lambda \int_0^\infty t dF(t) < 1$ is the occupation rate in the "usual" $M/G/1$-type system (without vacations) and $\mathbf{E}(V)$ stands for the mean vacation duration.

The first factor on the right side of (4) is the PGF of the number of customers at a service completion epoch in the "usual" $M/G/1$-type queue (a well-known Pollaczek-Khintchine formula). In consequence, $\widehat{X}$ can be represented as a sum of two independent random variables, one of them is the number of customers at the service completion epoch in the corresponding "usual" $M/G/1$ system without vacations. This conclusion is known as the so called decomposition property. As it turns out, such a property can be proved not only for a model with single vacation.

For the $M/G/1$-type queue operating under multiple vacation policy we have, similarly,

$$\widehat{P}^{MV}(z) = \widehat{P}(z) \cdot \frac{1 - g(\lambda - \lambda z)}{\lambda \mathbf{E}(V)(1-z)}. \tag{5}$$

Moreover, in 1986 in [10] Fuhrmann and Cooper showed that the stationary distribution of the queue size (number of jobs) in the $M/G/1$-type queueing system with generalized server vacation is a convolution of the distribution functions of two independent positive random variables. One of them is the stationary queue-size distribution of the number of jobs in the ordinary $M/G/1$-type queueing system (without server vacations). Some other results for the stationary state of queueing models with server vacations can be found e.g. in [4] and [5].

Let us investigate now a finite-buffer $M/G/1/K$-type queue with Poisson arrivals with rate λ and generally distributed service times with CDF $F(\cdot)$ with LST $f(\cdot)$. The system size is assumed to be K, i.e. we have a buffer with $K-1$ places and one place in service facility. Assume, moreover, that the multiple vacation policy is implemented in that one (single) server vacation has general distribution with CDF $G(\cdot)$ with LST $g(\cdot)$. Denote by $X(t)$ the number of packets present in the system at time t. As one can note, in the literature most of results obtained for different queueing systems relate to the stationary state of the system (as $t\to\infty$). However, transient analysis (for fixed time t) is often desired, e.g. just after the start of the system operation or after application of a new control mechanism, or in the case of low traffic intensity (in this case the convergence rate of transient characteristics to the stationary may be slow).

Introduce the the following notation:

$$\widehat{Q}_n(s,m) \stackrel{def}{=} \int_0^\infty e^{-st}\mathbf{P}\{X(t)=m\,|\,X(0)=K-n\}dt, \quad s>0,\, 0\le m,n\le K. \tag{6}$$

It can be proved (see Kempa, 2015 [16]) that the following system of equations is true:

$$\sum_{i=-1}^{n} a_{i+1}(s)\widehat{Q}_{n-i}(s,m)-\widehat{Q}_n(s,m)=\psi_n(s,m), \quad 0\le n\le K-1, \tag{7}$$

and

$$\widehat{Q}_K(s,m)=\sum_{i=1}^{K-1} b_i(s)\widehat{Q}_{K-i}(s,m)+\widehat{Q}_0(s,m)\sum_{i=K}^{\infty} b_i(s)+d(s,m)+\frac{\delta_{m,0}}{s+\lambda}, \tag{8}$$

where

$$a_k(s)\stackrel{def}{=}\int_0^\infty e^{-(s+\lambda)y}\frac{(\lambda y)^k}{k!}dF(y), \tag{9}$$

$$b_k(s)\stackrel{def}{=}\big(1-g(s+\lambda)\big)^{-1}\int_0^\infty e^{-(s+\lambda)y}\frac{(\lambda y)^k}{k!}dG(y), \tag{10}$$

$$d(s,m)\stackrel{def}{=}\big(1-g(s+\lambda)\big)^{-1}\Big(I\{1\le m\le K-1\}\varphi_{G,m}(s)+\delta_{m,K}\sum_{i=K-1}^{\infty}\varphi_{G,i+1}\Big)(s), \tag{11}$$

$$\psi_n(s,m)\stackrel{def}{=}a_{n+1}(s)\widehat{Q}_0(s,m)-\widehat{Q}_1(s,m)\sum_{k=n+1}^{\infty}a_k(s)-h_{K-n}(s,m), \tag{12}$$

$$h_k(s,m)\stackrel{def}{=}I\{k\le m\le K-1\}\varphi_{F,m-k}(s)+\delta_{m,K}\sum_{i=K-k}^{\infty}\varphi_{F,i}(s), \tag{13}$$

where for arbitrary CDF $H(\cdot)$

$$\varphi_{H,k}(s)\stackrel{def}{=}\int_0^\infty e^{-(s+\lambda)t}\frac{(\lambda t)^k}{k!}[1-H(t)]dt. \tag{14}$$

There is proved in Korolyuk, 1975 [20] that each solution of the infinite-sized system of type (7), written for $n \geq 0$, can be stated as

$$\widehat{Q}_n(s,m) = C(s,m)R_{n+1}(s) + \sum_{k=0}^{n} R_{n-k}(s)\psi_k(s,m), \quad n \geq 0, \tag{15}$$

where $C(s,m)$ is independent on n and successive terms of the sequence $(R_k(s))$ (called Korolyuk's potential) can be computed as follows:

$$R_k(s) = \lim_{z \to 0} \frac{1}{k!} \frac{\partial^k Q(s,z)}{\partial z}, \tag{16}$$

where $Q(s,z) \stackrel{def}{=} \frac{z}{A(s,z)-z}$ and

$$A(s,z) \stackrel{def}{=} \sum_{k=0}^{\infty} z^k a_k(s), \quad s > 0, |z| < 1. \tag{17}$$

As it turns out (see [20]) the sequence $\big(R_k(s)\big)$ can also be defined recursively as

$$R_0(s) = 0, \quad R_1(s) = \frac{1}{a_0(s)}, \quad R_{k+1}(s) = R_1(s)\Big(R_k(s) - \sum_{i=0}^{k} a_{i+1}(s)R_{k-i}(s)\Big), \tag{18}$$

where $k \geq 1$.

Treating the last equation (8) as a specific-type boundary condition (see [16]), we can find the representation for $C(s,m)$ explicitly and, utilizing (15), we obtain the following formula for the LT of the conditional transient queue-size distribution in the considered model:

$$\begin{aligned} &\int_0^{\infty} e^{-st}\mathbf{P}\{X(t) = m \,|\, X(0) = n\}dt = \Phi_{K-n}(s,m) \\ &+ \frac{\sum_{i=1}^{K-1} b_{K-i}(s)\Phi_i(s,m) + d(s,m) + \delta_{m,0}(s+\lambda)^{-1} - \Phi_K(s,m)}{\Theta_K(s) - \sum_{i=1}^{K-1} b_{K-i}(s)\Theta_i(s) - \sum_{i=K}^{\infty} b_i(s)}\Theta_{K-n}(s), \end{aligned} \tag{19}$$

where $0 \leq m,n \leq K$ and

$$\Theta_n(s) \stackrel{def}{=} a_0(s)R_{n+1}(s) + \sum_{k=0}^{n} R_{n-k}(s)\Big(a_{k+1}(s) - f^{-1}(s)\sum_{i=k+1}^{\infty} a_i(s)\Big), \tag{20}$$

$$\Phi_n(s,m) \stackrel{def}{=} \sum_{k=0}^{n} R_{n-k}(s)\Big(h_K(s,m)f^{-1}(s)\sum_{i=k+1}^{\infty} a_i(s) - h_{K-k}(s,m)\Big). \tag{21}$$

Transient results for main stochastic characteristics of queueing models with single/multiple vacation policies can also be found in [11], [12], [13] and [17].

4 Threshold-type *N*-policy

In this section we will present main analytical results for some exemplary queueing models with the mechanism of N-policy. Yadin and Naor in 1963 [27] were the first who studied the $M/G/1$-type queue with this type of threshold control policy.

Let us consider, firstly, the $M/G/1$-type infinite-buffer model. We will apply the mean-value approach to find the representation for the mean waiting time $\mathbf{E}(W)$ in the stationary state of the system. Assume that jobs arrive according to a Poisson process with intensity λ and that the service time of individual job is generally distributed with finite mean $\mathbf{E}(B)$ and the second moment $\mathbf{E}(B^2)$. Moreover, the first service after reaching the threshold level N is preceded by a generally distributed setup time with finite two first moments $\mathbf{E}(S)$ and $\mathbf{E}(S^2)$.

Let us note that we have (see e.g. Adan and Resing, 2001 [1])

$$\begin{aligned}\mathbf{E}(W) =& \mathbf{E}(X_Q)\mathbf{E}(B)+\rho\mathbf{E}(B_R)+\sum_{i=1}^{N}\mathbf{P}\{\text{Number of arriving job is i}\}\left[\frac{N-i}{\lambda}+\mathbf{E}(S)\right]\\ &+\mathbf{P}\{\text{Server is during setup time on arrival}\}\mathbf{E}(S_R),\end{aligned} \tag{22}$$

where X_Q, B_R and S_R stand for the number of jobs waiting in the queue (buffer), residual service time and residual setup time, respectively. Obviously $\rho = \lambda\mathbf{E}(B) < 1$.

Observe that the probability that a job enters the system during the accumulation period (when the processing is suspended) equals to $1-\rho$. Besides, the accumulation period consists of N successive interarrival times plus the following setup time. In consequence, the probability that the arriving job is the ith one during the accumulation period is equal to the product of $1-\rho$ and the proportion of means: of one interarrival time and the duration of the whole accumulation period. So, we have

$$\mathbf{P}\{\text{Number of arriving job is i}\} = (1-\rho)\frac{\lambda^{-1}}{N\lambda^{-1}+\mathbf{E}(S)}, \tag{23}$$

where $i \in \{1,...,N\}$.

Similarly, we have

$$\mathbf{P}\{\text{Server is during setup time on arrival}\} = (1-\rho)\frac{\mathbf{E}(S)}{N\lambda^{-1}+\mathbf{E}(S)}. \tag{24}$$

Substituting these two relationships into (22), after simplification, we obtain

$$\mathbf{E}(W) = \frac{\rho\mathbf{E}(B_R)}{1-\rho}+\frac{N\lambda^{-1}}{N\lambda^{-1}+\mathbf{E}(S)}\left[\frac{N-1}{2\lambda}+\mathbf{E}(S)\right]+\frac{\mathbf{E}(S)}{N\lambda^{-1}+\mathbf{E}(S)}\mathbf{E}(S_R). \tag{25}$$

Now, let us deal with the finite-buffer $M/G/1/K$-type queueing model operating under the N-policy, with Poisson arrival stream with rate λ and generally distributed service times with CDF $F(\cdot)$ with LST $f(\cdot)$ in that, as previously, the maximum system state is K. As usually, the service process is organized according to the FIFO discipline. We are interested in the transient analysis of the queue-size distribution (see Kempa and Kurzyk [18] for more detailed analysis). Observe that the operation of the system can be observed on successive buffer loading periods BL_1, BL_2, ... followed by busy periods BP_1, BP_2, ..., during which the system becomes idle. From the memoryless property of interarrival times follows that initial and completion epochs of successive busy periods are Markov moments. Hence (BL_k) and (BP_k), $k \geq 1$, are sequences of independent random variables with the same CDFs in each sequence separately. For simplicity we identify BL_k, BP_k, $k \geq 1$, with their durations.

Let $X(t)$ be the number of jobs present in the system at time t, including the one being served at this moment (if any). Investigate the queue-size distribution during the first buffer loading period BL_1 starting at time $t=0$. We get

$$\mathbf{P}\{\big(X(t)=m\big)\cap\big(t\in BL_1\big)\}=I\{0\le m\le N-1\}\frac{(\lambda t)^m}{m!}e^{-\lambda t}, \tag{26}$$

where $t\ge 0$. Introducing the following notation:

$$\widetilde{q}^{BL}(s,m)\stackrel{def}{=}\int_0^\infty e^{-st}\mathbf{P}\{\big(X(t)=m\big)\cap\big(t\in BL_1\big)\}dt, \tag{27}$$

where $s>0$, we have from (26)

$$\widetilde{q}^{BL}(s,m)=I\{0\le m\le N-1\}\int_0^\infty e^{-(s+\lambda)t}\frac{(\lambda t)^m}{m!}dt=I\{0\le m\le N-1\}\frac{\lambda^m}{(\lambda+s)^{m+1}}. \tag{28}$$

Obviously, each buffer loading period duration has the N-Erlang distribution with parameter λ, so we obtain

(29)

$$\widetilde{g}^{BL}(s)=\int_0^\infty e^{-st}dG^{BL}(t)\stackrel{def}{=}\int_0^\infty e^{-st}d\mathbf{P}\{BL_k<t\}=\int_0^\infty e^{-(s+\lambda)t}\frac{\lambda^N}{(N-1)!}t^{N-1}dt=\Big(\frac{\lambda}{\lambda+s}\Big)^N.$$

Let us consider now the evolution of the system during a busy period. Assume temporarily that the service process can be started with arbitrary possible level n of buffer state, where $1\le n\le K$ (not only at $n=N$). Let $Q_n^{BP}(t,m)$ be transient conditional queue-size distribution at time $t\in BP_1$, namely

$$Q_n^{BP}(t,m)\stackrel{def}{=}\mathbf{P}\{\big(X(t)=m\big)\cap\big(t\in BP_1\big)|X(0)=n\}, \tag{30}$$

where $t>0$ and $1\le m,n\le K$. For simplicity let us assume that BP_1 begins at time $t=0$. Since successive departure epochs are Markov moments then, applying the total probability law with respect to the first departure moment after $t=0$, we obtain the following system of integral equations:

$$Q_1^{BP}(t,m)=\sum_{i=1}^{K-2}\int_0^t\frac{(\lambda x)^i}{i!}e^{-\lambda x}Q_i^{BP}(t-x,m)dF(x)+\sum_{i=K-1}^{\infty}\int_0^t\frac{(\lambda x)^i}{i!}e^{-\lambda x}Q_{K-1}^{BP}(t-x,m)dF(x)$$
$$+\overline{F}(t)e^{-\lambda t}\Big[I\{1\le m\le K-1\}\frac{(\lambda t)^{m-1}}{(m-1)!}+I\{m=K\}\sum_{i=K-1}^{\infty}\frac{(\lambda t)^i}{i!}\Big], \tag{31}$$

and, for $2\le n\le K$,

$$Q_n^{BP}(t,m)=\sum_{i=0}^{K-n-1}\int_0^t\frac{(\lambda x)^i}{i!}e^{-\lambda x}Q_{n+i-1}^{BP}(t-x,m)dF(x)+\sum_{i=K-n}^{\infty}\int_0^t\frac{(\lambda x)^i}{i!}e^{-\lambda x}Q_{K-1}^{BP}(t-x,m)dF(x)$$
$$+\overline{F}(t)e^{-\lambda t}\Big[I\{n\le m\le K-1\}\frac{(\lambda t)^{m-n}}{(m-n)!}+I\{m=K\}\sum_{i=K-n}^{\infty}\frac{(\lambda t)^i}{i!}\Big], \tag{32}$$

where $\overline{F}(t)\stackrel{def}{=}1-F(t)$.

Introducing the following nomenclature:

(33) $\widetilde{q}_n^{BP}(s,m) \stackrel{def}{=} \int_0^\infty e^{-st} Q_{N-n}^{BP}(t,m)dt,$

(34) $a_n(s) \stackrel{def}{=} \int_0^\infty e^{-(s+\lambda)t} \frac{(\lambda t)^n}{n!} dF(t),$

(35) $\theta_n(s,m) \stackrel{def}{=} \int_0^\infty e^{-(\lambda+s)t} \overline{F}(t) \Big[I\{n \le m \le K-1\} \frac{(\lambda t)^{m-n}}{(m-n)!} + I\{m=K\} \sum_{i=K-n}^{\infty} \frac{(\lambda t)^i}{i!} \Big] dt,$

where $s > 0$, we can transform the equations of the system (31)–(32) to the following ones:

$$\widetilde{q}_{K-1}^{BP}(s,m) = \sum_{i=1}^{K-2} a_i(s) \widetilde{q}_{K-i}^{BP}(s,m) + \widetilde{q}_1^{BP}(s,m) \sum_{i=K-1}^{\infty} a_i(s) + \theta_1(s,m), \tag{36}$$

$$\sum_{k=-1}^{n} a_{k+1}(s) \widetilde{q}_{n-k}^{BP}(s,m) - \widetilde{q}_n^{BP}(s,m) = \phi_n(s,m), \tag{37}$$

where $0 \le n \le K-2$ and

$$\phi_n(s,m) \stackrel{def}{=} a_{n+1}(s) \widetilde{q}_0^{BP}(s,m) - \widetilde{q}_1^{BP}(s,m) \sum_{i=n+1}^{\infty} a_i(s) - \theta_{K-n}(s,m). \tag{38}$$

Using the similar method as described in Section 3, we get the representation for $\widetilde{q}_0^{BP}$ (we need only this one) in the form

$$\widetilde{q}_0^{BP}(s,m) = \Pi_1(s,m) \Pi_2^{-1}(s), \tag{39}$$

where

(40) $\Pi_1(s,m) \stackrel{def}{=} \sum_{i=1}^{K-2} a_i(s) \eta_{K-i}(s,m) - \frac{\theta_K(s,m)}{f(s)} \sum_{i=K-1}^{\infty} a_i(s) + \theta_1(s,m) - \eta_{K-1}(s,m),$

(41) $\Pi_2(s) \stackrel{def}{=} \gamma_{K-1}(s) - \sum_{i=1}^{K-2} a_i(s) \gamma_{K-i}(s) - \frac{1}{f(s)} \sum_{i=K-1}^{\infty} a_i(s),$

(42) $\gamma_n(s) \stackrel{def}{=} a_0(s) R_{n+1}(s) + \sum_{i=0}^{n} R_{n-i}(s) \Big[a_{i+1}(s) - \frac{1}{f(s)} \sum_{j=i+1}^{\infty} a_j(s) \Big]$

and

$$\eta_n(s,m) \stackrel{def}{=} \sum_{i=0}^{n} R_{n-i}(s) \Big[\frac{\theta_K(s,m)}{f(s)} \sum_{j=i+1}^{\infty} a_j(s) - \theta_{K-i}(s,m) \Big] \tag{43}$$

and the functional sequence $(R_k(s))$ is defined in (18).

Denoting now by $g_n^{BP}(\cdot)$ the LST of CDF of busy period duration in the system that starts working with $1 \le n \le K$ packets present in the buffer queue, we can formulate a system of equations for $g_1^{BP}(s), ..., g_K^{BP}(s)$ similar to (36)–(37) and find, in particular, the following representation:

$$\widetilde{g}^{BP}(s) \stackrel{def}{=} \widetilde{g}_N^{BP}(s) = \gamma_{K-N}(s) \widetilde{\Pi}_1(s) \Pi_2^{-1}(s) + \widetilde{\eta}_{K-N}(s), \quad n \ge 0, \tag{44}$$

where

(45) $$\widetilde{\eta}_n(s) \stackrel{def}{=} \sum_{i=0}^{n} R_{n-i}(s)\Big[a_0^{-1}(s)\big(1+a_0^{-1}(s)\big)^{-1}\widetilde{\theta}_K(s) - \widetilde{\theta}_{K-i}(s)\Big],$$

(46) $$\widetilde{\theta}_n(s) \stackrel{def}{=} \begin{cases} f(\lambda+s), & n=1, \\ 0, & n \geq 2, \end{cases}$$

and

(47) $$\widetilde{\Pi}_1(s) \stackrel{def}{=} \sum_{i=1}^{K-2} a_i(s)\widetilde{\eta}_{K-i}(s) + \widetilde{\eta}_1(s)\sum_{i=K-1}^{\infty} a_i(s) + \widetilde{\theta}_1(s) - \widetilde{\eta}_{K-1}(s)$$

and $\Pi_2(s)$ and $\gamma_n(s)$ were defined in (41) and (42), respectively.

Furthermore, the law of total probability gives

(48) $$\mathbf{P}\{X(t)=m\} = \sum_{i=1}^{\infty}\Big(\mathbf{P}\{\big(X(t)=m\big)\cap\big(t\in BL_i\big)\} + \mathbf{P}\{\big(X(t)=m\big)\cap\big(t\in BP_i\big)\}\Big).$$

Since BL_i and BP_i, $i \geq 1$, are independent and have identical distributions (in each sequence separately), we get

(49)
$$\mathbf{P}\{\big(X(t)=m\big)\cap\big(t\in BL_i\big)\} = \int_0^t \mathbf{P}\{\big(X(t-y)=m\big)\cap\big(t-y\in BL_1\big)\}d\big(G^{BL}*G^{BP}\big)^{(i-1)*}(y),$$

(50)
$$\mathbf{P}\{\big(X(t)=m\big)\cap\big(t\in BP_i\big)\} = \int_0^t \mathbf{P}\{\big(X(t-y)=m\big)\cap\big(t-y\in BP_1\big)\}d\big[\big(G^{BL}\big)^{i*}*\big(G^{BP}\big)^{(i-1)*}\big](y).$$

Taking LTs of (49)–(50), referring to (48), we obtain now, as main result, the formula for the LT of the queue-size distribution of in the $M/G/1/K$-type system with threshold-type N-policy:

(51) $$\int_0^{\infty} e^{-st}\mathbf{P}\{X(t)=m\}dt = \frac{\widetilde{q}^{BL}(s,m) + \widetilde{g}^{BL}(s)\widetilde{q}_0^{BP}(s,m)}{1-\widetilde{g}^{BP}(s)\widetilde{g}^{BL}(s)}.$$

In [15] a similar model with infinite buffer and batch arrivals is studied in transient state, where additionally setup times are implemented.

5 Active Queue Management - preventional "filtering" of arriving customers

A special-type group of solutions which may result in limitting the access to the service station is associated with the so called Active Queue Management (AQM). The main idea of AQM is in the intervention in the process of qualification for service the arriving customers. In other words, an AQM-type approach allows for dropping the arriving job even when there is a place in the accumulating buffer, differently than in the classical Tail Drop discipline (the accumulation of the buffer is undisturbed until the buffer becomes saturated; then all the arriving customers are being lost). Obviously, such a mechanism may also be called "balking" and may be considered as a type of customer "unpatience", which is known in queueing theory. However, AQM is originally motivated by networking studies and assumes that the "roles" are reversed here. It is not the customer who decides whether to join the system, but the system "manager" that have much more knowledge about the system operation now and in the past. AQM algorithms dedicated for IP routers have been investigated widely since the paper [9] by Floyd and Jacobson was published in 1993. Usually an arriving packet is rejected (dropped) with probability depending on the actual queue state (as in the classical “balking” discipline: from the point of view of the arriving customer, typically, the only knowledge about the system is the actual queue size at the arriving moment). However, the probability of dropping may also depend on other, even very complex, stochastic characteristics, like the “history” of customer losses or the frequency of empty buffer and full buffer occurrences. In the Internet routers AQM-type packet dropping is used mainly to prevent buffer queues from growing too long, keeping high link utilizations simultaneously. Other goals are stable and predictable buffer queues with low variances of queue sizes and, moreover, desynchronization of TCP sources (see e.g. Chydziński and Chróst, 2011 [6]). Indeed, in the classical Tail Drop approach, in the case of buffer overflow, according to the TCP protocol, all the sources (hosts sending packets directed ot the node) may reduce the intensity of sending the packets (synchronization).

Let us assume the $M/M/1/K$-type finite-buffer queueing system in which the entering job which finds i jobs present in the system is dropped by the “manager” with probability d_i, $i = 0,1,...,K$, $d_K = 1$ (the so called “dropping function”). If λ is the Poisson arrival rate and μ^{-1} denotes mean service time, the system of equilibrium equations for the steady-state queue-size probabilities $p_k \stackrel{def}{=} \mathbf{P}\{X = k\}$, where X is the number of packets present in the system in the stationary state, has the following form (see Kempa, 2011 [14]):

(52) $$\left\{ \begin{array}{rcll} \lambda(1-d_0)p_0 & = & p_1, & \\ \left[\lambda(1-d_k)+\mu\right]p_k & = & \lambda(1-d_{k-1})p_{k-1}+\mu p_{k+1}, & 1 \le k \le K-1, \\ \mu p_K & = & \lambda(1-d_{K-1})p_{K-1}. & \end{array} \right.$$

Hence we get

(53) $$p_k = \frac{\lambda^k}{\mu^k}\prod_{i=0}^{k-1}(1-d_i)p_0, \;\; k = 1,2,...K.$$

From the normalization condition

$$\sum_{i=0}^{K} p_i = 1$$

we find

(54) $$p_0 = \Big(1+\sum_{k=1}^{K}\rho^k\prod_{i=0}^{k-1}(1-d_i)\Big)^{-1},$$

where $\rho = \frac{\lambda}{\mu}$ is the occupation rate of the system, and hence, finally,

(55) $$p_k = \frac{\rho^k\prod_{i=0}^{k-1}(1-d_i)}{1+\sum_{j=1}^{K}\rho^j\prod_{i=0}^{j-1}(1-d_i)}, \qquad k = 0,1,...,K.$$

Obviously, if $d_i \equiv 0$, we get the well-known solution for the “classical” $M/M/1/K$-type queue.

Let us note that after the implementation of the dropping function we get not a usual “thinning” of the Poisson arrival process: starting with n jobs present, the first arriving one is dropped with probability d_n but the second one either with probability d_{n+1} (if the previous one was qualified for service), or with probability d_n (otherwise).

Now let us consider a more general queueing system in which the service time of individual customer is generally distributed with a CDF $F(\cdot)$ ($M/G/1/K$-type model). Denote by X_n the number of jobs present in the system after departing the nth job, hence $0 \leq X_n \leq K-1$. We are interested in the stationary distribution of X_n, namely $\widehat{p}_k = \lim_{n\to\infty}\{X_n = k\}$, $k = 0,...,K-1$. Obviously, (X_n) is an ergodic Markov chain, so probabilities $\widehat{p}_k$ satisfy the following system of equations (see [6]):

$$\widehat{p}_j = \sum_{i=0}^{K-1} \widehat{p}_i p_{i,j}, \quad 0 \leq j \leq K-1, \tag{56}$$

and $\sum_{j=0}^{K-1} \widehat{p}_j = 1$, where

$$p_{i,j} = \mathbf{P}\{X_{n+1} = j \,|\, X_n = i\}, \;\; 0 \leq i, j \leq K-1, \tag{57}$$

denote one-step transition probabilities of the Markov chain (X_n).

The following representation is true:

$$p_{i,j} = \begin{cases} q_{1,j} & \text{if} \quad i = 0, 0 \leq j \leq K-1, \\ q_{i,j-i+1} & \text{if} \quad 1 \leq i \leq K-1, \\ 0 & \text{otherwise}, \end{cases} \tag{58}$$

where

$$q_{i,j} = \int_0^\infty A_{i,j}(x)dF(x) \tag{59}$$

and $A_{i,j}(x)$ denotes here the probability that up to the fixed time x exactly j jobs are "qualified" for service on condition that initially (at time $t = 0$) we have exactly i jobs present. It is shown in [6] that

$$\int_0^\infty e^{-sx} A_{i,j}(x)dx = \frac{\lambda^j \prod_{k=0}^{j-1}(1-d_{i+k})}{\prod_{k=0}^{j}[s+\lambda(1-d_{i+k})]}, \quad s > 0, \tag{60}$$

where $i, j \geq 0$ and we accept the agreement that $\prod_{k=0}^{-1} = 1$.

In practice, in the next step, $A_{i,j}(x)$ (and hence $q_{i,j}$) are usually found inverting the right side of (60) which is not difficult.

An interesting study on AQM-type model with continuous job volumes and processor sharing can be found in [26].

REFERENCES:

[1] Adan, I., Resing, J., Queueing theory. Eindhoven University of Technology (2001).

[2] Bocharov, P.P., D'Apice, C., Pechinkin, A.V., Salerno, S., Queueing theory. VSP (2004).

[3] Borovkov, A.A., Probabilistic processes in the queueing theory. Nauka (1972).

[4] Choudhury, G., An $M^X/G/1$ queueing system with a setup period and a vacation period. Queueing Syst. **36** (2000), 23–38.

[5] Choudhury, G., A batch arrival queue with a vacation time under single vacation policy. Comp. Oper. Res. **29** (14) (2002), 1941–1955.

[6] Chydziński, A., Chróst, Ł., Analysis of AQM queues with queue size based packet dropping. Int. J. Appl. Math. Comput. Sci. **21** (3) (2011), 567–577.

[7] Cohen, J.W., The single server queue. North-Holland (1982).

[8] Doshi, B.T., Queueing systems with vacations - a survey, Queueing Syst. **1** (1) (1986), 19–66.

[9] Floyd, S., Jacobson, V., Random early detection gateways for congestion avoidance. IEEE/ACM Trans. Netw. **1** (4) (1993), 397–413.

[10] Fuhrmann, S.W., Cooper, R.B., Stochastic decompositions in the $M/G/1$ queue with generalized vacations. Oper. Res. **33** (5) (1985), 1117–1129.

[11] Kempa, W.M., $GI/G/1/\infty$ batch arrival queueing system with a single exponential vacation. Math. Methods Oper. Res. **69** (1) (2009), 81–97.

[12] Kempa, W.M., Some new results for departure process in the $M^X/G/1$ queueing system with a single vacation and exhaustive service. Stoch. Anal. Appl. **28** (1) (2010), 26–43.

[13] Kempa, W.M., Analysis of departure process in batch arrival queue with multiple vacations and exhaustive service. Commun. Stat. - Theory Methods **40** (16) (2011), 2856–2865.

[14] Kempa, W.M., On main characteristics of the $M/M/1/N$ queue with single and batch arrivals and the queue size controlled by AQM algorithms. Kybernetika **47** (6) (2011), 930–943.

[15] Kempa, W.M., On transient queue-size distribution in the batch arrival system with the N-policy and setup times. Math. Commun. **17** (1) (2012), 285–302.

[16] Kempa, W.M., Queue-size distribution in energy-saving model based on multiple vacation policy. In: Current Trends in Analysis and Applications. Proceeding of the 9th ISAAC Congress 2015, 733–740.

[17] Kempa, W.M., Transient workload distribution in the $M/G/1$ finite-buffer queue with single and multiple vacations. Ann. Oper. Res. **239** (2) (2016), 381–400.

[18] Kempa, W.M., Kurzyk, D., Queue-size distribution in a WSN node with power saving algorithm based on N-Policy (submitted).

[19] Kleinrock, L., Queueing systems. Volumes 1 and 2. John Wiley & Sons (1975).

[20] Korolyuk, V.S., Boundary-value problems for compound Poisson processes. Naukova Dumka (1975).

[21] Levy, Y., Yechiali, U., Utilization of idle time in an $M/G/1$ queueing. Management Sci. **22** (2) (1975), 202–211.

[22] Mancuso, V., Alouf, S., Analysis of power saving with continuous connectivity. Comput. Netw. **56** (2012), 2481–2493.

[23] Miller, L.W., Alternating priorities in multi-class queue. Ph.D. dissertation, Cornell University, Ithaca, N.Y. (1964).

[24] Takacs, L., Introduction to the theory of queues. Oxford University Press (1962).

[25] Takagi, H., Queueing analysis. Volume 1: Vacation and priority systems. Volume 2: Finite systems. North-Holland (1991–Vol. 1, 1993–Vol. 2).

[26] Tikhonenko, O., Kempa, W.M., Queueing system with processor sharing and limited memory under control of the AQM mechanism. Autom. Remote Control **76** (10) (2015), 1784–1796.

[27] Yadin, M., Naor, P., Queueing systems with a removable service station. Oper. Res. Quarterly **14** (1963), 393–405.

Wojciech Kempa
Silesian University of Technology, Faculty of Applied Mathematics, Institute of Mathematics,
Gliwice, Poland
E-mail:wojciech.kempa@polsl.pl

MODEL REPRESENTATIONS OF QUADRATIC NON-SELFADJOINT OPERATOR PENCILS

VLADIMIR A. ZOLOTAREV

B. VERKIN INSTITUTE FOR LOW TEMPERATURE PHYSICS AND ENGINEERING OF THE NATIONAL ACADEMY OF SCIENCES OF UKRAINE

ABSTRACT: *For quadratic operator pencils of the type* $L(\lambda)=\lambda^2 I+\lambda B+A$, *where* $B=B^*$, $A\neq A^*$, *the problem of factorization,* $L(\lambda)=(\lambda I-Y)(\lambda I-X)$, *where* X, Y *are selfadxjoint operators, is studied. Model representations for the operator roots* X, Y *are obtained. They are expressed via the Hilbert and Sieltjes transforms in weighted spaces.*

KEYWORDS: *non-selfadjoint operator pencils, Hilbert transform, Stieltjes transform.*

Vladimir A. Zolotarev
B. Verkin Institute for Low Temperature Physics and Engineering of
The National academy of Sciences of Ukraine;
V. N. Karazin Kharkiv National University
E-mail: vazolotarev@gmail.com

STABILITY FOR SOLITARY WAVES*

ATANAS G. STEFANOV
DEPARTMENT OF MATHEMATICS,
THE UNIVERSITY OF KANSAS, USA

ABSTRACT: *Solitary waves (or solitons or coherent structures) are ubiquitous objects that appear in most physically relevant time evolution models. More precisely, they are solutions of partial differential equations, depending in a special way on the time variable. There has been a lot of activity in the last fifty years regarding their existence, functional properties (such as smoothness, rate of decay etc.) and most importantly spectral/linear/nonlinear stability as solutions to the general time dependent models. In this talk, I will first review some general principles of the stability theory, and then I will discuss some recent results of mine and collaborators. .*

KEYWORDS: *Solitary waves, stability, Dispersive models*

Many equations in classical physics, mechanics and engineering are derived through their Hamiltonian formulation - namely they minimize an appropriate action functional. As such they enjoy special structure, which makes their theory especially rich, with specific properties and important dynamical consequences.

We consider, by way of introducing the important ideas, the classical non-linear Schrödinger equation (NLS),

$$iu_t + \Delta u + F(|u|^2)u = 0, x \in \mathbb{R}^n, t > 0 \tag{1}$$

and the (generalized) Korteweg-deVries equation (KdV)

$$v_t + v_{xxx} + F(|v|^2)v = 0. \tag{2}$$

for some appropriate function F, which specifies the nonlinearity. These are known to enjoy conservation of energy/Hamiltonian and particle number

$$\mathscr{E}(u) = \frac{1}{2}\int_{\mathbb{R}^n} |\nabla u|^2 - \int_{\mathbb{R}^n} G(u^2(t,x))dx = \mathscr{E}(u_0), \mathscr{P}(u) = \int_{\mathbb{R}^n} |u(t,x)|^2 dx = \mathscr{P}(u_0)$$

where $G(0) = 0, G'(z) = F(z)$.

Furthermore, in may of these problems, there are soliton solutions. Namely for NLS, $e^{i\omega t}\varphi(x)$ and traveling waves for KdV equation $\varphi(x - \omega t)$. These satisfy non-linear elliptic PDE's and do not in general solve explicitly. Their properties are however crucial for the applications, that we have in mind. One of the central questions of interest is the stability of these waves.

To introduce this notion, one takes the ansatz $u(x,t) = e^{i\omega t}(\varphi + v(t,x))$ and expands in powers of v, by then ignoring $O(v^2)$ terms and a similar construction is applied to the traveling waves for KdV. The resulting linearized equation, say

$$v_t = \mathscr{A} v$$

*Partially supported by NSF-DMS 1614734.

is studied in detail. The operator $\mathscr{A}$ is generally unbounded, and it has a number of important properties. The stability of the wave φ means absence of spectrum of $\mathscr{A}$ in the open right-half plane.

We discuss several recent results, pertaining to stability of said solitons. For example, for the related Hartree model - we present a full spectral characterization for all solitary waves, [1]. For the reduced Ostrovsky model, we construct traveling waves in the periodic case, and we identify the spectrally stable one, [2]. For the full Ostrovsky/short pulse model, in the whole line case, we construct and characterize traveling waves, [3]. Finally, we present a stability result for waves of the NLS model, with mixed power non-linearities, [4].

REFERENCES:

[1] V. Georgiev, A. Stefanov, *On the classification of the spectrally stable standing waves of the Hartree problem*, *Physica D*, **370** (2018), p. 29–39.

[2] S. Hakkaev, M. Stanislavova, A. Stefanov, *Spectral stability for classical periodic waves of the Ostrovsky and short pulse models*, *Stud. Appl. Math.*, **139**, (2017), p. 405–433.

[3] I. Posukhovskyi, A. Stefanov, *On the ground states for the generalized Ostrovskyi equation and their stability*, submitted.

[4] A. Stefanov, *On the normalized ground states of second order PDE's with mixed power non-linearities*, submitted

Atanas Stefanov:
The University of Kansas,
Department of Mathematics, USA
E-mail: stefanov@ku.edu

ON GLOBAL IN TIME ASYMPTOTIC BEHAVIOR OF SOLUTIONS OF STRONGLY DAMPED NONLINEAR WAVE EQUATIONS

VARGA K. KALANTAROV

DEPARTMENT OF MATHEMATICS
KOÇ UNIVERSITY ISTANBUL, TURKEY

ABSTRACT: *The talk will be devoted to the initial boundary value problem for nonlinear strongly damped wave equations and related systems under homogeneous Dirichlet's boundary condition in a bounded domain. It will be shown that the asymptotic behavior in time of solutions of considered nonlinear wave equations are completely determined by dynamics of the first N Fourier modes, when N is large enough. Recent results on existence of an exponential attractors of the semigroups generated by the strongly damped and structurally damped nonlinear wave equations will be also discussed.*

Varga Kalantarov
Department of Mathematics,
Koç University Istanbul, Turkey
E-mail: vkalantarov@ku.edu.tr

ON THE MULTIVARIATE SARMANOV DISTRIBUTION AND ITS ACTUARIAL APPLICATIONS

RALUCA VERNIC
FACULTY OF MATHEMATICS AND INFORMATICS,
OVIDIUS UNIVERSITY OF CONSTANTA, ROMANIA

ABSTRACT: *Sarmanov's family of multivariate distributions recently gained the interest of researchers in various domains due to its flexible structure that can model a large range of dependencies starting from given marginals. In particular, this property motivates the consideration of Sarmanov's distribution in the fields of insurance and finance, from which we will present several applications. Therefore, we start by presenting the distribution's main characteristics and then we discuss its fitting to some real insurance data. More precisely, as a first example, we shall see how the bivariate Sarmanov distribution with different types of truncated marginal distributions could serve as a good model for bivariate losses (i.e., we fitted it to a random data sample of motor insurance claims consisting of the costs of property damage and medical expenses). As a second example, we introduce some trivariate Sarmanov distributions with Generalized Linear Models for marginals with the aim to incorporate some individual characteristics of the policyholders when modeling a real trivariate data set of claims frequencies (i.e., we modeled a count data set corresponding to three types of accident risks, two for motor insurance and one for home insurance). Finally, we consider the capital allocation problem, which consists in fairly allocating the capital needed to cover the aggregate loss of a company (e.g., insurance company) among its various lines of business. Risk measures are well-known tools used for this purpose, and one of the most popular such risk measure is the Tail-Value-at-Risk (TVaR). Based on this risk measure, we present some closed-type allocation formulas for risks modeled by Sarmanov's distribution.*

KEYWORDS: *Sarmanov multivariate distribution, Insurance applications, Finance applications, Risk measures, Capital allocation problem.*

REFERENCES:

[1] Bahraoui, Z., Bolance, C., Pelican, E., Vernic, R., On the bivariate Sarmanov distribution and copula. An application on insurance data using truncated marginal distributions. SORT **39** (2) (2015), 209–230.

[2] Bolance, C., Vernic, R., Multivariate count data generalized linear models: Two approaches based on Sarmanov's distribution. 21 International Congress on Insurance: Mathematics and Economics, July 3-5, Vienna (2017).

[3] Kotz, S., Balakrishnan, N., Johnson, N.L., Continuous Multivariate Distributions. Vol. 1: Models and applications. Wiley (2000).

[4] Ratovomirija, G., Tamraz, M., Vernic, R., On some multivariate Sarmanov mixed Erlang reinsurance risks: aggregation and capital allocation. Insurance Mathematics and Economics **74** (2017), 197–209.

[5] Sarmanov, O.V., Generalized normal correlation and two-dimensional Frechet classes. Doclady (Soviet Mathematics) **168** (1966), 596–599.

[6] Vernic, R., On the distribution of a sum of Sarmanov distributed random variables. Journal of Theoretical Probability **29** (1) (2016), 118-142.

[7] Vernic, R., Capital allocation for Sarmanov's class of distributions. Methodology and Computing in Applied Probability **19** (1) (2017), 311-330.

Raluca Vernic:
Faculty of Mathematics and Informatics,
Ovidius University of Constanta, Romania
E-mail: rvernic@univ-ovidius.ro

TENDENCIES IN DEVELOPMENT OF DDOS ATTACKS AND THE DESIGN OF A SOFTWARE APPLICATION FOR DETECTING AND COUNTERING DDOS ATTACKS

VLADIMIR S. GALYAEV, E.A. ZYKOVA, A. E. KRASNOV, D. N. NIKOL'SKII, P. V. PANOV, D.S. REPIN

THE FEDERAL STATE AUTONOMOUS EDUCATIONAL ESTABLISHMENT OF ADDITIONAL PROFESSIONAL EDUCATION "CENTER OF REALIZATION OF STATE EDUCATIONAL POLICY AND INFORMATIONAL TECHNOLOGIES", MOSCOW, RUSSIA

ABSTRACT: *The report is devoted to the problem of protection against Distributed Denial of Service cyberattacks (DDoS attacks). Analysis of reports of several top vendors specializing in the field of information security is presented. It revealed that in recent years there are tendencies in growth of attacks' volume, duration and total number of DDoS attacks per resource. Conclusions are given on the possible further developments of DDoS attacks as the most massive of cyberthreats. Also, the report provides a brief review of scientific publications devoted to intellectual methods and algorithms for protection against DDoS attacks, which can be implemented in hardware and software solutions to counteract DDoS attacks. A general overview of the latest achievements and promising developments in the field of application of intellectual methods for detecting DDoS attacks is made. At the end of the report, the developed by our scientific group methods for detection and prevention of threats to information security aimed at denial of service are described.*

Vladimir Galyaev
Laboratory of Mathematical Modeling and Information Technologies,
Moscow, Russia

EXACT DISTRIBUTIONS OF LR TESTS AND THEIR APPLICATIONS

MILAN STEHLÍK
SCHOOL OF MATHEMATICAL AND STATISTICAL SCIENCES,
ARIZONA STATE UNIVERSITY, AZ, USA
UNIVERSITY OF VALPARAÍSO, VALPARAÍSO, CHILE
JOHANNES KEPLER UNIVERSITY LINZ, AUSTRIA

ABSTRACT: *During the talk we introduce exact statistical procedures based on likelihood ratio. Also practical examples will be given. We introduce exact likelihood ratio tests in exponential family and for a generalized gamma distribution and its properties. We will derive general forms of distributions for exact likelihood ratio test of the homogeneity and scale. Applications and illustrative examples (missing and censored data, mixtures) will be given. Geometry of life time data will be discussed and related to I-divergence decomposition. Small samples testing for frailty through homogeneity test will be discussed. We will provide the methodology for exact and robust test for normality.*

Acknowledgments. Milan Stehlík acknowledges the support of the projects FONDECYT Regular N1151441, LIT-2016-1-SEE-023 (modec) and WTZ Project No. BG 09/2017. The work was supported by the bilateral projects Bulgaria - Austria, 2016-2019, "Feasible statistical modelling for extremes in ecology and finance", Contract number 01/8, 23/08/2017.

REFERENCES:

[1] Balakrishnan N. and Stehlík M., "Likelihood testing with censored and missing duration data" *Journal of Statistical Theory and Practice*, volume 9(1), 2015, 2-22

[2] Stehlík, M., Economou, P., Kiselák, J. and Richter, W.-D. (2014) "Kullback-Leibler life time testing", *Applied Mathematics and Computation*, 240 122-139

[3] Stehlík M. (2003), "Distributions of exact tests in the exponential family", *Metrika* **57**: 145-164

[4] Stehlík M. (2008), Homogeneity and Scale Testing of Generalized Gamma Distribution, *Reliability Engineering & System Safety* 93, 1809-1813.

Milan Stehlík,
Email: milan.stehlik@jku.at