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;

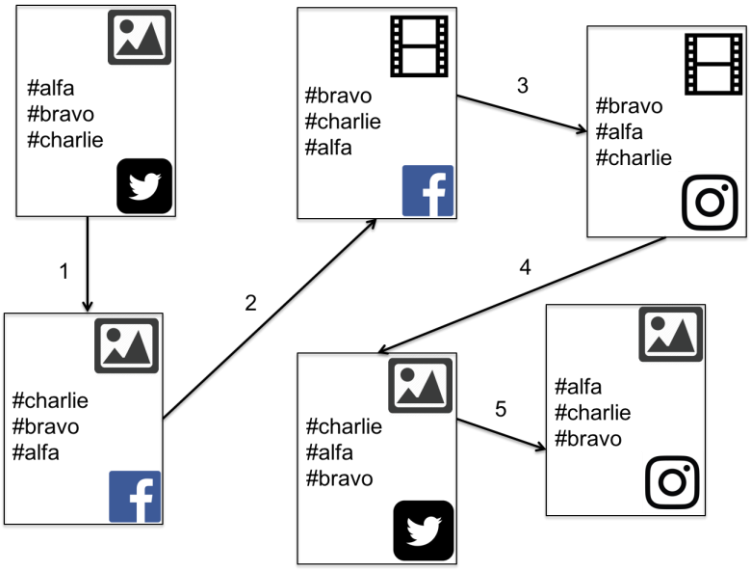

Figure 1 Example of StegHash method [6]

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.

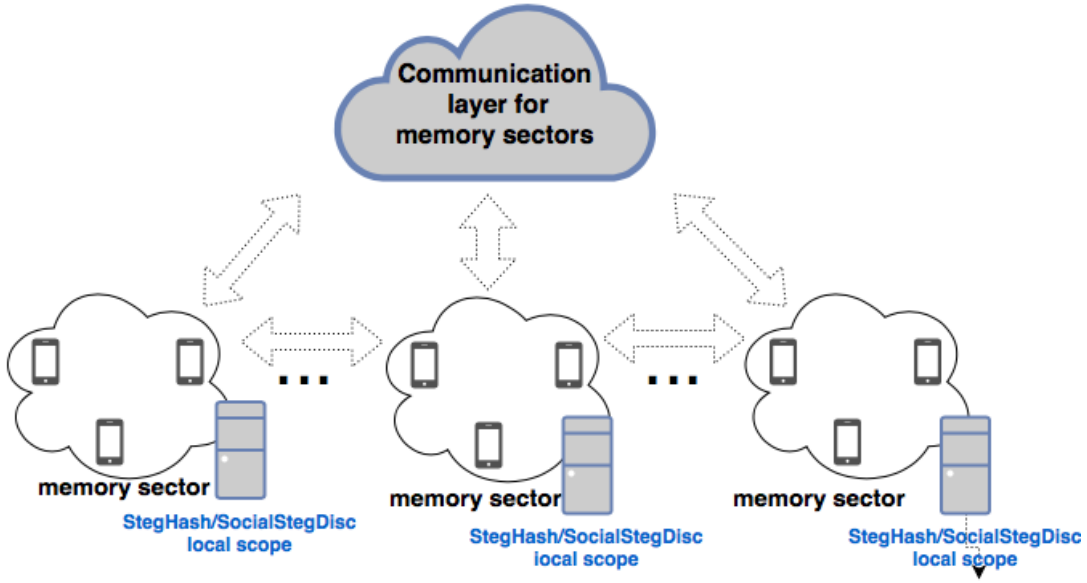

Figure 2 Applying StegHash/SocialStegDisc into the fog architecture

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.

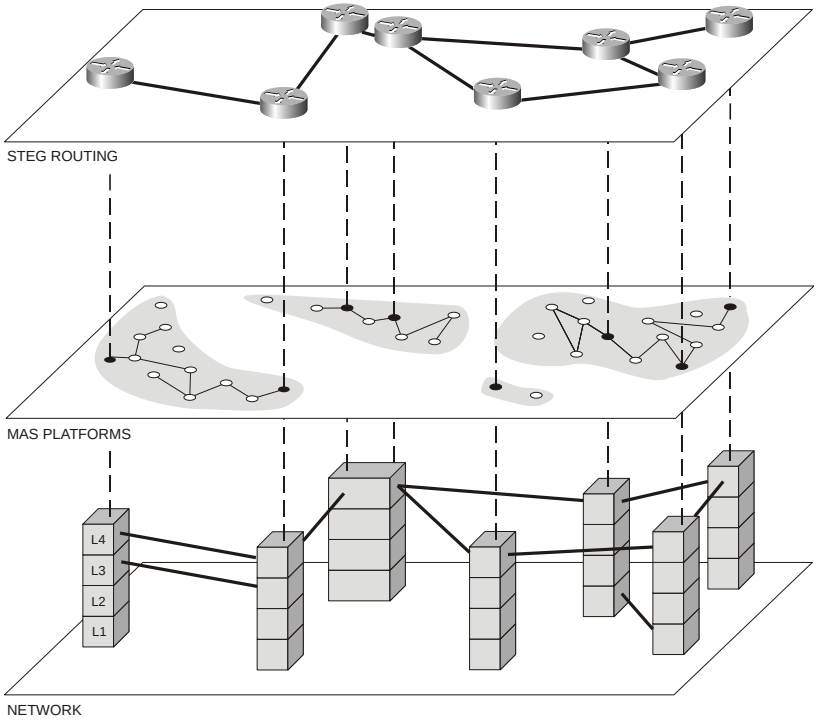

Figure 3 Three-layer architecture of TrustMAS [10]

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.

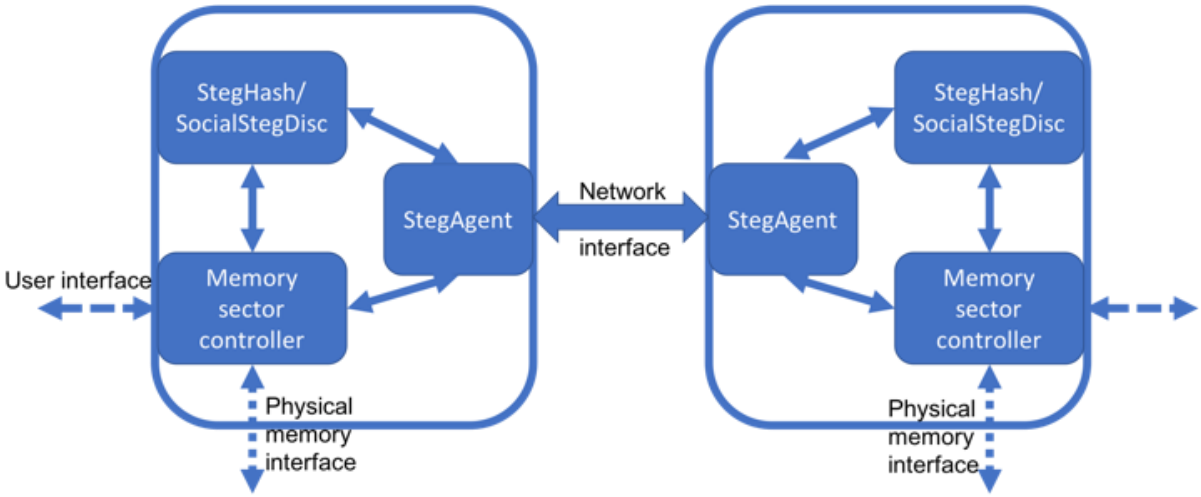

Figure 4 Components of the system and their interaction scheme

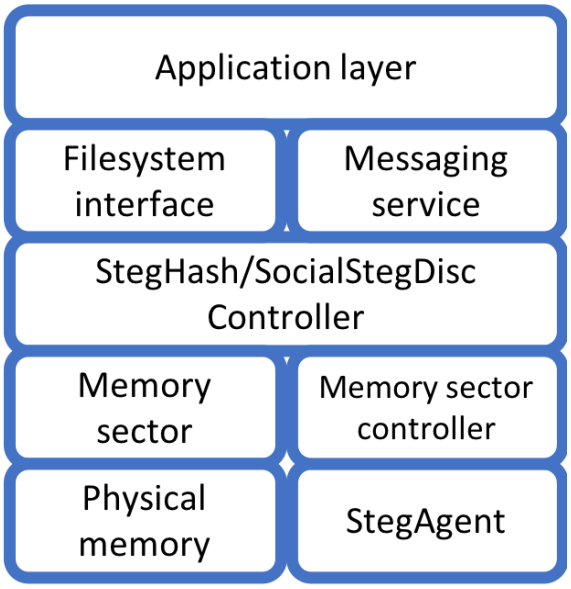

Figure 5 The layer view of the system

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