Journal of Theoretical and Applied Information Technology
31st August 2022. Vol.100. No 16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SYSTEM OF A SELF-ORGANIZING VIRTUAL SECURE COMMUNICATION CHANNEL BASED ON STOCHASTIC MULTI-LAYER ENCRYPTION AND OVERLAY TECHNOLOGIES

[1]E.A.BASINYA, [2,3]ZH.B.AKHAYEVA, [2]A.B.ZAKIROVA, [4]D.ZH. OMARKHANOVA, [2,3]G.B.TOLEGENOVA, [2]B.K.ABDURAIMOVA, [5]L.ALDASHEVA, [2]ZH.A.ZHANAYEVA

[1] National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, Russia

[2] L.N.Gumilyov Eurasian National University, Department of Information Systems, Astana, Kazakhstan

[3]Higher School of Information Technology and Engineering, Astana International University, Kazakhstan

[4] S.Seifullin Kazakh Agro Technical University, Department of Information Systems, Astana, Kazakhstan

[5]Astana IT University, Department of Information Systems, Astana, Kazakhstan

E-mail: [1]basinya@mail.ru, [2,3]ahaeva07@mail.ru, [2]alma_zakirova@mail.ru, [4]dinara.omarkhanova@mail.ru, [2,3]gulnaztolegenova@mail.ru,[2]aikosha_01@mail.ru,[5]aldasheva_ls@enu.kz, [2]zhanayeva.zhylduz@gmail.com

ABSTRACT

The paper considers the problem of ensuring information security of information flows in computer networks operating on the basis of the TCP/IP protocol stack (Transmission Control Protocol/Internet Protocol). one of possible solutions for providing the data transfer protection, proposed in this paper is to use a combination approach to dynamically build encapsulated virtual network tunnels using onion and garlic routing, and additional encryption layers. The algorithm of tunnel self-organization based on the author's modification of the port knocking technology, taking into account metadata about the previous connection (route tracing, duration of interaction, the sequence of port enumeration, protocols and other control information, including the client's computing device) is described. An analysis of the conceptual vulnerabilities of TOR (The Onion Router) and I2P (invisible internet project) overlay networks is made, and possible approaches to their elimination are reviewed. The experiment in identification of unauthorized access to transit traffic on the output nodes of the TOR network is described. Under the experiment " traps" were used in the form of sending authentication data of own test servers through all active outgoing nodes in an open form , the redirection from top 100 _ foreign sites to other information resources (including falsification of DNS <Domain Name System> replies) was tracked . The duration of the experiment was two years. An analysis of the dynamics of the development of TOR network is further described, and the conclusions about the credibility of these networks are outlined. The results of the developed and software - implemented solution are described, as well as the results of testing. In conclusion, findings and recommendations on the use of the proposed system with various parameters are presented.

Keywords: *Traffic Management, Network Attacks, Virtual Secure Communication Channels, Overlay Networks, Onion Traffic Routing, Garlic Traffic Routing, Traffic Analysers, Multilayer Encryption, Encapsulation Of Virtual Tunnels, DPI, TOR, I2P.*

1. INTRODUCTION

One of the main trends in world development has become the fundamental and applied scientific research aimed at obtaining new knowledge in the field of information and communication technologies. Stack TCP / IP protocols is an integral technical part of the operation of most companies and government agencies connected to the global network Internet. At the same time, the spectrum of cyber risks and threats is rapidly developing. The attackers aim at

a personal enrichment, and a number of states are connected with unscrupulous activities to obtain total control over the global information interaction. Mandatory integration of root certificates on user devices, implementation of automatic and automated analysis and decryption of traffic, artificial legal restrictions on bit depth and civil encryption algorithms, pressure on software vendors for integrating backdoors (tools using for hidden operating and collection statistics/metadata) are just a small part of the threats.

Ensuring information security of network interaction is becoming a high priority of the scientific community. It should be noted that the stack TCP / IP protocols has a wide range of vulnerabilities: from the possibility of falsification address space, the imposition of false routes to the substitution of sessions. Specialists try to neutralize these problems with intellectual functions of managed network equipment: by profiling an access with the basis on a bunch of ports, MAC addresses and IP addresses (port security and ip-mac binding), access control lists (English ACL, Access Control List), authentication and authorization servers (RADIUS) and etc. However, the situation is aggravated by errors in software products (operational systems and applied solutions). Developers do not always promptly fix bags in products. Vulnerability CVE-2016-5195 have been existing in the Linux kernel for more than 10 years from the first mention. It allows any unprivileged user to get full access to the system, which is associated with the processing of the copy-on-write (COW) algorithm by the subsystem of kernel memory. Vulnerability of IPsec (IP Security) tunnel CVE-2013-4350 survived for several years without providing SCTP encryption traffic (when the following options are selected: AH + ESP + IPv6). It is worth noting that most instances of managed network equipment are implemented on hardware base of MIPS (Microprocessor without Interlocked pipeline Stages) or ARM (Advanced RISC Machine) microprocessors. Consequently, there are not always fresh Linux kernels with integrated patches to fix recent vulnerabilities.

Ensuring confidentiality and anonymity of network interaction formed the trends of scientific research on the use of complex encryption at all levels of the TCP / IP protocol stack and the development of overlay technologies, functioning on the basis of the existing stack.

This issue is investigated by Russian and foreign scientists: Avdoshin S.M., Lazarenko A.V., Pavlenko E.Yu., Moskvin D.A., Reshetov D.V., He Y., Li X., Salcedo P., Jansen R., Chen C., Danezis G., Goldberg I., Haraty R., Assi M. and other colleagues. [1-8].

It is worth noting the high practical significance of the works presented researchers. Unfortunately, the proposed approaches are not focused on protection against global watcher attacks and malicious insider actions by providers, and also do not take into account the imperfection of the used software.

The key research topic is the solution of the problem of secure network communication in the global network Internet. The area of research covers the convergence of circuit and packet switching technologies, overlay technologies operating on the basis of the existing TCP/IP protocol stack. It should be noted that anonymization technologies are considered as an integral tool for ensuring the security of information resources in remote communication.

The study includes methods and technologies for managing information flows in computer networks operating on the basis of the TCP / IP protocol stack version 4. The technologies of proxying, cascading, virtual secure communication channels are considered. Traffic masking technologies that can eliminate the risk of identifying the fact of secure communication of objects by automatic and automated traffic analysis tools on the provider's side, including on backbone communication lines are analyzed. The study includes the hiding of virtual private network (VPN) traffic under the HTTPS (HyperText Transfer Protocol Secure) protocol or the extensible messaging and presence protocol (XMPP - eXtensible Messaging and Presence Protocol).

It is worth highlighting the limitations of the study: within the framework of multilayer encapsulation of information flows using onion routing, datagram transmission based on the UDP transport layer protocol (User Datagram Protocol) is not considered. Instead, a protocol with the establishment of a logical connection TCP (Transmission Control Protocol) is used. This limitation is due to the algorithmic limitations of the TOR network. But within the framework of composite combinations to promote traffic without onion routing, the UDP protocol can be used. Therefore, instead of Socks 4, which works only with TCP connections, Socks 5 is used with support for TCP and UDP, login and password

authorization, and the ability to remotely resolve symbolic names (DNS queries).

The sixth version of the TCP/IP stack is not considered. However, the solution proposed in the work can be ported from the fourth to the sixth version of the stack without any problems. Relay cascading technologies are also beyond the scope of the study, the speed of rebuilding the traffic promotion chains of which is more than two seconds, as well as the average throughput of which does not exceed 1 Mbps.

2. PURPOSE OF THE STUDY

The purpose of this work was to study the existing overlay technologies and virtual secure communication channels, development and implementation the original system for ensuring the information security of the network interaction of devices in the global Internet. During the research, the final task was narrowed down to the development and implementation of a system of self-organizing virtual secure communication channel based on stochastic multilayer encryption and overlay technologies.

As part of the decomposition of the overall goal using a multi-level approach, key tasks should be identified:

1) research of the subject area;
2) development of an algorithm for a self-organizing virtual secure communication channel based on stochastic multilayer encryption and overlay technologies;
3) designing an original system for ensuring information security of network interaction of devices in the global Internet, functioning on the basis of the proposed algorithm;
4) software implementation of the system;
5) testing and research of the proposed solution.

It should be noted that the development of the software product was carried out in accordance with a comprehensive methodology for the organization and maintenance of the technological pipeline for the development of DevSecOps software and hardware-software products.

3. THEORY AND PRACTICE

Currently, there are various ways to provide information interworking security. As noted earlier, one of the leading trends in the development of the information technology industry is the formation and improvement of overlay technologies. There is a wide range of projects worth considering: RetroShare, Morphmix/Tarzan, Mixminion/Mixmaster, JAP, I2P, MUTE/AntsP2P, Haystack , TOR, StealthNet, giFT and many others.

Today, the most common overlay networks are based on onion and garlic routing principles, namely TOR (The Onion Router) and I2P (Invisible Internet Project).

TOR is a distributed system of proxy servers, providing data transmission in encrypted form through virtual tunnels. According to the onion routing mechanism, the transmitted data after pre-encryption with three different keys pass through three TOR nodes, chosen randomly. Each node finds out the destination address of the transmitted data fragment only after removing the corresponding layer encryption. The distributed anonymous I2P overlay network is based on combined tunnelling principle: outgoing tunnel - established by sender - is intended to deliver datagrams to the sender, and an incoming tunnel – created by receiver - is used to deliver datagrams from the sender. Transferred data are encrypted on the sender's side and decrypted on the recipient's side. This chain is one-way. Thus, the identification of the participants in the interaction is excluded.

The main vulnerabilities of overlay technologies are:

1) datagram time delay correlation. It is implemented by an attacker when accessing the traffic of an intermediate network segment. Deliberate delay of datagrams is introduced to establish correlation, delimitation and personification of information flows. An additional tool may be an attack on the deviation of timestamp TCP;
2) identification of connections to these networks. Most of the IP addresses of the interconnect hosts (nodes) as well as the parent servers are in public access. In order to bypass the blocks, you have to resort to the use of intermediaries («bridges»);
3) automodel of network traffic. Clients of different types of overlay networks make homogeneous fragmentation of packages, which allows the offender to uniquely identify the technology used through tools of probability theory and mathematical statistics;
4) attacks by a global observer with passive surveillance, impersonation, hacking into secure communications. The probability of successful

implementation is proportional to the proportion of monitored/observed network segments;
5) imperfection of software implemented by enthusiasts.

Stochastic delays, datagram fragmentation with traffic masking, and generation of "white" noise are one of the simple solutions to the described problems.

In addition to the development of overlay technologies, one of the fundamental trends is the software implementation of virtual private VPN (Virtual Private Network) technologies, which involve the creation of logical networks on top of existing ones using encryption mechanisms, authentication, etc. Considering the implementation of VPN technology, the following solutions should be mentioned: n2n, SoftEther VPN, WireGuard, FreeLAN, GoVPN, OpenVPN, IPsec, GateVPN, Tinc, and others.

It is worth noting a number of key features inherent to each of them when analysing these products. SoftEther VPN differs from its counterparts in having support for a variety of VPN protocols, thus providing the ability to use a single application to work with clients with different operating systems. The fundamental characteristic of FreeLAN, Tinc, n2n is the ability to choose a client-server, peer-to-peer (P2P), or mixed tunnel architecture. GoVPN offers three modes of operation: the standard one that sends encrypted data, the noise one that completes packets with noise, and the constant speed mode that completes the noise mode by sending packets at a specified time interval. The IPsec protocol is a tried and tested civil crypt-resistant solution, but has less flexibility than OpenVPN technology, which is characterized by stable performance at two or more NAT (Network Address Translation). WireGuard shows higher performance and better bandwidth than OpenVPN and IPsec, while being easier to configure.

These technologies are most often prone to attacks of under-encryption level, false authentication/authorization, deanomization, intermediary attacks, software errors and algorithms of «hard logic» behaviour.

Accordingly, the task was to neutralize the potential damage from:
1) previously described vulnerabilities;
2) tools for active and passive traffic analysis (including automatic, automated analysis systems and decryption of traffic with using DPI <Deep packet inspection> technology);
3) software errors;
4) backdoors (intentionally integrated tools for indoor system management and hidden information gathering/sending);
5) vulnerabilities in encryption algorithms that have not previously been revealed to the public.

In order to ensure the information security of network interaction devices in the global Internet, it was decided to develop and implement a system of a self-organizing virtual secure communication channel based on stochastic multilayer encryption and overlay technologies, meeting the stated requirements.

Block diagram of the self-organizing virtual secure channel algorithm communications based on stochastic multilayer encryption and overlay technologies is presented in Figure 1.

A system functioning on the basis of the proposed algorithm involves the sequential execution of a number of operations. First of all, metadata on previously established connections are analysed. Further, if the algorithm is run for the first time, the private key is generated depending

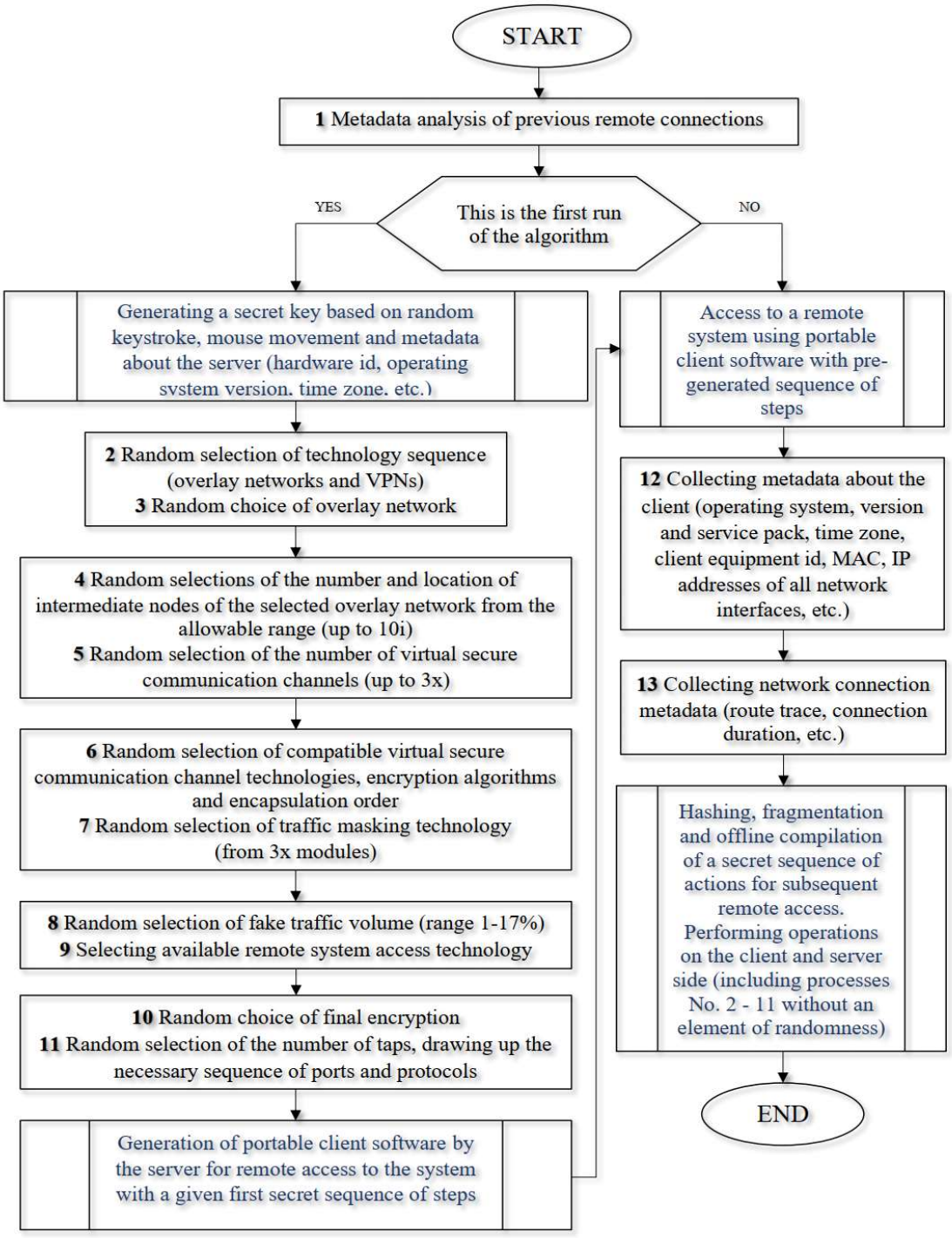

Figure 1: Algorithm Of A Self-Organizing Virtual Secure Communication Channel On Based On Stochastic Multilayer Encryption And Overlay Technologies

on a number of software and hardware parameters, and the sequence is randomly selected, according to which overlay networks and virtual secure communication channel technologies are used. After performing these steps, there is also a random selection of an overlay network with the subsequent determination of the number and location of its

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

intermediate nodes. In addition, the following parameters are randomly selected: the number of virtual secure communication channels, the technologies of virtual secure communication channels and encryption algorithms, methods of traffic masking and the volume of fake traffic. The next step is to select the available remote network access technology and the final encryption algorithm, as well as to determine the number and sequence of «taps» required for access to the system. At the end of the secretive sequence generation, the server releases the client software to provide remote network access using described sequence. The steps described are missing if the algorithm has already been run before - connection to the remote system is performed with pre-generated parameters. After remote access, metadata about the client and the network connection are collected with hashing and signed fragmentation. The sequence of steps for the next remote access session is generated by the signature approach on the client and server side.

It is worth noting that overlay networks and technologies of virtual secure communication channels are selected taking into account compatibility. An example is the TOR network, which only works with the TCP (Transmission Control Protocol) transport protocol and does not support the UDP (User Datagram Protocol).

Also, distinctive feature of the presented algorithm is the use of various encryption algorithms from the list of acceptable in the case of a match of these algorithms in the selected technologies in Figure 2.

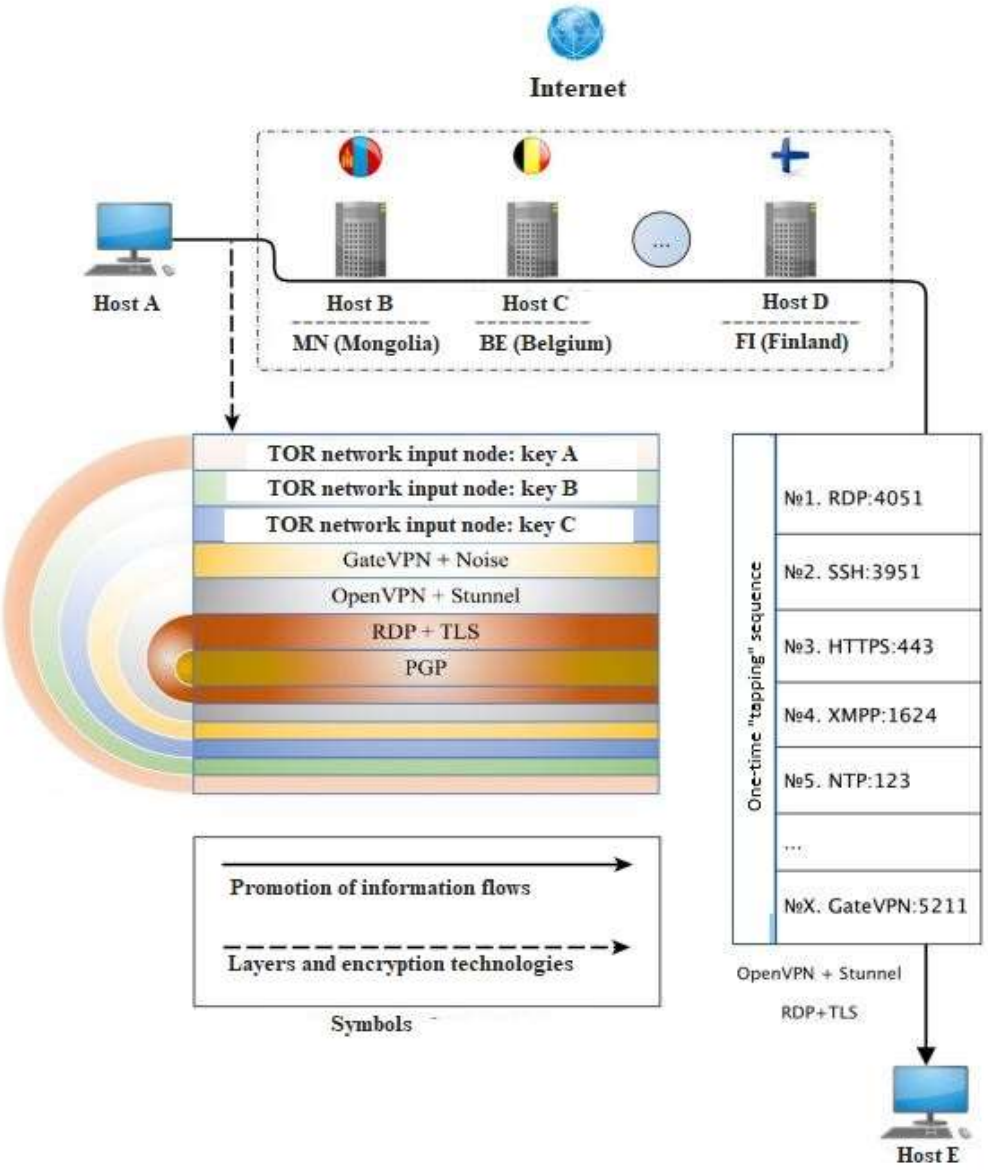

Figure 2: An Example Of A System Of Self-Organizing Virtual Secure Communication Channel Based On Stochastic Multi-Layer Encryption And Overlay Technologies

4. RESEARCH RESULTS AND DISCUSSION

To assess the level of trust in the TOR overlay network, an experiment was conducted for two years to identify unauthorized access to transit traffic on network output nodes. The experiment used «traps» in the form of sending authentication data of own test servers through all active outgoing nodes in open access via the hypertext transfer protocol HTTP (HyperText Transfer Protocol). Server solutions were isolated model objects in a hypervisor virtualization environment. The software implementation was done in Python. During the two years of the experiment 5 successful intruders were made to the system under the sent data. The hacker used VPNs to achieve anonymity. At the same time, it should be noted that the leak of information through other communication channels is excluded, no attacks on the server were not carried out (including brute force). Accordingly, it is possible to assume the malicious actions on the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

part of the owner of one of the output nodes of the TOR network (or Internet providers from this node to the model object). To confirm this hypothesis and identify a specific culprit, 1000 accounts (with a reserve for the number of nodes) were created on another model object to fix the «traps» and the output node. At the same time, hidden redirection from the top 100 Russian and foreign sites (according to Google and Yandex search engines) to other information resources (including falsification of DNS responses) is tracked. Redirect facts has not yet been recorded. The experiment continues.

The previous support of the US National Security Agency undermines the credibility of this network. Another important argument is network development dynamics: the number of exit nodes has increased by only 263 nodes since 2015 (Figure 3).

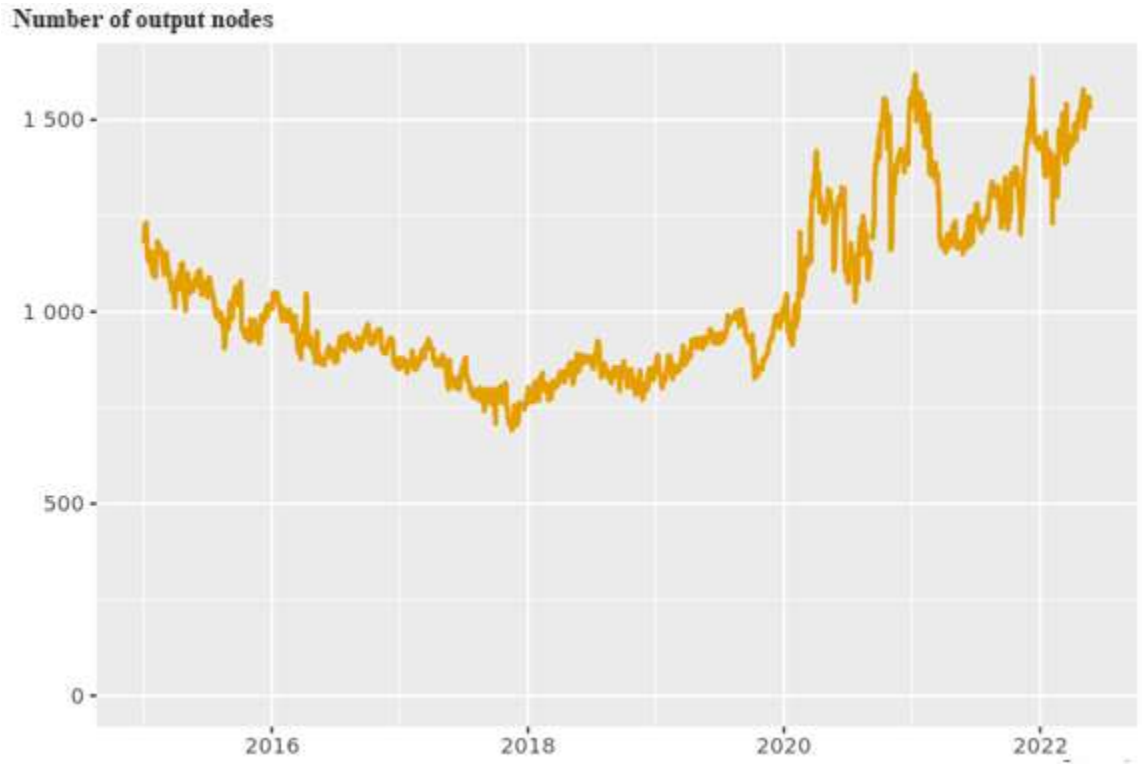

Figure 3: Dynamics Of Changes In The Number Of Exit Nodes Of The TOR Network From Mid-2015 To May 2022

Cost of VPS (Virtual Private Server) or VDS (Virtual Dedicated Server) to organize such an output node is $2, which puts a significant question about the weak development of the infrastructure of this network and the potential possibility of concentrating most of its resources in "one hand". At the same time, it is necessary to once again express respect to its developers for the knowledge-intensive decision.

These shortcomings, as well as the previously described vulnerabilities of the used technologies, are offset by the developed and software implemented system of self-organizing virtual secure communication channel on the basis of stochastic multilayer encryption and overlay technologies. To check the functionality of the proposed solution, manual and automated testing was carried out. Further, the information flows of interacting servers were iteratively studied by the tools of passive and active traffic analysis (sniffers, network analysers, systems of automatic and automated analysis and interpretation of traffic) on the proposed system. The following software products were used: Wireshark, Trisul, CommView, SpyNet, Linux, sniffer, decryptor, self-similarity, etc. Traffic was identified only by the top-level protocol (and its encryption layer) depending on the sequence of

encapsulation technology, performed by the algorithm. For example, the sniffer identified the HTTPS protocol with almost successfully masked VPN traffic. Accordingly, the result demonstrated the successful achievement of the purpose of the research. The shortcomings of the system include a rather low transmission rate, with an average of 850 kbit/s. First of all, this is due to the limitations of chains of overlay networks, where each intermediate host has its own quotas for the bandwidth capacity. Taking into account the number of encapsulated virtual communication channels with multilayer encryption, it is necessary to note the high requirements of the system to the computing resources of the end hosts: an x86 platform is required with the extension of the AES command system (Intel Advanced Encryption Standard New Instructions; AES-NI) and 2 GB of RAM.

5. CONCLUSION

During the research, a study of existing overlay technologies and virtual secure communication channels was conducted. Conceptual vulnerabilities and possible approaches to their elimination were highlighted. An experiment to identify unauthorized access to transit traffic on the output nodes of the TOR network is described. As a result, malicious actions of the owner of one of the output nodes of the TOR network (or Internet providers from this node to the model object) were detected. The experiment continues to this day to identify the culprit. The analysis of dynamics of development of TOR network have been made, the conclusions about impossibility of trust to this network in the initial form have been drown.

The scientific novelty of the work and its contribution to the development of the subject area of information technology lies in the proposal of a new algorithm for a self-organizing virtual secure communication channel based on stochastic multilayer encryption and overlay technologies that ensures the security of network interaction between devices in the global Internet. Unlike existing methods, the risk of global observability and the possibility of identifying the fact of secure communication of objects by automatic and automated traffic analysis tools on the provider's side, including those on backbone communication lines, are excluded.

Methods of deanomization (disclosure of anonymity) of users, presented by Pavlenko E.Yu., Moskvin D.A., Reshetov D.V., Avdoshin S.M., Lazarenko A.V. cannot be applied to the developed algorithm and system, because the initial selection will be incorrect due to the use of masking information flows under third-party protocols. Additionally, it is worth noting that stochastic delays and traffic fragmentation will eliminate attack timing and communication identification through onion routing, which will bypass the rules of analysis by the methods of He Y., Li X., Chen M., Wang W.

The proposed solution improves previously known methods (including Salcedo P., Haraty R.): the reliability and fault tolerance of communication is increased due to the multilayer encapsulated combination of virtual secure communication channels and overlay technologies with remote network access protocols. Even if one of the layers is compromised (including the presence of a backdoor or a tool for decrypting traffic), this will not have a significant impact on the functioning of the communication channel as a whole, it will continue to operate normally while ensuring the confidentiality of the transmitted data.

Unlike the existing methods proposed by the authors of Chen C. et al., the throughput of the communication channel will not be less than 1.5 Mbps due to the preparation of redundant communication lines, balancing and optimizing the load of the communication channel when building multiple chains.

Modification of port tapping with autonomous compilation of secret sequences of actions on the client and server side based on meta-information and secret parameters with further hashing and signature matching allows to eliminate the shortcomings of the known anonymization systems described by Haraty R., Assi M., Rahal I.

But it is worth highlighting the shortcomings and limitations of the proposed solution: within the framework of multilayer encapsulation of information flows using onion routing, datagram transmission based on the UDP transport layer protocol is not considered (due to the algorithmic limitations of the TOR network). The methods of cascading repeaters, the speed of rebuilding the traffic promotion chains of which is more than two seconds, and also the average throughput of which does not exceed 1 Mbps, were not considered and applied. The use of such technologies could increase the level of information security, but would make the use of the system extremely non-user-friendly for ordinary users, limiting them, among other things,

in terms of functionality and the use of a wide range of protocols. The sixth version of the TCP / IP stack was not affected either. However, the solution proposed in the work can be ported from the fourth to the sixth version of the stack without any problems.

This software was tested manually and automatically, and the system was successfully tested with passive and active traffic analysis tools. This system can be used to protect the data transfer process between servers in the global Internet, assembled on the x86 platform with the availability of an extension of the AES command system (Intel Advanced Encryption Standard New Instructions; AES-NI) and 2 GB of RAM. Hardware support of the AES algorithm is a limitation on the application of the method, which leads to an increase in the cost of complete sets for the hardware configurations of the object. It should be noted that these are small costs in the context of the task of ensuring secure and anonymous communication.

The work submitted for review was used as one of the modules ensuring network information security of the intelligent adaptive system for managing the network infrastructure of enterprise, developed by the author.

REFERENCES:

[1] Avdoshin S.M., Lazarenko A.V. Texnologiya anonimny`x setej // Informacionny`e texnologii. 2016. V. 22. № 4. pp. 284-291. (in Russian)

[2] Pavlenko E.Yu., Moskvin D.A., Reshetov D.V. Raskry`tie anonimnosti pol`zovatelej seti Tor metodom analiza potokov peredachi danny`x // Problemy` informacionnoj bezopasnosti. Komp`yuterny`e sistemy`. 2015. no. 4. pp. 13–16. (in Russian)

[3] Avdoshin S.M., Lazarenko A.V. Metody` deanonimizacii pol`zovatelej Tor // Informacionny`e texnologii. 2016. V. 22, no. 5. pp. 362–372. (in Russian)

[4] He Y., Li X., Chen M., Wang W. Identification of Tor Anonymous Communication with Cloud Traffic Obfuscation // Gonchcheng Kexue Yu Jishu/Advanced Engineering Science. 2017. Vol. 49(2). P. 177–194.

[5] Bandwidth optimization in tor nodes through frame configuration in hidden services / Salcedo P. [and oth.] // Journal of Engineering and Applied Sciences. 2017. Vol. 12(13). P. 3338–3344.

[6] Jansen R., Johnson A. Safely Measuring Tor // In Pro- ceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016, pp. 1553–1567.

[7] High-speed Onion Routing at the Net- work Layer / Chen C. [and oth.] // In Proceedings of the 22nd ACM SIGSAC Confer- ence on Computer and Communications Security, Denver, CO, USA, October 12-6, 2015, pp. 1441–1454.

[8] Haraty R., Assi M., Rahal I. A Systematic Review of Anonymous Communication Systems // In Proceedings of the 19th International Conference on Enterprise Information Systems - Volume 2: ICEIS, ISBN 978-989-758-248-6, pages 211-220.