CYBERSECURITY AND CRITICAL INFRASTRUCTURE PROTECTION

DOI 10.20535/2411-1031.2020.8.2.222612
UDC 004[413.3+738.5.057.4]

VITALII ZUBOK

DETERMINATION OF COMPONENTS OF ROUTE HIJACK RISK BY INTERNET CONNECTIONS TOPOLOGY ANALYSIS

The possibility of dynamic routes change between nodes that are not physically connected is a key feature of the Internet routing. The exterior gateway protocol BGP-4 has been developed to deliver this feature, along with policies and procedures of inter-domain routing. Developed for the network of hundreds nodes that rely on information from each other, after decades BGP-4 is still the same with tens thousands nodes and its crucial lack of routing data integrity. One of the most significant problems deriving from its weaknesses is route leaks and route hijacks. None of the proposed and partially implemented upgrades and add-ons like MANRS and RPKI can not deliver reliable defense against those types of attacks. In this paper, the approach of risk assessment via internetworking links analysis is developed. Although modern information security is based on risk management, in this paper it is proposed to mitigate route hijack risks by enhancing links topology. Estimating the risks of route hijack requires quantitative measurement of the impact of an attack on the routing distortion, and therefore, the loss of information security breach. For this assessment, this paper proposes to use knowledge of the features of the Internet topology at the layer of global routing, which is determined by the interaction of autonomous systems - groups of subnets under common control - according to the routing protocol BGP-4. Based on our formal representation of IP routing, the relationship between topology and the risk of route hijack is shown. A new approach to quantifying information risk using a new risk-oriented model of global routing, which will reflect the properties of Internet nodes in terms of the risk of routes hijack.

Keywords: global routing, Internet, route hijacking, routing model, cybersecurity, risk assessment.

Problem statement. Nowadays there are over 80000 nodes called Autonomous Systems (ASes) interconnected in some way and thus building the telecommunication network – the Internet [1]. Such a big figure of transit nodes and even much bigger number of links moves us from the theory of graphs to the theory of complex networks, where the study of the general properties of topology is preferred to the study of specific connections between nodes. [2], [3]. This is the start point of route forges, route hijacks and other frauds with global impact each [4].

There are proposed proactive mechanisms such as Resource Public Key Infrastructure (RPKI) [5]. It's part of the Internet Routing Registry system. This service provides a collective method to allow one network to filter another networks routes. The method begins with cryptographic signing the route origin. A Route Origin Authorisation (ROA) is a cryptographically signed object that states which AS are authorized to originate a certain prefix. A ROA contains three informational elements: the AS Number that is authorized, the prefix that may be originated from the AS, and the maximum length of the prefix. However, such techniques are fully effective only in global deployment, and operators are reluctant to deploy them because of the associated technical and financial costs. For example, Telia, one of the Tier-I Internet backbone operators, announced that it's using RPKI for security in its internet routing infrastructure since only September 2019.

In the face of the impossibility of reliable protection against damage associated with an attack, it is necessary to learn how to manage risks arising from cyberattacks on global routing. For this purpose, we must use well-studied topological peculiarities of the Internet to find methods of routing attacks mitigation by a forehead improvement of the connections between Internet nodes.

© V. Zubok, 2020

Analysis of recent research and publications. Anti-hijack protection consists of two steps: detection and mitigation. RPKI mechanism with route origin validation is not sufficient to mitigate AS hijacking. An analysis of the mechanisms of the attack, depending on its objectives and options for its implementation is described in detail in [6]. Detection is mainly provided by third-party services such as BGPMon. They notify the network administrator of suspicious events related to their prefixes based on routing information. They track worldwide routes by tracing and keep track of route announcements in BGP. In the event of an incident, the affected networks begin to mitigate the consequences of the event, for example by announcing more specific prefixes to their networks or by requesting other ASs to filter out false announcements. There are some other studies, which offer mechanisms for route attack detection such as ARTEMIS [7] and Peerlock [8].

However, due to the combination of technological and practical deployment issues, existing reactive approaches are largely inadequate. In particular, the most advanced technologies have the following major problems:

- the variety of types of routing attacks and combinations of methods lead to the lack of a reliable method for detecting route interception;
- operators should be informed in advance of legitimate changes to their routing policy (new interactions between AS, announcement of a new prefix, etc.) so that such changes are not considered suspicious events for conditional third party detection systems. Otherwise, adopting a less rigorous policy to compensate for the lack of updated information and reducing the number of false positives carries the risk of neglecting real events and not detecting false negatives;
- only a few minutes of unauthorized traffic diversion can result in heavy financial losses due to the unavailability of service or security breaches. At the same time, the response time to incidents is slow in any case, as a current practice requires the need to manually check alerts coming from monitoring systems and third-party services. The duration of widely known incidents ranged from several hours to months.

At the risk identification stage of the risk assessment process, specific requirements for the quality of information are raised. There is a requirement of the highest possible level of completeness, accuracy, and conformity at the time of its receipt. Quality requirements are also raised to the quality of information sources [9].

The aim of this paper is, assuming the information above, to develop the methodology of finding the relationship between topology and routing vulnerability to obtain certain methods for quantifying information risk using a formal global routing model.

The main material of the research. Briefly overview the *basics of global Internet routing and the nature of route hijack.* Existence of links between Internet nodes is determined by existence of border interaction between groups of network communication equipment. With relation to border interaction, we suppose these groups as a node, or as an autonomous system. An Autonomous System (AS) is a group of IP networks having a single clearly defined routing policy that is run by one or more network operators. ASes exchange routing information with other ASes using Border Gateway Protocol BGP-4. Exterior routing decisions are frequently based on policy-based rules rather than purely on technical parameters [10]. A model of 4 AS interconnection is represented in Fig. 1.

Each AS provides network prefixes to which it is ready to accept traffic, to a connected AS (it is called peer AS). So, AS4 has peering with AS3 using link *d* and announces it's prefixed to AS3. It means that AS3 now "knows" at least one way to transfer packets addressed to networks in which prefixes are announced by AS3. At the same time, AS3 announces to AS4 its prefixes too. As it's shown in Fig. 1, AS3 also is peering with AS1 and AS2 using links *b* and *c*. Due to this, AS3 can reannounce AS4 prefixes accepted from AS1 and AS2, and vice versa, reannounce AS1's and AS2's prefixes to AS4. This ability comes from gateway protocols features, and its presence is subject to a *routing policy*. Also, we can see that AS1 and AS2 are peering (a), so in the AS1-AS2-AS3 triangle, they can be a transit node for each other. However, AS4 has only one peer and it can't provide any transit. It's called "stub" AS.

Pretend AS3 and AS4 are not linked, however AS3 by misconfiguration or maliciously announces to AS1 and (or) AS2 prefixes originated by AS4. Due to the lack of integrity inherit to BGP-4, AS1 and AS2 have no mechanisms to automatically verify and authorize those routes. More complex network of ASes is shown in Fig. 2. AS6 is legitimate origin for 12.34.0.0./16 route, however, if AS1 also announces this route, even in such easy network map we can see the nodes (AS2 and AS3) which accept this route as the best (shortest) path. Being aware that according to [BGP RFC] each BGP system can announce only one path –the best, i.e. shortest route for each prefix, we understand that AS2 and AS3 will use and propagate forged route to all their peers.

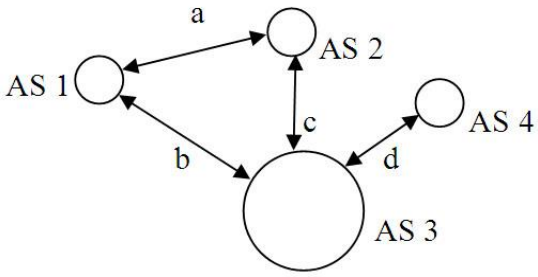

Figure 1 – Autonomous systems interconnection:
AS 1,2,3,4 – autonomous systems; a, b, c, d – links

And let's look at Fig. 3 where the route hijack has become a prefix hijack due to (erroneous or malicious) deaggregation of 12.34.0.0/16 prefix to more specific 12.34.0.0/17 + 12.34.128.0/17, in result all other nodes will not use the route to 12.34.0.0/16 because of the existence of more specific ones. In this case affected ASes will not stop to announce legal route to whole prefix 12.34.0.0./16, although it can be used only if more specific /17 prefixes are not accepted by some AS for any reason.

When (or rather "if") the RPKI is implemented on 100% Internet providers including the biggest Tier I networks, such hijack will not be possible due to route origin validation procedure, complementary to global routing. But there's nothing to counteract to man-in-the-middle attack with AS path forgery, when origin keeps looking valid (Fig.4). Any ideas of registering and validating complete set of legitimate Internet routes do not look realistic neither now nor in the future. That's why we suppose global routing will be vulnerable for a long time.

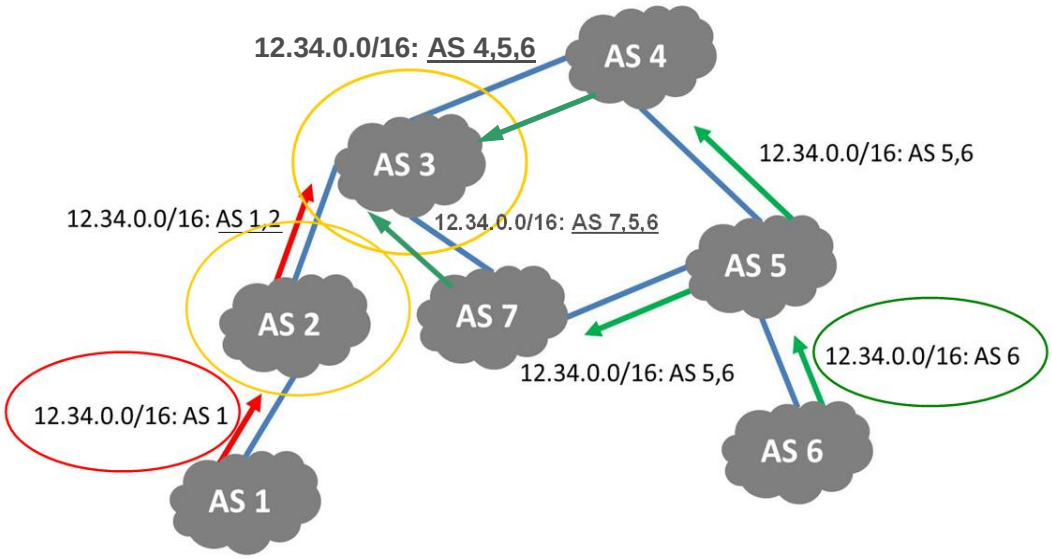

Figure 2 – AS1 performs hijack of the route to 12.34.0.0/16 belonging to AS6

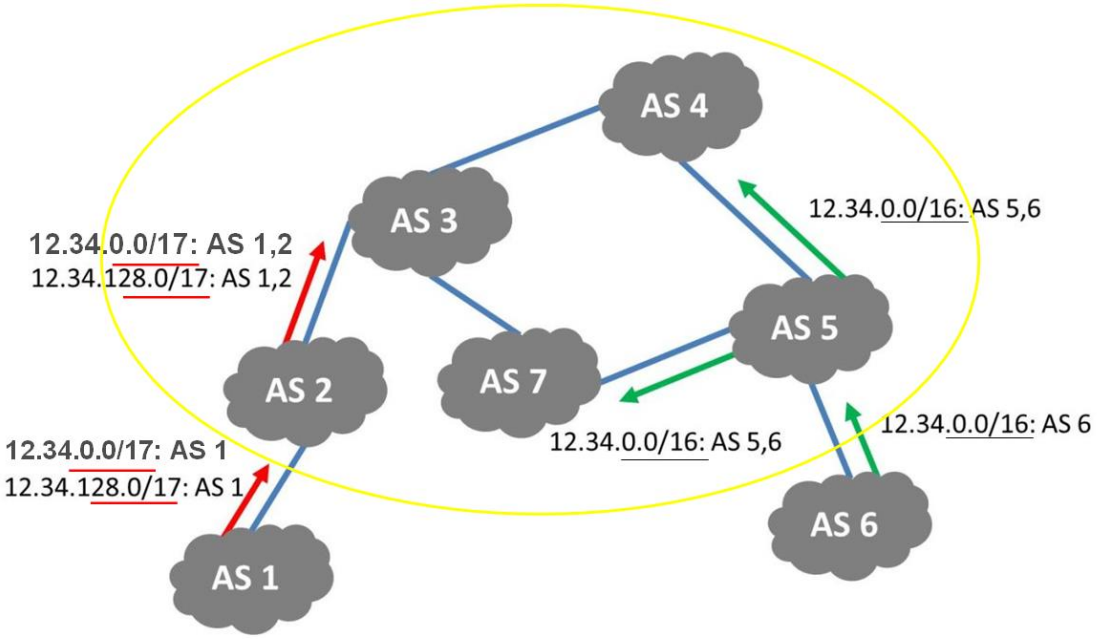

Figure 3 – AS1 uses deaggregatin to hijack the route to 12.34.0.0/16 belonging to AS6

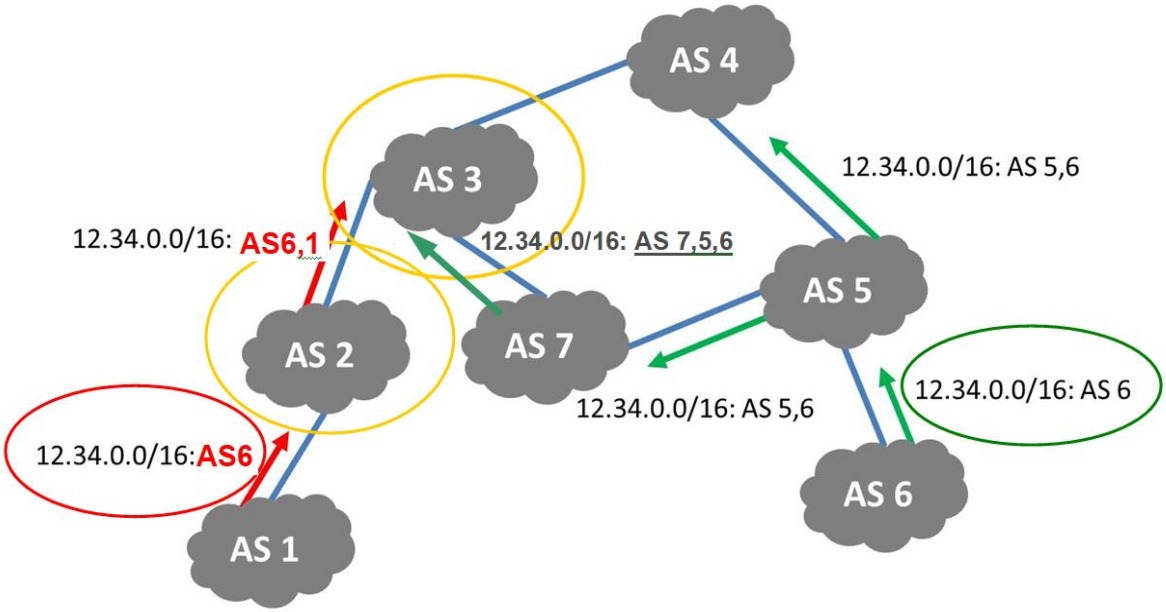

Figure 4 – AS1 forges route origin while AS2 is not using appropriate BGP filtering

Authors of major *Internet topology studies* have confirmed that the Internet is typically to the naturally created networks has scale-free topology [3]. The main feature of this topology is power-law degree distribution of nodes:

$$P(k) \sim 1/k^{\gamma},$$

where k – number of links, or "power" of random node;

γ – scaling exponent, $2 < \gamma < 3$;

$P(k)$ – node degree distribution, i.e. probability that a randomly selected node has k links.

Scale-free networks are characterized by the presence of nodes with lots of relationships (called *hubs*), while the average number of connections per node is relatively small. In this sense, such networks are called scale-free. Measurements have shown that any network that develops and in which the number of nodes grows, eventually turns into a scale-free network. The reason is that

the older nodes in the network have more opportunities to receive links to new nodes at their formation, as the Barabasi-Albert model shows. For the Internet, this phenomenon is familiar to the average user on an example of World Wide Web (WWW): the longer web resource stays in WWW without changing its address (URL), the more references to it calculated in web search engines. Those two factors – the expansion of the network and the preferential connectivity of its new members – explain the existence of hubs.

Distance between nodes can be defined as the number of steps required to get from one node to another. Naturally, the nodes can be connected directly (via a single edge), or indirectly through several other nodes edges. The shortest distance between nodes is called the *path*. Some authors use the definition of "*geodesic path*". There is an *average geodesic path l* value for the whole network:

$$l = \frac{2}{n(n-1)} \sum_{i>j} d_{ij},$$

where d_{ij} – geodesic path between nodes *i* and *j*.

Due to hubs, the average path between random ASes on the Internet is between 3 and 4, although the number of ASes exceeds 80000. And there are metrics for Internet distance [11]:

$$d(v,u) = \min_i \left(d(v,i) + d(i,u)\right). \quad (1)$$

where i, u, v – random network nodes.

Then proceed from metric distance (1) to *formal description of route hijack.* Distance is the parameter routing attacks are tampering. From a practical point of view, this means that if route is hijacked only if the distance through the fictitious route will be less than through the real route. Then let's find the formula for affecting the node with forged route. The task of finding the best route is complicated and non-linear. Therefore, the TCP/IP stack has adopted the so-called one-step approach to optimizing the packet route (next-hop routing) - each router and destination node only have to choose one step forward of packet transmission. A formal description of the Internet global routing objects and processes is described in [12]. Here are formulated the process of choosing a prefix $p(a)$ by destination IP address and then choosing a route with the shortest path $\pi(p)$:

$$\begin{cases} p(a) = \{\min_j (p_j) \ : \ a \in p \subset A, \ 0 < j \le |A|\} \\ \pi_v(p) = \{\min_v (m_v(p)) \ : \ \pi \in M_p, \ v \in V_p\} \end{cases} \quad (2)$$

For common case, we assume that our network is connected, that is, at least one route to any prefix is known at each node. If there are two or more prefixes on a particular node *u*, BGP chooses one of them based on known criteria, the most important of which is path length. After that, this route is in use at this node, and will it be announced to neighboring nodes. If at some node two or more routes have the same path length, the decision will be made according to secondary criteria. After passing each transit node, the route is extended by 1 node.

Consider at this stage the case of hijacking a route without deaggregation. Pretend there is a prefix p_v legitimately originated from node *v*, and it is hijacked by a spoofed route $\pi'(p_v)$, which is created in the result of illegal, announce appeared to the network typically from one particular malicious node [4]. False route $\pi'(p_v)$ is competing with true route $\pi(p_v)$ according to the 2nd part of (2).

In Fig. 1, we can see that $\pi'(p_v)$ will capture the nodes AS2 and AS3. On the other hand, AS4 and AS7 will receive a false route $\pi'(p_v)$ but it will lose $\pi(p_v)$. These nodes will not pass it on to their other neighbors. In more complex topology we could see that on some hubs route hijack with initially one forged route can significantly increase the number of competing routes on some network hubs.

In more complex topology, we could see that on some hubs route hijack with initially one forged route can significantly increase the number of competing routes on some network hubs. In our opinion, the most plausible way to model route distribution is method of cellular automation. However, the forged route leads to information risk only in two cases:

a) if it changes the route of IP packets through malicious node;
b) if it changes the final destination of IP packets.

Let's *explore relation between distance and risk assessment*. As described below, likelihood of inequality $\pi'(p_v) < \pi(p_v)$ seen on particular node *u*, the more likely the bigger is $d(v,u)$. The extreme value $d(v,u) = 1$ leads to impossibility to provide forged route $\pi'(p_v)$ through the node *u*. So this should also eliminate for node *v* the risk of data loss on node *u*.

It is easier to manipulate the path length if the path is longer. In the long way in the middle, there are more nodes through which you can announce a forged route. Therefore, the probability *P* of interception between nodes *u, v* increases for distant nodes and decreases for close ones:

$$P(v,u) \sim d(v,u).$$

And also information losses increase with increasing number of affected nodes. So does the risk, and we reasonably assume that risk is proportional to distance:

$$R_v \sim \sum_{i=1}^{|V|} d(v,u);\ u \in V\ . \tag{3}$$

The expression (3) is relative quantity of route hijack risk for node *v* regarding the target group of network nodes *V*.

Conclusions. The most significant problem deriving from Border Gateway Protocol weaknesses and vulnerabilities is route leak and route hijack threats. An important step towards assessing the risk posed by attacks on global routing is to predict the impact of the attack, namely to assess the scale of the attack (distribution routes, impact area, number of "damaged" routes). Estimating the risks of route hijack requires quantitative measurement of the impact of an attack on the routing distortion, and therefore, the loss of information security breach. There is a relationship between the topology of the Internet and routing vulnerability. In this paper, formulated and proposed a new approach for quantifying information risk using a metric distance based on formal global routing model.

REFERENCE

[1] Internet Mapping and Annotation. Center for Applied Internet Data Analysis. [Online]. Available: https://www.caida.org/research/topology/internet_mapping/. Accessed on: June 28, 2020.

[2] M. Newman, "The structure and function of networks", *Computer Physics Communications*, vol. 147, iss. 1-2, 2002, pp. 40-45, doi: 10.1016/S0010-4655(02)00201-1.

[3] M. Faloutsos, P. Faloutsos, and C.Faloutsos, "On Power-Law Relationships of the Internet Topology", *Computers and Communications Review*, no. 29, 1999, pp. 251-263, doi: 10.1145/ 316194.316229.

[4] V. Zubok. "Retrospective Analysis of Cyber Incidents related to Attacs on Global Routing", *Modeling and Information Technologies*, iss. 86, 2019, pp.42-49.

[5] RIPE NCC's Implementation of Resource Public Key Infrastructure (RPKI) Certificate Tree Validation. [Online]. Available: https://tools.ietf.org/html/rfc8488. Accessed on: May 25, 2020.

[6] V. Zubok, "Metric Approach to Risk Evaluation of Cyberattacks on Global Routing", *CEUR Workshop Proceedings*, vol. 2318, pp. 251-260. [Online]. Available: http://ceur-ws.org/Vol-2318/. Accessed on: June 28, 2020.

[7] P. Sermpezis et al., “ARTEMIS: Neutralizing BGP Hijacking within a Minute”. [Online]. Available: https://arxiv.org/abs/1801.01085. Accessed on: June 27, 2020.
[8] T. McDaniel, M. Smith, and M. Schuchard, “Peerlock: Flexsealing BGP”. [Online]. Available: https://arxiv.org/abs/2006.06576. Accessed on: July 17, 2020.
[9] International Organization for Standardization. (2009, Nov. 13). ISO Guide 73, Risk management. Vocabulary. [Online]. Available: https://www.iso.org/standard/44651.html. Accessed on: Aug. 20, 2019.
[10] Y. Rekhter, and P. Gross, “RFC 1772. Application of the Border Gateway Protocol in the Internet”. [Online]. Available: http://tools.ietf.org/html/rfc1772. Accessed on: June 20, 2020.
[11] V. Mokhor, and V. Zubok, *Forming of Internode Connectionsin the Internet Using the Theory of Complex Networks*. Kyiv, Ukraine: Prometey, 2017.
[12] V. Zubok, “Formal Description of Global Internet Global Routing Objects for Assessing the Risks of Attacks on Global Routing”, *Data Recording, Storing and Processing*, vol. 21, no. 4, 2020, pp. 67-74, doi: 10.35681/1560-9189.2019.21.4.199409.

The article was received 10.08.2020.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Internet Mapping and Annotation. Center for Applied Internet Data Analysis. [Електронний ресурс]. Доступно: https://www.caida.org/research/topology/internet_mapping/. Дата звернення: Чер. 28, 2020.
[2] M. Newman, “The structure and function of complex networks”, *Computer Physics Communications*, vol. 147, iss. 1-2, pp. 40-45, 2002, doi: 10.1016/S0010-4655(02)00201-1.
[3] M. Faloutsos, P. Faloutsos, and C.Faloutsos, “On Power-Law Relationships of the Internet Topology”. *Computers and Communications Review.* №.29, 1999, pp.251-263, doi: 10.1145/ 316194.316229.
[4] В. Зубок, “Ретроспективний аналіз інцидентів кібербезпеки, пов’язаних з атаками на глобальну маршрутизацію”, *Моделюванняя та інформаційні технології,* том 86, 2019, с. 41-49.
[5] RIPE NCC’s Implementation of Resource Public Key Infrastructure (RPKI) Certificate Tree Validation. [Електронний ресурс]. Доступно: https://tools.ietf.org/html/rfc8488. Дата звернення: Трав. 25, 2020.
[6] V. Zubok, “Metric Approach to Risk Evaluation of Cyberattacks on Global Routing”, *CEUR Workshop Proceedings*, vol. 2318, pp. 251-260. [Online]. Available: http://ceur-ws.org/Vol-2318/. Accessed on: June 28, 2020.
[7] P. Sermpezis, et al., “ARTEMIS: Neutralizing BGP Hijacking within a Minute”. [Online]. Available: https://arxiv.org/abs/1801.01085. Accessed on: June 27, 2020.
[8] T. McDaniel, M. Smith, and M. Schuchard, “Peerlock: Flexsealing BGP”. [Online]. Available: https://arxiv.org/abs/2006.06576. Accessed on: July 17, 2020.
[9] International Organization for Standardization. (2009, Nov. 13). ISO Guide 73, Risk management. Vocabulary. [Online]. Available: https://www.iso.org/standard/44651.html. Accessed on: Aug. 20, 2019.
[10] Y. Rekhter, and P. Gross, “RFC 1772. Application of the Border Gateway Protocol in the Internet”. [Online]. Available: http://tools.ietf.org/html/rfc1772. Accessed on: June 20, 2020.
[11] В. Мохор, та В. Зубок, *Формування міжвузлових зв’язків в Інтернет з використанням методів теорії складних мереж.* Київ, Україна: Прометей, 2017.
[12] В. Зубок, “Формальний опис об’єктів і процесів глобальної маршрутизації у мережі Інтернет для оцінки впливу кібератак на маршрутизацію”, *Реєстрація, зберігання і обробка даних*, том 21, № 4, с. 67-74, 2020, doi: 10.35681/1560-9189.2019.21.4.199409.

ВІТАЛІЙ ЗУБОК

ВИЗНАЧЕННЯ СКЛАДОВИХ РИЗИКУ ПЕРЕХОПЛЕННЯ МАРШРУТУ ШЛЯХОМ АНАЛІЗУ ТОПОЛОГІЇ ІНТЕРНЕТ-ЗВ'ЯЗКІВ

Можливість зміни динамічних маршрутів між вузлами, які фізично не пов'язані, є ключовою особливістю маршрутизації в Інтернеті. Для забезпечення цієї функції був розроблений протокол зовнішнього шлюзу BGP-4, а також політики та процедури маршрутизації між доменами. Розроблений для мережі сотень вузлів, які покладаються на інформацію один від одного. Сьогодні BGP-4 зв'язує мережу з десятками тисяч вузлів попри його критичні вразливості, насамперед відсутність цілісності даних маршрутизації. Однією з найбільш значущих проблем, що виникає унаслідок його вразливостей, є витоки та викрадення маршрутів. Жодне із запропонованих та частково реалізованих оновлень та доповнень до глобальної маршрутизації як MANRS та RPKI не може забезпечити надійний захист від таких типів атак. Отже, існує проблема зменшення наслідків реалізацій загроз через ці вразливості і вона потребує принципово нової методології. В роботі розвивається напрям оцінки ризику перехоплення маршрутів шляхом аналізу міжмережевих зв'язків. Оскільки сучасна методологія управління інформаційною безпекою базується на управлінні ризиками, в даній роботі поводження з ризиком (зменшення ймовірності настання тригера ризику, зменшення максимального збитку) пропонується шляхом вдосконалення топології зв'язків. Оцінка ризиків перехоплення маршрутів вимагає кількісного вимірювання впливу атаки, який матеріалізується як спотворення легітимних маршрутів проходження інформації та призводить до втрати інформаційної безпеки. Для такої оцінки в даній роботі запропоновано використати знання особливостей топології Інтернету на рівні глобальної маршрутизації, яка фактично визначається взаємодією автономних систем – груп підмереж під загальним керуванням – по протоколу маршрутизації BGP-4. В роботі на основі власного формального представлення IP-маршрутизації показано зв'язок між топологією та ризиком перехоплення маршруту. Запропоновано новий підхід до кількісного визначення інформаційного ризику за допомогою нової ризик-орієнтованої моделі глобальної маршрутизації, яка відображатиме властивості Інтернет-вузлів з точки зору ризику перехоплення маршрутів.

Ключові слова: глобальна маршрутизація, Інтернет, перехоплення маршруту, модель маршрутизації, кібербезпека, оцінка ризиків.

Zubok Vitalii, candidate of technical sciences, doctoral student, Pukhov Institute for Modelling in Energy Engineering of National Academy of Sciences of Ukraine, Kyiv, Ukraine.

ORCID: 0000-0002-6315-5259.

E-mail: vitaly.zubok@gmail.com.

Зубок Віталій Юрійович, кандидат технічних наук, докторант, Інститут проблем моделювання в енергетиці ім. Г.Є. Пухова Національної академії наук України, Київ, Україна.