Mathematical Support for the Assessment and Regulation of The Successful Implementation of Virus Attacks on Information Networks

Alexander Ostapenko [1], Evgeniya Shvartskopf [1], Vladimir Pitolin [1], Oleg Makarov [1], Nikolay Tikhomirov [1] and Yuri Pasternak [1]

[1] *Voronezh State Technical University, 20 letiya Oktyabrya St., 84, Voronezh, 394000, Russian Federation*

Abstract

In this paper, we propose mathematical support for assessing and regulating the risks of successful implementation of virus attacks on network information structures. In this regard, binomial and hypergeometric distributions of discrete random variables are used. As a result, analytical expressions have been obtained that contribute to both parametric and structural risk regulation in the context of a networked viral confrontation. The proposed software can be used to increase the epidemic resistance of network information structures.

Keywords

Risk regulation, information networks, virus attacks

1. Introduction

In the context of modern digital transformation, the problem of ensuring the security of information systems and networks for various purposes [1-7] is of particular importance. For the most adequate mathematical tool for assessing security should consider risk analysis [1-7], which is successfully applied both in direct analytical calculations of the probabilities of expected damage [1-3, 5], and in their expert measurements [6] for corporate [1-4] and social networks [5, 7]. However, from a practical point of view, it is extremely important not only to assess the risk, but also to try to manage its magnitude in the context of information confrontation. This is what this work is about in relation to network virus attacks.

From the works [1-7], the authors gleaned a conceptual basis, which is so necessary for an adequate formulation of research objectives. This concept allowed the authors to consider the security of information networks as their state, in which the risks of a virus invasion do not exceed the permissible value. In this case, the risk is understood as the possibility of damage as a result of the virus affecting the elements of the analyzed network. Risk analysis in this case is considered as a process of assessing and regulating the risk of a successful virus attack.

In homogeneous networks [1], the damage is determined by the number of elements affected by the virus. To calculate the probability of damage, various approximating distributions of a random variable are used [1, 2, 6], which in this case is the power of the set of affected components.

From works [1, 2, 5-7], the authors drew a measure of risk in the form of the product of the amount of damage and the probability of its occurrence, which is quite convenient for the corresponding analytical calculations. This measure is quite effectively used [6] and in the space of fuzzy logic and expert assessments, used in case of difficulties in the analytical expression of the above parameters.

AISMA-2021: International Workshop on Advanced in Information Security Management and Applications, October 1, 2021, Stavropol, Krasnoyarsk, Russia

EMAIL: alexostap123@gmail.com (Alexander Ostapenko), ogeina@gmail.com (Evgeniya Shvartskopf), v.pitolin@mail.ru (Vladimir Pitolin), o.y.makarov@bk.ru (Oleg Makarov), ni.tikhomirov@mail.ru (Nikolay Tikhomirov), yuripasternak@mail.ru (Yuri Pasternak)
ORCID: 0000-0002-9610-0852 (Alexander Ostapenko); 0000-0002-2295-6738 (Evgeniya Shvartskopf); 0000-0002-5225-8981 (Vladimir Pitolin); 0000-0003-2795-419X (Oleg Makarov); 0000-0002-0261-7325 (Nikolay Tikhomirov); 0000-0002-1127-1864 (Yuri Pasternak).

CEUR Workshop Proceedings (CEUR-WS.org)

The successful implementation of this measure in the risk analysis of attacked corporate [2, 3, 4] and even social [5, 7] networks convinced the authors of its effectiveness in relation to computer [1, 4] viruses of network structures. Epidemic processes [1] arising in this case require an adequate assessment of the epistability of networks, which are now actually the basis of digital transformation. However, this transformation is exposed to significant danger from cyberattacks for various purposes [2-4,6], among which viruses and worms [1] are the most harmful. Assessment of damages and risks from attacks by these malware can be carried out [1] through the moments of their statistical distribution. It is this message that prompted the authors of this work to turn to an attempt to build software for risk analysis of information and telecommunication networks and their clusters in the context of total virus attacks.

2. Risk assessment and management methodology

In the context of the development of this study, it seems appropriate to consider heterogeneous networks. In this case, the expected damage from a virus infection for different network nodes is unequal and the application of the binomial probability distribution will be incorrect. Here, apparently, you should use a polynomial distribution and talk about the risk analysis of several random variables. Appropriate analytics exist for this. It only needs to be adapted for the risk analysis methodology.

The possibilities of continuous probability density distributions should also not be neglected. With appropriate discretization, one can pass from them to purely probabilistic estimates. The variety of continuous distributions here opens up sample opportunities for approximating random processes in attacks on information networks, including for protecting them from viral intrusions.

In terms of direct calculations of the margin of network stability (epistability) by the risk function, it is possible to propose taking into account its variance. In other words, it seems possible to be limited only by the difference between the fatal value and the expected risk, and to introduce a dispersion correction. In this case, it is proposed to count from the sum of the expectation and the standard deviation of the risk function. Here it is also necessary to determine the range of admissible values and propose adequate control algorithms.

Regarding the regulation of the above reference point (for assessing the epistability), obviously, the criterion is the minimization of its coordinate value along the damage axis. This is not difficult for the first of the two design cases considered in the work. However, for the second case (hypergeometric distribution), this approach is not obvious. Here, the analytical estimate (binomial distribution) will have to be replaced by a numerical calculation, possibly even with subsequent optimization. An even more complicated situation will take place in the case of using polynomial distribution and performing risk analysis of heterogeneous networks exposed to virus attacks.

In this case, the infection of these elements occurs randomly with the probability of a single infection p estimating the expected x number of damaged elements and minimizing damage to the network by setting up antivirus tools (adjusting the p parameter).

For the risk assessment in this problem, it is appropriate to use the binomial probability distribution:

$$P(x,n,p) = C_n^x \, p^x \, (1-p)^{n-x}, \tag{1}$$

where $0 \leq p \leq 1$ – the probability of a single infection by a virus of one network node;

n – a positive integer equal to the number of hosts;

x – the expected number of nodes affected by the virus.

Based on the risk measure taken [1-7], its value will be (x, u_0). When normalizing (2) in terms of damage to one node), can be determined from (1) as follows:

$$Risk\ (x,n,p) = (x,u_0)\, P(x,n,p). \tag{2}$$

Risk function shown on Figure 1.

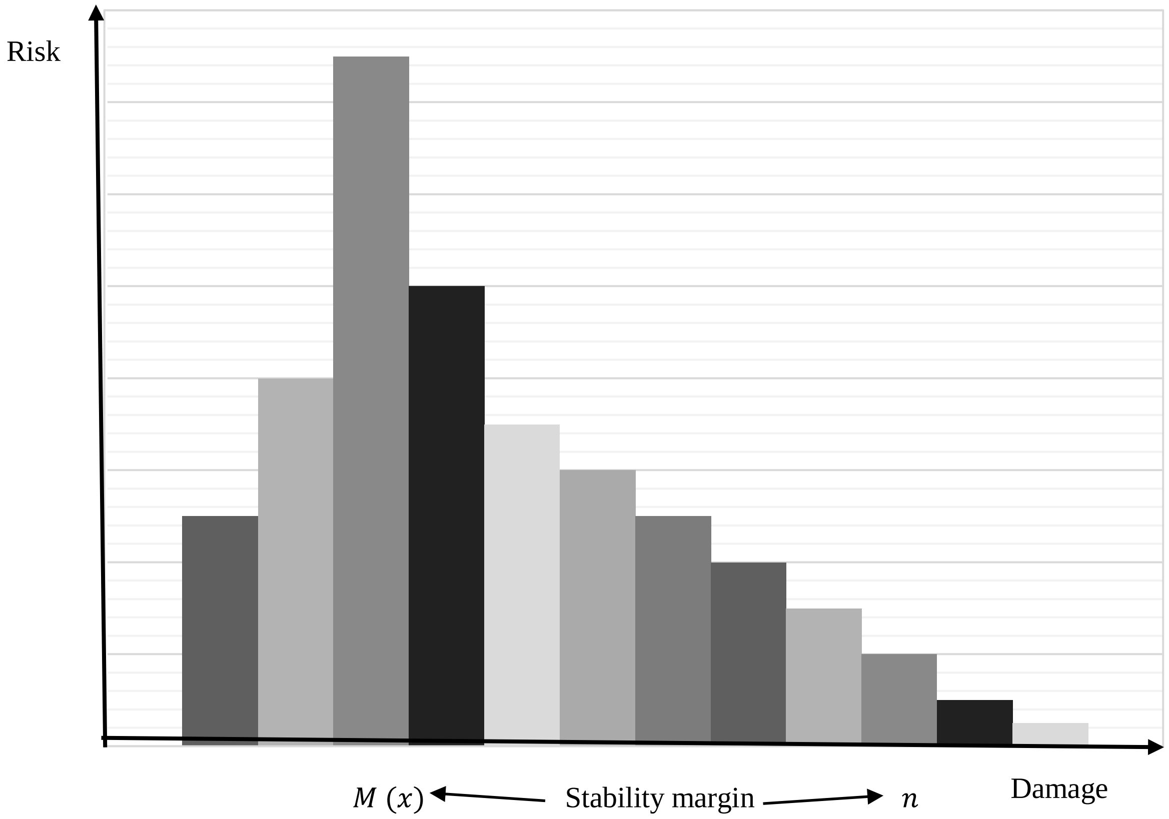

Figure 1: Risk function

When normalizing (2) in terms of damage to one node u_0 we have:

$$\underline{Risk}\ (x,n,p)\ =\ x\ C_n^x\ p^x\ (1-p)^{n-x}. \tag{3}$$

For a random risk variable (3), the expected value is found:

$$\mathrm{M(x)}\ =\ np\ [1-p(n+1)]. \tag{4}$$

From expression (4), it is possible to estimate the margin of stability of the virus-attacked network in relation to its lethal damage to all nodes n:

$$z(\underline{Risk})\ =\ n - M(x). \tag{5}$$

When normalizing expression (5) with respect to n, it has the value of normalized risk stability:

$$\underline{z}(\underline{Risk}) = \frac{n-M(x)}{n} = 1-\frac{M(x)}{n}. \tag{6}$$

Taking into account (4), the last expression (6) can be reduced to the following quadratic equation:

$$p^2(n+1)-p+(1-\underline{z}) = 0. \tag{7}$$

The solution to (7) with respect to p is two roots:

$$p_{1,2} = \frac{1 \pm\sqrt{1-4(n+1)(1-\underline{z})}}{2(n+1)}. \tag{8}$$

Let us find their (8) range of admissible values of p:

$$\frac{1}{2(n+1)} \leq p \leq \frac{1}{n+1}. \tag{9}$$

It is within these (9) limits of probability p that the inductive anti-virus systems of homogeneous elements of the attacked network are tuned by moving away from the fatal edge of risk resistance:

$$\underline{z} = \frac{4n+3}{4n+4}. \tag{10}$$

The set of expressions (8) - (10) is the main one for this regulation. Let's complicate the task. The above methodology provides parametric (using p) risk management. However, in practice, a more successful increase in the epidemic resilience of the network can be obtained through effective clustering of the network. With this in mind, let's take a look at the following problem.

There is an information network with fairly uniform n elements. In this case, the unacceptable damage to the network is the destruction of k elements by the virus. In order to increase the epidemic security of the network, its administrator seeks to create clusters of dimension m elements in it. Hence, it becomes necessary to assess the number of surviving operational elements in such clusters (risk analysis) and try to manage it in the course of information warfare.

For the risk assessment in this problem, it is appropriate to use the hypergeometric probability distribution:

$$P(x,m,n,k) = \frac{C_x^k C_{n-x}^{m-x}}{C_n^m}, \tag{11}$$

where $x \leq k;\ n - x \leq m - k;\ m, n, k, x$ – are integers.

Based on the expression (11), by analogy with (2) and (3), we have a normalized risk:

$$\underline{Risk}(x,m,n,k) = xP(x,m,n,k). \tag{12}$$

For a random value of risk (12), the expected value is found:

$$M(x) = \frac{nk}{m}\left[\frac{(m-k)(n-m)}{m(m-1)} - \frac{nk}{m}\right]. \tag{13}$$

From expression (4), by analogy with (5) and (6), the stability of a virus-attacked cluster in relation to lethal damage to all of its m nodes:

$$\underline{z}(\underline{Risk}) = 1 - \frac{M(x)}{n}. \tag{14}$$

Using (14), we can construct the equation:

$$\frac{nk}{m^2}\,\frac{(m-k)(n-m)-nk(m-1)}{m(m-1)} = 1 - \underline{z}. \tag{15}$$

The last expression (15) can be reduced to a fourth-order control:

$$m^4\left(1 - \underline{z}\right) - m^3\left(1 - \underline{z}\right) + m^2(nk) + m(nk^2) - (nk^2) = 0, \tag{16}$$

which solution with respect to m is appropriate to implement in an automated mode. In this case, it is possible to determine the dimension of the clusters into which the network should be partitioned to ensure their specified epidemic resistance. This is the structural regulation of the network architecture in order to counteract virus attacks on its users.

3. Conclusion

In this paper, we propose mathematical support for parametric and structural regulation of network virus attacks. Algorithmization and software implementation of the proposed methods have been carried out outside the scope of the present. In the order of software development, it is possible to describe network structures with heterogeneity of their elements. In terms of the practical application of the proposed software, it is pertinent to note that it can be used to counter not only computer viruses, but also viral content on social networks.

The above can be considered as a proposed "roadmap" for further research to expand the listed contradictions.

In conclusion, it should be noted that along with the purely cybernetic application of the developed methods, their educational and methodological use is visible. In particular, at the Faculty of Information Technologies and Computer Security of the Voronezh State Technical University for students of the specialties "Computer security of telecommunication systems", a methodology of risk analysis of viral intrusions into network structures has been developed, which can be introduced into the following disciplines of the curriculum:

- "Mathematical foundations of risk analysis";
- "Information Risk Management";
- "Research work of students".

In this case, the integration of design and training activities within the framework of the Regional Educational and Scientific Center for Information Security Problems (Voronezh, Russian Federation), coordinated by the Institute of Management Problems of the Russian Academy of Sciences, will bring substantially positive results as part of creating the necessary staffing for information security, and in terms of increasing the security of information and telecommunication systems and networks of the region and the state.

Therefore, research in the direction declared in this work is advisable to continue and increase their intensity in accordance with the plan proposed in this conclusion.

4. Acknowledgements

The authors are deeply grateful to:

1. The faculty of the Department of Information Security Systems of the Voronezh State Technical University for useful advice within the main scientific direction of the Department of Information Risk Management and Security of Information and Telecommunication Networks.

2. Administration of the strategic project "Safe Internet", initiated by the Department of Education and Science of the Voronezh Region, for practical recommendations on the creation of methodological support for risk analysis of information networks in the interests of the region.

3. To the management of the Regional Educational and Scientific Center for Information Security Problems (Voronezh, Russian Federation) for objective and benevolent review of the research results, useful recommendations for their improvement.

4. Head of the Department of Information Security Systems of Voronezh State Technical University, Professor A. Ostapenko. for the creation of comfortable conditions for conducting research by the team of authors with a promising further development of this scientific direction

5. References

[1] Islamgulova V.V., Ostapenko A.G., Radko N.M., Babadzhanov R.K., Ostapenko O.A. Discreet risk-models of the process of the development of virus epidemics in non-uniform networks. Journal of Theoretical and Applied Information Technology, 2016, vol. 86, no. 2, pp. 306-315.

[2] Butuzov V.V., Ostapenko A.G., Parinov P.A., Ostapenko G.A. Email-flooder attacks: The estimation and regulation of damage. Life Science Journal, 2014, vol. 11, no. 7s, pp. 213-218.

[3] Ostapenko A.G., Bursa M.V., Ostapenko G.A., Butrik D.O. Flood-attacks within the hypertext information transfer protocol: damage assessment and management. Biosciences Biotechnology Research Asia, 2014, vol. 11, pp. 173-176.

[4] Tsaregorodtsev A.V., Kravets O.Ja., Choporov O.N., Zelenina A.N. Information Security Risk Estimation for Cloud Infrastructure. International Journal on Information Technologies and Security, 2018, vol. 10, no. 4, pp. 67-76.

[5] Schwarzkopf E.A., Choporov O.N., Razinkin K.A., Yurasov V.G., Mazalov A.N. Mathematical and algorithmic support for the early detection process automation of potentially dangerous content in internet resources. IOP Conference Series: Materials Science and Engineering, 2020, P. 52037. doi:10.1088/1757-899X/862/5/052037

[6] Ermakov S.A., Zavorykin A.S., Kolenbet N.S., Ostapenko A.G, Kalashnikov A.O. Optimization of expert methods used to analyze information security risk in modern wireless networks. Life Science Journal, 2014, no 11(10s), pp. 511-514.

[7] Eshchenko A., Ostapenko G., Bataronov I., Tolstykh N. The automated networks and regional users: risk analysis of their reactions to the attacks of different destructive orientation, IOP Conf. Series: Materials Science and Engineering Bristol, Krasnoyarsk, Russia, 2019, vol. 537, i. 5. P. 1727. Doi:10.1088/1757-899X/537/5/052020