Akhramovych, Volodymyr

Article

Development of a method for calculation of information protection from the clustering coefficient and information flow in social networks

Technology audit and production reserves

Reference: Akhramovych, Volodymyr (2022). Development of a method for calculation of information protection from the clustering coefficient and information flow in social networks. In: Technology audit and production reserves 2 (2/64), S. 31 - 37.
http://journals.uran.ua/tarp/article/download/255962/253244/589617.
doi:10.15587/2706-5448.2022.255962.

UDC 004.738.5
DOI: 10.15587/2706-5448.2022.255962
Article type «Reports on Research Projects»

Volodymyr Akhramovych

DEVELOPMENT OF A METHOD FOR CALCULATION OF INFORMATION PROTECTION FROM THE CLUSTERING COEFFICIENT AND INFORMATION FLOW IN SOCIAL NETWORKS

The object of research is the system of information protection of the social network. The article investigates the dynamic models of the information protection system in social networks taking into account the clustering coefficient, and also analyzes the stability of the protection system. In graph theory, the clustering factor is a measure of the degree to which nodes in a graph tend to group together. The available data suggest that in most real networks, and in particular in social networks, nodes tend to form closely related groups with a relatively high density of connections. It is probability is greater than the average probability of a random connection between two nodes. There are two variants of this term: global and local. The global version was created for a general idea of network clustering, while the local one describes the nesting of individual nodes. There is a practical interest in studying the behavior of the system of protection of social networks from the value of the clustering factor.

Dynamic systems of information protection in social networks in the mathematical sense of this term are considered. A dynamic system is understood as any object or process for which the concept of state as a set of some quantities at a given moment of time is unambiguously defined and a given law is described that describes the change (evolution) of the initial state over time. This law allows the initial state to predict the future state of a dynamic system. It is called the law of evolution.

The study is based on the nonlinearity of the social network protection system. To solve the system of nonlinear equations used: the method of exceptions, the joint solution of the corresponding homogeneous characteristic equation. Since the differential of the protection function has a positive value in some data domains (the requirement of Lyapunov's theorem for this domain is not fulfilled), an additional study of the stability of the protection system within the operating parameters is required. Phase portraits of the data protection system in MatLab/Multisim are determined, which indicate the stability of the protection system in the operating range of parameters even at the maximum value of influences.

Keywords: *dynamic models, information protection system, exception method, homogeneous characteristic equation, system stability.*

Received date: 14.03.2022
Accepted date: 20.04.2022
Published date: 30.04.2022

How to cite
Akhramovych, V. (2022). Development of a method for calculation of information protection from the clustering coefficient and information flow in social networks. Technology Audit and Production Reserves, 2 (2 (64)), 31–37. doi: http://doi.org/10.15587/2706-5448.2022.255962

1. Introduction

Descriptions of dynamical systems for various problems depending on the law of evolution are also various: with the help of differential equations, discrete mappings, graph theory, Markov chain theory, and so on. The choice of one of the methods of description determines the specific form of the mathematical model of the corresponding dynamic system [1].

The mathematical model of a dynamic system is considered to be given if the parameters (coordinates) of the system are introduced, which unambiguously determine its state, and the law of evolution is specified. Depending on the degree of approximation to the same system, different mathematical models can be matched.

Theoretical study of the dynamic behavior of a real object requires the creation of its mathematical model. In many cases, the procedure for developing a model is to compile mathematical equations based on physical laws. Usually these laws are formulated in the language of differential equations. As a result, the coordinates of the state of the system and its parameters are interconnected, which allows to begin to solve differential equations under different initial conditions and parameters.

In the article [2] the definition of the clustering coefficient in the case of (binary and weighted) directional networks is extended and the expected value for random graphs is calculated. In [3], it is noted that the properties of the small world of neighboring connections are higher than in comparative random networks. If a node has one or no neighbors, in such cases the local clustering is traditionally set to zero, and this value affects the global clustering factor. It is proposed to include the coefficient θ for isolated nodes in order to estimate the clustering coefficient, except in cases from the determination of Watts and Strogats. In [1] a method of determining trust and protection of personal data in social networks was developed. In research [4, 5] the clustering coefficients for social networks, including power ones, are considered. In [6] the system of protection of the social network from the indicators of the relationship between users and the amount of information flow is studied. In the article [7] the analysis of the clustering coefficient on the social network twitter is carried out. In [8], an analysis of the clustering coefficient through triads of connections was performed. In the article [9] the dependence between the clustering coefficient and the average path length in a social network is investigated. In [10, 11] the use of clustering methods of social networks for personalization of educational content is investigated. In articles [12, 13] discusses the behavior of the clustering coefficient for complex networks. In [14], it was concluded that based on the results of the experiment, it can be concluded that among the clustering algorithms there is no universal algorithm that would be significantly ahead of others on all data sets. The leaders of benchmarking are the algorithms Spinglass and Walktrap.

In paper [15], mathematical modeling of nonlinear dependences of personal data security on the parameters of trust among users was carried out. However, there remained unresolved issues related to the impact of parameters of users' interaction on the security system. The study is important in terms of determining quantitative indicators of the impact on the security system of a specific parameter of the social network, which is trust among users, but it does not consider users' interaction.

Paper [16] considered the functioning of social networks (SN) in accordance with four main user properties: the geographical location of a user, the weight of a user, the number of interactions with a user, and the life expectancy of a user. In the research, the ratio between the specific parameters of a social network, for example, interaction and information security indicator, was not brought to qualitative characteristics.

Therefore, since no studies of the nonlinear system of protection of the social network from the coefficient of clustering and information flow, the studies presented in this article are relevant.

Thus, *the object of research* is the system of information protection of the social network. And *the aim of research* is to study the dependences of the parameters of the information protection system in the social network on the parameters of its clustering, the flow of information and study the stability of the protection system under different types of influences.

2. Research methodology

Analysis of graphical dependences of a linear system [1] indicates the nonlinearity of the system. Therefore, in the system of equations (1) let's introduce nonlinear components (2):

$$\begin{cases} \dfrac{dI}{dt} = Z_p Z + (C_v + C_K) I, \\ \dfrac{dZ}{dt} = -\left(\dfrac{\sum_{v \in V} C_{v1}}{N^2}\right) - I(C_{d2} + C_{d1}), \end{cases} \quad (1)$$

where $\sum_{v \in V} C_{v1}$ – the total number of connections in the network; N – the number of vertices in the network.

$$\begin{cases} \dfrac{dI}{dt} = Z_p Z + (C_v + C_K) I + L_2(I^2) + L_3(I^3) + \ldots, \\ \dfrac{dZ}{dt} = -\left(\dfrac{\sum_{v \in V} C_{v1}}{N^2}\right) - I(C_{d2} + C_{d1}) + K_2(Z^2) + K_3(Z^3) + \ldots, \end{cases} \quad (2)$$

where L_2, L_3, ..., L_i, and K_2, K_3, ..., K_i – some linear operators. Let's consider the nonlinearity of the system to be weak, which allows to find a solution for each equation of the system (2) by the method of successive approximation, putting:

$$I = I_1 + I_2 + I_3 \ldots$$

Since the differential (Fig. 1, 2) of the protection function has a positive value in some data domains (the requirement of Lyapunov's theorem for this domain is not fulfilled), an additional study of the stability of the protection system within the operating parameters is required:

$$Z = Z_1 + Z_2 + Z_3 + \ldots$$

Let at:

$$dI = 0,\ \frac{dI}{dt} = 0, \text{ and } dZ = 0,\ \frac{dZ}{dt} = 0,$$
$$I = I_0 s \text{in}\,\omega t,\ Z = Z_0 s \text{in}\,\omega t.$$

Let's obtain a system of equations:

$$\begin{cases} \dfrac{dI}{dt} = Z_p Z + (C_v + C_K) I - L_2(I_0^2 s \text{in}^2\,\omega t) - \\ -L_3(I_0^3 s \text{in}^3\,\omega t) - \ldots, \\ \dfrac{dZ}{dt} = -\left(\dfrac{\sum_{v \in V} C_{v1}}{N^2}\right) - I(C_{d2} + C_{d1}) - K_2(Z_0^2 s \text{in}^2\,\omega t) - \\ -K_3(Z_0^3 s \text{in}^3\,\omega t) - \ldots, \end{cases} \quad (3)$$

Let's rewrite the system and present it as follows:

$$\begin{cases} \dfrac{dI}{dt} = \alpha Z + \beta_1 I - \sum_{k=2}^{\infty} L_k I_0^k \sin^k \omega t, \\ \dfrac{dZ}{dt} = \beta_2 I + \gamma - \sum_{k=2}^{\infty} K_k Z_0^k \sin^k \omega t, \end{cases} \quad (4)$$

where

$$\alpha = Z_p,\ \beta_1 = C_v + C_K,\ \beta_2 = -\left(C_{d2} + C_{d1}\right),\ \gamma = -\left(\frac{\sum_{v \in V} C_{v1}}{N^2}\right).$$

The results of the calculation are presented in Fig. 3.

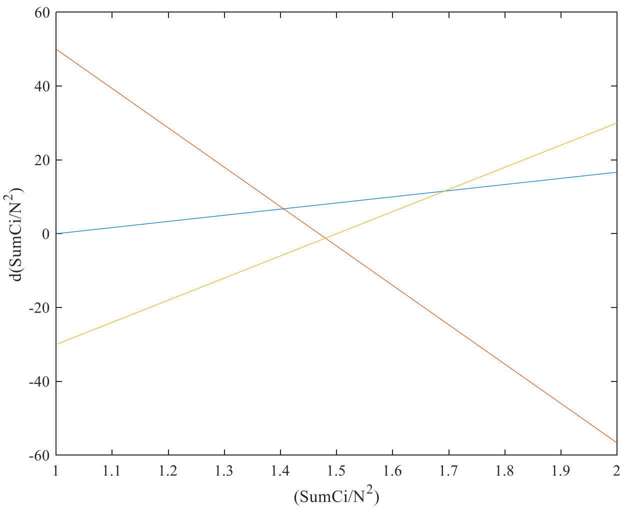


Fig. 1. Differential of the clustering coefficient function

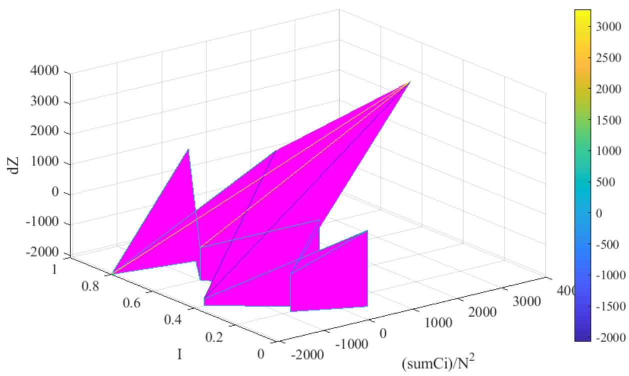


Fig. 2. Differential of protection function

Next, use the exception method:

$$\frac{dZ}{dt} = \beta_2 I + \gamma - \sum_{k=2}^{\infty} K_k Z_0^k \sin^k \omega t \Rightarrow$$

$$\Rightarrow I = \frac{1}{\beta_2}\left(\frac{dZ}{dt} - \gamma + \sum_{k=2}^{\infty} K_k Z_0^k \sin^k \omega t\right) \Rightarrow$$

$$\Rightarrow \frac{dI}{dt} = \frac{1}{\beta_2}\left(\frac{d^2 Z}{dt^2} + \frac{1}{\omega}\sum_{k=2}^{\infty}\left(kK_k Z_0^k \sin^{k-1} \omega t \cos \omega t\right)\right). \quad (5)$$

The results of the calculation are presented in Fig. 4, 5.

Substitute all the found expressions (5) in the first equation of system (4):

$$\frac{1}{\beta_2}\left(\frac{d^2 Z}{dt^2} + \frac{1}{\omega}\sum_{k=2}^{\infty}\left(kK_k Z_0^k \sin^{k-1} \omega t \cos \omega t\right)\right) =$$

$$= \alpha Z + \frac{\beta_1}{\beta_2}\left(\frac{dZ}{dt} - \gamma + \sum_{k=2}^{\infty} K_k Z_0^k \sin^k \omega t\right) -$$

$$- \sum_{k=2}^{\infty} L_k I_0^k \sin^k \omega t, \quad (6)$$

or

$$\frac{d^2 Z}{dt^2} - \beta_1 \frac{dZ}{dt} - \alpha\beta_2 Z =$$

$$= -\frac{1}{\omega}\sum_{k=2}^{\infty}\left(kK_k Z_0^k \sin^{k-1} \omega t \cos \omega t\right) -$$

$$- \beta_1\gamma + \beta_1 \sum_{k=2}^{\infty} K_k Z_0^k \sin^k \omega t - \beta_2 \sum_{k=2}^{\infty} L_k I_0^k \sin^k \omega t. \quad (7)$$

Now let's find a common solution of the corresponding homogeneous equation:

$$Z'' - \beta_1 Z' - \alpha\beta_2 Z = 0. \quad (8)$$

The characteristic equation has the form:

$$\lambda^2 - \beta_1\lambda - \alpha\beta_2 = 0.$$

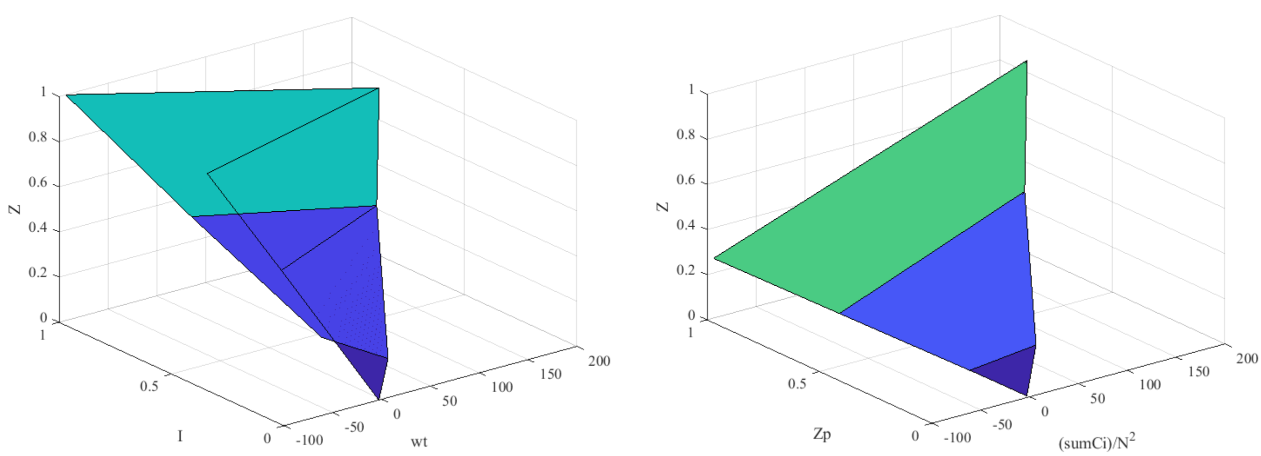
Let's consider the case of the positive discriminant of this equation:

$$D = \beta_1^2 + 4\alpha\beta_2 > 0 \Rightarrow \lambda_{1,2} = \frac{\beta_1 \pm \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}. \quad (9)$$

From:

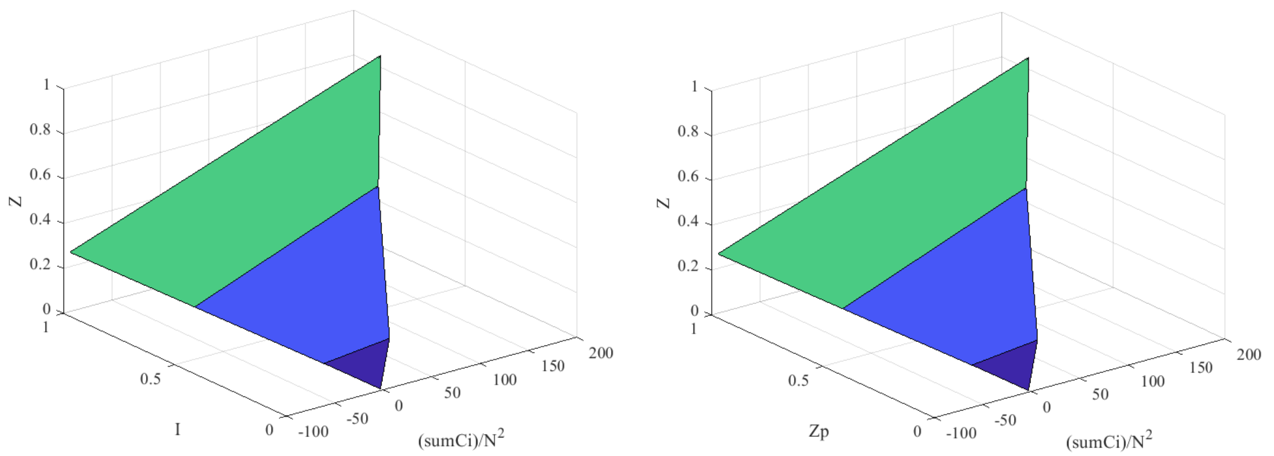
$$Z_{hom}(t) = c_1 e^{\frac{\beta_1 + \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2} t} + c_2 e^{\frac{\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2} t}.$$

It is joint solution of a homogeneous equation.



a *b*

Fig. 3. Graphs by dependence (4): *a* – the dependence of the protection rate on the amount of information flow and frequency; *b* – the dependence of the protection index on the value of the system security indicator and the clustering factor



a *b*

Fig. 4. Graphs by dependence (5): *a* – the dependence of the protection rate on the amount of information flow and frequency; *b* – the dependence of the protection index on the value of the system security indicator and the clustering factor

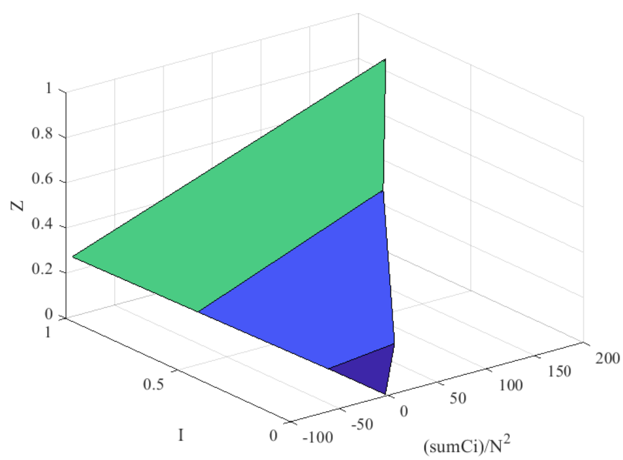


Fig. 5. Graphs by dependence (7)

To find the general solution of the inhomogeneous equation let's use the method of variation of arbitrary constants:

$$Z_{hom}(t)=c_1'(t)e^{\frac{\beta_1+\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}+c_2'(t)e^{\frac{\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}, \qquad (10)$$

where $c_1'(t), c_2'(t)$ are from the system:

$$\begin{cases} c_1'(t)e^{\frac{\beta_1+\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}+c_2'(t)e^{\frac{\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}=0, \\ c_1'(t)\dfrac{\beta_1+\sqrt{\beta_1^2+4\alpha\beta_2}}{2}e^{\frac{\beta_1+\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}+ \\ +c_2'(t)\dfrac{\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}e^{\frac{\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}=N(t), \end{cases}$$

where

$$N(t)=-\frac{1}{\omega}\sum_{k=2}^{\infty}\left(kK_kZ_0^k\sin^{k-1}\omega t\cos\omega t\right)-\beta_1\gamma+$$

$$+\beta_1\sum_{k=2}^{\infty}K_kZ_0^k\sin^k\omega t. \qquad (11)$$

From equations (10), (11) let's obtain:

$$c_1'(t)e^{\frac{\beta_1+\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}=-c_2'(t)e^{\frac{\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}\Rightarrow$$

$$\Rightarrow c_2'(t)e^{\frac{\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}\left(\begin{array}{l}-\dfrac{\beta_1+\sqrt{\beta_1^2+4\alpha\beta_2}}{2}+ \\ +\dfrac{\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}\end{array}\right)=N(t), \qquad (12)$$

or

$$c_2'(t)e^{\frac{\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}\sqrt{\beta_1^2+4\alpha\beta_2}=-N(t), \qquad (13)$$

where will get:

$$c_2(t)=-\frac{1}{\sqrt{\beta_1^2+4\alpha\beta_2}}\int N(t)e^{\frac{-\beta_1+\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}\mathrm{dt}, \qquad (14)$$

$$c_1(t)=\frac{1}{\sqrt{\beta_1^2+4\alpha\beta_2}}\int N(t)e^{\frac{-\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}\mathrm{dt}. \qquad (15)$$

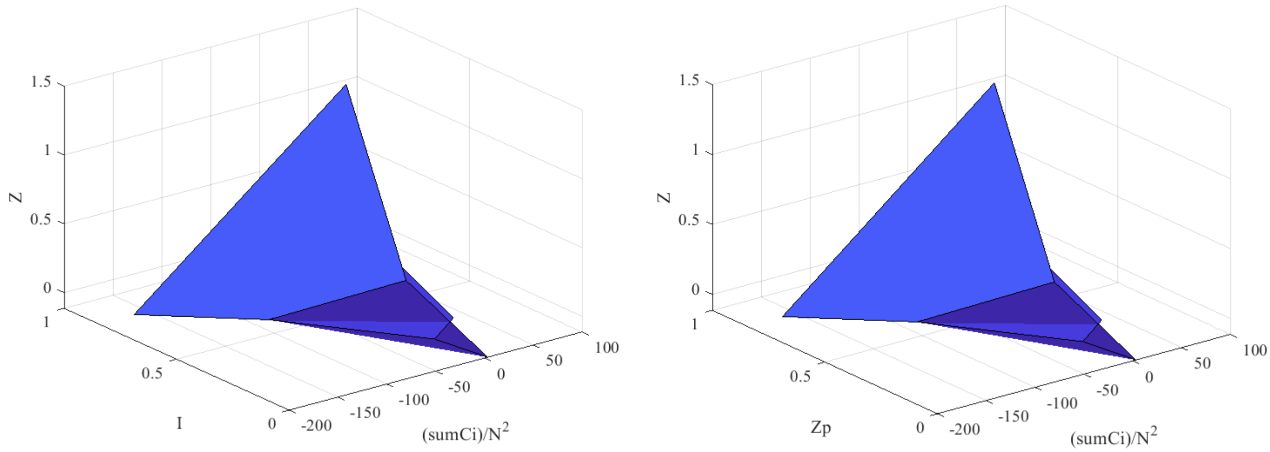
Given (13)–(15):

$$Z(t)=\int\left(N(t)-e^{\frac{-\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}\frac{e^{\frac{\beta_1+\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}}{\sqrt{\beta_1^2+4\alpha\beta_2}}\right)\mathrm{dt}-$$

$$-\int\left(N(t)-e^{\frac{-\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}\frac{e^{\frac{\beta_1-\sqrt{\beta_1^2+4\alpha\beta_2}}{2}t}}{\sqrt{\beta_1^2-4\alpha\beta_2}}\right)\mathrm{dt}. \qquad (16)$$

Graphical dependences are presented in Fig. 6.

Thus, mathematical and graphical dependences of the change in the indicator of social network information protection from clustering parameters and other specific parameters are obtained.



a *b*

Fig. 6. Graphs by dependence (16): *a* – the dependence of the protection rate on the amount of information flow and frequency; *b* – the dependence of the protection index on the value of the system security indicator and the clustering factor

3. Research results and discussion

Initial equation:

$$\frac{d^2Z}{dt^2}-\beta_1\frac{dZ}{dt}-\alpha\beta_2 Z=$$
$$=-\frac{1}{\omega}\sum_{k=2}^{\infty}\left(kK_kZ_0^k\sin^{k-1}\omega t\cos\omega t\right)-\beta_1\gamma+$$
$$+\beta_1\sum_{k=2}^{\infty}K_kZ_0^k\sin^k\omega t-\beta_2\sum_{k=2}^{\infty}L_kI_0^k\sin^k\omega t. \quad (17)$$

The solution will be implemented in the program MatLab/Multisim. Let's make the scheme (Fig. 7).

The phase portrait is presented in the form of an ellipse, which indicates the stability of the personal data protection system. The results of the program are presented in Fig. 8–11.

In contrast to previous research by scientists, it has been proven that the SN protection system is stable even from external maximum influences and a specific parameter of the clustering coefficient in the operating range of parameters.

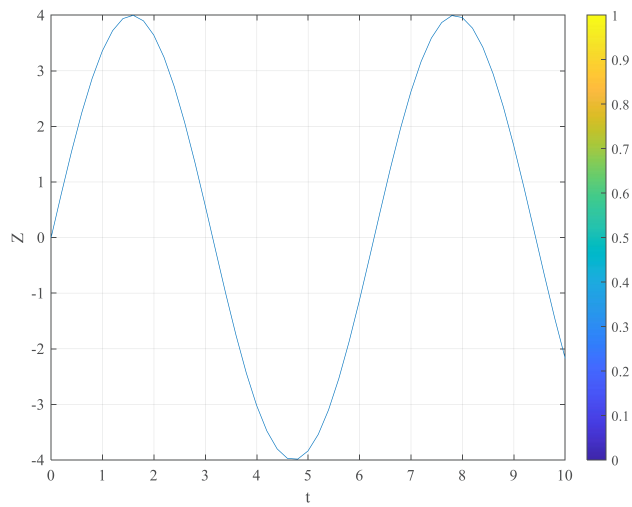


Fig. 8. Harmonic oscillations of the protection system on time $Z=f(t)$

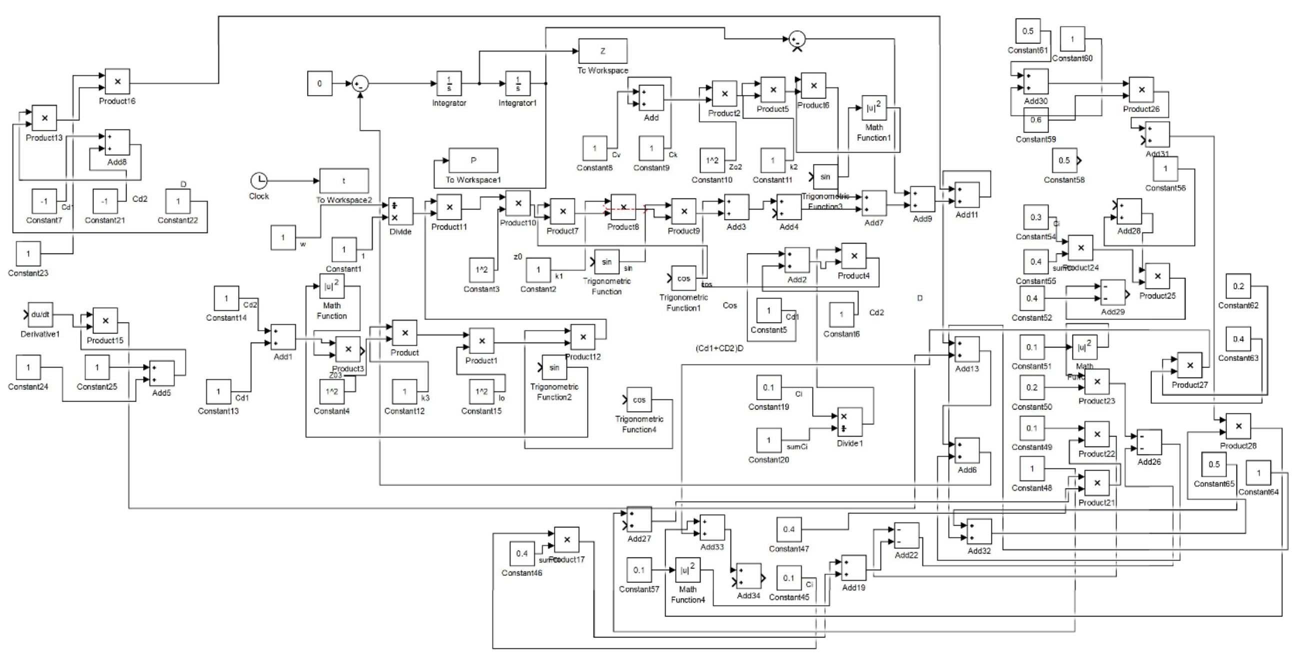


Fig. 7. Block diagram of the phase portrait program in the Multisim program, taking into account the attack block

ISSN 2664-9969

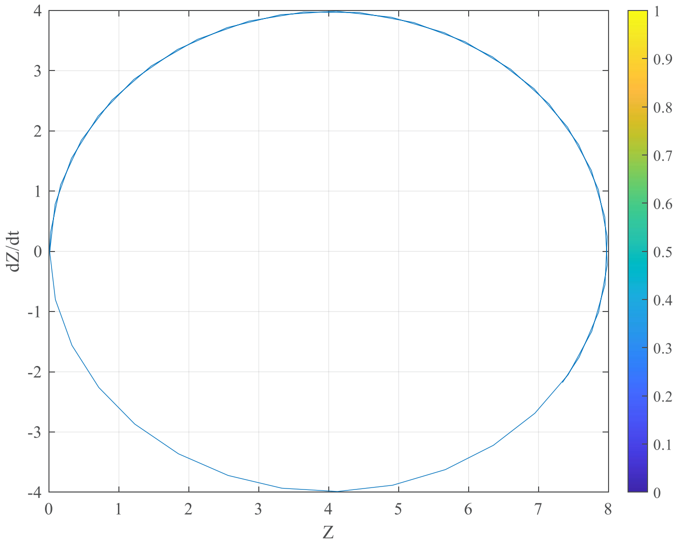


Fig. 9. Phase portrait of the protection system on clustering factor

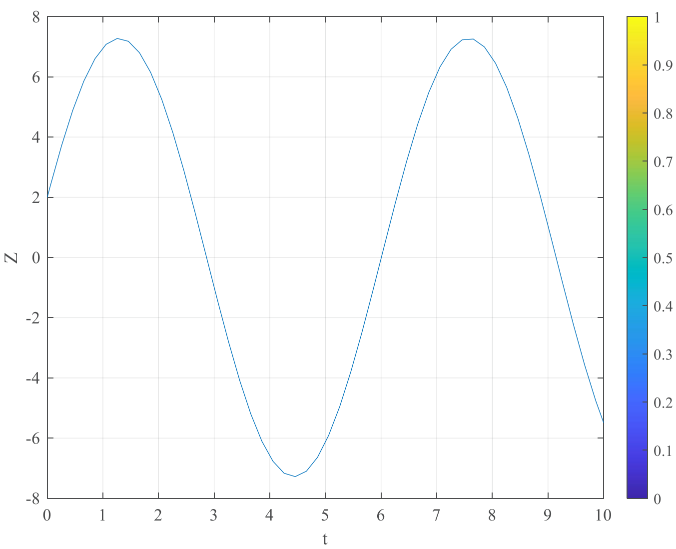


Fig. 10. Harmonic oscillations of the protection system on time $Z=f(t)$ taking into account the attacks

The results of the study of the stability of the protection system in SM make it possible to say that it is stable even at the maximum value of the amplitudes of influences in the operating range of parameters and at different values of specific network parameters – clustering factor. That is, it is possible to ensure reliable protection of information.

Further development of this study is to use known specific parameters of social networks (interaction, average distance between users, network expansion, interaction, dissemination of information, network centrality, etc.) to identify new factors and parameters.

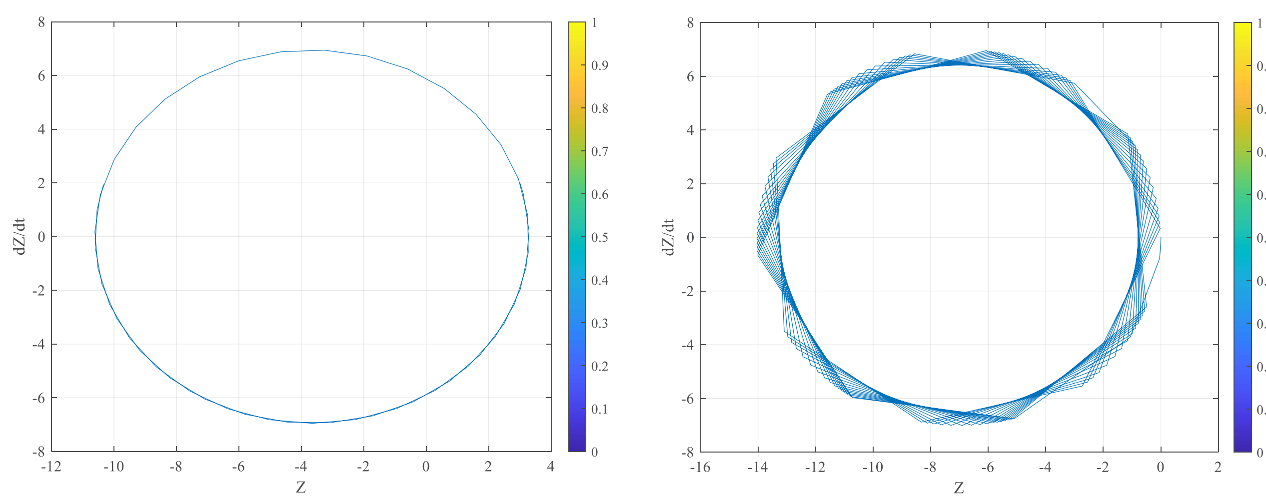
4. Conclusions

A system of nonlinear equations was obtained, which allowed obtaining mathematical quantitative results of the influence of specific parameters and parameters of user interaction on the social network protection system and their graphical interpretation by analysis and solution. The parameters of the impact of user interaction on the protection system show a significant impact on the protection rate of up to 100 %. This approach in research has allowed moving to the study of the stability of the protection system.

The study of the stability of the protection system in the social network without influences and in the presence of influences on the protection system was conducted using a flowchart created in the program MatLab/Multisim. Due to the obtained graphical interpretations of the oscillations of the protection system and phase portraits, the stability of the protection system is proved even in the presence of maximum influences.

References

1. Laptiev, O., Savchenko, V., Kotenko, A., Akhramovych, V., Samosyuk, V., Shuklin, G., Biehun, A. (2021). Method of Determining Trust and Protection of Personal Data in Social Networks. *International Journal of Communication Networks and Information Security, 1,* 15–21. doi: http://doi.org/10.1109/atit54053.2021.9678599



a *b*

Fig. 11. Phase portrait of the protection system on clustering factor taking into account the attacks: *a* – in the absence of influence; *b* – at the maximum value of the impact on the protection system

2. Fagiolo, G. (2007). Clustering in complex directed networks. *Physical Review E, 76 (2)*. doi: http://doi.org/10.1103/physreve.76.026107
3. Kaiser, M. (2008). Mean clustering coefficients: the role of isolated nodes and leafs on clustering measures for small-world networks. *New Journal of Physics, 10 (8),* 083042. doi: http://doi.org/10.1088/1367-2630/10/8/083042
4. Akhramovych, V., Shuklin, G., Pepa, Y., Muzhanova, T., Zozulia, S. (2022). Devising a procedure to determine the level of informational space security in social networks considering interrelations among users. *Eastern-European Journal of Enterprise Technologies, 1 (9 (115)),* 63–74. doi: http://doi.org/10.15587/1729-4061.2022.252135
5. Savchenko, V., Akhramovych, V., Tushych, A., Sribna, I., Vlasov, I. (2020). Analysis of Social Network Parameters and the Likelihood of its Construction. *International Journal of Emerging Trends in Engineering Research, 8 (2),* 271–276. doi: http://doi.org/10.30534/ijeter/2020/05822020
6. Wang, Y., Ghumare, E., Vandenberghe, R., Dupont, P. (2017). Comparison of Different Generalizations of Clustering Coefficient and Local Efficiency for Weighted Undirected Graphs. *Neural Computation, 29 (2),* 313–331. doi: http://doi.org/10.1162/neco_a_00914
7. Gracheva, M., Iakobi, Y., Stepanenko, V., Luneva, Y. (1950). Communication Patterns in Task-Oriented Groups. *The Journal of the Acoustical Society of America, 22 (6)*, 725–730. doi: http://doi.org/10.1121/1.1906679
8. Yevin, I. A. (2010). Introduction to the theory of complex networks. *Computer Research and Modeling, 2 (2),* 121–141. doi: http://doi.org/10.20537/2076-7633-2010-2-2-121-141
9. Zharinov, I., Krylov, V. (2008). Constructing graphs with minimum average path length. *Bulletin of Izhevsk State Technical University, 4,* 164–169.
10. Pavlov, Y. L., Cheplyukova, I. A. (2019). On the asymptotics of degree structure of configuration graphs with bounded number of edges. *Discrete Mathematics and Applications, 29 (4),* 219–232. doi: http://doi.org/10.1515/dma-2019-0020
11. Mamedova, H., Agaev, F., Zeinalova, L. (2019). Using social media to personalize e-learning. *İnformasiya texnologiyaları problemləri, 1,* 27–34.
12. Starodubtsev, V. (2015). Personalization of the virtual educational environment. *Pedagogical Education in Russia, 7,* 24–29.
13. Pavlov, J. (2019). On the asymptotics of the cluster coefficient of a configuration graph with an unknown distribution of vertex degrees. *Information and its application, 13 (3),* 9–13. doi: http://doi.org/10.14357/19922264190302
14. Firsov, M. (2017). Benchmarking graph clustering algorithms for decision-making problems. *Basic research, 12 (1),* 138–142.
15. Mahmoudi, A., Yaakub, M. R., Bakar, A. A. (2019). The Relationship between Online Social Network Ties and User Attributes. *ACM Transactions on Knowledge Discovery from Data, 13 (3),* 1–15. doi: http://doi.org/10.1145/3314204
16. Mahmoudi, A., Yaakub, M. R., Abu Bakar, A. (2018). New time-based model to identify the influential users in online social networks. *Data Technologies and Applications, 52 (2),* 278–290. doi: http://doi.org/10.1108/dta-08-2017-0056

Volodymyr Akhramovych, *Doctor of Technical Sciences, Senior Research Fellow, Professor, Department of Information and Cyber Defense Systems, State University of Telecommunications, Kyiv, Ukraine, e-mail: 12z@ukr.net, ORCID: https://orcid.org/0000-0002-6174-5300*