MODELING THE PROTECTION OF PERSONAL DATA FROM TRUST AND THE AMOUNT OF INFORMATION ON SOCIAL NETWORKS

Serhii Yevseiev
Department of Cyber Security and Information Technology
Simon Kuznets Kharkiv National University of Economics
9-A Nauky ave., Kharkiv, Ukraine, 61166
serhii.yevseiev@hneu.net

Oleksandr Laptiev
Department of Information and Cybersecurity Systems
State University of Telecommunications
7 Solomenska str., Kyiv, Ukraine, 03110
alaptev64@ukr.net

Sergii Lazarenko
Department of Information Security
zi.i.lazarenko@nau.edu.ua

Anna Korchenko
Department of IT-Security
annakor@ukr.net

Іryna Manzhul
Special department No. 2
National Academy of the Security Service of Ukraine
22 Maksimovich str., Kyiv, Ukraine, 03022
vassa00@ukr.net

1National Aviation University
1 Liubomyra Huzara ave., Kyiv, Ukraine, 03058

Abstract
The article analyzes the parameters of social networks. The analysis is performed to identify critical threats. Threats may lead to leakage or damage to personal data. The complexity of this issue lies in the ever-increasing volume of data. Analysts note that the main causes of incidents in Internet resources are related to the action of the human factor, the mass hacking of IoT devices and cloud services. This problem is especially exacerbated by the strengthening of the digital humanistic nature of education, the growing role of social networks in human life in general. Therefore, the issue of personal information protection is constantly growing. To address this issue, let’s propose a method of assessing the dependence of personal data protection on the amount of information in the system and trust in social networks. The method is based on a mathematical model to determine the protection of personal data from trust in social networks. Based on the results of the proposed model, modeling was performed for different types of changes in confidence parameters and the amount of information in the system.

As a result of mathematical modeling in the MatLab environment, graphical materials were obtained, which showed that the protection of personal data increases with increasing factors of trust in information. The dependence of personal data protection on trust is proportional to other data protection parameters. The protection of personal data is growing from growing factors of trust in information.

Mathematical modeling of the proposed models of dependence of personal data protection on trust confirmed the reliability of the developed model and proved that the protection of personal data is proportional to reliability and trust.

Keywords: social network, protection, user, parameter, transfer, information, metric, density, cycle.

DOI: 10.21303/2461-4262.2021.001615

1. Introduction

In today’s world, information needs reliable protection: from unauthorized access and distribution, accidental deletion or alteration. All developed European countries are concerned about
the problem of information security, as well as the protection of personal data of citizens. This is due to the fact that informatization and digitization of information have become widespread in all areas of human activity, including the storage of personal and work data.

Social networks are one of the main methods of communication, search for connections and exchange of both publicly available and confidential information. Social networks make up an ever-growing share of shared networks. The network itself acquires new properties, acting as an independent factor.

Because information in the global network exists outside of space and time, the network itself becomes an active agent of influence on the person, keeping, above all, large amounts of data publicly available. In recent years, the vision of the problem of cybersecurity has begun to change significantly, as people increasingly cease to be the subject of cybercrime, becoming an object in itself, and not only its financial and economic interests and capabilities.

This problem is especially exacerbated by the strengthening of the digital humanistic nature of education, the growing role of social networks in human life in general.

The protection of personal data in today’s information life is perhaps the most important aspect in meeting the safe use of all the capabilities of current technologies. Therefore, the problem of studying the parameters of social networks for their further use in solving problems of information and personal data protection is very relevant.

2. Literature review and problem statement

The exchange of structural and thematic data potentially allows the use of social networks to address a wide range of information security issues.

In [1] discusses social networks that can track user actions and control data for future use. A study of 45 social networks found that approximately 90% of sites unreasonably require personal information to, for example, allow permission to join them; 85% of sites do not use standard encryption protocols to protect data from cybercrime attacks; 72% of sites transfer information about users to third parties.

In [2] the standards, attributes and characteristics of the profile are considered and a method of detecting signs of public opinion manipulation in social networks based on the construction of information security profiles of social Internet services, based on gradient boosting of binary trees, which automates early detection procedures.

In [3–5] it is indicated that the dissemination of personal data through social networks is much faster than in real life. It is most dangerous when personal information comes to people for whom it is not intended. Social media users are often unaware that they can change personal privacy settings to protect their data.

In [6, 7] the mechanism of application of correlation of potential crisis situations for an estimation of average and total level of criticality of a current situation in information sphere is considered. The mechanism is based on methods of expert evaluation and fuzzy logic. A correlation mechanism is proposed to determine the correlation coefficient of each dependent identification of potential crisis situations with the main one, which determines the interdependencies between them. The obtained correlation coefficients can be used to calculate the average and total levels of criticality of a situation that has arisen under the influence of several interrelated and simultaneous potential crisis situations. Only information correlation problems are considered.

In [8, 9] developed a structural-parametric model of information security risk assessment system which, due to the structural components of subsystems, the formation of primary and secondary data, as well as their components modules of initialization of input data, formation and conversion of reference values, weighing evaluation parameters and their adjustment, risk assessment and report generation, which implement the proposed methods, assessment based on databases of vulnerabilities, incrementing and decrementing the order of linguistic variables, allows for high flexibility and convenience in assessing information security risks without the participation of experts in the subject area. But only the problems of information security, which are presented in local databases, are considered.

In [10] the qualitative-quantitative method of analysis and assessment of information security risks by modifying the procedures for determining many parameters of risk assessment and
assessment of current values of parameters with the possibility of integrating the values of indicators presented in the relevant databases. To do this, it is proposed to use appropriate databases of vulnerabilities, which present their quantitative estimates. In [14], only information security issues are considered, which are presented in the databases of CVSS (Common Vulnerability Scoring System is the framework for rating the severity of security) indicators.

In [11, 12] the methodology of construction of the system of information security of banking information in automated banking systems (ABS) is considered, which is based on the first proposed three-level model of strategic management of information technology security. In [13] only information security problems are considered, without taking into account technical problems.

In addition, in the literature, the study of certain issues of this issue at different times paid attention to such specialists. Consideration of this issue is carried out by prominent.

The aim of the article is to study the whole set of parameters of threats in social networks from the loss of trust between users for their further use in solving problems of information and data protection.

3. Formulation of the problem

The exchange of personal data potentially allows the use of social networks to solve a wide range of information problems, but there is a problem of data protection. Therefore, the question of developing new mathematical models for assessing the dependence of personal data protection on trust and the amount of information on social networks is very relevant.

The aim of this research is to develop a new mathematical model for estimating the dependence of personal data protection on trust and the amount of information on social networks.

4. The main section

In the classical approach to the problem of personal data protection, there are many threats of loss of trust between users, which can be represented as a function:

\[ T_j = F(D_j, D_n, D_m, D_k), \]

where \( T_j \) – the set of threats of loss of trust between users; \( D_j \) – trust in the provision of services (a person trusts the party in the provision of quality services or resources by the provider); \( D_n \) – delegation trust describes the trust in the user (representative), who acts and makes decisions on behalf of the party he trusts; \( D_m \) – access trust describes the trust on the part (provider) to the user who is granted access to resources. This is access control. Used in authentication systems; \( D_k \) – contextual trust determines the degree of faith of the participant in the necessary systems and institutional mechanisms that support transactions and ensure network security [15–17].

Loss of such a quality as trust is a process that has a time interval. Denote the amount of information in the system as \( I \). The flow of information outside the information system through \( df \), the rate of change of this flow as \( df/dt \). It is logical that if the flow and the rate of change of flow are zero, then there is no leakage of information:

\[ df = 0; \quad df/dt = 0. \]

Leakage of information depends on the security of the system and the measures taken to neutralize threats to the security of personal data [18–21].

Let \( Z \) be an indicator of information system security. Let’s make the equation:

\[ \begin{align*}
\frac{df}{dt} &= Z_f Z + (C_1 + C_2)I - L_2(I_3 \sin^2 \omega t) - L_3(I_3 \sin^3 \omega t) - \ldots; \\
\frac{dZ}{dt} &= D_1 - I(C_1 + C_2) - K_2(Z_2 \sin^2 \omega t) - K_3(Z_2 \sin^3 \omega t) - \ldots.
\end{align*} \]

26
To solve the system of equations (3) let’s write system (4) in the form:

\[
\frac{dI}{dt} = \alpha Z + \beta_1 I - \sum_{k=2}^{\infty} L_k I_k \sin^k \omega t; \\
\frac{dZ}{dt} = \beta_2 I + \gamma - \sum_{k=2}^{\infty} K_k Z_k \sin^k \omega t,
\]

where \( \alpha = Z_p, \beta_1 = C_v + C_K, \beta_2 = -(C_{d2} + C_{e1}), \gamma = D_i. \)

Next, let’s use the exclusion method:

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

then:

\[
\frac{dI}{dt} = \frac{1}{\beta_2} \left( \frac{d^2Z}{dt^2} + \frac{1}{\omega} \sum_{k=2}^{\infty} \left( kK_k Z_k^1 \sin^{k-1} \omega t \cos \omega t \right) \right).
\]

Substitute in the first equation of the system:

\[
\frac{1}{\beta_2} \left( \frac{d^2Z}{dt^2} + \frac{1}{\omega} \sum_{k=2}^{\infty} \left( kK_k Z_k^1 \sin^{k-1} \omega t \cos \omega t \right) \right) = \\
= \alpha Z + \frac{\beta_1}{\beta_2} \left( \frac{dZ}{dt} - \gamma - \sum_{k=2}^{\infty} K_k Z_k \sin^k \omega t \right) - \sum_{k=2}^{\infty} L_k I_k \sin^k \omega t.
\]

Or

\[
\frac{d^2Z}{dt^2} - \beta_1 \frac{dZ}{dt} - \alpha \beta_2 Z = -\sum_{k=2}^{\infty} (kK_k Z_k^1 \sin^{k-1} \omega t \cos \omega t) - \beta_1 \gamma + \\
+ \beta_1 \sum_{k=2}^{\infty} K_k Z_k \sin^k \omega t - \beta_2 \sum_{k=2}^{\infty} L_k I_k \sin^k \omega t.
\]

Find the solution of the corresponding equation:

\[
Z'' - \beta_1 Z' - \alpha \beta_2 Z = 0.
\]

The characteristic equation has the form:

\[
\lambda^2 - \beta_1 \lambda - \alpha \beta_2 = 0.
\]

Let’s consider only the case for a positive discriminant of this equation:

\[
D = \beta_1^2 + 4 \alpha \beta_2 > 0 \Rightarrow \lambda_{1,2} = \frac{\beta_1 \pm \sqrt{\beta_1^2 + 4 \alpha \beta_2}}{2}.
\]

And

\[
Z_{out}(t) = c_1 e^{\frac{\beta_1 + \sqrt{\beta_1^2 + 4 \alpha \beta_2}}{2}} + c_2 e^{\frac{-\beta_1 + \sqrt{\beta_1^2 + 4 \alpha \beta_2}}{2}}
\]

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

\[ Z_{\text{inh}}(t) = c_1(t)e^{\frac{\beta_1 + \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} + c_2(t)e^{\frac{\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}}, \]  

(10)

where \( c_1(t), c_2(t) \) will be found from the system:

\[
\begin{align*}
\left\{ \begin{array}{l}
\frac{\beta_1 + \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}c_1(t)e^{\frac{\beta_1 + \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} + c_2(t)e^{\frac{\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} = 0, \\
c_1(t)\beta_1 + \sqrt{\beta_1^2 + 4\alpha\beta_2}e^{\frac{\beta_1 + \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} \beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}e^{\frac{\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} = N(t),
\end{array} \right.
\]

(11)

where

\[
N(t) = -\frac{1}{\omega} \sum_{k=2}^{\infty} (kK_kZ_0^2 \sin^{k-1}\omega t \cos \omega t) - \beta_1 \gamma + \beta_1 \sum_{k=2}^{\infty} K_kZ_0^2 \sin^k \omega t - \beta_2 \sum_{k=2}^{\infty} L_k \bar{t}_0^k \sin^k \omega t.
\]  

(12)

Let’s get:

\[
c_1(t)e^{\frac{\beta_1 + \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} = -c_2(t)e^{\frac{\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} \Rightarrow
\]

\[
\Rightarrow c_2(t)e^{\frac{\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} = -c_1(t) = \frac{\beta_1 + \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2} + \beta_1 - \frac{\sqrt{\beta_1^2 + 4\alpha\beta_2}}{2} = N(t),
\]

(13)

or

\[
c_2(t)e^{\frac{\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} = -N(t).
\]  

(14)

Then:

\[
c_2(t) = -\frac{1}{\sqrt{\beta_1^2 + 4\alpha\beta_2}} \int N(t)e^{\frac{-\beta_1 + \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} dt.
\]  

(15)

And

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

(16)

The mathematical model in the final form will look like:

\[
Z(t) = e^{\frac{-\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} \int N(t)e^{\frac{-\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} dt - e^{\frac{-\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} \int N(t)e^{\frac{-\beta_1 - \sqrt{\beta_1^2 + 4\alpha\beta_2}}{2}} dt. \]

(17)

In general, results are obtained in general: the dependence of personal data protection on trust is proportional to the constant parameters of protection.

In order to confirm the obtained results, let’s perform modeling in the MatLab environment.

In Fig. 1, the dependence of personal data protection (in relative units) on the amount of information in the system is the main parameter, and the parameter of trust in information.

In Fig. 2, the dependence of personal data protection (in relative units) on the parameter of trust in information is the main parameter and the amount of information in the system.
As it is possible to see from the simulation results, the protection of personal data directly depends on the amount of information and the parameters of trust in this information. The protection of personal data increases with the amount of reliable information and the amount of general information, which fully confirms the accuracy of the proposed method of assessing the protection of personal data.

5. Discussion of experimental results

The peculiarity of the method is that in addition to estimating the amount of information in the system, we use the rate of change of information flow, or rather the rate of leakage of information outside the information system of personal data exchange.

When there is no information leakage, the information leakage rate is also zero. The occurrence of information leakage directly depends on the information security indicator.
The resulting system consists of two equations. One equation is the dependence of the information security indicator, the second is the dependence of the amount of information leakage on the information security parameters.

The solution of these differential equations shows the dependences of the parameters of information protection from information leakage.

An additional feature of the proposed method is using the parameter of trust in personal information and the quantitative parameter of information leakage.

But this feature is based on the accepted restrictions. It isn’t taking into account the more detailed parameters of personal data protection, which in some cases may lead to an error in determining data protection.

Mathematical modeling of the proposed model proved that the protection of personal data is proportional to the reliability and trust with constant protection parameters.

The results of mathematical modeling have shown that with the increase of its quantity in the system, the trust in the information decreases, not according to the linear, but practically according to the exposition law. The simulation results showed that the confidence in the information after reaching a relative confidence factor of 0.45, is significantly reduced, which requires taking into account the greater parameters of information security.

Additional confirmation of the simulation results is shown in Fig. 2. The graphical results prove the statement that the confidence parameter in relative units from 0.85 to 0.99 has little effect on the parameter of increasing the amount of reliable information received.

In general, the protection of personal data increases with increasing trust parameters. This proves the adequacy of the model and is quite a favorable result.

Further development of the proposed method is a more detailed consideration of the parameters of information security.

6. Conclusions

A method for assessing the dependence of personal data protection on the amount of information in the system and trust in social networks is proposed.

Simulations for various types of changes in trust parameters and the amount of information in the system are conducted. All variants of solving the equation near the steady state of the system proved that, based on the conditions of the ratio of dissipation and natural frequency of confidence, and the attenuation of the latter to a certain value is carried out periodically, with decaying amplitude, or exponentially decaying law.

The obtained graphic materials fully showed that the protection of personal data increases with increasing factors of trust in information. Dependence of protection of personal data on trust is proportional at constant other parameters of protection.

The simulation results showed that the trust in information after reaching a relative confidence factor of 0.45, is significantly reduced, which requires taking into account the greater parameters of information security. When the parameters of trust in relative units from 0.85 to 0.99, the quantitative parameter of information has little effect on the parameter of increasing the amount of reliable information.

With the growth of the amount of information in the system and trust in information, the overall rate of information protection in modeling by the proposed method increases by 9% than modeling by old methods, which is a quite acceptable result.

References


