# An approach to building a probabilistic model of spreading a social engineering attack between two users

Anastasiia Khlobystova [1], Maxim Abramov[1] and Alexander Tulupyev[1,2]

[1] *St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), 14-th Linia, VI, № 39, St. Petersburg, 199178, Russia, dscs.pro*
[2] *St. Petersburg State University, Mathematics and Mechanics Faculty, Universitetskaya Emb., 7, St. Petersburg, 199034, Russia*

**Abstract**

The article proposes an approach to building probabilistic estimates of the spread of a social engineering attack between two users, which takes into account the informational influence of users. The theoretical significance of the study lies in creating a basis for the subsequent modeling of multi-step social engineering attacks. The practical significance lies in the formation of a base for software automation of identification of the most probable and critical scenarios for the spread of social engineering attacks, and, as a consequence, the creation of a tool that helps decision-makers to take effective measures to eliminate vulnerabilities.

**Keywords**

Social engineering attacks, probabilistic model, informational influence, attack spread estimate.

## 1. Introduction

Currently, social engineering remains the main trend in the field of information security threats [1]. In doing so, the worldwide move to digital technologies forces most organizations to expand their presence on the Internet, which leads, among other things, to greater vulnerabilities [2], and also encourages malefactors to invent new methods of social engineering attacks that can compromise even the most knowledgeable users [3, 4]. In this regard, there is a need to create a tool that helps to increase the level of security of employees and, indirectly, information systems from social engineering attacks. An integrated approach to the creation of such a tool, based on the construction of estimates of the security/vulnerability of users of information systems, was proposed in [5]. Also in their work, the authors propose approaches to modeling multistep social engineering attacks — attacks in which not one user is involved, but a chain. Such attacks represent a particular threat to the information security of an organization, as they are aimed at obtaining particularly valuable information and important documents [6].

To simulate multistep social engineering attacks, it is convenient to use the social graph of the organization's employees [5]. The nodes of such a graph are associated with users of information systems, and the edges are connections between them. According to [5] each nodes of the graph is associated with a probabilistic estimate of the success of a one-way social engineering attack on a user with whom this node is associated. When calculating the success scores for multistep social engineering attacks, it is necessary to take into account both the success scores of a one-way attack and the strength of the relationships between users. Online social networks can serve as an open source of information about the relationships between users, and one of the ways to estimates the intensity of relationships can be to estimate the informational influence of one user on another [7]. The goal of this study was to

Russian Advances in Fuzzy Systems and Soft Computing: Selected Contributions to the 10th International Conference «Integrated Models and Soft Computing in Artificial Intelligence» (IMSC-2021), May 17–20, 2021, Kolomna, Russian Federation
EMAIL: aok@dscs.pro (A. 1); mva@ dscs.pro (A. 2); alt@ dscs.pro (A. 3)
ORCID: 0000-0002-9811-5476 (A. 1); 0000-0002-5476-3025 (A. 2); 0000-0003-1814-4646 (A. 3)


improve the approach to constructing probabilistic estimates of the spread of social engineering attacks between two users by taking into account informational influence.

The novelty of the study lies in the development of a new approach to estimating the spread of a social engineering attack between two users. The significance of the study lies in creating the basis for the subsequent modeling of multistep social engineering attacks and laying the foundation for further software automation.

## 2. Related works

A model describing the process of disseminating multiple information is proposed in [8]. It is based on a population dynamics model. The distinctive advantages include taking into account the interaction between the disseminated information. The model is well suited for numerical simulation, but it is inapplicable when looking for specific scenarios of attack propagation.

The authors of [9] propose an approach based on the detection of overlapping communities to identify the nodes of a social graph, the dissemination of information through which would maximize the number of target notes that are influenced. In [10] the problem of maximizing the influence is also considered, while an adaptable probabilistic independent cascade model models the dissemination of information. These approaches can be useful in the subsequent stages of this study. In addition, in this section, the study [11] seems to be interesting, the purpose of which is to find a set of intermediate nodes that would help to maximize the probability of the influence of the initial node on the target one; to achieve this goal, two different diffusion models are proposed. An approach to quantifying social influence based on sociological theories and the theory of probabilistic reasoning was proposed in [12]. This approach is interesting in the context of identifying the most influential user on a social network, but it is not very useful for modeling the spread of a social engineering attack.

The model of information influence of one user on another, developed in [7], is based on public actions that users perform. For example, it can be posting a post, leaving a comment, sending a gift, etc. This model is most suitable for this study, since it can be used to obtain estimates of the interaction between two users, and generalize to a chain of users.

## 3. Model for obtaining probabilistic estimates of multistep social engineering attacks

According to [5], the estimate of the probability that a social engineering attack will spread between users is calculated by the formula $P = 1 - Q$, where $Q$ is estimate of the probability that a social engineering attack will not spread between users, taking into account the intensities of various types of communication. $Q = \prod_i (1 - q_i)^{n_t}$, where $q_i$ is estimate of the probability of success of a social engineering attack in one episode of interaction, $n_t$ is the number of episodes. $Q$ was proposed to calculate as follows: $Q = (1 - p_{\text{rel}}) \cdot (1 - p_{\text{likes}})^{\text{count\_likes}} \cdot (1 - p_{\text{reposts}})^{\text{count\_reposts}} \cdot (1 - p_{\text{com\_photos}})^{\text{count\_photos}} \cdot \cdot (1 - p_{\text{com\_groups}})^{\text{count\_groups}}$, where $p_{\text{rel}}$ is the probability of success of the spread of an attack from employee to employee, based on the type of communication declared in the social network. As $p_{\text{rel}}$ can be used the estimates obtained in [5]. $p_{\text{likes}}$, $p_{\text{reposts}}$, $p_{\text{com\_photos}}$, $p_{\text{com\_groups}}$ are estimates of the probability of success of a social engineering attack, obtained based on likes, reposts, shared photos and communities, $\text{count\_likes}, \text{count\_reports}, \text{count\_photos}, \text{count\_groups}$ is the number of such episodes

## 4. Model for calculating information influence of the user

Consider the model proposed in [7] in the context of social engineering attacks.

First of all, we introduce a set of agents $N$ as a set of users of an information network $U = \{U_1,\ldots,U_n\}$, which will consist of employees of the analyzed organization and their friends, information about which can be obtained from open sources, namely from online social networks.

The set $K$ consists of valid types of actions (for example, posting, liking). For unambiguous identification of actions performed by users, we introduce $\Delta$ a set of actions, which will consist of elements $a \in \Delta,\ a = (U_i, K_i, T_i),$ where $U_i$ is the user $(U_i \in U),$ who performed the action of the type $K_j$, $(K_j \in K)$ at the moment of time $T_k$, $(T_k \in T)$. We will assume that one user cannot perform an action more often than 1 time per second, and the set $T$ is discrete.

Let's define a binary partial order relation — a cause relation $a \to b$. For example, let action $a$ is the publication of a post on online social network by a user $U_i$ at a moment in time $T_i$, an action $b$ is commenting to this post by a user $U_j$ at a moment in time $T_j$ $(T_i < T_j),$ then we can say that the action $a$ is the cause of the action $b$ and for them we can set the relation $a \to b$.

For a subset of actions $A \subseteq \Delta$ we define the set of all its consequences $\pi(A) = \{b \in \Delta \mid \exists a \in A : a \to b\}$. The introduction of such a set is required to further compare all the consequences of the actions of a particular user to the set of all actions that he performed, and is one of the ways to obtain a quantitative assessment of the influence of one user on another. Actions that are not a consequence of any other actions can be distinguished into a separate set $\Delta^0 = \{b \in \Delta \mid \exists a \in A : a \to b \Rightarrow a = b\}$ — a set of initial actions.

Let us introduce a significance function $\Phi : 2^{\Delta} \to [0;+\infty)$ such that $\Phi(\Delta) > 0$, $\Phi$ is monotonic function, if $A, B \subseteq \Delta, A \subset B,$ then $\Phi(A) \le \Phi(B),$ $\Phi$ is additive, $\Phi(A \cup B) = \Phi(A) + \Phi(B),\ A, B, A \cup B \subseteq \Delta.$ In other words, domain of the significance function is all subsets of the set of actions, and this function assigns a non-negative number to a subset of the set of actions.

Function $\alpha : \Delta \to U$ matches the action of the agent who performed it.

According to [7] the function of the influence user $U_i$ on the user $U_j$, on the one hand, can be based only on the initial actions of the agent, then it is set as follows:

$$\chi(U_i, U_j) = \begin{cases} \dfrac{\Phi(\pi(\delta_{U_i}^0) \cap \delta_{U_j})}{\Phi(\delta_{U_j})}, & \Phi(\delta_{U_j}) > 0 \\ 0, & \Phi(\delta_{U_j}) = 0 \end{cases}, \qquad (1)$$

where $\delta_{U_j} = \{a \in \Delta \mid \exists\ \alpha(a) \in U_j\}$ is a set of agent actions $U_j \in U$, $\pi(\delta_{U_i}^0) = \{b \in \Delta^0 \mid \exists\ a \in \delta_{U_i} :\ a \to b\}$ is a set of actions that are the consequences of the initial actions of the agent $U_i \in U$.

On the other hand, the function of influencing the user $U_i$ on the user $U_j$ can be based on all the actions of the agent, in which case it will be defined as follows:

$$\chi(U_i, U_j) = \begin{cases} \dfrac{\Phi(\pi(\delta_{U_i}) \cap \delta_{U_j})}{\Phi(\delta_{U_j})}, & \Phi(\delta_{U_j}) > 0 \\ 0, & \Phi(\delta_{U_j}) = 0 \end{cases}, \qquad (2)$$

where $\delta_{U_j} = \{a \in \Delta \mid \exists\ \alpha(a) \in U_j\}$ is a set of agent actions $U_j \in U$, $\pi(\delta_{U_i}) = \{b \in \Delta \mid \exists\ a \in \delta_{U_i} :\ a \to b\}$ many actions that are the consequences of all actions of the agent $U_i \in U$.

## 5. Probabilistic model of a multistep social engineering attack, taking into account informational influence

To apply models (1) and (2) of information influence, we first of all expand the representation of the action as follows $a=\left(U_i,K_j,T_k,a\right)$, where $U_i$ is the user $(U_i\in U)$, who performed an action of type $K_j$, a point in time $T_k$, $\left(T_k\in T\right)$ in response to the action $a=(U_{l_1},K_{l_2},T_{l_3})$, $U_{l_1}\in U$, $K_{l_2}\in K$, $T_{l_3}\in T, T_k<T_{l_3}$. If the action $a$ is not a consequence of any other action: $a=\varnothing$. Then the set of initial actions $\Delta^0$ can be represented as $\Delta^0=\left\{a\in\Delta\middle|pr_4(a)=\varnothing\right\}$, where $pr_i(a)$ is a projection operation that selects the $i$ component of an element $a$. The set of all consequences of a subset of actions $A\subseteq\Delta$ look like $\pi(A)=\left\{b\in\Delta\middle|\exists a\in A: pr_4(b)=(pr_1(a),pr_2(a),pr_3(a))\right\}$. The inclusion of the action $a$ the element $a$ is necessary for better formalization of the considered models, as well as to simplify in the future the programmatic specification of these objects. At the same time, such a task does not create a situation of infinite nesting of actions.

Let's set the significance function $\Phi:2^{\Delta}\to[0;+\infty)$, as the cardinality of the set to which it is applied $\Phi(A)=\left|A\right|, A\subset\Delta$. Then (2) takes the following form.

$$\chi(U_i,U_j)=\begin{cases}\dfrac{\left|\pi(\delta_{U_i})\cap\delta_{U_j}\right|}{\left|\delta_{U_j}\right|}, & \delta_{U_j}\neq\varnothing\\ 0, & \delta_{U_j}=\varnothing\end{cases}. \qquad (3)$$

Let's reduce $\chi(U_i,U_j)$ to probability estimates. We will assume that space $\Omega$ consists of elementary events $\omega_{ij}$ (the user $U_i$ has an impact on the user $U_j$). Let's introduce the following restriction: $\Omega$ is finite, $\left|\Omega\right|\le m$, where $m=2^n$, $n$ is the number of peaks in the considered social graph of the organization's employees. σ-algebra $\mathfrak{A}$ is the set of subsets $\Omega$, $\mathfrak{A}=2^{\Omega}$. Let's check that the function of influence of one user on another is a probabilistic measure:

- Non-negativity is evident from the assignment.
- Additivity was proved in [7] the proposition 3.2.
- Finiteness follows from additivity and the fact that $\left|\Omega\right|\le m$.

Thus, $\chi(U_i,U_j)$ can be used to estimate the probability that a social engineering attack will spread between users. The sought formula for the propagation of a social engineering attack from user $U_i$ to $U_j$ will be as follows:

$$P_{ij}=1-(1-p_{rel})\cdot(1-\chi(U_i,U_j)),$$

where $p_{rel}$ is the probability of success of the spread of an attack from employee to employee, based on the type of communication declared in the online social network, $\chi(U_i,U_j)$ is the estimate of the probability of the influence of one user on another.

Let us give an example of the calculation according to the resulting model. For clarity of presentation, we will restrict ourselves to three considered users and three permissible types of their actions. The set of users $U=\left\{U_1,U_2,U_3\right\}$. The set of users permissible actions $K=\left\{K_1,K_2,K_3\right\}$, where $K_1$ is posting, $K_2$ is post commenting, $K_3$ is post liking. User action sets are defined as follows:

$\delta_{U_1}=\left\{\left(U_1,K_1,T_1,\varnothing\right),\left(U_1,K_3,T_4,(U_2,K_1,T_2)\right),\left(U_1,K_1,T_7,\varnothing\right)\right\}$,

$\delta_{U_2}=\left\{\left(U_2,K_1,T_2,\varnothing\right),\left(U_2,K_2,T_6,(U_3,K_1,T_5)\right),\left(U_2,K_3,T_8,(U_1,K_1,T_7)\right)\right\}$,

$\delta_{U_3} = \{(U_3, K_1, T_3, \varnothing), (U_3, K_1, T_5, \varnothing)\}$. Let the user $U_2$ marked $U_1$ as a "best friend", then, according to [13] $p_{rel}^{12} = 0,7838,$ other types of connections are characterized by being in the list of friends $p_{rel} = 0,2938.$

Let us calculate the probability of success of the spread of a social engineering attack from user $U_1$ to user $U_2$.

$$P_{12} = 1-(1-0,7838)\cdot\left(1-\frac{\left|\pi(\delta_{U_1})\cap\delta_{U_2}\right|}{\left|\delta_{U_2}\right|}\right) = 1-0,2162\cdot\left(1-\frac{1}{3}\right)\approx 0,8559.$$

Let us calculate the probability of success of the spread of a social engineering attack from user $U_1$ to user $U_3$.

$$P_{13} = 1-(1-0,2938)\cdot\left(1-\frac{\left|\pi(\delta_{U_1})\cap\delta_{U_3}\right|}{\left|\delta_{U_3}\right|}\right) = 1-0,7062\cdot\left(1-\frac{0}{2}\right)\approx 0,2938.$$

That is, according to the estimates obtained, the probability that a social engineering attack will spread from user $U_1$ to user $U_2$ is higher than from $U_1$ to $U_3$.

## 6. Conclusion

Thus, the article proposes an approach to constructing probabilistic estimates of the spread of a social engineering attack between two users by taking into account the informational influence of users. The theoretical significance of the study lies in creating a basis for the subsequent modeling of multistep social engineering attacks. The practical significance lies in the formation of a base for software automation of identification of the most probable and critical scenarios for the development of social engineering attacks, and, as a consequence, the creation of a tool that helps decision-makers to take effective measures to eliminate vulnerabilities.

## 7. Acknowledgements

The research was carried out in the framework of the project on state assignment SPC RAS № 0073-2019-0003; supported by St. Petersburg University, project № 73555239; with the financial support of the RFBR №20-07-00839.

## 8. References

[1] Information security threats of 2021: The Top List, 2021. URL: https://itglobal.com/company/blog/infosec-threats-2021/

[2] Top 10 Cybersecurity Trends in 2021, 2021. URL: https://www.securebrain.co.jp/eng/blog/cybersecurity-trends-2021/

[3] N. Klimburg-Witjes, A.Wentland, Hacking Humans? Social Engineering and the Construction of the “Deficient User” in Cybersecurity Discourses, Science, Technology, & Human Values, 2021, 0162243921992844. doi:10.1177/0162243921992844

[4] Z. Wang, H. Zhu, L. Sun, Social Engineering in Cybersecurity: Effect Mechanisms, Human Vulnerabilities and Attack Methods, IEEE Access 9 (2021) 11895–11910. doi:10.1109/ACCESS.2021.3051633

[5] M. Abramov, T. Tulupyeva, A. Tulupyev, Social Engineering Attacks: social networks and user security estimates, SUAI, St. Petersburg, 2018.

[6] T. Koide, D. Chiba, M. Akiyama, K. Yoshioka, T. Matsumoto, To Get Lost is to Learn the Way: An Analysis of Multi-Step Social Engineering Attacks on the Web, IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences 104, 1 (2021) 162–181. doi:10.1587/transfun.2020CIP0005

[7] D.A. Gubanov, A.G. Chkhartishvili, Influence levels of users and meta-users of a social network, Automation and Remote Control 79, 3 (2018) 545–553. doi:10.1134/S0005117918030128
[8] A. Chen, X. Ni, H. Zhu, G. Su, Model of warning information diffusion on online social networks based on population dynamics, Physica A: Statistical Mechanics and its Application 567 (2021) 125709. doi:10.1016/j.physa.2020.125709
[9] Z. Wang, C. Sun, J. Xi, X. Li, Influence maximization in social graphs based on community structure and node coverage gain, Future Generation Computer Systems 118 (2021), 327–338. doi:10.1016/j.future.2021.01.025
[10] M. Kahr, M. Leitner, M. Ruthmair, M. Sinnl, Benders decomposition for competitive influence maximization in (social) networks, Omega 100 (2021) 102264. doi:10.1016/j.omega.2020.102264
[11] Y. Zhang, J. Guo, W. Yang, W. Wu, Targeted Activation Probability Maximization Problem in Online Social Networks, IEEE Transactions on Network Science and Engineering 8, 1 (2021) 294–304. doi:10.1109/TNSE.2020.3037106
[12] L. Vega, A. Mendez-Vazquez, A. Lopez-Cuevas, Probabilistic reasoning system for social influence analysis in online social networks, Social Network Analysis and Mining, 11, 1 (2021), 1–20. doi:10.1007/s13278-020-00705-z
[13] A.O. Khlobystova, T.V. Tulupyeva, A.G. Maksimov, A.A. Korepanova, An approach to quantification of relationship types between users based on the frequency of combinations of non-numeric evaluations, Proceedings of the 4th International Scientific Conference “Intelligent Information Technologies for Industry” (IITI’19), Advances in Intelligent Systems and Computing, Springer, Cham 1156, (2019) 206–213. doi:10.1007/978-3-030-50097-9_21