World Applied Sciences Journal 25 (3): 410-415, 2013
ISSN 1818-4952

DOI: 10.5829/idosi.wasj.2013.25.03.13319

Analytical Models of Information-Psychological Impact of Social Information Networks on Users

Grigoriy Aleksandrovich Ostapenko Larisa Vladimirovna Parinova, Vladimir Ivanovich Belonozhkin, Igor Leonidovich Bataronov and Konstantin Vladimirovich Simonov

Voronezh State Technical University, Voronezh, Russia

Submitted: Aug 21, 2013; **Accepted:** Sep 29, 2013; **Published:** Oct 3, 2013

Abstract: The article is devoted to the modeling of information-psychological impacts (IPI) of social information networks (SIN) on users, introduces the concept of IPI, shows a risk of IPI and discusses the possible types of information-psychological threats for the users of the SIN. The article substantiates the probabilistic model of IPI impact on the users of the SIN describing an approach to determination of the number of the most visited sites of SIN and further, the total probability of successful attacks on the most visited sites of the SIN of the total number of sites - the subjects of IPI. We introduce and assess the concept of detriment of the IPI for SIN users, used further for the building of the risk-model of IPI for SIN users, which includes the analytical expressions for elementary, total and normalized risk of IPI, absolute and relative indicators of protection of the SIN and approximate estimation of the risk mode of IPI.

Key words: Information-psychological impact · Risk · Detriment · Risk-model · Probabilistic model · Social information network

INTRODUCTION

Negative information-psychological impact (IPI) is the influence on mentality of a man or group of the persons (i.e. in spite of their will), carried out using the special means and methods of influence and resulting in negative consequences for the individual, society and state [1].

The risk of IPI is value equals to the product of the probability of detriment to SIN users from the IPI on the amount of detriment [2].

Information-psychological danger for SIN users implies the activities of the various groups and associations of people, in particular, some political parties, socio-political movements, nationalist and religious organizations, financial-economic and commercial structures, lobbying and mafia groups, etc.

General: The visiting of the SIN portals supposes a trust technological and psychological relationships establishes between the user and the server. The user expects that the website will provide him a safe content. In addition, the user does not expect any attacks at particular website [3]. The attackers can easily achieve their goals exploiting this trust by placing the specific target IPI information on the websites.

The threats of IPI for SIN users can be divided into the following types:

- Posting of information through the cross-site scripting (XSS). The cross-site scripting allows an attacker to send executable code to the server, which will be forwarded to the user's browser;
- Posting of information in OS commands. The attacks of this class are aimed at executing of operating system commands on the Web-server by manipulating the input data;
- Posting of information through the introduction of server scripts. The attacks of this class allow an attacker to send executable code that further will be executed on the Web-server;
- Posting of information by substitution of the content; using this technique, the attacker forces the user to believe that the webpages have been generated by the Web-server and are not transferred from an external source;

Corresponding Author: Ostepenko, Voronezh State Technical University, Moskovsky prospect, 14, 394026, Voronezh, Russia.

- Posting of information by HTTP response splitting; using this vulnerability, the attacker sends the specific request to the Web-server which response is interpreted by attack target as two different responses [4].

If denote the probability of posting of required information for IPI on the SIN website as p_i caused by realization of i threat then the total probability of the posting of this types of information can be determined by the equation [5]:

$$p = 1 - \prod_i (1 - p_i) .$$

In practical calculations, p_i is substituted by particular numerical values used in equation [5] for calculation of p. Further, for modeling of SIN attacks, we have accepted that $p = const$ and assessed if the risk of detriment of the posting of information of SIN websites depends on different values of p.

The number of attacked portals n is a random variable because the certain number of attacked websites is unknown.

To determine the probability of attack of n number of websites, we used a discrete analogue of gamma distribution:

$$P_n = \frac{n^{a-1} e^{-bn}}{\sum_{m=0}^{N} m^{a-1} e^{-bm}}$$

where $m \in N, a \geq 1, b > 0$, N - total number of websites.

Therefore the hackers attacking SIN, are aimed at maximum distribution of information required for the IPI among the users, it is apparent that the most visited SIN sites first will undergone the attacks.

According to the statistics of visit of the SIN sites, we can determine the most visited among them. To calculate the probability of a visit to a website, we introduce the following designations:

M_i = Number of visits to the i website over a given period;
M = Total number of visits over a given period.

Then the estimate of the probability of visiting the i website can be determined using the equation:

$$p_i = \frac{M_i}{M} .$$

The number of the most visited sites (n_0) can be determined by equation:

$$n_0 = \sum_{i=1}^{N} \alpha_i p_i ,$$

where $\alpha_i = \begin{cases} 0, p_i < l \\ \frac{1}{p_i}, p_i \geq l \end{cases}$, l - the minimum probability of portal visits to consider this website as the most visited.

The attacks most visited sites are the most dangerous. Their number is denoted by the random variable - r, therefore the number of attacked sites (n) among most visited (n_0) is unknown. The hypergeometric distribution was used to describe r:

$$P_r(r, N, n_0, n) = \frac{C_{n_0}^{r} C_{N-n_0}^{n-r}}{C_N^n} = \\ = \frac{n_0! \cdot n! \cdot (N - n_0)! \cdot (N - n)!}{r! \cdot N! \cdot (n_0 - r)! \cdot (n - r)! \cdot ((N - n_0) - (n - r))!} ,$$

Since every SIN web-server stores the various confidential information its security can be ensured by implementation of appropriate means and protection measures [6] and as a result, the probabilities of information posting at particular website will be different. However, we will be able to assess the risk for the worst scenario, when all websites are poorly protected, accepting that the probability of placing of the information on each portal is equal to the probability of placing information on the less secure site. Administrating the data of maximum risk, we are able to manage a actual risk.

If assume that an attack of any particular SIN website is independent of the attacks of other websites and the probability of placing of the information on the portals is equal for each website, it is possible to implement the binomial distribution for determination of the probability of the successful attacks k [7] at the attack r of the most visited sites (k successful postings of information on r different websites):

$$P_r(k) = C_r^k p^k q^{r-k} ,$$

where $C_r^k = \frac{r!}{k!(r-k)!}, q = 1 - p$.

Total probability of successful attacks on r number of most visited websites at the total number of sites under attack n was obtained by equation of full probability:

$$P_k = \sum_{n=k}^{N} P_n \sum_{r=k}^{n} P_r \cdot P_r(k) =$$

$$= \frac{p^k \cdot n_0! \cdot (N-n_0)!}{k! \cdot N! \cdot \sum_{m=0}^{N} m^{a-1} e^{-bm}} \cdot \sum_{n=k}^{N} \left(\frac{n! \cdot (N-n)!}{e^{bn} \cdot n^{1-a}} \times\right.$$

$$\left. \times \sum_{r=k}^{\min(n_0;n)} \left(\frac{(1-p)^{r-k}}{(n_0-r)! \cdot (n-r)! \cdot (r-k)! \cdot ((N-n_0)-(n-r))!}\right)\right).$$

In this case, the detriment of IPI is the number of attacked websites. Modern technology of surveys allows the approximate evaluation of the percentage of successful attacks, which depends on many factors such as professional implementation of the methods of influence, intensity and duration of influence, the nature of susceptible information, psychological conditions and the general awareness of a user, its age and sex. Sociological polls, which proportionally represent all strata of the population, reveal that the 30-40% of the entire adult population is susceptible to suggestion. Thus, u_i - detriment of the attack of the i SIN resource, is the average number of people succumbed to suggestion from viewing information used for IPI.

Hence, the detriment of the successful posting of information on k attacked the most visited websites of the SIN, will be as follows [8]:

$$U_k = \sum_{i=1}^{k} \beta \cdot \int_{T_N}^{T_K = T_N + T_D} \theta_i(t) \cdot \gamma_i \cdot h_i(t) dt,$$

where β- coefficient characterizing the susceptibility of visitors to the SIN websites to the harmful information;

$\theta_i(t)$ = Duration of stay of user at i resource;
γ_i = The weight function of periodicity of the visit bursts of the i resource depending on the observation period and the period of predictions;
$h_i(t)$ = Frequency of visits of the i SIN resource over considered time period;
T_N = Duration of posting of harmful information on i SIN resource;
T_K = Duration of neutralize (removal) of harmful information from i SIN resource;
T_D = Duration of existence of the harmful information on SIN resource.

Only detriment of the attacks of the most visited sites was considered as elements of the detriment since the detriment of attacks of the websites with minimum visits is insignificant.

The mean and maximum detriments of IPI are:

$$U_{av} = \frac{\sum_{i=1}^{n_0} \beta \cdot \int_{T_N}^{T_K = T_N + T_D} \theta_i(t) \cdot \gamma_i \cdot h_i(t) dt}{n_0};$$

$$U_{\max} = \sum_{i=1}^{n_0} \beta \cdot \int_{T_N}^{T_K = T_N + T_D} \theta_i(t) \cdot \gamma_i \cdot h_i(t) dt \cdot$$

The value of the average detriment was used for further calculations and normalization [9]. The average detriment of k number of successful attacks was divided by the maximum possible detriment on condition that the most visited websites have been attacked:

$$\overline{u}_k = \frac{k \cdot \sum_{i=1}^{n_0} \int_{T_N}^{T_K = T_N + T_D} \theta_i(t) \cdot \gamma_i \cdot h_i(t) dt}{n_0 \cdot \sum_{i=1}^{n_0} \int_{T_N}^{T_K = T_N + T_D} \theta_i(t) \cdot \gamma_i \cdot h_i(t) dt} = \frac{k}{n_0}.$$

Therefore, the equation for assessment of elementary risk is [8]:

$$risk(\overline{u_k}) = \overline{u_k} P_k =$$

$$\frac{k}{n_0} \cdot \frac{p^k \cdot n_0! \cdot (N-n_0)!}{k! \cdot N! \cdot \sum_{m=0}^{N} m^{a-1} e^{-bm}} \cdot \sum_{n=k}^{N} \left(\frac{n! \cdot (N-n)!}{e^{bn} \cdot n^{1-a}} \times\right.$$

$$\left. \times \sum_{r=k}^{\min(n;n_0)} \left(\frac{(1-p)^{r-k}}{(n_0-r)! \cdot (n-r)! \cdot (r-k)! \cdot ((N-n_0)-(n-r))!}\right)\right)$$

The total risk was determined (on condition of independent IPI) by the following equation:

$$Risk \sum = \sum_{k=1}^{n_0} risk(\overline{u_k}) \cdot$$

If we introduce M_{Pk}- the mathematical expectation of the probability of distribution P_k then:

$$M_{P_k} = \sum_{l=0}^{n_0} \left(\frac{p^l \cdot (n_0-1)! \cdot (N-n_0)!}{(l-1)! \cdot N! \cdot \sum_{m=0}^{N} m^{a-1} e^{-bm}} \cdot \sum_{\eta=l}^{N} \left(\frac{\eta! \cdot (N-\eta)!}{e^{b\eta} \cdot \eta^{1-a}} \times\right.\right.$$

$$\left.\left. \times \sum_{\delta=l}^{\min(\eta;n_0)} \left(\frac{(1-p)^{\delta-l}}{(n_0-\delta)! \cdot (\eta-\delta)! \cdot (\delta-l)! \cdot ((N-n_0)-(\eta-\delta))!}\right)\right)\right) \cdot$$

Thus, the analytical expression for the normalized risk is [8]:

$$\overline{Risk(\overline{u_k})}) = \frac{1}{M_{P_k}} risk(\overline{u_k}) =$$

$$= \frac{p^k \cdot \sum_{n=k}^{N} (\frac{n! \cdot (N-n)!}{e^{bn} \cdot n^{1-a}}}{(k-1)! \cdot \sum_{l=0}^{n_0} \frac{p^l}{(l-1)!} \cdot \sum_{\eta=l}^{N} (\frac{\eta! \cdot (N-\eta)!}{e^{b\eta} \cdot \eta^{1-a}}} \times$$

$$\times \frac{\sum_{r=k}^{\min(n;;n_0)} (\frac{(1-p)^{r-k}}{(n_0-r)! \cdot (n-r)! \cdot (r-k)! \cdot ((N-n_0)-(n-r))!}))}{\sum_{\delta=l}^{\min(n;;n_0)} (\frac{(1-p)^{\delta-l}}{(n_0-\delta)! \cdot (\eta-\delta)! \cdot (\delta-l)! \cdot ((N-n_0)-(\eta-\delta))!})))} .$$

Hence, the analytic expression for the absolute and the relative value of protection is [9]:

$$E_{abs} = 1 - \sum_{k=1}^{n_0} risk(\overline{u_k}) =$$

$$= 1 - \sum_{k=1}^{n_0} (\frac{p^k \cdot (n_0-1)! \cdot (N-n_0)!}{(k-1)! \cdot N! \cdot \sum_{m=0}^{N} m^{a-1} e^{-bm}} \cdot \sum_{n=k}^{N} (\frac{n! \cdot (N-n)!}{e^{bn} \cdot n^{1-a}} \times$$

$$\times \sum_{r=k}^{\min(n;n_0)} (\frac{(1-p)^{r-k}}{(n_0-r)! \cdot (n-r)! \cdot (r-k)! \cdot ((N-n_0)-(n-r))!})) .$$

$$E_{rel} = \frac{\sum_{k=1}^{n_0} (1 - risk(\overline{u_k}))}{\sum_{k=1}^{n_0} risk(\overline{u_k})} = \frac{\sum_{k=1}^{n_0} (1 - \frac{p^k \cdot (n_0-1)! \cdot (N-n_0)!}{(k-1)! \cdot N! \cdot \sum_{m=0}^{N} m^{a-1} e^{-bm}}}{\sum_{k=1}^{n_0} (\frac{p^k \cdot (n_0-1)! \cdot (N-n_0)!}{(k-1)! \cdot N! \cdot \sum_{m=0}^{N} m^{a-1} e^{-bm}}} \times$$

$$\times \frac{\sum_{n=k}^{N} (\frac{n! \cdot (N-n)!}{e^{bn} \cdot n^{1-a}} \cdot \sum_{r=k}^{\min(n;n_0)} (\frac{(1-p)^{r-k}}{(n_0-r)! \cdot (n-r)! \cdot (r-k)! \cdot ((N-n_0)-(n-r))!})))}{\sum_{n=k}^{N} (\frac{n! \cdot (N-n)!}{e^{bn} \cdot n^{1-a}} \cdot \sum_{r=k}^{\min(n;n_0)} (\frac{(1-p)^{r-k}}{(n_0-r)! \cdot (n-r)! \cdot (r-k)! \cdot ((N-n_0)-(n-r))!})))} .$$

Further, the mode and maximal risk were calculated by solution of the system of inequalities:

$$\begin{cases} \overline{Risk(\overline{u_k})}_k \geq \overline{Risk(\overline{u_k})}_{k+1} \\ \overline{Risk(\overline{u_k})}_k \geq \overline{Risk(\overline{u_k})}_{k-1} \end{cases} .$$

The system of equations was received by substitution of the values of risk into the system, which can be hardly analytically solved in general form regards to the parameter k. Therefore, the approximate evaluation of risk mode based on theoretical calculations was used [10].

Distribution of the probability of detriment as a result of the placed information is the sum of the product of probability, which is distributed according to the law of the discrete analog of the gamma distribution and the sum of the product of two probabilities, one of which is distributed according to hypergeometric distribution law and the second to binomial law [9]. Each distribution was considered separately and the possible types of the products of these distributions at different values of k were assessed. We have considered the maximum of distribution $P_n^*(k) = \sum_{r=k}^{n} P_r \cdot P_r(k)$, at $n = const$, as well as hyper geometric distribution: $P_r(r, N, n_0, n) = \frac{C_{n_0}^r C_{N-n_0}^{n-r}}{C_N^n}$.

Its mathematical expectation is:

$$M_{P_n(k)} = \sum_{r=1}^{n} (r \cdot \frac{C_{n_0}^r C_{N-n_0}^{n-r}}{C_N^n}) =$$

$$\sum_{r=1}^{n} (r \cdot \frac{n_0! \cdot n! \cdot (N-n_0)! \cdot (N-n)!}{r! \cdot N! \cdot (n_0-r)! \cdot ((N-n_0)-(n-r))!}) =$$

$$= \frac{n_0! \cdot n! \cdot (N-n_0)! \cdot (N-n)!}{N!} \cdot$$

$$\cdot \sum_{r=1}^{n} (\frac{1}{(r-1)! \cdot (n_0-r)! \cdot (n-r)! \cdot ((N-n_0)-(n-r))!}) = \frac{n \cdot n_0}{N} .$$

Thus, hypergeometric distribution has an extreme near point $n_0^* = \frac{n \cdot n_0}{N}$.

The mathematical expectation of binomial distribution $P_r(k) = C_r^k p^k q^{r-k}$ is [10]: $M_{P_n(k)} = r \cdot p(p+q)^{r-1} = r \cdot p$ that is a whole part of $[r.p]$, with the an extreme near point $k_0^* = r \cdot p$.

Whereas, the summation was carried out by parameters n, k, n_0, $N = const$, we have determined inverse relationship $P_k(r)$, with an extreme near point $n_0^{**} = \frac{k_0^*}{p}$.

At multiplication of this dependence and hypergeometric distribution, the maximum will be reached if the extremes of these dependencies are congruent i.e. $n_0^* = n_0^{**}$ or:

$$\frac{n \cdot n_0}{N} \approx \frac{k_0^*}{p} .$$

When extremes are not congruent, the reminder of first dependence compensates the extreme of the second dependence that results in reduction of the value of the product and an increase of deviation between extremes tends to zero.

Thus, the maximum of the distribution $P_n^*(k) = \sum_{r=k}^{n} P_r \cdot P_r(k)$ is near the point k_o [10] and equal to $k_0^* \approx \frac{p \cdot n \cdot n_0}{N}$ and the maximum of probability of the detriment as a result of posting of information is $P_k = \sum_{n=k}^{N} P_n \sum_{r=k}^{n} P_r \cdot P_r(k)$. First, we have considered normalized discrete analogue of gamma distribution:

$$P_n = \frac{n^{a-1} e^{-bn}}{\sum_{m=0}^{N} m^{a-1} e^{-bm}}.$$

Its mathematical expectation and general equation of this distribution were calculated as:

$$M_{P_n} = \sum_{m=0}^{N} m \frac{m^{a-1} e^{-bm}}{\sum_{l=0}^{N} l^{a-1} e^{-bl}} = \sum_{m=0}^{N} \frac{m^{a} e^{-bm}}{\sum_{l=0}^{N} l^{a-1} e^{-bl}}.$$

General equation of distribution will have a maximum near the point

$$n_0^* = \sum_{m=0}^{N} \frac{m^{a} e^{-bm}}{\sum_{l=0}^{N} l^{a-1} e^{-bl}}$$

Therefore, the summation was carried out by the parameter, *a, b,* n_0, $N = const$, we have determined the inverse relationship $P_k^*(n)$, which the maximum near the point $n_0^{**} = \frac{k_0^* \cdot N}{p \cdot n}$.

At the multiply of this dependence and a discrete analog of the gamma distribution, the maximum will be reached, when the extremes of these dependencies are congruent $n_0^* = n_0^{**}$, or:

$$\sum_{m=0}^{N} \frac{m^{a} e^{-bm}}{\sum_{l=0}^{N} l^{a-1} e^{-bl}} \approx \frac{N \cdot k_0^*}{n_0 \cdot p}.$$

If extremes are not congruent, the reminder of first dependence compensates the maximum another dependence that results in reduction of the product and tends to zero at an increase of deviations between extremes.

Hence, the risk mode is located near the point k_o and equal to:

$$k_0 \approx p \frac{n_0}{N} \sum_{m=0}^{N} \frac{m^{a} e^{-bm}}{\sum_{l=0}^{N} l^{a-1} e^{-bl}}.$$

Knowing the risk mode, we have calculated the limiting risk of IPI on the SIN users, which will enabled us to manage a risk, in particular, calculate the dynamics of movement of the risks of detriments caused by posting of information on the SIN websites [11].

Thus, the SIN imposes the millions of computers and their users a risk of infection by malicious software. Implementation of these threats can result in the material and moral detriments of users, which evidences the importance of the research of the problems of the SIN to improve the protection of the SIN users by building a probabilistic and risk-models of IPI.

CONCLUSIONS

The article describes the problem of information-psychological impacts on users of social information networks, substantiates the probabilistic model and modeling of the risk-model of IPI for the SIN users and its basic parameters have been calculated. The study results provide the basis for further studies of this subject, in particular for the calculation of dynamics of the risks of detriments caused by posting of the information on the SIN websites used for the protection control.

ACKNOWLEDGMENTS

This study was performed within the framework of the program “Management of information risks and ensuring of security of information and communication systems” at the Voronezh State Technical University. The authors are grateful to the Department of Information Security Systems of the Voronezh State Technical University for assistance in this study.

REFERENCES

1. Gubanov, D.A., 2009. A Review of On-line Systems of Reputation and Trust. Moscow: IPU RAN, pp: 25.
2. Introduction to the Social Internet Networks. www.social-networking.ru.
3. Wu, F., B. Huberman, L. Adsmic and J. Tyler, 2004. Information Flow in Social Groups. Statistical and Theoretical Physics, 337: 327-335.
4. Ventsel, E.S., 1969. The Theory of Probability. Moscow: Nauka, pp: 576.
5. Watts, D., 2007. Networks, Influentials and Public Opinion Formation. Journal of Consumer Research, pp: 123-134.

6. Albert, R., H. Jeong and A.L. Barabasi, 1999. Diameter of the World-Wide Web. Nature, pp: 130-131.
7. A Low of Russian Federation "On Safety" Approved on March 5, 1992. Rossiiskaya Gazeta, May 1992, pp: 2446-1.
8. Axelrod, R., 1997. The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. Princeton: Princeton University Press, pp: 232.
9. Watts, D.J., 2002. A Simple Model of Information Cascades on Random Networks. Proc. Natl. Acad. Sci. U.S.A. 99(5766): 71.
10. Osmolovskiy, S.A., 2004. Stochastic Methods of Information Protection. Moscow: Radio I Svyaz, pp: 320.
11. Potterat, J.J., R.B. Rothenberg and S.Q. Muth, 1999. Network Structural Dynamics Acid Infectious Disease Propagation. Int. J. STD AIDS, 10(182): 85.