METHODOLOGY FOR ASSESSING THE RISK ASSOCIATED WITH INFORMATION AND KNOWLEDGE LOSS MANAGEMENT

Hanna YAROVENKO1*, Yuriy BILAN2, 3, Serhiy LYEOVONOV1, Grzegorz MENTEL4

1Economic Cybernetics Department, Sumy State University, Sumy, Ukraine
2Alexander Dubceck University of Trencin, Trencin, Slovakia
3Sumy State University, Sumy, Ukraine
4Faculty of Management, Rzeszow University of Technology, Rzeszów, Poland

Received 04 August 2020; accepted 04 November 2020

Abstract. In practice, there is a massive time lag between data loss and its cause identification. The existing techniques perform it comprehensively, but they consume too much time, so there is a need for fast and reliable methods. The article’s purpose is to develop a rapid methodology to assess the risk of information and knowledge loss management. It provides the implementation of eight steps and combines a risk mapping method modified by assessments based on risk factors and incidents as elements from set theory and using formalization via binary estimates. The methodology includes five significant events caused by the company staff, technical problems, software, cybercriminals, viral attacks, and 66 factors influencing company incidents. As a result, a risk map of 9 groups was built for a Ukrainian enterprise. Only two groups with the minimum number of incidents and low losses are represented by all five incidents. The defined overall level of each risk group ranges from 0.14 to 0.26, which indicates a low probability of all happenings in the group. In general, the resulting map shows the existence of specific security problems of the company under investigation. The proposed assessment allows us to interpret the level of risk in the company quickly, identify weaknesses in the information security system, and predict future losses.

Keywords: risk, information loss, knowledge loss management, factor and incident, binary estimate, risks map.

JEL Classification: C13, C60, C80, D81, D83, G32.

Introduction

Today, when information flows and scientific-technological progress are increasing, a company is interested in providing its information security at the highest level. The main reason for this is the information and knowledge loss management. Access to information opens the way to financial flows of the company, its documentation, contracts, employees, tech-

*Corresponding author. E-mail: a.yarovenko@uabs.sumdu.edu.ua
nologies, products, personal data, etc. Today, companies depend entirely on information and knowledge, so accidental or non-accidental loss of any information can have negative consequences for the entrepreneurs. It will relate not only to the cost to recover information but also to the financial losses – results from substantial information loss.

According to the research conducted by the Ponemon Institute commissioned by IBM Security, the average financial loss from hacking and leakage of information in June 2019 for medium-sized businesses in the world was about $ 3.92 million (Ponemon Institute, 2019). This sum has been increased by 1.55% ($ 3.86 million) from 2018, by 8.29% ($ 3.62 million) in 2017, and by 12% ($ 3.50 million) over the last five years (Ponemon Institute, 2018, 2017, 2014). The leader in this area is the United States, the companies of which lost an average of $ 8.19 million in 2019. One can also point out companies in the Middle East ($ 5.97 million), Germany ($ 4.78 million), Canada ($ 4.44 million), and France ($ 4.33 million). Enterprises in India and Brazil received the lowest average losses of $ 1.83 million and $ 1.33 million, respectively Ponemon Institute (2019).

Analysing industry losses, companies in health ($ 6.45 million), financial ($ 5.86 million), energy ($ 5.60 million), industrial ($ 5.20 million) and pharma ($ 5.20 million) suffered the most considerable average losses (Ponemon Institute, 2019). According to Breach Level, more than 18 million records are lost every day, which means 214 records every second. In the first half of 2018, the record number was 3,353,172,708 records (Gemalto, 2018). It means that the situation in the whole world is unfavourable since there is a tendency to increase financial losses as a result of leaks, break-ins, theft, and other types of information loss.

Loss of information and knowledge can lead to a loss of the company’s reputation and customer trust since the information may be publicly available. Data of millions of customers of the company Microsoft have become available on the Internet. The reason was the incorrect setup of the Elasticsearch database. Two hundred fifty million records were publicly available from 05/12/2019 to 31/12/2019 (Riley, 2020). A similar situation was in February 2020 at Decathlon, the information about customers of which was also available online. The reason was the poor security of the Elasticsearch server (Targett, 2020). Also, in February 2020, it was reported that hackers stole data of more than 10.6 million customers of MGM Reports in 2019 during a hacking attack (Cimpanu, 2020).

In 2019, many companies faced the problem of information loss, which concerned not only the personal data of individuals but also banking information – credit and debit cards. Such famous firms as Mastercard, Wyze, Honda, Toyota, Lexus, Yves Rocher, the financial holding company Capital One, several Iranian banks have suffered losses. The companies do not only lose customers in such a way; they often have to pay fines. Thus, for the leakage of data of 9.4 million customers, Cathay Pacific has to pay a fine of about $ 642,000 (Lee, 2020). Unfortunately, there are many cases where companies are obliged to pay fines when they lose their information.

Since the problem of information and knowledge loss management is relevant, this study will solve the issue of assessing the risk of information and knowledge loss for companies, because identifying risks enables the company to predict not only the probable loss but also to identify security issues. In practice, such techniques as COBRA, RA Software Tool, CRAMM, MethodWare, etc. are used for risk assessment. Their advantages include a comprehensive
approach to risk identification, which involves the collection of large data amounts, the calculation of particular methods, the security standards maintenance. The use of these techniques takes considerable time. The companies need an average of 206 days to find information and 73 days to recover (Ponemon Institute, 2019). That is why the study focuses on the development of a rapid methodology that will quickly assess the risk of loss of information and knowledge. Its practical application will reduce time and labour costs.

This paper is structured as follows. Section “Literature review” shows different approaches of international scientists to study the problem of the risk associated with information and knowledge loss management. Part “Risk incidents and risk factors” makes a list of specific incidents and influence factors to assess risk, and explains their concept. Section “Research methodology” presents the developed methodology for the risk assessment, which includes eight stages. Part “Results” demonstrates the results of applying the methodology for one Ukrainian company. Section “Conclusions” contains brief conclusions, limitations, recommendations for the implementation of a measures’ set to reduce the risk of information loss, further possible research.

1. Literature review

The problems associated with the study of the information and knowledge loss risks are quite common in the world. The main reason is the growing trend in the level of informatization and computerization of society. Scientists are exploring various aspects of information loss in different areas: banks (Aryani & Hussainey, 2017; Limba et al., 2019), entrepreneurship (Vasa et al., 2014; Brahmana & Tan, 2018), stock market (Leonov et al., 2012), agriculture (Podaras, 2017), national and global economics (Bilan et al., 2019c; Leonov et al., 2017; Kendidukhov & Tvaronavičienė, 2017). Separately, we can highlight the methodology proposed for systemic risk identification in the banking system of Ukraine (Vasylyeva et al., 2014; Vasa & Angeloska, 2020), which allows us to reduce the risk of information loss in the process of bank consolidation. Also, the risk assessment proposed in an article (Boyko & Roienko, 2014) is interesting for assessment of the insurance companies used in suspicious transactions, which affects a change in the approach to maintaining knowledge in the insurance industry.

One of the main reasons for the information loss is a fraud, which is carried out by employees, company management, external criminals. Morsher et al. (2017) identified that the cause is the unlimited availability of information, especially financial information. To resist this phenomenon, Lyulyov and Shvindina (2017) proposed the Pentagon theory, which can be one of the methods of reducing information leakage from the company. Kostyuchenko et al. (2018), Leonov et al. (2019) also proposed to use monitoring systems to fight against fraud that affects the information and knowledge loss management. Kollár et al. (2017) developed the transformation model as one of the possible tools to increase the level of information security and reduce the risk of information loss. One of another reason for the information loss is the unintentional implementation of errors by employees due to their insufficient experience or lack of required professional knowledge (Gupta, 2017). Therefore, some researchers emphasize the importance of developing innovative approaches to creating and using training systems in companies to solve this problem (Kolomiets & Petrushenko, 2017).
Many studies pay attention to the fact that the problems associated with the collection, processing, storage of information at a high and safe level are increasing with the growth of the level of the scientific and technical process, the informatization of society and enterprises. These aspects are addressed in researches by groups of authors (Bilan et al., 2019a, 2019b; Hrytsenko et al., 2019; Karaoulanis, 2018). Levchenko et al. (2019), Lyoenov et al. (2019) raised the issue of information security in banks to protect anti-money laundering. Along with it, the impact of big data on company informatization and corporate social responsibility is investigated by Hammerström et al. (2019). Creating corporate databases has an important impact on the state of information and knowledge in the company, therefore, we need effective tools to reduce the risk of their losses in the conditions of Big Data functioning. To ensure this aspect, Vasyl' eva et al. (2017) proposed the use of the innovation’s diffusion theory, which allows to reduce the risk of data loss during the process of data integration.

Another approach is to increase the effectiveness of management methods that affect risks in the activities of companies, including information (Grenčíková et al., 2019). Nasr et al. (2019) suggested to create the integrated risk management framework for firms.

Specialists use various techniques and methods to assess risks. Kuzmenko and Bozhenko (2014) considered the optimization models of bank risks for a quantitative assessment of market risks. Berzin et al. (2018) used an approach to assess the risks of business activity, based on creating a quadrangle of factors and determining the centre of mass, which allows us to predict the likelihood of stability in the level of business activity. Dmytrov and Medvid (2017) developed the approach of quantifying indexed information for risk assessment, that fits the needs of the National Risk Assessment of Money Laundering and Terrorist Financing Risks. Lazaroiu et al. (2018) proposed measures to maintain data confidentiality to ensure the General Data Protection Regulation. In researches of risk issues, there are quite popular statistical methods of risk assessment (Hudákova et al., 2018), panel cointegration and causality analysis (Bilan et al., 2020), system dynamics (Jin, 2019), probabilistic methods (Polak, 2019), econometric methods (Bilan et al., 2019d; Mura et al., 2018). Hudáková and Dvorský (2018) proposed assessing the risks in dependence on the rate of implementing the risk management process in the SMEs. Other researchers suggest using the neural network apparatus (Subeh & Yarovenko, 2017); sectoral analysis (Nocoń & Pyka, 2019); bifurcation theory (Vasilyeva et al., 2019).

The issue of information security and the risk of information loss is widely discussed at international conferences. So, the issues of critical (information) infrastructures protection, to solve significant problems of resilience and societal safety, were presented at the conference “The 15th International Conference on Critical Information Infrastructures Security” on 2–3 September 2020 in Bristol, UK (University of Bristol, 2020). Scientists discussed the most crucial directions in data security, cyber-espionage, cyber-terrorism, opportunities of risk mitigation, using cloud computing, machine learning to improve resilience data protection, etc., at the 19th Annual AusCERT Cyber Security Conference on 15–18 September 2020 in Australia (AusCERT, 2020). Specialists in the field of computer and information security debated about security guarantees against arbitrary attacks, biometric backdoors, Deep Learning, Neural Networks, Genetic Testing for detection breaches in information security, development software for malware detection, etc., at the 5th IEEE European Symposium on Security and Privacy on 7–11 September 2020 (IEEE, 2020).
The analysis of the achievements described in the researches shows different areas that need to solve the problem of assessing the risk associated with information and knowledge loss management. There are no universal approaches, but the assessment process must be fast and efficient. Thus, this article will focus on developing a rapid risk assessment methodology.

2. Risk incidents and risk factors

The risk of information and knowledge loss is a possible danger, a threat to the company, which leads to the loss of the most valuable resource – information and knowledge. This risk depends on certain conditions – incidents, which company staff can cause by actions, technical problems, software, illegal actions of cybercriminals, virus attacks. On the other hand, the occurrence of such an incident may be due to various factors. When an employee unknowingly did not save the results of his or her work, the information was lost. As a result, additional time and additional resources were necessary to recover, i.e., the company lost not only information but also financial support, the size of which is usually measured by information loss in the company. Based on the example, a specific employee’s action is a factor that has affected the loss of information, i.e., a generated risk. Since the initiator was a person, this factor refers to an incident caused by human actions.

To determine the level of information and knowledge loss risk, we identify five incidents (causes) and 66 factors of influence that cause the incident in the company.

1. “Human Error Incident” (HE), caused by the misconduct of staff. Thus, users’ errors, their careless use of the computer, and the software can lead to information loss. According to statistics, the human factor causes about 32% of loss (Karabuto, 2020). The following twelve factors influencing the occurrence of this incident were selected: Intentional deletion of data files or sections of text; Unintentional deletion of data files or parts of the text; Intentional non-saving information; Unintentional non-saving information; Overwriting important files; Accidental formatting of your hard drive; Liquid spills; Intentional making a mistake; Unintentional making a mistake; Using of other usernames and passwords; Theft of information by employees; Violation of the rules and procedures for working with data.

2. “Viruses and Malware Incident” (VM), related to virus attacks and antivirus programs. Companies often face a situation where, due to the malicious action of an antivirus program or the appearance of a new virus, the virus enters the system, which leads to the information loss and blocking the work of the entire company. About 7% of information loss is due to a VM incident (Karabuto, 2020). To assess the risk ten factors causing this incident were used: Lack of antivirus updates; Lack of scanning by antivirus; Loss of information due to a virus; Corruption virus; Intentional activating a virus email by a user; Unintentional activating a virus email by a user; Intentional disabling antivirus software; Unintentional disabling antivirus software; False Signal Antivirus; Removing important information by antivirus.

3. “Technical risk” (TR), resulting from a technical, mechanical malfunction. The equipment failure, mechanical damage, wear of the media, improper use cause 44% of the information and knowledge loss management (Karabuto, 2020). This incident includes twelve factors that lead to the malfunctioning of the equipment or its physical destruction, such as Hard disk mechanical failure; Computer damage due to overheating; Computer damage
due to dust accumulation in the computer; Intentional dropping or jostling a computer; Unintentional dropping or jostling a computer; Tornadoes, earthquakes, and other natural disasters; Fire; Planned power outage; Unplanned power outage; Intentional turning off the computer without saving information; Unintentional turning off the computer without saving information; Conflict between devices.

4. “Criminal risk” (CR), caused by the illegal actions of cybercriminals against a company to steal data or knowledge. Today, this incident occurs in 4% of cases (Karabuto, 2020), but it is difficult to predict because it is the cause of actions that are external to the company. As a rule, criminals are interested in information about the company’s financial flows, new technologies and developments. Theft or spoof of this information causes the incomparably significant loss, possible bankruptcy of the company. Often, competitors use this type of crime to harm other companies. Nineteen potential factors causing CR were selected: Logging in with someone else’s login; Computer theft; Computer loss; Copying information to removable media; Sending information to an external email address; Information theft; Information substitution; Unauthorized using of administrator rights; Social engineering; DoS attack; Smurf attack; UDP Storm; UDP Bomb; Sniffing; IP Hijack; Dummy DNS Server; IP-Spoofing; Information Loss due to encryption/decryption; Hacking encryption keys.

5. An incident involving the incorrect work of software in the company, “Software Corruption” (SC). 14% of information and knowledge loss is in the company’s software (Karabuto, 2020). It is the result of improper settings of operating and application programs, non-use of job descriptions, violation of license terms, inadequate testing, errors in the program code. Thirteen following factors influencing the formation of SC were defined, i.e. Unexpected or improper software shutdowns; Lack of software updates; Reformatting during system updates; Errors in Windows registers; The program is not responding; Inaccurate removal or installation of the software; Errors in drivers; Calculation errors; Logical errors; Data I / O Errors; Data manipulation errors; Compatibility errors; Pairing errors.

Each company can expand the own list of factors, but this study highlights the most typical factors. Each risk factor is characterized by the number of cases over some time and the number of monetary losses the company has spent on recovering information and lost profits. This article provided calculations using the given methodology and the information on the number of cases and amounts of loss for the month for the selected factors from one Ukrainian company (there is no name of the company due to its trade secret).

3. Research methodology

It is advisable to use rapid techniques that will quickly identify its level to assess the risk of information and knowledge loss management. Such one method is to build a risk map, which is common in practice because it visually assesses the various dangers posed by economic agents. It is built on a plane, one side of which is the probability of an occurring event, and the other side is the amount that the company may lose when the event occurs. Usually, this area is divided into sectors. The company sets the number of industries, depending on what level of detail of risk it wants to receive. Then the subject sphere is analysed to determine in
which sector event will occur at a given probability and relates to a given level of loss. The disadvantage of this risk map is that in its formation, managers often use subjectivity judgments, which are supported only by their own experience. It mainly concerns the probability of being determined in practice using one’s consideration. This approach is only appropriate for quick decision making.

In this study, we use the approach of constructing a risk map, modifying the construction process by mathematically determining risk estimations based on the factors and incidents as elements of the theory, and using formalization through the binary estimates. They were only used to identify the operational risk of banks, which formed the basis to develop the methodology for the National Bank of Ukraine (Dmitrov et al., 2010). Let several incidents determine the risk of loss of information and knowledge from 1 to k (this paper deals with five incidents – HE, VM, TR, CR, SC (incident designations are represented by the capital letters of their names, according to Chapter 2), i.e., \( k = 5 \)). Several factors \( (n \text{ factors, although 66 factors are calculated in Chapter 2 of this paper}), \) which we consider as the number of cases that occur in a company and cause the information loss, and the amount of financial loss related to the information and knowledge loss influence the formation of every incident. The sets of incidents \( M_{i,j} = \{g_{p,j} \mid p=1+k\} \) (where, \( 1+k \) means that \( i \) goes from 1 to \( k \)), caused by every \( j \) factor for \( p \)-th group of a risk map, can intersect, forming the set \( M_{i,j} \cap M_{j,i} = \{g_{p,j} \mid p=1+n, i=j\} \) (where, \( 1+n \) means that \( i \) goes from 1 to \( n \)). Besides, each of the factors causes the formation of only one incident. On this basis, the methodology for assessing the risk of information loss consists of the following steps.

**Stage 1.** It is necessary to build a table of the number of cases that form five identified incidents (Table 1) and a table of losses related to the implementation of cases in the company’s activities (Table 2).

### Table 1. The cases of factors that form the five incidents (source: compiled by author)

<table>
<thead>
<tr>
<th>№</th>
<th>The name of the agents</th>
<th>Incidents related to the risk of information loss</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Human Error Incident</td>
</tr>
<tr>
<td>1</td>
<td>Agent 1</td>
<td>( a_{11} )</td>
</tr>
<tr>
<td>2</td>
<td>Agent 2</td>
<td>( a_{21} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \ldots )</td>
</tr>
<tr>
<td>j</td>
<td>Agent ( j )</td>
<td>( a_{j1} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \ldots )</td>
</tr>
<tr>
<td>n</td>
<td>Agent ( n )</td>
<td>( a_{n1} )</td>
</tr>
</tbody>
</table>

*Note: \( n \) is the maximum value of the number of factors, \( a_{ji} \) is the number of cases of element \( j \), that affect the formation of the incident \( i \).*

Then, it is necessary to divide the operations into groups that take into account the principle of risk map construction. It means that it is necessary to select operations based on the number of cases and the number of losses. That is why we propose the logic of selection, which will occur according to formula (1):
Table 2. Amounts of losses related to the cases in the business of the company (source: compiled by author)

<table>
<thead>
<tr>
<th>№</th>
<th>The name of the agents</th>
<th>Incidents related to the risk of information loss</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Human Error Incident</td>
</tr>
<tr>
<td>1</td>
<td>Agent 1</td>
<td>( s_{11} )</td>
</tr>
<tr>
<td>2</td>
<td>Agent 2</td>
<td>( s_{21} )</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>j</td>
<td>Agent ( j )</td>
<td>( s_{j1} )</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>n</td>
<td>Agent ( n )</td>
<td>( s_{n1} )</td>
</tr>
</tbody>
</table>

Note: \( s_{ji} \) is the amount of loss by factor \( j \), that affect the formation of the incident \( i \).

\[
\begin{align*}
ap_{ji} = \begin{cases} 
  a_{1ji}, & \text{if } a_{ji} \leq m \land s_{ji} \leq h \\
  a_{2ji}, & \text{if } a_{ji} > m \land s_{ji} \leq h \\
  \ldots & \\
  a_{aji}, & \text{if } a_{ji} > m \land s_{ji} > h
\end{cases}
\end{align*}
\]  

(1)

where \( a_{pji} \) is a selected factor value \( j \) for the incident \( i \), which corresponds to \( p \) \((p = 1 \div t; \quad t = 4 \lor t = 9 \lor t = 25 \lor \ldots)\) (where, \( 1\div t \) means that \( i \) goes from 1 to \( t \)) in the risk map group, \( m \) is the threshold for the number of cases of factors that the company establishes independently, based on case statistics for previous periods, \( h \) is the limit for the amount of loss that the company sets itself based on its policies. It can be a sum that is equal to a percentage of the company’s profits or a rate of its cash flow, which is not significant to the company.

**Stage 2.** Suppose that risk is the probability of an event occurring under unfavourable circumstances. We need to formalize the significance of the factors. It means that it is necessary to calculate the number of cases for \( p \)-group factors using binary properties. If there is a case that causes information loss and, consequently, financial loss, it is a negative phenomenon for the company to deal with, regardless of the number of such cases. Therefore, regardless of the number of cases, the factor will be equal to “1”, which will mean the fact of realization of the case of information and knowledge loss management. If the value is “0”, the company does not have any cases of information loss due to a certain factor. We use formula (2) for formalization:

\[
a_{pji} = \begin{cases} 
  1, & \text{if } a_{ji} > 0 \\
  0, & \text{if } a_{ji} = 0
\end{cases}
\]

(2)

It is necessary to define the sum of binary peculiarities for the \( i \)-th incident for each \( p \) group of risk card according to Eq. (3) to determine the total number of cases for each incident, considering the sampling of data for each group of risk card:

\[
A_{pi} = \sum_{j=1}^{n} a_{pji}.
\]

(3)
The value of $A_{pi}$ shows us the effect of the impact of factors on a risk incident. If $A_{pi} = 0$, there are no cases of impact factor on the $i$-th risk incident. If $A_{pi} = 1$, we have one case of impact factor which may be accidental, but if $A_{pi} > 1$, it can be argued that the company has problems in the security system that have an additional impact on the risk incident. Therefore, two components should be identified to define the level of risk. The first component will reflect the underlying set of risk incident values, which will take into account only that we have a fact or no influence of the factor on the incident, or the existence of the influence of the factor regardless of the number of cases of such impact. The second component will reflect the additional impact on the risk incident, which considers the fact that the number of cases for each risk incident may be greater than “1”, which also takes into account the impact of loss on the risk incident.

The value of the first component is calculated by formula (4):

$$\sum_{i=1}^{k} Z_{pi} | A_{pi} \geq 1.$$  

(4)

In this case, $Z_{pi}$ is the basic set of values of risk incidents:

$$Z_{pi} = \begin{cases} 1, & \text{if } A_{pi} > 0 \\ 0, & \text{if } A_{pi} = 0. \end{cases}$$  

(5)

The value of the second component for risk assessment is in the third stage.

**Stage 3.** The calculated characteristics of $A_{pi}$ reflect the total number of negative cases for each incident. Still, one should take into account that these cases can lead to the loss of different amounts of information and therefore cause various losses to the company. For example, one case involving “DoS attacking” could result in a loss of $1,000,000$, and several incidents involving “Liquid spills” could result in a loss of $10,000$. Thus, it is necessary to consider the impact of factors not only taking into account the number of cases of their implementation but also considering the impact of the loss amount on risk incidents as a set $f(M_{i=1+k} \cup M_{j=1+n}) = \{d_{p,p=1+i}\}$.

Therefore, it is necessary to adjust the binary values of $a_{pji}$ using Eq. (6):

$$a_{pji}^* = a_{pji} \times r_{pi},$$  

(6)

where $a_{pji}^*$ is an adjusted value of $a_{pji}$, $r_{pi}$ are weighting coefficients calculated as $\sum_{j=1}^{n} S_{pji}$ and then ranked from 1 to $i$. We receive the sum of loss for each incident and assign the rank as follows – the highest sum is equal to rank “1”, the smallest sum is equal to rank “$i$”.

The adjustments will allow us to identify the second component to the risk assessment, which reflects the additional impact on the incident. It will be as follows:

$$\frac{1}{n} \sum_{j=1}^{n} a_{pji} | A_{pi} \geq 2.$$
\[ B_{pi} = Z_{pi} \left\lfloor A_{pi} \right\rfloor \geq 1 + \frac{1}{n} \sum_{j=1}^{n} r_{pj} \] \tag{7} \]

where \( B_{pi} \) is the number of factors’ occurrences that affect the incident and that take into account the underlying set of risk incident values and the additional impact on the incident, \([\cdot]\) is the integer part of the number.

**Stage 5.** It is also necessary to consider the situation when the company has all possible instances of impact factors on risk incidents for the risk level identification. It means that the security service has identified at least one fact of such factor impact on each event. For this purpose, a matrix is constructed (Table 3), the elements of which take values equal to “1”. It means that the impact of \( i \)-th factor generates every \( j \)-th \((j = 1 \div k)\) risk incident.

<table>
<thead>
<tr>
<th>№</th>
<th>The name of the agents</th>
<th>Incidents associated with the risk of information loss</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Human Error Incident</td>
</tr>
<tr>
<td>1</td>
<td>Agent 1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Agent 2</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>j</td>
<td>Agent ( j )</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>n</td>
<td>Agent ( n )</td>
<td>1</td>
</tr>
<tr>
<td>( \Sigma )</td>
<td></td>
<td>( n )</td>
</tr>
</tbody>
</table>

Using this approach, we specify the number of all possible factors’ occurrences that affect the incident, and which consider the additional impact on the incident, depending on the loss amount:

\[ B_{pi}^{*} = Z_{pi} + \left\lfloor \frac{1}{n} \sum_{j=1}^{n} r_{pj} \right\rfloor \] \tag{8} \]

where \( B_{pi}^{*} \) are all possible occurrences of the factors affecting the incident and taking into account the additional impact on the incident depending on the loss amount, \( Z_{pi} \) is the basic set of values of risk incidents calculated by formula (5), \( r_{pj} \) is the rank of the \( j \)-th factor affecting the \( i \)-th risk incident which was selected depending on the \( p \)-group of the risk map, \([\cdot]\) is the integer part of the number.

**Stage 6.** At this stage, we calculate the level of risk by formula (9) based on the ratio between the number of cases of factors affecting the incident, considering the basic set of risk incidents values and the additional impact on the incident, and the number of all possible cases of factors affecting the incident, taking into account the additional effects on the incident depending on the extent of losses:
where \( R_{pi} \) is an assessment of the risk level for each incident, the value of which is from “0” to “1”. A value closer to “1” indicates an increased risk of incident \( i \), meaning the information loss will be significant for the company. If the risk value approaches “0”, the incident generates a low level of risk, i.e., the information loss will be negligible or acceptable to the company.

**Stage 7.** At the second last stage, we identify the overall risk level by formula (10) for each of the \( p \)-groups that correspond to the information distribution on the risk map:

\[
R_p = \frac{\sum_{i=1}^{k} B_{pi}}{\sum_{i=1}^{k} B'_{pi}},
\]

where \( R_p \) is the overall level of risk for each of the \( p \)-th group.

**Stage 8.** Finally, there is a company’s risk-loss map that displays the level of risk for each incident, depending on the factors affiliation to one of the map-sectors.

### 4. Results

Using the information on the selected 66 factors relating to the cases that are the leading causes of information loss in the company, the data were divided into nine groups for the future risk map. This number was chosen because nine sectors have basic risk maps for companies. Another reason is that increasing the number of groups requires more data sampling. Given that the primary information we possess is the number of cases and the amount of loss for each factor, we classified the data into nine groups (Figure 1).

Figure 1 shows the classification of factors according to the number of cases and the loss amount. The graphs color changes depending on the increase in the risk level for each of the groups, where a light tone corresponds to a low level, a dark tone shows a high risk. Groups 1, 2, and 4 form a safe risk zone, in which cases of information loss are rare, and the amount of loss is negligible. Groups 3, 5, and 7 form a tolerable risk zone, i.e., such cases occur very often, but the amount of loss is small, or the examples are quite rare. They generate significant information loss for the company. Groups 6, 8, and 9 are at risk because there is a significant loss for the company, and such cases occur quite often. There are groups according to criteria, the values of which were selected on the basis of the analysis results carried out in the process of study preparation (Table 4).

Companies can decide the value of the number of cases and the loss amount which they can set to find the risk of information and knowledge loss.

Figure 1 demonstrates that factors from 1, 4, and 6 groups are the most common; there are single factors in other groups. In conclusion, it is necessary to define the level of information and knowledge loss risk. The steps of the proposed methodology help to calculate the risk level for each incident and each group. The risk map in Figure 2 presents the results.
Table 4. Criteria for factors classification (source: compiled by author)

<table>
<thead>
<tr>
<th>Number of groups</th>
<th>Axis Y</th>
<th>Axis X</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum number of agent cases</td>
<td>Maximum number of agent cases</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>+∞</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>+∞</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>+∞</td>
</tr>
</tbody>
</table>

Note: X-axis is the loss amount in dollars, Y-axis is the number of cases.

Figure 1. Classification of factors forming five risk incidents into nine groups
(source: compiled by author)
At first, we consider “Criminal Risk”, presented in groups 1, 2, 4, and 7 (Figure 2). It means that there are few cases of factors that form this type of risk since they are related to external interference with the company's information system. Still, the risk of their existence is moderate and equals to 0.5 and 0.67. It suggests that the company is likely to have some problems with its cyber defence system, which allows situations where the company loses information and knowledge through external sources, such as scams, hackers, etc. Inclusion of this category to Group 7 also indicates that, with a certain amount of probability, a company may lose significant amounts of money, which can eventually lead to enormous losses. Therefore, it is worth paying attention to those situations that lead to an increase in “Criminal Risk” in the company.

“Software Corruption” occurs in groups of 1, 2, 4, 5, 6, 9 (Figure 2), which indicates the prevalence of this type of risk in cases of information and knowledge loss management. Particularly this type of risk is critical in groups 5 and 9. That is, the cases of information loss with a high degree of probability occur in the company, and they are caused by factors that

---

Figure 2. The risk map (source: compiled by author)
form the incident “Software Corruption”. The company should review the software usage, and setup instructions and protocols since information in this area may be lost due to incorrect operating system settings and custom software that distorts information, reduces computer performance, time loss, etc.

“Human Error” is present in 1, 3, 4, 6, 7, and 8 groups (Figure 2), which indicates a large number of human-initiated information loss cases. Particularly the risk in Group 8 (equal to 1.00) is critical, which means that high levels of employee action are likely to result in relevant information and financial loss. Both “Software Corruption” and “Human Error” can contribute to cybersecurity issues and deadtime. As a result, it will lead not only to information and knowledge loss but also to financial loss. “Viruses and Malware” occurs in groups 1, 4, 6, and 9. The risk of information loss about this incident is moderate. Its value (0.5) in group 9 may be the result of a virus attack, which indicates the atypical impact factors of this incident on the information loss. The fact that this incident occurred in Group 9 signals companies that they should take additional antivirus protection measures. “Technical Risk” occurs in groups 1, 3, 4, 6, 7 (Figure 2). The risk level does not exceed 0.50, and in most cases, it is low despite its prevalence. That is, cases of information loss, which cause technical problems, occur in the company, but do not lead to significant losses.

An overall risk level was determined for each group (Figure 2), which shows the likelihood of loss due to the entire set of incidents. On the whole, one should note that it is insignificant and ranges from 0.14 to 0.26. The probability of risk occurrence for the 6th group is maximum. This value suggests that a situation is possible in the studied company with a possibility of 0.26 when the event of factors will be repeated very often (10–100 times) for each incident and lead to significant losses ($20,000–70,000). In other situations, factors of five incidents may influence the information loss is hardly probable. However, such conditions are also possible for groups 1 and 4 of Figure 2.

Conclusions

Thus, the problem of information and knowledge loss management is quite relevant for different companies, because by losing data, the company loses money. Internal and external factors related to user error, external virus and hacking attacks influence the innovative technology, supernew software, and hardware. Therefore, timely response to the company’s management by predicting the harmful incidents will reduce losses. The proposed methodology will allow timely and rapid assessment and identification of the risk of information and knowledge loss in general and in the context of incidents. This approach will just avoid subjectivism in the methods of companies since actual data on the number of cases, the amounts of loss for each factor form it. There is also real experience in applying such approaches in the operational risk assessment process of banks used by the National Bank of Ukraine.

A positive fact is the visual interpretation of the risk of information and knowledge loss in the form of a risk map, which considers the number of cases and loss, outputs information by groups of factors with the determination of risk for each incident, and the overall level by group. Analysing such a map, you can identify the problematic places in the company, which cause the information and knowledge loss management. The results of the map enable
to predict the consequences for the company with the obtained risk level. For this purpose, it is advisable to determine the scenarios when such risk is present and there are fundamental decisions in the information security system of the company, which scenarios the company will receive if the number of cases and the number of losses increase (decrease). The proposed methodology can be used for companies regardless of the ownership form and activity type. Its main limitation is associated with the criteria for the factors classification for which there are no reasonable measurements. To overcome this disadvantage, the methodology can be refined by developing a statistical estimate of the minimum and maximum boundaries for the number of agent cases and the allowable amount of losses for each risk group.

The next limitation is that the approach proposed in the paper will not replace the set of measures that need to be implemented to reduce the risk of information loss in the company. Companies should conduct regular training sessions to increase computer literacy for users, especially for young and inexperienced employees, to reduce the risk of human factors. It is also necessary to provide employees with information regarding the procedures for dealing with data. It is essential to ensure that users have access rights to job descriptions. This measure reduces the amount of fraud that staff can commit by having an expanded amount of administrator privileges or passwords. Regular monitoring of user actions will help to detect errors in their work. More constructive measures should be taken, such as the use of solid-state drives instead of hard drives, the use of surge protectors, generators, backup batteries, systematic cleaning of computers, and the keeping of devices in specially equipped rooms, use of dust and waterproof enclosures, to reduce the risk of technical incident factors. It is necessary to have a secure lock/unlock procedure, to shut down the software after each use, to perform several software testing procedures, to use systematic backup and archive of information on additional servers or external media, for the reduction of the “Software Corruption” incident impact. Companies should implement anti-theft software on laptops, regularly update antivirus software and scan files, verify access rights and roles of employees in the company information system to reduce the risk of “Viruses and Malware” and “Criminal Risk” incidents.

There are plans to develop the proposed methodology, considering the impact of predicted results of implementation measures to prevent situations of information and knowledge loss on reducing the risk level in the future. It is possible to add identification of factors depending on the degree of their control by company and determination of loss from the respective groups.

Funding

This work is carried out within the taxpayer-funded researches: No. 0118U003574 “Cybersecurity in the banking frauds enforcement: protection of financial service consumers and the financial and economic security growth in Ukraine”, No. 0118U003569 “Modeling and forecasting socio-economic and political road reform map in Ukraine for the transition to the model of modern business”, No. 0120U104798 “Quadrocentric recursive model of Ukrainian unshadow economy to increase its macroeconomic stability” and No. 0120U104810 “Optimization and automation of financial monitoring processes to increase information security of Ukraine”.
Author contribution
All the authors contributed equally to the elaboration of this research.

Disclosure statement
The authors declare no conflict of interest.

References


