Risk of undesired changes to significant information quality criteria

Krzysztof LIDERMAN

Institute of Teleinformatics and Cybersecurity, Faculty of Cybernetics, MUT
ul. gen. Sylwestra Kaliskiego 2, 00-908 Warsaw, Poland
krzysztof.liderman@wat.edu.pl

ABSTRACT: The article presents a method of estimating the risk of an undesirable change in the information quality criterion of secrecy, meaning estimating the risk of a certain class of information security incidents. The qualitative risk estimation method is adopted and the impact of a descriptive grade composition method on the results is discussed. The paper also shows considerations on the possibilities of interpreting the variables used in risk estimation and establishing the range of their actual values. It also describes how the identified range of actual variable values translates into grades used in risk estimation.

KEYWORDS: incident, risk estimation, information security

1. Introduction

The problem presented herein concerns the estimation of the risk of a specific class of information security incidents. In this article, an information security incident means an event or a series of events (resulting in threat execution) that causes or may cause an undesired change in the value of significant information quality criteria[1]. The issue of the incident and its

[1] Only events that have caused an undesired change in the value of significant information quality criteria are considered in a risk analysis (also in this article). This means that attacks that were stopped by IPS and did not cause damage are not taken into consideration. However, according to the rules of the art, such events are also classified as incidents by IT specialists and security departments.

handling is described in paper [1], among others. Since 2018, the perception and handling of incidents has been significantly impacted by the Act on the National Cyber Security System [10], with six regulations assigned to it, of which [6] and [7] are most important from the perspective of the subject matter described herein. The said Act and the accompanying regulations are an implementation under Polish law of the EU NIS Directive (*Network and Information Systems Directive* [3]). The incident occurrence and handling process can be illustrated through the so-called ICOM (**I**nput, **C**ontrol **O**utput, **M**echanism) cube - see Figure 1.

Which of the information quality criteria, mentioned in the title, from among the elementary criteria are significant to the information resources of a given organisation and what their required values are should be:

- determined in a risk analysis,
- approved by the organisation's management,
- entered into the appropriate documents, such as the security policy.

Secrecy, integrity and availability are usually the basic set of criteria from which significant criteria are selected. Further criteria are also those regarding actions on information resources, such as accountability, non-repudiation, etc. The above criteria set out the basic classes of incidents - this article presents the issue of estimating the risk of undesired changes in the value of one of these elementary criteria - secrecy.

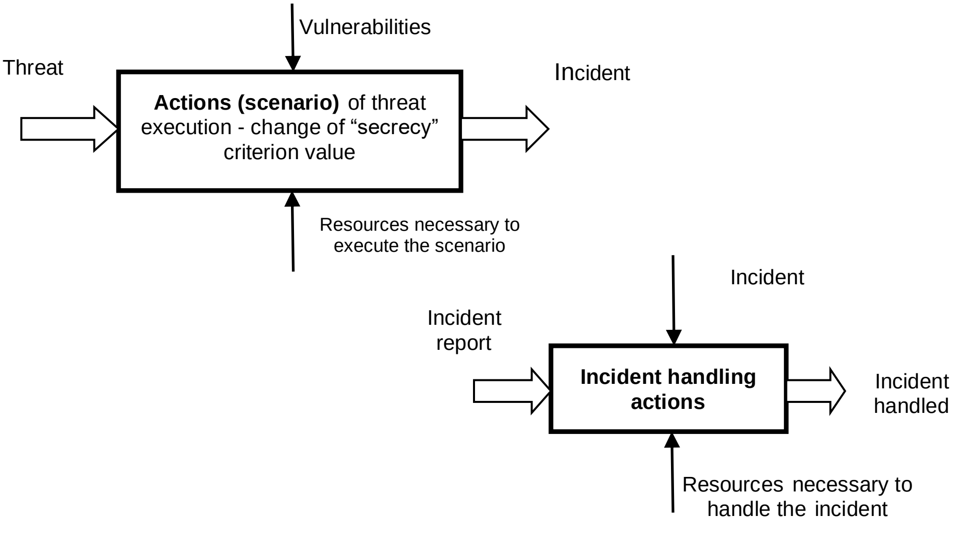
Vulnerabilities
Threat
Incident
Actions (scenario) of threat execution - change of "secrecy" criterion value
Resources necessary to execute the scenario
Incident
Incident report
Incident handling actions
Incident handled
Resources necessary to handle the incident

Fig. 1. Incident and incident handling

Secrecy indicates the required level (strength) of information resource protection against any information being obtained about these resources in an unauthorised manner. This level is agreed by the entities that exchange information. An example of the issue of information resource secrecy in legal frameworks is the Act [9] and the Regulation [8], where the value set of the security level is established as:

$$\{top\_secret, secret, confidential, classified\}$$

2. Descriptive grades in risk estimation and grade composition

The existence of a risk of an information security incident for (any) information resource means the occurrence of a threat affecting the significant information quality criteria (secrecy, integrity, availability, etc.) for that resource; the magnitude of the risk is established by combining the value the assessment of the possibility of events caused by the threat, and the value of the assessment of damage resulting from these events, i.e. the incident effects.

Risk estimation consists in determining the possibility of threat execution (familiarity is also needed for this purpose, as shown in Figure 1, vulnerabilities) and potential losses. Such estimation includes two basic parameters: the possibility or probability of threat execution and the measurable effects of such an event. In risk estimation, the choice is basically limited to two methods [1]:

1. **Quantitative** risk analysis method, where a random event measure is applied - probability expressed by a number in the range of [0, 1].
2. **Qualitative** risk analysis method, which uses descriptive, arbitrarily selected measures expressing the possibility of the event occurring.

This article assumes that risk estimation is carried out using a qualitative method. The detailed assumptions are as follows:

- information resource $z_i \in Z$, where $Z=\{z_i |$ and$\in[1, n]\}$ is a set of resources subject to risk analysis;
- z_i has vulnerabilities $p_j \in P$, where $P=\{p_j | j\in[1, m]\}$ is a set of vulnerabilities of the resources in set Z;
- vulnerability can be used by threat $d_k \in D$, where $D=\{d_k | k\in[1, r]\}$ is a set of identified threats that may affect resources Z.
- risk analysis will be carried out in the resource variant.

The following should be specified:

1) Uniform symbolic grades for properties (attributes) of: threats $d_k \in D$, vulnerability $p_j \in P$, damage, losses and risk. These features describing risk components are as follows:
   - possibility of threat execution, hereinafter referred to as PTE,
   - degree of vulnerability, hereinafter referred to as DV,
   - loss value, hereinafter referred to as LV,
   - magnitude of risk, hereinafter referred to as RISK.
2) The method for assigning grades and calculating generalised grades.

It is proposed to adopt the following system K for assigning descriptive grades from the GRADE set to selected features in the FEATURE set:

$$K = \langle \text{FEATURE, GRADE, } \textit{Procedure} \rangle$$

where:

- FEATURE - a set of features: PTE, DV, LV, RISK.
- GRADE - a set of descriptive grades. This article assumes that it is a three-element set {L, M, H}, where:

  H - HIGH possibility, degree or loss,

  M - MEDIUM possibility, degree or loss,

  L - LOW possibility, degree or loss.
- *Procedure* - a method for assigning descriptive grades from the GRADE set to features of the FEATURE set (e.g. by a decision of experts after a "brainstorming session").

The following method for descriptive grade composition (Algorithm 1) is recommended (as stated in [2]):

ALGORITHM 1

```
Assuming the following:
    1) set A of grades in ascending order, i.e.
       $A=\{q_1, \ldots, q_i, \ldots, q_n\}$ where $i \in N$ is the (position)
       number of element $q_j$ in set A and $q_j < q_{j+1}$;
    2) "element_selected" – any arbitrarily indicated element
       of set A.
    3) number $r \in R$ equals the remainder of the average of
       position numbers of the selected elements.
         IF $r = 0$ THEN
             $q = q_k$ where k is the average of position
             numbers of elements selected from set A
         ELSE ($r \neq 0$):
             IF $r \geq 0.5$ THEN
```

```
        $q = q_{\lceil k \rceil}$
    ELSE ($r < 0.5$)
        $q = q_{\lfloor k \rfloor}$
END OF ALGORITHM 1
```

In the presented algorithm, the *floor, integer part, feature* or *entier* of real number *x*, marked as $\lfloor x \rfloor$, is the largest integer not greater than *x*. However, the *ceiling* or *upper feature* of real number *x* is the smallest integer not less than *x*, marked as $\lceil x \rceil$.

EXAMPLE 1

The symbol ☼ means the descriptive grade composition operation as per Algorithm 1. For example, if the following descriptive grade values are adopted:

N (negligible), L (low), M (medium), H (high), C (catastrophic), i.e.

$$A=\{N, L, M, H, C\}$$

Where $N< L< M< H< C$ i.e. $q_1=N$, $q_2=L$, $q_3=M$, $q_4=H$, $q_5=C$,

then for q=N☼C☼L☼C:

$$(1+5+2+5)/4=3.\mathbf{25}\text{, i.e. } r < 0.5,$$

therefore: q_1 ☼q_5 ☼q_2 ☼q_5 = $q_{\lfloor 3.25 \rfloor}$ =q_3 i.e. N☼C☼L☼C=M

******** (end of example)

The following should be done to estimate the risk of losses caused by an incident, for resource $z_i \in Z$, specific threat $d_k \in D$ and vulnerability $p_j \in P$ (Algorithm 2):

ALGORITHM 2

Assuming that *grade*: FEATURE→GRADE

Using the grades from the set {L, M, H} estimate the following:

1) possibility of threat execution (PTE) "as such"[2] for $d_k \in D$, i.e. give the *grade* (PTE) value;
2) degree of vulnerability DV $p_j \in P$ for resource $z_i \in Z$, which can be used by the threat, i.e. the *grade* value (DV) should be provided;

[2] The threat's "potentiality" is estimated at this stage. For example, when assessing the possibility of hardware theft from the organisation's office, the bars, locks, alarm systems etc. of the building are not taken into account (this affects the vulnerability to theft, as considered at the next stage); only the fact that the target is located in a district where there are a lot of thieves is considered.

3) as per formula (1) the RISK of occurrence of an event such that threat $d_k \in D$ uses vulnerability $p_j \in P$ to cause damage with loss value LV:

$$\mathit{grade}(\mathrm{RISK}) = \mathit{grade}(\mathrm{PTE})☼\mathit{grade}(\mathrm{DV})☼\mathit{grade}(\mathrm{LV}) \qquad (1)$$

END OF ALGORITHM 2

The interpretation of elements of Algorithm 2 for the issue of undesired changes in the significant information quality value, i.e. secrecy, is as follows:

- $d_k \in D$ is a threat to the information resource secrecy;
- $p_j \in P$ is the vulnerability to actions that violate the information resource secrecy;
- PTE is the possibility of executing a threat resulting in an undesired change in the value of the "secrecy" quality criterion;
- LV is the value of losses caused by damage relating to an undesired change in the value of the “secrecy” quality criterion;
- RISK is the risk of an incident of an undesired change in the value of the “secrecy” quality criterion of resource $z_i \in Z$.

The next chapter presents considerations on the possibility of determining (indicating and describing) elements of sets D and P as well as variables PTE, DV and LV and assigning values (descriptive grades) to these variables.

3. Setting variables and the values of their descriptive grades

According to the subject of the article, information resource secrecy is subject to security. Before estimating the risks, the meaning of "secrecy" should be clarified and agreed with all stakeholders[3]. As we can see in EXAMPLE 2, this will also have an impact on the magnitude of the estimated effects of threat execution.

EXAMPLE 2

In the sentence *"I cannot give you this document because it is secret"*, the word "secret" may mean one of the following options, depending on the circumstances:

1. The document is "secret" in the common sense of the word, if its owner does not wish to disseminate it for various reasons.

[3] “implicit” would be an equally good term, but following other publications, the term "secrecy" is also used herein.

2. The document is assigned the "secret" clause under the Act on the Protection of Classified Information (JoL. of 2010, No. 182, item 1228).
3. The document is "secret" within the meaning of the regulations in force at the organisation (these might be only internal regulations) on the dissemination of information, for example, when it contains information constituting trade secrets (within the meaning of Art. 11(4) of the Act of 16/04/1993 *"on combating unfair competition"*; JoL.93.47.211).

**** (end of example)

It is also necessary to establish and reach a consensus among stakeholders as to the nature of "secrecy". Is secrecy something indivisible (something <u>is</u> or <u>is not</u> secret, there are no other options) or whether "secrecy" can be somehow graded. The globally prevailing view is that secrecy can be graded. This was significantly influenced by early works in the field of security, made in the 1970s, in particular those by D. Bell and L. La Padula (for example, see the descriptions of the results of these works in Chapter 2 of [1]]). The grading of "secrecy" has also been entered into the Polish Act on *the Protection of Classified Information* [9], which features four values for the "secrecy" criterion. This view is also adopted in this article.

However, such an approach raises some interpretative complications - how should the effects of a secrecy incident be estimated, assuming that grade values are established and possible levels of secrecy are determined? As already mentioned, "secrecy" specifies the required protection strength of information resource security against any information being obtained about these resources in an unauthorised manner. Levels (labels assigned to items) specify sets of requirements assigned to them. Assuming that the SECRET variable can take values from the set[4]:

{secret, confidential, classified}

the phrase "undesired change in the SECRECY information quality criterion" means a change from SECRECY to ~SECRECY[5]. For example, an information resource as a set on a disk array was secured in as per the requirements for secrecy at the "confidential" level (i.e. SECRECY=confidential), but after the threat execution, the condition is not met, i.e. ~(SECRECY=confidential), e.g. it has been demonstrated that the required safeguards can be bypassed or the safeguards preventing unauthorised access to the information have been broken. Does this mean that the intruder can read only information classified as

[4] This three-element set of the SECRET variable is used later in the article: tables and examples.

[5] The ~ symbol is a sentence-forming functor, "is false".

"confidential", but not "classified", or that they can read both "confidential" and "classified" information, classified below, assuming that the set of information classified in such a way is on this same disk as the confidential set? The answer to this question requires an analysis of possible threat scenarios, considering the actual allocation of the secured information resources and applied safeguards. Another practical consequence of adopting the possible grading of "secrecy" is the need to take into account the classification[6] of am information resource when assigning grades as part of risk estimation - which is best represented in Table 5 (last column) and Table 10.

This article does not consider the type of an information resource by its carrier (electronic, paper, microfilm tape, etc.), although such a preliminary classification of resources might be useful for comprehensive and detailed risk analysis. This would systematise the identification of possible methods of threat execution, e.g. by the tools necessary to execute the threat and possible vulnerabilities.

A uniform method for describing threats is also good for practical reasons. Table 1 presents a proposal for such a method. The provisions of standard ISO/IEC WD 29115 [4], for example, can be used when determining the estimated effects and the possibility of threat execution. Although this standard applies to identity and authentication, it includes some guidelines on what to look for when considering the impact of a violation of information resource secrecy. According to the standard, the potential impact of incorrect authentication applies to:

1. Discomfort, trouble or damage to reputation or position.
2. Financial loss or liability.
3. Damage to the entity, its plans or public interests.
4. Leakage of sensitive information or unauthorised access.
5. Personal security,
6. Violations of civil or criminal law.

The strength of each of the above factors is set on a scale of values: low, moderate, significant, high (i.e. the set of grades differs from the one adopted herein, but it does not matter for the considerations here). The organisation is to determine, based on the estimation of the risk specific to the organisation, what their interpretation is, e.g. what level of financial losses is low, moderate, etc. ISO/IEC WD 29115 does not specify how to carry out risk estimation - it can be done as recommended by PN-ISO/IEC 27005:2010 [5].

[6] A classified resource is a resource for which a security class has been determined by specifying the required level of significant security quality criterion and category. If SECRECY is a significant quality criterion, and the set of values for this variable consists of three labels {secret, confidential, classified}, a sample security class can be as follows: ⟨confidential, ABW_documents ⟩.

The situation is different in the case of entities (organisations) subject to the Act on the *National Cyber Security System* ([10], hereinafter referred to as NCS), which states the following:

Art. 6. The Council of Ministers shall determine, by regulation:

1) a list of key services referred to in Art. 5(2)(1) to assign the key service to a given sector, subsector and the type of entity listed in Annex 1 to the Act, and the importance of the service for maintaining critical social or economic activity;

2) **thresholds for the relevance of the disruptive incident effect** on the provision of key services provided in the list of key services, taking into account:

a) number of users dependent on the key service provided by the entity,

b) dependence of other sectors, referred to in Annex 1 to the Act, on the service provided by the entity,

c) impact the incident could have, due to its scale and duration, on economic and social activities or public security,

d) market share of the key service provider,

e) geographical scope of the area that could be affected by the incident,

f) entity's ability to maintain a sufficient level of the key service, taking into account the availability of alternative ways of providing it,

g) other factors specific to a given sector or subsector, if applicable **-** in order to provide protection against the threat to human life or health, significant property losses and reduction of the quality of the key service provided.

Tab. 1. Threat description sheet template (example)

SHEET No. ……. DESCRIPTION OF THREAT TO SECRECY OF INFORMATION RESOURCE $z_i \in Z$ **Threat ID:** [threat symbol]	
Threat:	[one-sentence descriptive name of threat, e.g. *actions of an intruder - employee of this organisation*]
Threat execution scenario:	[a few sentences of description in words or a block diagram]
Resource owner:	[identification data]
Possible (estimated) effects/damage if the threat is executed:	[a few sentences of description or a list specification]
Possible (estimated) losses if the threat is executed:	[amount in specified currency or description]
Threat potential:	[a few sentences of description]

To the entities specified in the NCS, Art. 6 shall mean the obligation to describe the incident using the method contained therein. It seems that to improve the handling of incidents on both the national and European scale, the proposed incident description method should also be used by organisations (entities) that are not subject to NCS regulations.

It should also be clarified in preliminary arrangements what may be a threat to the information resource secrecy. In this case[7], it is the so-called human factor, manifested as intentional or erroneous actions. This must be determined primarily to estimate the value of the possibility of threat execution (PTE factor) and other risk factors (DV, LV). This can be done using the following table, for example:

Tab. 2. Threats to the information resource secrecy (example description)

Type of action	Who	Motive
INTENTIONAL	intruder, employee	benefits (financial, ideological, psychological, etc.), revenge, curiosity, blackmail, etc.
ERRONEOUS	employee	none

In the case of risk estimation, the possibility of exposing the organisation and its information resources to the threat execution through intentional actions, requires an organisation description in terms of its attractiveness to the intruder. The description should include factors that affect the intruder's motivation. This can be done by adopting a certain set of features (hereinafter referred to as ZC) as ordered four values describing the organisation from this perspective:

$$ZC = \langle BS, OA, PM, AT \rangle$$

where:

- $BS=\{bs_i | i\in[1,m]\}$ is a non-empty set of features describing the organisation's "business" size;
- $OA=\{oa_j | j\in[1,n]\}$ is a non-empty set of features describing the area of the organisation's activities;
- $PM=\{pm_k | k\in[1,l]\}$ is a non-empty set of features describing the impact of the organisation's activities on public mood;
- $AT=\{at_p | p\in[1,r]\}$ is a non-empty set of features describing the industry's "attractiveness" for an intruder.

[7] Unlike threats to the availability of an information resource, where the most common are failures of infrastructure in which the information resource is embedded, and disasters and adverse natural phenomena affecting the infrastructure and the resource itself (such as floods or fires).

These sets of features should be specified by experts or imposed by significant regulations[8] to obtain analysis repeatability. Let’s assume that the above feature sets are specified as follows:

- BS = {large, medium, small, micro enterprise};
- OA = {global, local};
- PM = {significant, moderate, low, none};
- AT is set in the predefined Table 3:

Tab. 3. Type of organisation and "attractiveness" for an intruder (example)

Type of organisation	Attractiveness
Telecoms	**big**
Media companies	
Public administration	
….	
Chain stores	**moderate**
Defense industry	
Medicine	
….	
Other	**low**
….	

As already mentioned, determining the set of features (how many and what elements), their specification and assignment of grade values should be done by a group of experts, e.g. through brainstorming, or should refer to known official regulations. It should be noted that the description provided will not apply if the intruder's motivation is revenge or the intention to cause harm to a particular organisation (e.g. the intruder was paid to do so). In such cases, it should be assumed that the PTE value is high. It should also be taken into account that the "attractiveness" of the target is only one of the elements affecting the intruder's motivation. It is certainly reduced by high penalties for this type of crime, the effectiveness of their prosecution and the belief that there are strong safeguards to be broken (although this factor may very well be a motivation for an intruder who likes a challenge).

EXAMPLE 3

For a large media company operating on a global scale, a sample set ZC_p of feature values can be as follows:

[8] The question remains which entity would issue such regulations. Government Security Centre? Ministry of Digitisation?

$$ZC_p = \{\text{large, global, moderate, big}\}$$

This set of features describing the organisation should be translated into grades adopted for risk estimation, i.e. an interpretation table similar to Table 4 should be made.

Tab. 4. A set of features describing the organisation, their values and corresponding grades (example)

FEATURE	Possible values of FEATURE variable	Grade
BS	large	H
	medium	M
	small, micro-enterprise	L
OA	global	H
	local	L
PM	significant	H
	moderate	M
	low none	L
AT	"telecoms" class, "chain stores" class	H
	-	M
	"other" class	L

Then the PTE value - the possibility of the organisation being exposed by intentional actions of an intruder - is set by the formula:

$$\mathit{grade}[\text{PTE}(ZC_p)] = \mathit{grade}[\text{PTE}(\{\text{large, global, moderate, big}\})] =$$
$$= \text{PTE}(\{H, H, M, H\})$$

where *grade* is a function (usually heuristic) assigning grades from the set of grades (in this article - set {L, M, H}) to the elements of the set of features (here: {large, global, moderate, big}), i.e.

$$\mathit{grade}: \text{PTE}(ZC) \rightarrow \text{PTE}(GRADE)$$

where: $GRADE=\{grade_i | \text{ and} \in [1,n]\}=\{L, M, H\}$.

Adopting the method for grade composition as per Algorithm 1 in this example results in a high possibility of threat execution:

$$\text{alg}(\text{PTE}(\{H, H, M, H\}))=\text{PTE}(H)$$

∗∗∗∗ (end of example)

When estimating the risk of the organisation being exposed to erroneous actions of its employees, historical data should be available regarding the errors that resulted in the incident relating to the information resource secrecy in order to determine the PTE value. Such data, including both the type of error and its frequency, should be collected by the organisation. If there are no such data, the data from organisations of a similar company profile may be used, provided they are available. The third option is to use generalised statistical data published by various organisations involved in information security (e.g. CERT). Naturally, such records include only cases of detected errors and may not be adequate to the actual situation of the organisation for which the risk is estimated.

Errors may result in the disclosure of the information resource content to unauthorised entities, divulging the information about the existence of an undisclosed resource in the system, disclosure of all or some entities authorised to access such a resource, the possibility of leading to said situations by performing an unauthorised operation in the system, etc. If a table of possible effects has been prepared (for example, developed as a result of expert brainstorming), effects should be assigned a frequency of occurrence based on historical data and assigned significant grades based on interpretation tables (e.g. such as Tables 6-9), depending on the classification of the resource affected by the incident. An example of the description is shown in Table 5.

Assuming that disclosure of the content of an information resource classified as *secret* or *confidential*, or divulging of information about the existence of a resource classified as *secret* in the system is not permissible under any circumstances, and specifying the thresholds for the frequency of specific events, interpretation tables of descriptive grades for PTE can be made. This type of assumptions-decisions regarding the thresholds for the frequency of a specific event must be made by the management board of the organisation or its security department. Examples of such descriptions are shown in Tables 6-9. Having considered the contents of the interpretation tables, the last column of Table 5 may be filled in.

Considering the discussion of effects earlier in this chapter (when proposing the threat description), it can be assumed that damage caused by intentional or erroneous actions depend on the following factors:

1. Security class assigned to the information resource (designated *secret*, *confidential* and *classified* in this article).
2. Number of users depending on the resource/service affected by the incident.
3. Dependence of other organisations (or sectors within the meaning of NCS) on the resource/service affected by the incident.

Tab. 5. Effects of errors, empirical data on frequency and PTE grade according to interpretation Tables 6-9 (example)

No.	Effects of a secrecy error	Incident frequency	Grade for PTE (based on Tables 6-9)	
1	Disclosure of the content of an undisclosed resource to unauthorised entities	Once every two years	*secret*	H
			confidential	H
			classified	L
2	Divulging information about the existence of an undisclosed resource in the system	Once every three years	*secret*	H
			confidential	M
			classified	M
3	Disclosure of all or some entities authorised to access an undisclosed resource	Twice a year	*secret*	H
			confidential	M
			classified	L
4	<u>Possibility</u> of situations 1-3 occurring by performing an unauthorised operation in the system	Five times a year	*secret*	H
			confidential	H
			classified	H
5	….	….	….	

Tab. 6. Interpretation of descriptive grades for the possibility of threat execution (PTE) for the error of disclosing the content of an undisclosed information resource to unauthorised entities

GRADE	INTERPRETATION
H	Whenever the error is made for *secret* and *confidential*, more than once a year for *classified*
M	*Classified* once a year
L	*Classified* once every two years

Tab. 7. Interpretation of descriptive grades for the possibility of threat execution (PTE) for the error of divulging information about the existence of an undisclosed resource in the system

GRADE	INTERPRETATION
H	For *secret* whenever the error is made
M	For *confidential* and *classified* whenever the error is made
L	Never

Tab. 8. Interpretation of descriptive grades for the possibility of threat execution (PTE) for the error of disclosure of all or some entities authorised to access an undisclosed resource

GRADE	INTERPRETATION
H	For *secret* regardless of frequency
M	For *confidential* regardless of frequency
L	For *classified* regardless of frequency

Tab. 9. Interpretation of descriptive grades for the possibility of threat execution (PTE) for the error of possibility of situations 1-4 in Tab. 5 occurring by performing an unauthorised operation in the system

GRADE	INTERPRETATION
H	For *secret* regardless of frequency, for *confidential* when once a year or more, for *classified* when more than three times a year
M	For *confidential* when not more than once every two years, for *classified* when two or three times a year
L	For *confidential* when not more than once every three years, for *classified* when not more than once a year

4. Impact the incident could have, due to its scale and duration, on economic and social activities or public security.
5. Market share of the organisation affected by the incident.
6. Geographical scope of the area that could be affected by the incident.
7. Violations of civil or criminal law.
8. Impact on personal security.
9. Impact on the organisation's image.
10. Impact on the organisation's plans or public interests.

Unlike the estimates for the loss of availability of an information resource (see example 3.6 in [1], for example), in case of violation of its secrecy, damage is difficult to translate into losses measured in a particular currency. This leads to complications in determining the content of the interpretation table for losses (LV). Therefore, it is proposed not to include losses to interpret the LV factor in formula (1) as the extent of damage assessed by experts. Assuming that the list of factors affecting the extent of damage is limited to the ten said factors, an interpretation table similar to Table 10 should be developed, using normative and legal guidelines and expert opinions.

The generalised damage value should be estimated using Algorithm 1, following the example shown for PTE in example 3.

Tab. 10. Specification of possible damage (LV) and grade values (example)

No.	Factors affecting the extent of damage	Extent of damage	Grade
1	Security class of the resource affected by the incident	*secret*	H
		confidential	M
		classified	L
2	Number of users affected by the incident	Sector-specific actions of the organisation, according to NCS, for details see Regulation[9]	
3	Dependence of other organisations	as above	
4	Impact on economic and social activities or public security	as above	
5	Market share	as above	
6	Geographical scope of the incident	as above	
7	Violations of civil or criminal law	Assessment by the Legal Department	
8	Impact on personal security	Assessment by the Security Department	
9	Impact on the organisation's image	Assessment by the Management Board	
10	Impact on the organisation's plans or public interests	as above	

There must be a corresponding vulnerability for the threat to be executed. In practice, the set of vulnerabilities is based on the results of the operation of security scanners (detecting vulnerabilities in software and configuration files), penetration tests[10], local inspections, documentation reviews and expert consultations. Then, the degree of vulnerability for all elements of the above set is determined (usually using expert assessments). The results can be presented in tables, as shown in Tables 11 and 12[11].

[9] Regulation of the Council of Ministers of 31/10/2018 *on the thresholds for considering an incident serious* JoL. item 2180.

[10] They should also include testing the staff's resistance to social engineering and resistance to physical security penetration.

[11] Symbols $\wedge$, $\vee$, and ~ are sentence-forming functors "and", "or" and "is false", respectively.

Tab. 11. Interpretation of descriptive grades for vulnerability[12] to intentional actions of an intruder (example)

GRADE	INTERPRETATION
H	~ (safeguards required for "classified", "confidential", "secret" security levels) ∧ ~ (correct safeguard configuration)
M	[~ (safeguards required for "classified", "confidential", "secret" security levels) ∧ (correct safeguard configuration)] ∨ [(safeguards required for “classified”, “confidential”, “secret” security levels) ∧ ~ (correct safeguard configuration)]
L	(safeguards required for "classified", "confidential", "secret" security levels) ∧ (correct safeguard configuration)

Tab. 12. Interpretation of descriptive grades for vulnerability to erroneous actions of an employee (example)

GRADE	INTERPRETATION
H	~ (proper employee training in information resource security) ∧~(supervision over employee operations)
M	[~ (proper employee training in information resource security) ∧ (supervision over employee operations)] ∨ [(proper employee training in information resource security) ∧ ~(supervision over employee operations)]
L	(proper employee training in information resource security) ∧ (supervision over employee operations)

EXAMPLE 4

In the case of an organisation of the *gov.pl* domain, it was decided to estimate the risk of exposure of its information resources to intentional actions of intruders and errors by employees aimed at violating the secrecy of these resources, resulting in information security incidents. The estimates apply to resources classified as *secret*, *confidential* and *classified*. The organisation authorities, with employees of its Security Department and risk analysis specialists engaged under a contract of mandate[13], determined the following, based on historical data and expert estimates:

1. "Employee error" incidents usually resulted in two effects:
   a) disclosure of the content of a confidential information resource to unauthorised entities;

[12] In this and the next table, the number of vulnerabilities is limited to two. In practice, their number depends on the identification results, and the method for their composition to obtain the interpretation of the grades depend on the knowledge and decisions of a risk analyst or a supporting expert.

[13] Specialists proposed a three-element set of grades: high (H), medium (M), low (L) and grade composition using Algorithm 1.

   b) disclosure of some entities authorised to access a resource classified as "secret".

2. "Intentional action" incidents usually resulted in two effects:
   a) disclosure of all entities authorised to access a resource classified as "confidential".
   b) disclosure of the content of a classified information resource to unauthorised entities.
   c) there were no actions motivated by revenge or ordered actions. It was considered that such actions would also be unlikely in the future.
3. Damage caused by intentional actions of intruders and employee errors, incurred in the past and identified as possible in the future, relate to:
   a) impact on social activities or public security,
   b) violations of civil and/or criminal law,
   c) impact on personal security,
   d) impact on the organisation's image,
   e) impact on the organisation's public interests.

When estimating damage, the security class (KB in Table 13) of the resource affected by the incident was also taken into account. Risk estimators used the Interpretation Table 10. For damage in items a-e in the above lists, the organisation's Legal Department, the Security Department and the Management Board set the grade values as in Table 13. The resultant grade (last column in Table 13) was obtained using Algorithm 1 ─ the whole damage estimation can be presented as in Table 13.

4. On the basis of local inspections, analysis of documentation and review of safeguard configuration files, the specialists found that all safeguards required for information resources were used, but some of them were misconfigured. In addition, deficiencies were found in the supervision of employee operations on sensitive resources, although the the security training of employees was highly rated. Assuming that these were all the vulnerabilities found, and that the interpretation of grades for DV is given in Tables 11 and 12, the level of vulnerability for both types of incidents is at level M.
5. It was found that in the past, there were cases (errors) of disclosing information resource content classified as *confidential* to unauthorised persons. It was also found that over the past six years, information about who has access to an information resource classified as *secret* was disclosed twice to unauthorised persons due to error by an employee. Tables 6 and 8 were used to estimate the PTE values for these cases.

6. For intentional actions, the possibility of threat execution PTE was estimated based on the set of features ZC (see example 3). The set of feature values was set at {medium, local, significant "telecoms"}, which translates into the set of grades {M, L, H, H}, so the resultant grade of alg{M, L, H, H}=M.
7. The results of the risk estimation are shown in Table 14.

Tab. 13. LV damage estimation (for example 4)

INCIDENT	**INCIDENT TYPE**	**a**	**b**	**c**	**d**	**e**	**KB**	alg{.}
Intentional actions	disclosure of all entities authorised to access a resource classified as "confidential"	L	L	M	L	L	M	**L**
	disclosure of the content of a classified information resource to unauthorised entities	M	L	L	M	L	L	**L**
Employee error	disclosure of the content of a confidential information resource to unauthorised entities	M	M	L	H	M	M	**M**
	disclosure of some entities authorised to access a resource classified as "secret"	H	H	H	L	M	H	**H**

Tab. 14. Risk estimation (for example 4)

THREAT	**INCIDENT TYPE**	**PTE**	**DV**	**LV**	**RISK**
Intentional actions	disclosure of all entities authorised to access a resource classified as "confidential"	M	M	L	**M** (R'$_{515}$)
	disclosure of the content of a classified information resource to unauthorised entities	M	M	L	**M** (R'$_{515}$)
Employee error	disclosure of the content of a classified information resource to unauthorised entities	H	M	M	**M** (R'$_{205}$)
	disclosure of some entities authorised to access a resource classified as "secret"	H	M	H	**H** (R'$_{204}$)

∗∗∗∗ (end of example)

4. Methods for descriptive grade composition and risk interpretation

The analyst can choose any formally correct method for grade composition - currently there are no norms, standards or regulations that would explicitly impose or otherwise govern this issue. In many applications (for example, see NIST SP 800-53 [11]), it is recommended to apply the formula max{GRADE}, because of its simplicity and the fact that it is sufficient for many practical problems; it selects the maximum grade from the set of composed grades as the resultant. The disadvantage of this method for grade composition is the migration of resultant grades towards the highest grades, contrary to the formula min{GRADE}, where the resultant grades migrate towards the lowest grade - this issue is presented in Table 15.

Assuming that:

$$\otimes = \max\{grade_1, \ldots, grade_j, \ldots, grade_k\} \quad \text{for } j\in[1,k] \quad \text{where: } grade_j \in GRADE \tag{3}$$

$$\varnothing = \min\{grade_1, \ldots, grade_j, \ldots, grade_k\} \quad \text{for } j\in[1,k] \quad \text{where: } grade_j \in GRADE \tag{4}$$

$$\text{☼} = alg\{grade_1, \ldots, grade_j, \ldots, grade_k\} \quad \text{for } j\in[1,k] \quad \text{where: } grade_j \in GRADE \tag{5}$$

where:

- alg{...} means the grade composition as per Algorithm 1;
- ⊗, ∅, ☼ are symbols for descriptive grade composition as per specific formulas or algorithms.

Tab. 15. Resultant values for composition of two descriptive grades using different methods

No.	A	B	$C=A\otimes/\varnothing B$
1	H	H	**H** H
2	H	M	**H** M
3	H	L	**H** **L**
4	M	H	**H** M
5	M	M	M M
6	M	L	M **L**
7	L	H	**H** **L**
8	L	M	M **L**
9	L	L	L **L**

Tab. 16. Estimations of the risk value using composition operations ⊗ (column 5) **and composition operations** ☼ (column 6)

No.	PTE	DV	LV	RISK ⊗	RISK' ☼
1	2	3	4	5	6
1	H	H	H	$R_{101} \rightarrow H$	$R'_{101} \rightarrow H$
			M	$R_{102} \rightarrow H$	$R'_{102} \rightarrow H$
			L	$R_{103} \rightarrow H$	$R'_{103} \rightarrow M$
2	H	M	H	$R_{204} \rightarrow H$	$R'_{204} \rightarrow H$
			M	$R_{205} \rightarrow H$	$R'_{205} \rightarrow M$
			L	$R_{206} \rightarrow H$	$R'_{206} \rightarrow M$
3	H	L	H	$R_{307} \rightarrow H$	$R'_{307} \rightarrow M$
			M	$R_{308} \rightarrow H$	$R'_{308} \rightarrow M$
			L	$R_{309} \rightarrow H$	$R'_{309} \rightarrow M$
4	M	H	H	$R_{410} \rightarrow H$	$R'_{410} \rightarrow H$
			M	$R_{411} \rightarrow H$	$R'_{411} \rightarrow M$
			L	$R_{412} \rightarrow H$	$R'_{412} \rightarrow M$
5	M	M	H	$R_{513} \rightarrow H$	$R'_{513} \rightarrow M$
			M	$R_{514} \rightarrow M$	$R'_{514} \rightarrow M$
			L	$R_{515} \rightarrow M$	$R'_{515} \rightarrow M$
6	M	L	H	$R_{616} \rightarrow H$	$R'_{616} \rightarrow M$
			M	$R_{617} \rightarrow M$	$R'_{617} \rightarrow M$
			L	$R_{618} \rightarrow M$	$R'_{618} \rightarrow L$
7	L	H	H	$R_{719} \rightarrow H$	$R'_{719} \rightarrow M$
			M	$R_{720} \rightarrow H$	$R'_{720} \rightarrow M$
			L	$R_{721} \rightarrow H$	$R'_{721} \rightarrow M$
8	L	M	H	$R_{822} \rightarrow H$	$R'_{822} \rightarrow M$
			M	$R_{823} \rightarrow M$	$R'_{823} \rightarrow M$
			L	$R_{824} \rightarrow M$	$R'_{824} \rightarrow L$
9	L	L	H	$R_{925} \rightarrow H$	$R'_{925} \rightarrow M$
			M	$R_{926} \rightarrow M$	$R'_{926} \rightarrow L$
			L	$R_{927} \rightarrow L$	$R'_{927} \rightarrow L$

Comments on Table 16:

1. The number yy in subscript R_{xyy} is the ordinal number of risk.
2. The number x in subscript R_{xyy} is the row number in Table 16.
3. The apostrophe in symbol R'_{xyy} means that the risk was estimated using operation ☼. The absence of an apostrophe means that the risk was estimated using operation ⊗.

Further considerations on risk estimation refer to the results obtained by applying Algorithm 1 (see column 6 in Table 16). The following conclusions can be drawn from Table 16 regarding the theoretical risk minimisation options:

1. There are <u>three</u> options for minimising the risk value (through the impact on PTE, DV and LV) for the risk set:

$$\{R'_{101}, R'_{102}, R'_{204}, R'_{205}, R'_{410}, R'_{411}, R'_{513}, R'_{514}\}$$

2. Only <u>two</u> options for minimising the risk value exist for risk sets:
   - PTE and LV minimisation: $\{R'_{307}, R'_{308}, R'_{616}, R'_{617}\}$

- PTE and DV minimisation: $\{R'_{103}, R'_{206}, R'_{412}, R'_{515}\}$
- LV and DV minimisation: $\{R'_{719}, R'_{720}, R'_{822}, R'_{823}\}$

3. Only <u>one</u> option for minimising the risk value exist for risk sets:
   - PTE minimisation: $\{R'_{309}, R'_{618}\}$
   - DV minimisation: $\{R'_{721}, R'_{824}\}$
   - LV minimisation: $\{R'_{925}, R'_{926}\}$
4. <u>No</u> options to minimise the risk (the values of all components, i.e. PTE, DV and LV, are at a low level, meaning the risk is minimal) exist for $\{R'_{927}\}$

Assuming an acceptable risk value L, the set of acceptable risk for the grade composition method consists of the following: $\{R'_{618}, R'_{824}, R'_{926}, R'_{927}\}$. However, for elements $\{R'_{618}, R'_{824}, R'_{926}\}$, there are also potential options to reduce the risk value - see the shaded values in item 3. This situation does not occur when estimating the risk value as per formula max{.} - in this case there is only one acceptable risk, in which the value of all components is L (R_{927} in Table 16).

Sometimes, to make decisions on how to minimise risk, it is necessary to know what influences the increase in risk or, looking at the issue differently, which elements composing the risk are at a low level and can be ignored. And so, based on Table 16, it can be concluded that:

1. The set of risk values for which the possibility of threat execution is low, i.e. grade{PTE}=L, is composed of nine elements:

   $$\{R'_{719}, R'_{720}, R'_{721}, R'_{822}, R'_{823}, R'_{824}, R'_{925}, R'_{926}, R'_{927}\}$$

2. The set of risk values for which the level of vulnerability is low, i.e. grade{DV}=L, is composed of nine elements:

   $$\{R'_{307}, R'_{308}, R'_{309}, R'_{616}, R'_{617}, R'_{618}, R'_{925}, R'_{926}, R'_{927}\}$$

3. The set of risk values for which the level of damage is low, i.e. grade{LV}=L, is composed of nine elements:

   $$\{R'_{103}, R'_{206}, R'_{309}, R'_{412}, R'_{515}, R'_{618}, R'_{721}, R'_{824}, R'_{927}\}$$

<u>EXAMPLE 5</u>

Referring the considerations of this chapter to EXAMPLE 4 (see Table 14, last column), the following risk minimisation methods can be recommended for identified incidents:

1. Incident: *Disclosure of all entities authorised to access a resource classified as "confidential"* – **R'_{515}, PTE and DV minimisation**.

   <u>For PTE</u>: it is impossible to reduce the target's attractiveness for the intruder. Only the intruder's <u>motivation</u> can be weakened by setting high penalties and

through effective prosecution of this type of crime, but such undertakings are beyond the scope of ordinary organisations - they require action by government administration and (usually) changes in the law.

For DV: EXAMPLE 4 shows (see item 4 of the example - findings of specialists) that the vulnerability found was *safeguard misconfiguration.* Recommended actions: **improve safeguard configuration**.

2. Incident: *Disclosure of the content of a classified information resource to unauthorised entities* – **R'$_{515}$**, **PTE and DV minimisation**.

   For PTE: comment as in item 1.

   For DV: comment as in item 1.

3. Incident: *Disclosure of the content of a classified information resource to unauthorised entities* – **R'$_{205}$, PTE, DV and LV minimisation.**

   For PTE: improve the control system for operations on sensitive resources, improve training for persons who have access to sensitive information (despite being rated as good!), verify the rules for allowing employees to work with sensitive information.

   For DV: EXAMPLE 4 shows (see item 4 of the example - findings of specialists) that the vulnerability found was a *lack of proper supervision over employee operations on sensitive resources*. Therefore, the recommended action: **improve supervision of employee operations on sensitive resources**.

   For LV: the incident affects *social activities and public security, violation of civil and/or criminal law, organisation's public interests.* Minimising these damages requires coordinated actions by the organisation's Management Board, its lawyers and people responsible for PR.

4. Incident: *Disclosure of some entities authorised to access a resource classified as "secret"* – **R'$_{204}$**, **PTE, DV and LV minimisation**.

   For PTE: comment as in item 3.

   For DV: comment as in item 3.

   For LV: the incident affects *social activities and public security, violation of civil and/or criminal law, personal security and organisation's public interests.* Minimising these damages requires coordinated actions by the organisation's Management Board, its lawyers and people responsible for PR. In addition, the organisation's Security Department should provide personal security to those who have access to information classified as "secret".

******** (end of example)

5. Conclusion

The article presents a method of estimating the risk of an undesirable change in the information quality criterion of secrecy, meaning estimating the risk of a certain class of information security incidents. Knowledge about risk, its value, the value of components and their practical relevance (interpretation) is the basis for both building a security system (risk minimisation) and actions related to the handling of incidents caused by the execution of threats for which the risk was estimated.

In the case of risk estimation (or more broadly - risk analysis) of information security for specific organisations, the estimates usually apply to secrecy, integrity and availability of information resources. This article describes a proposal for such an estimate for the secrecy of information resources. An example of risk estimation for availability is shown in Chapter 3.4 in [1]. Risk estimations regarding the integrity of an information resource will be the subject of a separate article.

Literature

[1] LIDERMAN K.: *Bezpieczeństwo informacyjne*. *Nowe wyzwania*. PWN. 2017.

[2] MALIK A.: *Propozycja doboru i składania ocen opisowych w jakościowym szacowaniu ryzyka systemów informacyjnych*. Praca dyplomowa. Politechnika Warszawska. Podyplomowe Studium Bezpieczeństwa Systemów Informatycznych. 2011.

[3] Dyrektywa Parlamentu Europejskiego i Rady (UE) 2016/1148 z dn. 6.07.2016 r. *w sprawie środków na rzecz wysokiego wspólnego poziomu bezpieczeństwa sieci i systemów informatycznych na terytorium Unii* (NIS).

[4] ISO/IEC WD 29115:2019 Information technology – Security techniques – *Entity authentication assurance Framework.*

[5] PN-ISO/IEC 27005:2010 – Technika informatyczna – Techniki bezpieczeństwa – *Zarządzanie ryzykiem w bezpieczeństwie informacji*.

[6] Rozp. Rady Ministrów z dn. 31 października 2018 r. *w sprawie progów uznania incydentu za poważny*. Dz. U. poz. 2180.

[7] Rozporządzenie Rady Ministrów z dn. 11 września 2018 r. *w sprawie wykazu usług kluczowych oraz progów istotności skutku zakłócającego incydentu dla świadczenia usług kluczowych*. Dz. U. poz. 1806.

[8] Rozporządzenie Prezesa Rady Ministrów z dn. 20 lipca 2011 r. *w sprawie podstawowych wymagań bezpieczeństwa teleinformatycznego.* Dz. U. z 2011 r. nr 159, poz. 948.

[9] Ustawa z dn. 2.08.2010*: o ochronie informacji niejawnej*. Dz. U. 182/10 poz. 1228.

[10] Ustawa z dn. 05.07.2018 r. *o krajowym systemie cyberbezpieczeństwa.* Dz. U. poz. 1560.

[11] SP-800-53 Rev.4: *Recommended Security Controls for Federal Information System.* April 2013.

Ryzyko niepożądanej zmiany istotnych kryteriów jakości informacji

STRESZCZENIE: W artykule przedstawiono sposób szacowania ryzyka niepożądanej zmiany kryterium jakości informacji jakim jest tajność, czyli szacowania ryzyka wystąpienia pewnej klasy incydentów z zakresu bezpieczeństwa informacyjnego. Przyjęto jakościową metodę szacowania ryzyka i przedyskutowano wpływ wyboru metody składania ocen opisowych na uzyskane wyniki. Przedstawiono także rozważania na temat możliwości interpretacji zmiennych użytych w szacowaniu ryzyka oraz ustalenia zakresu ich rzeczywistych wartości. Opisano także, jak zidentyfikowany zakres rzeczywistych wartości tych zmiennych przełożyć na oceny użyte w szacowaniu ryzyka.

SŁOWA KLUCZOWE: incydent, szacowanie ryzyka, bezpieczeństwo informacyjne

Received by the editorial staff on: 12.11.2019