

Network model of risk analysis in the technical structures

Andrzej Walczak¹, Jarosław Napiórkowski^{1*}, Piotr Adamczyk¹, Grzegorz Kiryk¹

¹ Faculty of Cybernetics, Military University of Technology (WAT), Gen. Sylwestra Kaliskiego 2 Str., 00-908 Warsaw, Poland

Abstract. The aim of the article is to present the possibilities offered by network models. This article discusses the network model of risk analysis in technical structures. We are discussed some network centrality measures and their use in risk analysis. We have proposed a mixed, expert-formal model for analyzing the risk of asset risk in technical structures based on own methodology of network building. We shown that mathematically model of structure of the network and its corresponding physical are linked by bijection. This creates a unequivocal model for quantitative risk analysis.

1 Introduction

Generally known and common development and describing the concept of the risk assessment process is PN-EN 31010: 2010 (Risk management - risk assessment techniques. Standard is not intended for certification, regulatory or contractual use but provides guidance on the selection and application of systematic and methodological risk assessment techniques [1]. Risk assessment carried out in accordance with this standard supports other measures of risk management. It shows the use of certain techniques, referring to other international standards, which describe in more detail the concept and application of these techniques.

As examples of such techniques the norm indicates, for example: "brainstorming", Delphi technique, preliminary analysis of threats, Failure Mode and Effects Analysis, FMEA, Fault Tree Analysis, FTA, methods based on Bayesian statistics and Bayesian networks.

The techniques are classified according to their use at different stages:

- risk identification;
- analysis of the consequences at the stage of risk analysis;
- qualitative, quantitative or semi-quantitative estimation of the probability at the stage of risk analysis;
- analysis of control effectiveness at the stage of risk analysis;
- evaluation of risk level at the stage of risk analysis;
- risk assessment;

At each stage of risk assessment, it is possible to use various tools and methods.

This standard, Table A.1 (Applicability of tools used for risk assessment) contains special classification,

* Corresponding author: jaroslaw.napiorkowski@wat.edu.pl

proposal of the techniques that can be applied at each stage of risk assessment and their usefulness.

Methods based on Bayesian statistics and Bayesian networks are classified as non-applicable to risk identification of qualitative, quantitative or semi-quantitative estimation of the probability at the stage of risk analysis or risk assessment at the stage of risk analysis. At the same time, the standard shows them as having application to the analysis of the consequences at the stage of risk analysis and risk assessment.

According to the standard, one of the strong points of this approach is the fact that only the Bayes rule and knowledge of a priori probabilities are required.

What is more, the language is easy to understand and the method provides a mechanism for using subjective beliefs.

At the same time, it has some limitations, such as:

- defining dependency on the Bayesian network may not be feasible due to the complexity and costs related thereto;
- the Bayesian approach requires knowledge of a number of conditional probabilities, which are generally determined on the basis of expert knowledge. The software based on the Bayesian network can provide answers only on the basis of such assumptions.

The advantage of network models is the simplicity of their interpretation. Network models presented in graphical form are understandable not only for those involved in their creation. In case of risk assessment, the undoubted advantage of the network model is the ability to visualize the links between the effects of threats.

Taking an inventory of resources and identifying the processes of vulnerability and risks, the scale effect

should be taken into account [4]. The size of the network representing expertise knowledge can also be a problem.

However, for network models supported by the fact that if we build a network model is the ability to obtain estimates of the probability distributions which should be used on the network. Given the above, it is true to define dependencies in the Bayesian network is troublesome but developing methods for building the network security models may prove to be a truly cost-effective method of analysis of potential security despite the above-mentioned limitations.

2 What we get by building network model of technical infrastructure?

The construction of a network model of technical infrastructure should begin with the development of a methodology for the production of a network model of the examined technical structure, which will result in each expert product the same graphical image of the examined technical structure. As a result, we obtain a graphic form of the model examined structure with each node environment analysis. This will give us the opportunity:

- quantitative characterization of the effect of each node in this structure for asset security risk,
- simulate changes in the network structure of the so that through structural changes control the risk of failure. For example, in the analysis of the security of strictly law [15] regulated information we can simulate the impact of statutory changes or standards on selected elements of the information loss,
- build a simulator of real situations, which by their nature (crisis situations) are not repetitive, that in the simulation can occur under varying conditions of risk.

A particularly interesting element of the network model of technical infrastructure (including critical infrastructure) is the ability to enumerate characteristics that uniquely indicate the impact of the network structure and individual nodes on the propagation of threat in such a structure.

The impact of individual network elements has of both local features, resulting from the immediate environment of the test element and the global features (or rather non-local) resulting from the nature of the network structure in the wider environment of the examined element.

3 Network centrality measures

Centrality measure defines how important is a node in the network. Centrality measures serve to measure the intuitive feeling that in most real composite networks, some vertices or edges are more important / prestigious than others. Centrality can affect both nodes and the entire network. There are a vast number of different

centrality measures that have been proposed over the years.

In the network structure, the numerical values of selected characteristics of its components (in particular network nodes) which describe the influence of these elements in its structure can be determined. Defining centrality in networks that analysis security can help identify key vulnerabilities and vulnerabilities.

In the literature, these items are illustrated on basic network models such as linear structure, star structure, linear cyclic structure.

These will be respectively: centrality degree, betweenness centrality and closeness centrality. These three centrality measures are developed and tested in many different ways creating slightly different parametric characteristics of network elements. We will refer to their definitions and applications to the sample technical infrastructure. Each of these parameters will be called a centrality of node - according to the terminology accepted in the literature.

We will introduce a measure of degree centrality node, closeness and betweenness for work Borgatti, Everett, Freeman (2002). Similar measures have been introduced centrality of the analysis of complex networks. [8].

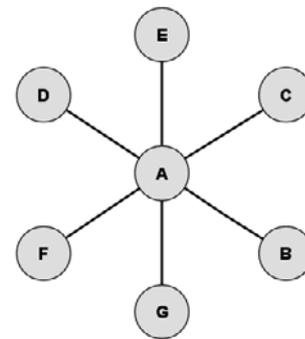


Fig. 1. Star structure. Source: own elaboration.



Fig. 2. Linear structure. Source: own elaboration.

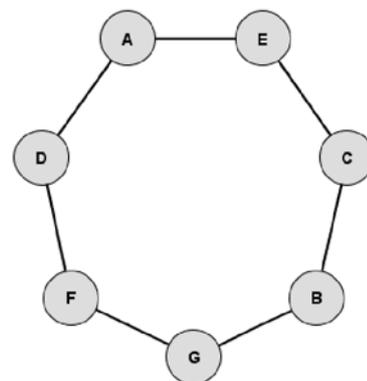


Fig. 3. Ring (linear cyclic) structure. Source: own elaboration.

Let us consider the typical characteristics computed for network nodes.

The degree (or valency) - $\text{deg}(v)$ - of a vertex of a graph is the number of edges incident to the vertex, with loops counted twice. [10, 12].

In directed graphs we can distinguish degree called in-degree and out-degree [5]. We call the number of edges entering or leaving the node (for non-directed graphs we use the vertex degree term).

In star topology network (Fig. 1) node called A has a degree higher than any other node. This gives it more possibilities to create different relationships than any other node in this structure.

3.1 Degree Centrality

Degree is a simple centrality measure that counts how many neighbours a node has. It can be used to illustrate the popularity or influence of nodes. It is useful for determining which nodes are critical (key node) for spreading information or influencing nodes located in the immediate neighbours.

Let us note that the simple enumeration of degree of node does not yet give the essence of this characteristic. Only when we add a descriptive creation of a large (more than the other nodes) number of relations in the analysed network is "node degree centrality" becomes a descriptive characteristic. It can be used for further analysis of the properties of network elements because it points to clearly defined properties in the real case of the network. Because the node degree is easy to calculate with a known network structure, assigning it a parametric measure of degree centrality allows for the transfer of the obtained numerical characteristic to descriptive characteristics of the network used in separate analyzes of its properties.

In linear networks, the degree of node A is not significantly higher than the degrees of other nodes, and the number of relationships created by node A with other nodes does not open a special preference for node A in its power to influence other network elements.

In the real world there are many interpersonal relationships and people think that a large number of relationships with others (key figure) are the most important.

For undirected graph degree C_d is:

$$C_d(v_i) = d_i \quad (1)$$

where

d_i – number of edges at v_i .

If the network is directed, we have two versions of the measure: in-degree is the number of in-coming links, or the number of predecessor nodes; out-degree is the number of out-going links, or the number of successor nodes.

3.2 Prestige and gregariousness

Prestige (popularity) and gregariousness of node are measures that showing which node is more important because it communicates with more nodes. Nodes with more outgoing edges are preferred., in this way we determine the publicity of the node. When we using in-degrees d_i^{in} , degree centrality measures how popular a node is and its value shows prominence or prestige.

$$C_d(v_i) = d_i^{\text{in}} \quad (2)$$

Similar calculations should be performed to calculate gregariousness node d_i^{out} .

$$C_d(v_i) = d_i^{\text{out}} \quad (3)$$

When we join the in-degree (d_i^{in}) and out-degree d_i^{out} as a result, we ignore the direction of the edges. In fact, when we remove the direction of the edge we get $C_d(v_i)=d_i$.

$$C_d(v_i) = d_i^{\text{in}} + d_i^{\text{out}} \quad (4)$$

Simple standardization methods include normalization by the maximum possible degree:

$$C_d^{\text{norm}}(v_i) = \frac{d_i}{n-1} \quad (5)$$

where n is the number of nodes.

We can also normalize by maximizing the degree of node:

$$C_d^{\text{max}}(v_i) = \frac{d_i}{\max_j d_j} \quad (6)$$

Finally, we can normalize by the sum of degrees:

$$C_d^{\text{sum}}(v_i) = \frac{d_i}{\sum_j d_j} \quad (7)$$

We can also normalize by the maximum degree. For the purpose of this work, we will adopt the following designation.

$$C_D = \frac{\text{deg}(v)}{n-1} \quad (8)$$

where

v - node index,

n - the number of nodes,

$\text{deg}(v)$ is number of adjacent edges

3.3 Closeness, reach

Closeness centrality is a measure of centrality in a network, calculated as the sum of the length of the shortest paths between the node and all other nodes in the graph. This is therefore the expected distance between the node and any other node. Thus the more central a node is, the closer it is to all other nodes.

In star type network (Fig. 1) node A is closer to each network node than any other node. Such a structural advantage of node A translates into the strength of its

impact on other nodes in the network. Parameter called „closeness” is , Just like the node degree calculated from the definition given in a later section of this article. Obtaining numerical characteristics describing the closeness of nodes gives us another ability to use this parameter in the structure of the network to analyze its properties.

Just like earlier, in a star-type network, the closeness of node A to other nodes in the network is much higher than in acyclic and cyclic (ring) linear structures. In ring structure of network each node other than A is actually closer to this node than in the acyclic network. A small change in the structure has a significant impact on this parameter because it is higher than in the acyclic linear structure but it is significantly lower than in the star structure.

$$C_C(v) = \frac{n-1}{\sum_{i=1, v \neq k}^{n-1} d(v, k_i)} \quad (9)$$

So we have to count all distances $d(v, k_i)$ from node v to all other nodes and calculate $C_C(v)$,

3.4 Betweenness

Especially important for us is the betweenness centrality.

Betweenness is the ability of a node in the network to create connections with other nodes. A node with a higher value of this parameter than other nodes in the network is often called a hub. In a star-type and a ring topology network, node A can connect much more nodes than a linear network - although the latter is also located between the other nodes. But the connection structure is in a cyclic network other than the star network.

In the descriptive definition, we mean that this measure shows the strength of the network node to the extent that it belongs to the roads in the network graph such that they join two other nodes. If we assign all the roads connecting each pair of nodes and calculate how often the node "x" lies on these roads, we define its betweenness centrality. To summarize, if a node is on all or nearly all the roads connecting two arbitrarily selected nodes in the network its betweenness is high. It is easy to see that this node property strongly depends on the network structure.

$$C_B(v) = \sum_{i \neq j \neq k} \frac{\sigma_{jk}(v)}{\sigma_{jk}} \quad (10)$$

In formula (10) j and k these are indexes of two arbitrary nodes in the network between which there is a path in the network graph. The numerator of an equation is the number of shortest paths from j to k that pass through v .

An example network built for the analyzed example of any technical infrastructure (in our example, the technical equipment of the office with automatic document flow control [17]) is based on constructing a triangle (Fig. 4) asset – vulnerability – threat.



Fig. 4. Network reference model. Source: [7]

The threat is passed on to the resource through its single susceptibility or a subset of individual, independent vulnerabilities. Individual nodes of susceptibility need not be interrelated, i.e. there is no obligatory relationship in the network graph between nodes of vulnerability.

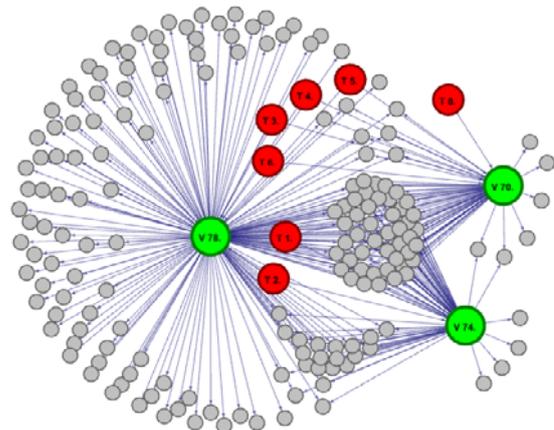


Fig. 5. Network graph [6, 7, 16] for security model described in the office - selected portion of the assets. (black - asset node, green - vulnerability node, red - threat node). Source: own elaboration, Gephi ver. 0.9.1

Table 1. Betweenness Centrality of sample nodes.

Node	Node degree	Betweenness Centrality	Normalized Betweenness Centrality [0,1]
V 70.	65	239.333333	0.007108
V 74	75	70.333333	0.002089
V 78	170	667.333333	0.019819

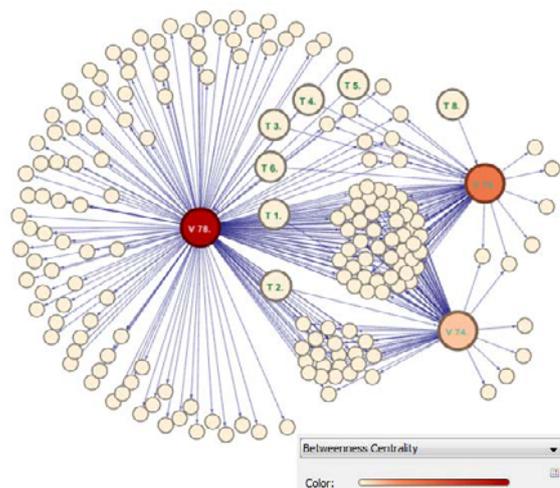


Fig. 6. Coloured illustration of Betweenness Centrality Distribution for example network. Source: own elaboration, Gephi ver. 0.9.1

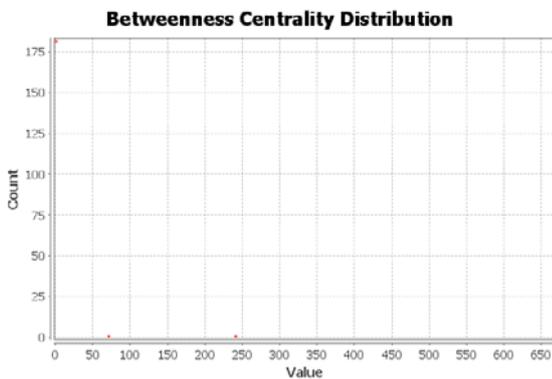


Fig. 7. Betweenness Centrality Distribution. Source: own elaboration, Gephi ver. 0.9.1

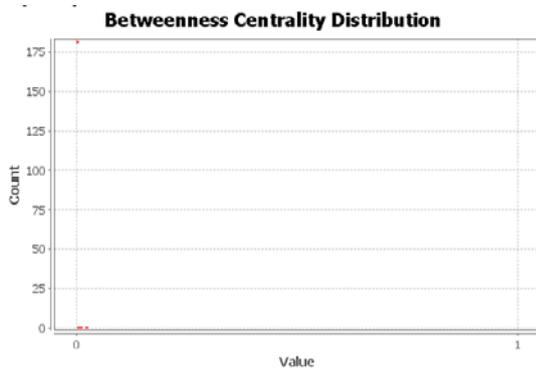


Fig. 8. Normalized Betweenness Centrality Distribution. Source: own elaboration, Gephi ver. 0.9.1

For our model are possible situations when for an asset to have one and only one susceptibility, or one common vulnerability has a group (subset) of assets (Fig. 5). If such single vulnerability is activated by a subset of threats, betweenness for the node illustrating this structure vulnerability will be equal to 1 because this node belongs to all designated routes linking assets and threats nodes. We do not take into account the routes theoretically possible but not created in the network graph.

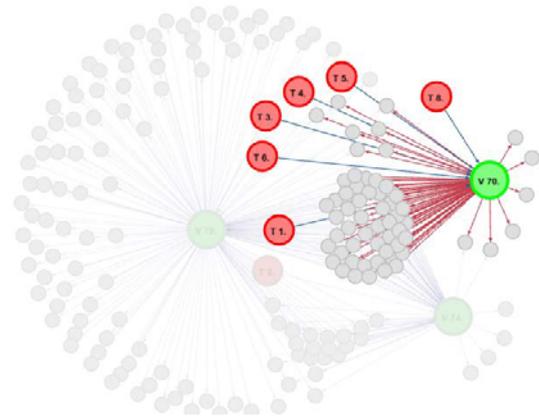


Fig. 9. Assets that may be affected by the occurrence of one of the six above-mentioned threats in the context of vulnerability V 70. - "uncontrolled remote access of third party users." Source: own elaboration, Gephi ver. 0.9.1

Table 2. Closeness and Betweenness Centrality of nodes from Fig. 9. Source: own elaboration, Gephi ver. 0.9.1

Node	Node Degree	Closness Centrality	Betweenness Centrality
T 1.	3	0.504249	0.000000
T 2.	2	0.502906	0.000000
T 3.	2	0.502906	0.000000
T 4.	2	0.502906	0.000000
T 5.	2	0.502906	0.000000
T 6.	1	0.504201	0.000000
T 8.	1	0.504201	0.000000
V 70.	65	1.000000	239.333333

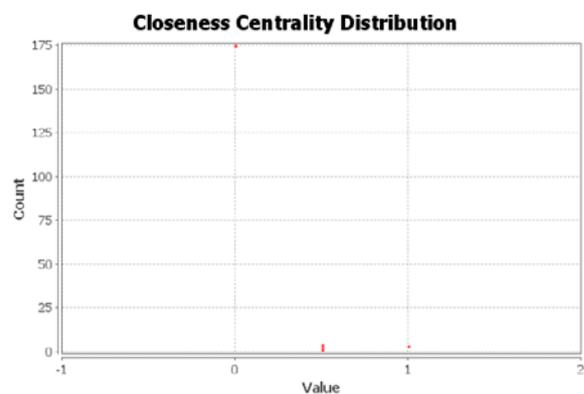


Fig. 10. Closeness Centrality Distribution. Source: own elaboration, Gephi ver. 0.9.1

The second situation is when an asset is characterized by several vulnerabilities shared with other assets (Fig. 6). Betweenness at the time will be less than 1 for each of the nodes describing vulnerability, because the number of all routes in the graph between pairs of nodes

will always be greater than the number of routes that pass through the selected single node of vulnerability.

So the model adopted for determining the risk enforce structure for created network in the convention "asset - vulnerability - threat" and by that on the range of values of the individual centrality measures for the vulnerability nodes in this network. We can easily see that the betweenness for vulnerability nodes is less or equal than 1 in constructed network models.

In our network, if the degree of vulnerability node is high (its connects many assets nodes with a large number of threat nodes), at the same time it will have closeness too high because a large number of nodes can achieve it in one step. Betweenness for each node of vulnerability will be depend only on the network structure. The closer to 1 will be value of betweenness for node that have high degree, it such vulnerability in the risk analysis structure will be a very important node with high impact.

But if the vulnerability node has a low degree, i.e. it connects a small number of assets nodes with threat nodes, then the closeness factor in this case will be not high, but betweenness factor may be closer to 1. This means that even with a small group of threats they are transferred to the asset nodes. So, the value of the betweenness parameter indicates the significance of the vulnerability node in the forwarding of threats to the asset. This means that the vulnerability node can be a hub even with a relatively small value of degree centrality or closeness centrality of that node.

4 Risk of exposure assets to the threat

This risk is defined in a classic form:

$$R = p(T) * con(A) \quad (11)$$

Probability of threat occurrence $p(T)$ is usually estimated based on experts' knowledge of threats. However, whether and to what extent the occurrence of a threat will mean the occurrence of a risk on a particular resource depends on the structure of the network and, in particular, on the location of the vulnerability nodes of the asset at risk. The first approximation will be the dependency of the asset threat function $F(A)$ on the vulnerability structure of the examined asset. With a certain probability of threat $p(T)$ its influence on asset may be described by following equation:

$$F(A) = B * p(T) \quad (12)$$

where B is the calculated value of the betweenness of vulnerability in the examined network for the vulnerability node that transport the threat to the resource. This equation does not raise doubts when there is a single vulnerability node between a group of assets and a threat node. We see here that the network structure model has a decisive influence on obtained values of B . But, if vulnerability nodes have relations in the network graph, then the proposed probability calculation is inadequate since each of the vulnerabilities will have in

that case their influence in relation together with the others.

Likewise, it will also be when there are no relationships between vulnerability nodes, but we arrive at a given asset node via several possible vulnerabilities activated by threat with its probability $p(T)$ (Fig.5). In this case each of the vulnerabilities has own influence, and the threat function $F(T)$ can be summed up after all vulnerabilities. The risk of an asset threat will then be the result of the aggregation of the impact of the threats.

The dependencies (relations) between the vulnerabilities theoretically could be of various types. However, let's note that in the adopted network build-up method, the asset always has the vulnerabilities as a set of own properties. We are also always starting to build a network model with assets and their vulnerabilities in the semantic model: noun-adjective (or in Polish - adjective participle or adjective phrase).

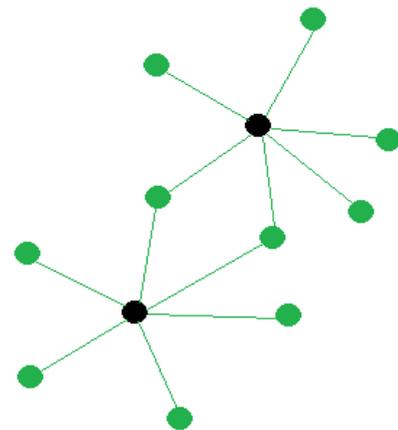


Fig. 11. Asset node (black) is always the center of the structure of the star formed from the nodes of his vulnerability (green). Source: own elaboration, Gephi ver. 0.9.1

This means that any possible relationship between vulnerabilities can result from their coexistence as in Fig. 8 created during the description of the asset. We can imagine a vulnerability such as „brake fluid leak” or „damaged seal in brake pump”, which can be related by the relationship in the network graph.

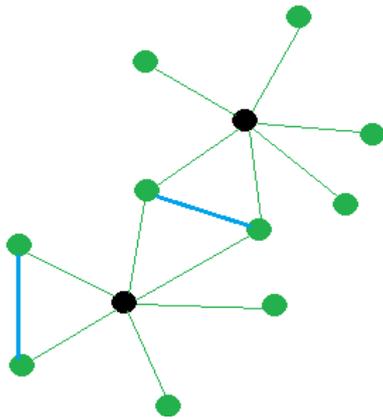


Fig. 12. In the star network some vulnerabilities may coexist in the examined technical structure. We will create at the time an edge illustrating this relation in the network graph (blue). Source: own elaboration, Gephi ver. 0.9.1

However, it should be clarified whether the application of the proposed model asset – vulnerability – threat in a version of single and independent nodes of vulnerability must be really extended to relationship-related vulnerabilities. We can't rule out this situation. However, one should strive for such a definition of vulnerability with adjective or adjective phrase [6, 7, 11, 16], so to eliminate build-up interdependent vulnerabilities. This is the role of a domain expert and therefore the presented model will always be in fact a mixed (formal-expert) model.

This is a leading feature of the presented network model of technical infrastructure. The vulnerabilities of assets in this model is independent of each other. However, it is possible that we can reach the selected asset from the activated threat through several vulnerabilities at the same time. This situation requires a separate probability model in the risk description. In such model of network we can (for a single vulnerability) we can describe it with a formula:

$$R_j = con(A) * F(z) = con(A) * \sum_i B_i * p(z_j) \quad (13)$$

The risk is assumed in this model of vector properties because we assign it to the threat node indexed by j cooperating by index i with sum of vulnerability nodes activated by the threat node indexed by j . The full risk in the infrastructure should be sum respectively by the indexes i and j , but we are practically interested in the risk from the selected threat indexed by j . Separating the risk associated with the occurrence of the selected threat j allows us to describe the effects attributed to the asset. In the model shown we can see that the structure of the relationship in the network graph selects the group of assets affected by the selected threat. This shows how complex is the process to define the effect. It also shows that the risk should be defined by the two-indicators function:

$$R_{jk} = con(A_k) * F(z) = con(A_k) * \sum_i B_i * p(z_j) \quad (14)$$

One indicator refers to the selected resource item and the other to the selected threat. The effects in this formula are a vector, but the components of this vector may be different for each resource k .

Impact models are always a feature of a particular system in which we investigate the risk and there is no universal impact model. In addition, more than one value or feature is required in many cases to determine the consequences and their probability for different periods, locations, groups, or situations.

The consequences can be expressed in material and non-material effects. In each system the effects measure may be different. A universal measure can usually be financial for an organization that is the owner of a system under test. At that point, however, we need to set the risk for all the resources and collect the total financial impact because, as shown above, even a single threat can affect many collections of resources [13, 14].

The effect is also treated as an immediate (operational) or postponed (business) effect, including financial and market consequences. Immediate (operational) effect can be direct or indirect [2]. The first of these is, for example, the financial value of replacing any asset or the cost of acquiring, configuring, and installing new assets [11]. It may also be the cost of suspended operations caused by an incident before the service provided by the asset is restored. Indirect effects include, for example, the cost of lost profits, interrupted operations, potentially misuse of information obtained through a security breach, or breach of regulatory obligations.

5 Summary

We have proposed a mixed, expert-formal model for analyzing the risk of asset risk in technical structures. The expert component of the model, is always present at the stage of creating a triad of "asset - vulnerability - threat". It results from the methodology of building a network model of the examined technical infrastructure [7] and principles of risk analysis. We get a mathematical network model that is typical for complex networks, where we introduce a new concept of asset-threat function. It depends on network structure. We also see that the accepted methodology in the task of risk analysis determines the structure of the network. The numerical value of risk is calculated with the specified probability of occurrence of hazards and effects on the resource determined by the expert description, but we show (what cannot be done without a network model) how it depends on the mathematical model of the network structure. We also showed that even when the threat is certain ($p(T)=1$). The importance of this threat depends strongly on the structure of the network of links

between the assets, its vulnerabilities and threats exploiting its vulnerabilities. We can calculate this importance accurately.

Mathematically the structure of the network and its corresponding physical are linked by bijection. This means that the build model of the technical structure corresponds to one and only one graph assuming that the model will be made using the given methodology [7]. This creates an unequivocal model for quantitative risk analysis.

References

1. *PN-EN 31010:2010 Risk management - Risk assessment techniques*
2. *PN ISO/IEC 27005 Information technology. Security techniques. Information security risk management*
3. M. Kiedrowicz, Uogólniony model danych w rozproszonych rejestrach ewidencyjnych, *Roczniki Kolegium Analiz Ekonomicznych*, vol. **33**, pp. 209-234, (2014).
4. K. Liderman, *Risk Analysis and protection of information in computer systems* (2008)
5. A. Fronczak, P. Fronczak, *Świat sieci złożonych. Od fizyki do Internetu*, (Wydawnictwo Naukowe PWN, 2009)
6. P. Adamczyk, G. Kiryk, J. Napiórkowski, A. Walczak, Sieciowy model systemu bezpieczeństwa., *Zarządzanie informacjami wrażliwymi. Bezpieczeństwo dokumentów, wykorzystanie technologii RFID*, 9-24, (Wojskowa Akademia Techniczna, 2016).
7. P. Adamczyk, G. Kiryk, J. Napiórkowski, A. Walczak, Network model of security system. *MATEC Web of Conferences*, vol. **76**, 02002, DOI: 10.1051/mateconf/20167602002, (2016)
8. C. Bartosiak, R. Kasprzyk, Z. Tarapata., Application of Graphs and Networks Similarity Measures for Analyzing Complex Networks, *Biuletyn Instytutu Systemów Informatycznych*, **7**, 1-7 (2011).
9. S.P. Borgatti, M.G. Everett, L.C. Freeman, *Ucinet 6.0 for Windows: Software for Social Network Analysis. Harvard: Analytic Technologies.*
10. R. Diestel, *Graph Theory* (3rd ed.), (Berlin, New York: Springer-Verlag, 2005).
11. M. Kiedrowicz M. (Editor). *Zarządzanie informacjami wrażliwymi. Wybrane aspekty organizacyjne, prawne i techniczne ochrony informacji niejawnych.* (Warszawa, Wojskowa Akademia Techniczna, 2015).
12. R.J. Wilson, *Wprowadzenie do teorii grafów.* (Warszawa, PWN, 2007)
13. R. Hoffmann, M. Kiedrowicz, J. Stanik, Risk management system as the basic paradigm of the information security management system in an organization, *20th International Conference on Circuits, Systems, Communications and Computers, MATEC Web of Conferences*, Sciences, vol. **76**, 04010, DOI: 10.1051/mateconf/20167604010, (2016).
14. R. Hoffmann, M. Kiedrowicz, J. Stanik, Evaluation of information safety as an element of improving the organization's safety management, *20th International Conference on Circuits, Systems, Communications and Computers, MATEC Web of Conferences*, vol. **76**, 04011 DOI: 10.1051/mateconf/20167604011, (2016).
15. M. Kiedrowicz, Objects identification in the informations models used by information systems, *GIS ODYSSEY 2016*, pp. 129-136, (2016).
16. M. Kiedrowicz (Editor). *Zarządzanie informacjami wrażliwymi: Bezpieczeństwo dokumentów, wykorzystanie technologii RFID*, (Warszawa, Wojskowa Akademia Techniczna, 2016).
17. M. Kiedrowicz, T. Nowicki, R. Waszkowski, Z. Wesolowski, and K. Worwa, Method for assessing software reliability of the document management system using the RFID technology, *20th International Conference on Circuits, Systems, Communications and Computers, MATEC Web of Conferences*, vol. **76**, 04009 DOI: 10.1051/mateconf/20167604009, (2016).