

APPLICATION OF THE BAYESIAN NETWORKS IN CONSTRUCTION ENGINEERING

Agnieszka LEŚNIAK*, Filip JANOWIEC

Department of Civil Engineering, Cracow University of Technology, Cracow, Poland

Abstract

Currently, significant development of methods supporting decision making under uncertainty conditions is observed. One of such methods includes Bayesian networks used in many fields of economy and science. The paper presents the use of the Bayesian network method in civil engineering problems with particular emphasis on construction engineering projects. In addition to the existing examples of the use of the method cited, the authors' method for the risk estimation of additional works is presented.

Keywords: Bayesian networks, construction engineering, risk management

1. INTRODUCTION

Undertakings in the construction engineering projects increasingly require the use of advanced tools and methods. The influence of new technologies is visible in contemporary works, including the implementation of BIM standards in the construction investment process [27], or the use of information from the "Big Data" [6]. The authors of many papers describe the implementation of the results of their research to the construction industry [9, 18, 19, 24], cost of construction works [13, 17, 21, 25], risk in construction [15, 16, 23].

One example combining the practical aspect and high research potential are the Bayesian Networks. This method, which is used for an effective inference in

*Corresponding author: Agnieszka Leśniak, Assoc. Prof., alesniak@L7.pk.edu.pl

conditions of uncertainty, is successfully used in areas where researchers use large amounts of data or attempt to support the processes of making various decisions (such as medicine and legal sciences). There are also works that confirm the applicability of the Bayes network method to problems of technical sciences [26].

In this review paper, the authors will present a brief theoretical outline of the Bayesian network, the process of creating individual network elements (vertices, relations), as well as network structures. Examples of the Bayes network application in engineering of construction projects will be presented, as well as proprietary proposal of their use to analyze the problem of occurrence of additional works in infrastructure construction.

2. THEORETICAL FOUNDATIONS OF BAYESIAN NETWORKS

Bayesian networks are a tool that combines statistical mechanisms with the use of the graph theory. In mathematical terms, Bayesian networks are characterized as directed acyclic graphs, where the vertices contain information about events, while the arches describe the causal relationships between the vertices [14]. An example of the BN is shown in Fig. 1.

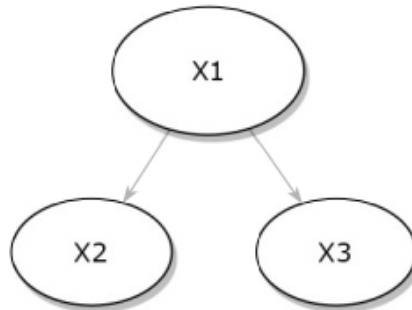


Fig. 1. An example of a Bayesian network. Source: own study based on [14]

The basic relationship used in the construction of the Bayesian network is the Thomas Bayes theorem, binding the conditional probabilities of two events conditioning each other. Bayes' theorem is presented in formula 2.1, while the conditional probability theorem is presented in formula 2.2 [3].

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (2.1)$$

where:

A and B are events and $P(B) > 0$;

$P(A|B)$ – probability of event A to occur if event B occurs;
 $P(B|A)$ – probability of event B to occur if event A occurs.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (2.2)$$

where:

A and B are events and $P(B) > 0$;

$P(A|B)$ – probability of event A to occur if event B occurs;

$P(B \cap A)$ – the probability of the joint part of events A and B.

By defining network nodes, it is necessary to understand the creation of an event describing certain circumstances (scenarios), which will be analyzed using the BN being built. Events are coded with the help of discrete values by assigning an appropriate level of probability of possible scenario variants. Depending on the adopted method of network modeling, events can be presented with two opposite variants or several optional scenarios that are a consequence of the described event. The development of scenarios in nodes results from the level of accuracy of the phenomena studied and the available input data. Examples of events with different ways of describing the consequences are presented in Table 1.

Table 1. Examples of events in Bayesian network nodes with a way of describing the consequences (discrete values). Source: own study based on [14]

Opposite events		Multi-variant events (optional events)	
Event Name	Possible variants	Event Name	Possible variants
Toss a coin	Head	The consequence of the accident at the construction site	Without consequences
	Tail		Temporary disability
The presence of a pathogen in the patient's body	Yes		Permanent disability
	No	Exceeding the duration of construction works	up to 3 months
Design documentation errors	Occur		from 3 to 6 months
	Do not occur		Over 6 months

The decisions taken may lead to further network nodes resulting from the occurrence of the appropriate scenario. Graph arches are used to give appropriate relationships between nodes. According to the assumption of an

acyclic graph, several arches may emerge from one node. Similarly, many arches can lead to one node. The arrangement of nodes together with the relations (consequences) given by means of graph arcs is called the structure of the Bayesian network. Examples of different BN structures are shown in Figure 2.

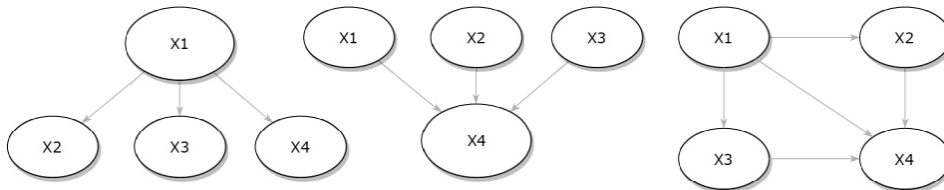


Fig. 2. Examples of various BBN structures. Source: own study based on [14]

Performing the decision-making process (analysis) using the BN mechanism requires that the user models events in the nodes. This is done by entering information on all possible combinations of events to the nodes and assigning probabilities to them. For each scenario, the level of probability must be determined, according to the general principles of the probability theory, the so-called "a priori" probabilities ("initial"). To this end, each network node must have an appropriate conditional probability table. An approach is also taken where events are described using a probability distribution [8], however in this article this type of event description will be omitted. An example of conditional probability table for one node with opposing events (node X2 of the network from Fig. 1) is presented in Table 2.

Table 2. Examples of conditional probability table of the X2 node of the Bayesian network presented in Fig. 1. Source: own study based on [14]

X2	P (X1 X2)	1- P (X1 X2)
True	0,8	0,2
False	0,6	0,4

After modeling, the resulting system of dependencies, nodes and graph arcs along with the given probabilities of events and consequences is a Bayesian network. In order to obtain conditional probability levels of related events, calculations should be made based on the adopted dependence tables. The introduced probabilities are treated as "a posteriori" and allow to start working with the created BN.

One of the main ways of using Bayesian networks is the decision-making process and examination of related elements, including criteria or alternatives of modeled scenarios [14]. The way of making decisions consists in the analysis of research hypotheses, which in the case of BN can be determined in several

ways, depending on the user's preferences and needs. The most useful ones include the following:

- Determining the sequence of events and examining its probability;
- Computing the probability level of certain events and observation of other nodes (for example, one event has a probability of 100%);
- Observation of the probabilities of the entire network without interfering with individual nodes.

Regardless of the adopted method of hypothesis analysis, the results obtained should be related to the previously determined comparative level. The combination of results and criteria allows the Bayesian network user to make or not make the planned decision, thus completing the decision-making process.

Creating a Bayesian network is a process consisting of several successive stages. The most important phases include [14, 22]:

- Defining variables (nodes);
- Establishing connections between nodes;
- Determination of conditional probabilities;
- Introduction of data to the network;
- Computing of conditional probabilities.

The process of creating a network should be repeated as often as possible, taking into account new information, updating events and checking the relationships created. The literature claims that the BN modeling is a closed cycle. The quality of the created decision tool depends on its frequency [20].

The basic source of obtaining information necessary for the construction of the Bayesian network are historical sources (such as an analysis of medical cases, an analysis of post-accident proceedings at the construction site, a statistical analysis of selected events or a review of archival documentation), or expert opinions (namely, expert interviews and form studies). Both methods allow modeling of nodes, dependencies and conditional probabilities of events. Using this approach, it should be remembered that each method has its drawbacks. In the case of BBN, the main questionable element is the lack of certainty as to the repeatability of the processes studied, as well as the large effort and time required to create the network, especially an extended one, regarding cases where the number of network nodes is more than 20 [14].

In order to optimize the process of creating Bayesian networks, research is carried out to minimize potential errors resulting from the nature of knowledge sources while facilitating network modeling. Currently, there are many modifications or approaches that allow more effective use of the method described [20].

Designing a Bayesian network involves frequent updating. It was noticed that this created a lot of problems which led to the proposition of methods of "learning" the network. Among the many proposals, two main branches of BN science are distinguished [20]:

- Parameter learning;
- Structure learning.

The number of the undertaken development directions of the Bayesian network method is significant. The proposed approaches include a number of individual conditions unique to the adopted research topic, which allows more accurate analyzes [4, 5]. It is impossible to list all the proposals created, and the review of the existing modifications of the BN method will be limited only to those described above.

3. EXAMPLES OF THE USE OF BAYESIAN NETWORKS

3.1 Selected domestic and foreign examples

The Bayesian network method has been used and developed since the 1990s [14]. Since then, the scope of application of the method has expanded significantly. Originally used for mathematical considerations, it was applied over time to other fields of science. The paper [26] presents possible approaches to using Bayesian networks to selected civil engineering issues. One example is their use to detect damage to buildings. In addition to the description of how to model the Bayesian network, an example of using it for determining the probabilities of events and a method of calculations was presented. The second example concerned the adaptation of the method to study the dependence of seismic vibration damping on buildings in the Tangshan city area of China.

From domestic studies addressing the issue of Bayesian networks in the research in the field of civil engineering, the study by [1] should be mentioned, describing the modification of the classic Bayesian network to the so-called "Object Oriented Bayesian Network", which is her special case. The author discussed the issue and presented the possibility of using this type of network for risk management problems during construction works. Additionally, the paper presents one way of analyzing the sensitivity of individual network nodes using a tornado type graph.

Bayesian networks can be a tool to support the decision making process. In [2], a decision problem was described consisting in the selection of the most suitable diagnostic test. According to the authors' assumption, the network should answer the question whether, due to the knowledge about the building, additional technical expertise should be performed. Three diagnostic tests were adopted for this problem to see which of them is the best source of knowledge.

For calculations, the "preposteriori" analysis algorithm was used before modeling the network, as well as the decision node in the structure of the Bayesian network. The next part presents the results of the analysis together with the final Bayesian network. For a better understanding, the authors discussed the studied problem using a decision tree.

Another issue to which Bayesian networks can be applied are research problems using probability as the main describing variable. These types of considerations include topics covering risk, risk management and the decision theory. The study [22] presents the application of the Bayesian network to analyze the risk of construction works due to undesirable circumstances during the implementation of tunnel works. For this purpose, the network was modeled together with individual nodes presenting possible events, the consequences of which could lead to stopping construction works. A decision node was added for the created Bayesian network, which aimed to indicate the most appropriate decision at the obtained probability of events. The next part of the work presents the method of determining the probability of events and connecting nodes and giving them conditional probabilities.

According to the adopted methodology [14], it is not recommended to create Bayesian networks with a large number of nodes. Despite this, there are studies, e.g. [27], in which a complex research problem is analyzed. In the example cited, the authors presented a comprehensive Bayesian network covering 27 risks that may occur during the implementation of construction projects. For such identified threats, they conducted a survey among enterprises conducting construction investments in Nigeria. Subsequently, the results obtained were quantified and then presented using the probability of individual events. Risks and probabilities of events have been linked to each other, resulting in a complete network of 34 nodes. The network was used by the authors to analyze the risks occurring during the implementation of construction investments in Nigeria.

3.2 Authors' example - a proposal to use Bayes network to analyze the risk of additional works in infrastructure construction

One of the methods using BN to research in the field of engineering of construction projects is the proprietary method of managing the risk of additional works in railway projects, presented in [16]. The Netica program was used to build the model [11], however the HUGIN program was used in the present work [10]. The risk analysis was carried out in four stages, which were related to the design of the BN in accordance with the phases presented in chapter 2 of this work.

Before starting to model BN, it was necessary to gather knowledge. For this purpose, data obtained from experts from the railway infrastructure construction

industry were used. Based on the information, nodes were coded, conditional probabilities of events were introduced, relations were given to the nodes, and calculations were made in the network. Which led to obtaining a full Bayesian network together with conditional probabilities. The network used for further analysis is shown in Figure 3.

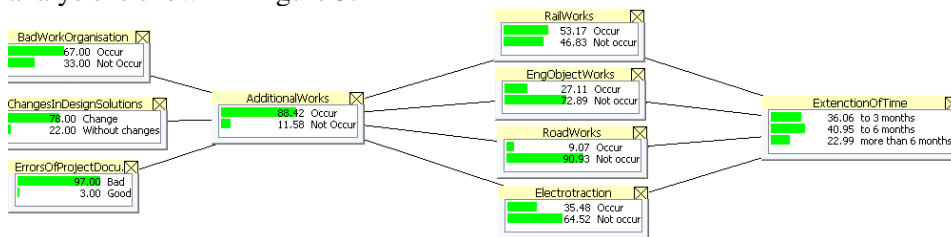


Fig. 3. Bayesian network used to analyze the risk of additional works in railway construction projects [16]

In this example, several factors causing additional work were identified. Several of them have been highlighted:

- Bad Organisation of works
- Changes in design solutions
- Errors in project documentation

One of them – “Errors in project documentation” - was selected for further analysis and subjected to detailed discussion.

The use of Bayesian networks has allowed to use the unique possibilities of this method in the form of analysis of selected event scenarios. A scenario should be understood as a series of consecutive events leading to a specific effect. For further analysis, the existence of a risk factor including the occurrence of project documentation errors and its consequence in the form of an impact on the duration of the investment was subjected to analysis. To examine the scenario, a 100% probability was asked in the node "Documentation Errors" in the event described as "Will occur". It was assumed that the impact of this factor may cause additional works leading to the extension of the investment in three time intervals (variants), with the following probability of occurrence:

- A - Extension of the duration of implementation to 3 months - 36.8%;
- B - Extension of the duration of implementation from 3 to 6 months - 40.2%;
- C - Extension of the duration of implementation above 6 months - 23.0%.

Therefore, in accordance with ISO 31000: 2018 [12], one of the suggested methods was used for risk assessment - the risk matrix. The risk matrix is a matrix in which the obtained level of probability of occurrence of an event is compared with the effect of that event. To read this relationship, numerical values had to be given to the probability ranges as well as the consequences

(effects). The probability ranges have values from 1 to 5, while the individual consequences of events have one of three numbers: 2, 5, 10.

The final element of creating the matrix was to classify the risk level. Based on the ISO 31000 standard, it was assumed that risk is a product of the probability and consequences of a given event. Therefore, for the needs of the matrix, all possible combinations were analyzed, and then four risk levels were identified and assigned numerical ranges. After entering the data, a complete matrix was obtained to assess the risk obtained. Risk ranges with point values for each of them are presented in Table 3.

Table 3. Risk ranges with point values [16]

Risk level	Low	Medium	High	Very high
Point value	from 2 to 6	from 7 to 10	from 11 to 29	from 30 to 50

On the basis of the created matrix, the authors assessed the level of risk resulting from the scenario that causes 100% design errors and its consequences for the duration of the investment for the variants (Table 4):

Table 4. Assessment of risk's level

Variant	Consequences for the duration of the investment	Risk level
A	Extending the duration of implementation to 3 months	low
B	Extending the duration of implementation from 3 to 6 months	high
C	Extending the duration of implementation above 6 months	high

Therefore, for a high level of risk, a risk mitigation method has been proposed, consisting in reducing the likelihood or effects of an adverse risk event to an acceptable level. According to this definition, the possible actions include: reducing the likelihood of faulty documentation, or reducing the duration of implementation time. It was assumed that it is more realistic to reduce the likelihood of faulty documentation as a result of various actions of the Investor. Examples of procedures include: checking created documentation by an additional design office. It was assumed that as a result of the procedure the Investor obtained a reduction in the level of probability of appearance of faulty design documentation to 50%. For such a reduced risk factor, the discussed scenario was reassessed and results were obtained for the options for extending the investment implementation (Table 5):

Table 5. Assessment of risk's level

Variant	Consequences for the duration of the investment	Risk level
A	Extending the duration of implementation to 3 months	low
B	Extending the duration of implementation from 3 to 6 months	medium
C	Extending the duration of implementation above 6 months	medium

Based on the results obtained, there was no change in the risk level for option A, while for options B and C a change from high to medium risk was obtained. In accordance with the adopted risk assessment methodology, this level may be accepted by the Investor.

4. SUMMARY

The use of Bayesian Belief Networks has gained popularity in recent years. The variety of possibilities to implement this method indicates its universality, implying that the method can handle the problems associated with managing a construction project. In the literature [14], one can find the criteria for using Bayesian networks, which include: well-defined variables, a properly set problem with identifiable cause-effect relationships and the benefits of their use, especially for solving repetitive problems, due to the effort to be made in building the model.

The implementation of the presented method supporting decision making is the author's example discussed in the content of the work. The use of Bayesian networks has been shown to be possible for complex civil engineering construction investments. It can be used to estimate the risk of additional works by analyzing possible factors causing additional works, such as "Errors in project documentation". Currently, this method is being developed by the authors.

In addition to the examples of using Bayesian networks mentioned in the paper, it should also be noted that the development of such methods broadens the spectrum of potential applications in the engineering of construction projects. In addition to the outlined directions for the implementation of Bayesian networks, new approaches should be sought or the methods developed so far should be experimented with. Many areas of engineering of construction projects are associated with decision making, the process optimization, scenario analysis or the search for the cheapest and the most effective technological solutions.

Therefore, it seems reasonable to attempt to make a greater use of the method in question in the research into engineering construction projects.

As the authors of the paper have shown, one of the directions of effective use of Bayesian networks may be issues related to risk management in construction projects. In the example presented by the authors, the Bayes network was a useful tool for studying the impact of individual event scenarios. Using a properly built network, it is possible to compare different ways to reduce risk, to check how reducing the risk of factors causing additional works affects the other tops of the network or determine a satisfactory level of effects.

REFERENCES

1. Apollo, M 2017. Prognostic and diagnostic capabilities of OOBN in assessing investment risk of complex construction projects. *Procedia engineering*, **196**, 236-243.
2. Apollo, M, Grzyl, B and Miszewska-Urbańska, E 2017. Application of BN in risk diagnostics arising from the degree of urban regeneration area degradation. In 2017 Baltic Geodetic Congress (BGC Geomatics), 83-88. IEEE.
3. Bayes T, An Essay towards solving a Problem in the Doctrine of Changes. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, M. A., and F. R. S., *Philosophical Transactions of the Royal Society of London*, 53 (1763), 370-418.
4. Chickering, D M 1996. Learning Bayesian networks is NP-complete. In *Learning from data* (pp. 121-130). Springer, New York, NY, 121-130.
5. Ghahramani, Z and Beal, MJ 2001. Propagation algorithms for variational Bayesian learning. In *Advances in neural information processing systems*, 507-513.
6. Guo, S, Luo, H and Yong, L 2015. A big data-based workers behavior observation in China metro construction. *Procedia Engineering*, **123**, 190-197.
7. Grzyl, B, Migda, W and Apollo, M 2019. Building Information Modeling in Small and Middle Sized Buildings–Case Study. In *IOP Conference Series: Materials Science and Engineering*, 603(3), 032077. IOP Publishing.
8. Heckerman, D and Wellman, MP 1995. Bayesian networks. *Communications of the ACM*, 38(3), 27-31.
9. Hoła, B and Nowobilski, T 2019. Analysis of the influence of socio-economic factors on occupational safety in the construction industry. *Sustainability*, 11(16), 4469.
10. <https://www.hugin.com> – HUGIN Researcher version, HUGIN EXPERT A/S, 2020.

11. <https://www.norsys.com/netica.html> - Netica free version, Norsys Software Corp. 2020.
12. ISO 31000:2018. Risk management – Principles and guidelines.
13. Juszczak, M 2019. Early Cost Estimates of Bridge Structures Aided by Artificial Neural Networks. In International Scientific Siberian Transport Forum (10-20). Springer, Cham.
14. Kjaerulff, UB and Madsen, AL 2008. Bayesian networks and influence diagrams. Springer Science+ Business Media, 200, 114.
15. Kowacka, M, Skorupka, D, Duchaczek, A and Zagrodnik, P 2019. Identification of geodetic risk factors occurring at the construction project preparation stage. *Open Engineering*, **9(1)**, 14-17.
16. Leśniak, A and Janowiec, F 2019. Risk Assessment of Additional Works in Railway Construction Investments Using the Bayes Network. *Sustainability*, **11(19)**, 5388.
17. Leśniak, A, Górka, M and Wiczorek, D 2019. Identification of factors shaping tender prices for lightweight. *Scientific Review – Engineering and Environmental Sciences*, **28 (2)**, 171–182.
18. Mrówczyńska, M, Skiba, M, Bazan-Krzywoszańska, A and Sztubecka, M 2020. Household standards and socio-economic aspects as a factor determining energy consumption in the city. *Applied Energy*, **264**, 114680.
19. Mrówczyńska, M and Gibowski, S 2016. Indicating vertical deviation of historical buildings using geodetic methods-case study of brick and wood tower in Nowe Miasteczko. *Civil and Environmental Engineering Reports*, **22(3)**, 127-136.
20. Neapolitan, RE 2004. Learning bayesian networks (Vol. 38). Upper Saddle River, NJ: Pearson Prentice Hall.
21. Nowogońska, B and Korentz, J 2020. Value of Technical Wear and Costs of Restoring Performance Characteristics to Residential Buildings. *Buildings*, **10(1)**, 9.
22. Odimabo, OO, Oduoza, C and Suresh, S 2017. Methodology for project risk assessment of building construction projects using Bayesian belief networks. *International Journal of Construction Engineering and Management*, **6(6)**, 221-234.
23. Siemaszko, A, Grzyl, B and Kristowski, A 2018. Evidence-Based Risk Management for Civil Engineering Projects Using Bayesian Belief Networks (BBN). In 2018 Baltic Geodetic Congress (BGC Geomatics) 191-195, IEEE.
24. Švajlenka, J, Kozlovská, M and Spišáková, M 2017. The benefits of modern method of construction based on wood in the context of sustainability. *International Journal of Environmental Science and Technology*, **14(8)**, 1591-1602.

25. Wieczorek, D, Plebankiewicz, E and Zima, K 2019. Model estimation of the whole life cost of a building with respect to risk factors. *Technological and Economic Development of Economy*, **25(1)**, 20-38.
26. Yuen, KV 2010. Recent developments of Bayesian model class selection and applications in civil engineering. *Structural Safety*, **32(5)**, 338-346.
27. Zima, K, Plebankiewicz, E and Wieczorek, D 2020. A SWOT Analysis of the Use of BIM Technology in the Polish Construction Industry. *Buildings*, **10(1)**, 16.

Editor received the manuscript: 08.06.2020