International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 2, April 2025, pp. 2181~2191
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i2.pp2181-2191

Markov processes in Bayesian network computation

Assem Shayakhmetova[1], Nurbolat Tasbolatuly[2,3], Ardak Akhmetova[1], Assel Abdildayeva[1], Marat Shurenov[4], Anar Sultangaziyeva[1]
[1]Department of Artificial Intelligence and Big Data Al-Farabi Kazakh National University, Almaty, Republic of Kazakhstan
[2]Higher School of Information Technology and Engineering, Astana International University, Astana, Republic of Kazakhstan
[3]Department of Computer Engineering, Astana IT University, Astana, Republic of Kazakhstan
[4]Department “IT and Service”, “Q” University, Almaty, Republic of Kazakhstan

Article Info

Article history:

Received Aug 22, 2024
Revised Oct 26, 2024
Accepted Nov 20, 2024

Keywords:

Bayesian network
Markov network
Markov random field
Markov properties
Graphical model
Evidence
Belief propagation

ABSTRACT

The article examines the influence of Markov processes on computations in Bayesian networks (BN), an important area of research within probabilistic graphical models. The concept of Bayesian Markov networks (BMN) is introduced, an extension of traditional Bayesian networks with the addition of a Markov constraint, according to which the probability in a node can only depend on the state of neighboring nodes. This constraint makes the model more realistic for many practical tasks, as most graphical models that reflect real-world processes possess the Markov property. The article also discusses that Bayesian networks, in the absence of evidence, actually exhibit the Markov property. However, when evidence (additional information) is introduced into the model, challenges arise that require more complex computational methods. In response, the article proposes algorithms adapted for working with Bayesian Markov networks in the presence of evidence. These algorithms are aimed at optimizing computations and reducing computational complexity. Additionally, a comparative analysis of calculations in Bayesian networks without Markov constraints and with them is conducted, highlighting the advantages and disadvantages of each approach. Special attention is paid to the practical applications of the proposed methods and their effectiveness in various scenarios.

Corresponding Author:

Nurbolat Tasbolatuly
Higher School of Information Technology and Engineering, Astana International University
010000 Astana, Republic of Kazakhstan
Email: tasbolatuly@gmail.com

1. INTRODUCTION

In the theory of Bayesian networks [1]–[3], the ideas of Markov [4]–[6] play a key role, especially in developing efficient algorithms for data processing under uncertainty. Markov processes are of particular interest to researchers for several reasons. First, they are intuitively understandable and natural for many applied tasks, making them attractive for various fields. Second, incorporating Markov ideas into Bayesian networks (BNs) [7]–[9] can significantly simplify computational algorithms, thereby speeding up the probability calculation process and expanding the possibilities of working with larger networks. Finally, Markov processes are well-studied within the framework of Markov chain theory [10]–[12], which allows for adapting existing developments to work with Bayesian networks. Intelligent decision support systems [13], [14], especially in various types of uncertainties, are often described using Bayesian networks. One of the main challenges in working with BNs is the need to perform calculations considering the presence of evidence-information that can alter the probability distribution in the network. Standard methods, such as algorithms based on junction trees [15], [16], require complex computational procedures. The most well-known method for exact inference in BNs is the probability propagation method [17]–[19] in cluster trees (PPTC), proposed by Lauritzen [20] and refined by Jensen. This method transforms the belief network into a more computation-friendly structure; however, its application can be time-consuming and slow.

Shayakhmetova *et al.* [21] discusses critical issues related to using Bayesian networks for solving practical problems in which graphical models contain directed cycles. The strict requirement for the acyclicity of the directed graph representing a Bayesian network does not allow for the effective solution of most tasks that involve directed cycles. Modern Bayesian network theory prohibits the use of directed cycles. The requirement for graph acyclicity can significantly simplify the general theory of Bayesian networks and simplify the development of algorithms and their implementation in code for computations in Bayesian networks. Akhmetova *et al.* [22] examines the urgent problem of maximizing human survival during a fire and their immediate evacuation. Various methods are being proposed worldwide, including intelligent devices and wireless systems, concerning this issue. The main goal is to efficiently utilize time, regularly inform people, and guide them out of danger zones, thereby reducing the number of casualties. The article also describes the concept of an emergency and its causes. Modern methods of detecting people and navigation during a fire, as well as the use of wireless equipment and intelligent algorithms, are discussed in detail. The system's architecture, goals, objectives, and solutions are presented. Examples are provided to explain the importance and usefulness of this research for people. The article includes information on the use of sensors, with tables illustrating their application, interconnections, and wireless communication with the central server. As a result, a system was developed that operates based on electronic indicators showing the direction to a safe place in the event of an emergency, which receives signals from the central server. Alimhan *et al.* [23] addresses the problem of global practical output tracking for a class of high-order nonlinear systems with time delays (using state feedback). Under moderate growth conditions of system nonlinearities, including time delays, a state feedback regulator design with an adjustable scaling factor is proposed. Using the Lyapunov-Krasovskii functional, this scaling factor is adjusted to dominate the time-delay nonlinearities, bounded by growth conditions, and to make the tracking error arbitrarily small while all states of the closed-loop system remain bounded. A simulation example is provided to illustrate the effectiveness of the proposed tracking regulator.

Alimhan *et al.* [24] discusses the problem of global practical output tracking using state feedback for a class of uncertain nonlinear systems with time delays. Under moderate conditions on the system nonlinearities, including time delays, a homogeneous state feedback regulator design with an adjustable scaling factor is proposed. Through the homogeneous Lyapunov-Krasovskii functional method, this scaling factor is adjusted to dominate the time-delay nonlinearities, bounded by homogeneous growth conditions, and to make the tracking error arbitrarily small while all states of the closed-loop system remain bounded. Alimhan *et al.* [25] explores the problem of global practical output tracking using state feedback for a class of uncertain high-order nonlinear systems with time delays. A homogeneous state feedback regulator with an adjustable scaling factor is then developed under moderate conditions on the system nonlinearities, including time delays. Through the homogeneous Lyapunov-Krasovskii functional method, this scaling factor is adjusted to dominate the time-delay nonlinearities, bounded by homogeneous growth conditions, and to make the tracking error arbitrarily small while all states of the closed-loop system remain bounded.

To address this problem, Markov processes need more efficient use, which allows for simpler and faster algorithms. The main idea is that the probability in a node depends only on the state of neighboring nodes (parents and children), which is a less stringent requirement than the concept of the Markov blanket. This simplification can significantly speed up computations, avoiding the complexities of building junction trees. This work proposes a new approach that divides the vertices of a Bayesian network into levels to simplify calculations. The vertices that have received evidence form the zero level, while the remaining vertices are divided into subsequent levels based on their connections to the previous levels. This approach allows for more efficient probability calculations in the network nodes using already known values from previous levels. The examples provided illustrate the proposed methods using educational data in the HUGIN EXPERT environment, which visually demonstrates the operation of the algorithms. The examples considered in the article are educational and do not require the search for real meanings, but they demonstrate the main principles and advantages of the proposed approach.

2. METHOD

Considering an arbitrary Bayesian network (BN) [26]–[28], one might question the dependency of a selected node on other nodes within the network. The primary computational complexity lies in the fact that the selected node generally depends on many other nodes, which becomes particularly challenging in the presence of evidence. Various approaches to solving this problem exist [29]–[31], which can be categorized into several groups. First, one method involves breaking down computations in a BN into blocks using the concept of d-separation. This approach allows the network to be divided into independent blocks, where computations can be performed separately, significantly reducing overall computational complexity. Second, the BN can be transformed into another, more simplified network where computations become easier and less resource-intensive. After performing the necessary calculations in this simplified network, the results can be transferred back to the main Bayesian network. The third approach is the elimination of specific nodes from the BN using the theory of potentials. This theory is well-developed both in general terms and specifically for Bayesian networks. Although this approach can formalize many computations, it often loses the intuitive clarity inherent in the framework of BNs. The fourth method suggests simplifying computations in a BN by assuming that calculations in the selected node depend only on its neighboring nodes and that dependencies on other nodes are mediated only through these neighboring nodes. While this may lead to a reduction in calculation accuracy within the context of Bayesian ideas, this method essentially represents a shift in the ideology of Bayesian networks. This work examines a case where computations in the current node depend only on its neighboring nodes, possibly not all of them. The main computational challenge arises when evidence is present in the network. In the absence of proof, the Markov property is automatically fulfilled, and calculations in the node depend only on a subset of neighboring nodes, making them relatively simple. The article explores the possibility of extending this ideology to computations in Bayesian networks with evidence in some nodes, aiming to simplify calculations and improve their efficiency. Let a graph $\mathrm{G} = < \mathrm{V}, \mathrm{E} >$ be given, where V is the set of vertices of the graph, and E is the set of edges. A directed acyclic graph G is called a Bayesian network if each vertex $v \in V$ is associated with a random variable $\mathrm{X_v}$, and each edge $= (\mathrm{u}, \mathrm{v}) \in \mathrm{E}$ represents a probabilistic dependence of the random variable $\mathrm{X_v}$on the random variable $\mathrm{X_u}$u. The random variable Xj contains several independent states $\{x_1, x_2, x_3, \dots, x_n\}$. A probability characterizes each state. The number of states for each random variable may vary. It is assumed that (1) (2).

$$P(x_1) + P(x_2) + P(x_3) + \cdots + P(x_n) = 1 \tag{1}$$

$$P(x_1 + x_2) = P(x_1) + P(x_2) \tag{2}$$

The following basic rules are used when working with Bayesian networks:
- Multiplication: $P(X, Y \mid Z) = P(X \mid Y, Z) \ * \ P(Y \mid Z) = P(Y \mid X, Z) \ * \ P(X \mid Z)$
- Summation: $(\mathrm{X} \mid \mathrm{Z}) = \sum_y P(X, Y = y \mid Z) = \sum_y P(X \mid Y = y, Z) * P(Y = y \mid Z)$
- Bayes' theorem: $P(\mathrm{X} \mid \mathrm{Y}, \mathrm{Z}) = \frac{\mathrm{P(Y \mid X,Z) * P(X \mid Z)}}{\mathrm{P(Y \mid Z)}} = \frac{P(\mathrm{Y \mid X,Z}) * \mathrm{P(X \mid Z)}}{\sum_x P(\mathrm{Y \mid X{=}x,Z}) * \mathrm{P(X{=}x \mid Z)}}$

Independence of random variables: $P\ (X, Y) = P(X) \ * \ P(Y)$. Initially, conditional probabilities and probabilities for some nodes are set in the Bayesian network. Based on this data, any combination of events in the network can be determined. A vertex u is a parent of vertex v if edge $(\mathrm{u}, \mathrm{v}) \ \in \mathrm{E}$. The set of all parent vertices of v is denoted by $\mathrm{parents}\ (\mathrm{X_v})$.To calculate the probability, there is a chain rule for Bayesian networks (3):

$$P(X_1, X_2, X_3, \dots, X_n) = \prod_{i=1}^{n} P(X_i \mid parents(X_v)) \tag{3}$$

The chain rule allows for decomposing a joint distribution into a product of conditional distributions. It is evident that during the calculation, the Markov principle is observed: a variable depends only on its parent variables, regardless of how the parent variables obtained their values. It is also clear that the number of parents for a vertex should not be large due to the limitations of computational resources. However, theoretically, the number of parents can be arbitrary. It is important to emphasize that specifying the conditional probability for parent-child pairs is not enough. This information does not define the relationships between the parents. Let the variables X, Y, and Z be represented as independent events, with X and Y being the parents of Z (4).

$$X = \{x_1, x_2, x_3, \dots, x_n\};\ Y = \{y_1, y_2, y_3, \dots, y_m\};\ Z = \{z_1, z_2, z_3, \dots, z_r\} \tag{4}$$

To uniquely determine the probability distribution, the following set of conditional probabilities must be specified: $\mathrm{P}(\mathrm{z_i} \mid \mathrm{x_j}, \mathrm{y_k})$. The primary task of computations in a Bayesian Network (BN) is to calculate the probability of independent events when conditional probabilities are specified at each vertex. This process is called initial propagation. If the BN is structured as a polytree, propagation algorithms are significantly simplified and yield a unique result. This is because, in this case, any two vertices in the BN graph are connected by a single path. Consequently, the influence of evidence can also propagate along a single path. If the BN graph contains undirected cycles, it has been shown in [3] that a marginal probability distribution can be constructed. The main goal (either initial or evidence propagation) is to calculate the distribution over a subset of variables within the Bayesian network. If the distribution is fully specified, it is simply a matter of summing this distribution over all possible assignments of the variables not included in the subset for which we want to obtain the distribution. However, the size of the problem grows exponentially with the number of nodes in the BN, and due to the data volume, these calculations may be practically impossible. The probability distribution for an arbitrary node with multiple pieces of evidence can be calculated using formula (5):

$$P(X_1X_2X_3\ \dots X_n) = \sum_{y_1,y_2,y_3,\dots,y_m} P(X_1X_2X_3\ \dots X_nY_1 = y_1Y_2 = y_2Y_3 = y_3 \dots Y_m = y_m) \tag{5}$$

The specified method defines local distributions. Based on these local distributions, conditional probabilities can be calculated (6):

$$P(X_1 \dots X_n|Z_1 = a_1 \dots Z_m = a_m) = \frac{\sum_{y_1,\dots,y_m} P(X_1 \dots X_nY_1 = y_1 \dots Y_m = y_mZ_1 = a_1 \dots Z_m = a_m)}{\sum_{y_1,\dots,y_mz_1,\dots,z_m} P(X_1 \dots X_nY_1 = y_1 \dots Y_m = y_mZ_1 = z_1 \dots Z_m = z_m)} \tag{6}$$

where the variables Z_k – are the variables that have received evidence. When evidence is present, the above formula no longer works. It is necessary to use more general formulas (7)-(10):

$$P(X \mid Y) = \frac{P(XY)}{P(Y)} \tag{7}$$

$$P(X \mid Y) = \frac{P(Y \mid X)P(X)}{P(Y)} \tag{8}$$

$$\sum_a P(X = a| Y) = 1 \tag{9}$$

$$P(X = a \mid Y = b) = \frac{P(Y=b \mid X=a)P(X=a)}{\sum_{a_j} P(Y=b \mid X=a_j)P(X=a_j)} \tag{10}$$

3. RESULTS AND DISCUSSION

Let's consider the educational example ASIA in Figure 1, widely known in the literature on Bayesian networks, which analyzes various scenarios of setting evidence and compares two approaches to algorithm construction considering Markov properties. In this example, the vertex *A* represents the fact of the subject's visit to Asia in Figure 2(a), the vertex *S* represents the fact of smoking in Figure 2(b), the vertex *T* represents the presence of tuberculosis, which significantly depends on the visit to Asia in Figure 2(c), and the vertex *L* indicates the presence of lung cancer in Figure 2(d). The vertex *B* represents the presence of bronchitis in Figure 2(e), the vertex *E* reflects the fact of either lung cancer or tuberculosis, the vertex *X* corresponds to the results of an X-ray, and the vertex *D* indicates the presence of dyspnea in the subject.

In this analysis, a comparison of solutions and methods for algorithm construction is conducted, taking into account the Markov property, which helps identify the differences and advantages of each approach in the context of this example. A visit to Asia triggers the risk of contracting tuberculosis. Smoking can contribute to developing lung cancer (L), bronchitis (B), or both simultaneously. The result of a chest X-ray (X) may indicate the presence of lung cancer or tuberculosis with a certain probability, but it cannot distinguish between these diseases. Lung cancer, tuberculosis, and bronchitis can cause dyspnea with a certain probability. Figure 2 shows a Bayesian network describing the relationships between five variables involved in diagnosing lung diseases. Figure 2(a) shows the probability of a subject traveling to Asia (A), where diseases such as tuberculosis can be contracted. The probability that the subject has visited Asia is 0.01 for “yes” and 0.99 for “no.” Figure 2(c) shows the probability of a subject having tuberculosis (T), depending on whether the subject has visited Asia. If the subject has visited Asia, the probability of having tuberculosis is 0.05 for “yes” and 0.95 for “no.” If the subject has not visited Asia, the probability of having the disease is 0.01 for “yes” and 0.99 for “no.” Figure 2(b) shows the probability that the subject is a smoker (S), which is equally divided: 0.5 for “yes” and 0.5 for “no.” Figure 2(d) illustrates the probability of a subject having lung cancer (L), depending on whether the subject smokes. If the subject smokes, the probability of lung cancer is 0.1 for “yes” and 0.9 for “no.” If the subject does not smoke, the probability of cancer decreases to 0.01 for “yes” and 0.99 for “no.” Figure 2(e) shows the probability that the subject will have bronchitis (B), also dependent on smoking. For smokers, the probability of bronchitis is 0.6 for “yes” and 0.4 for “no,” whereas for nonsmokers, these values are 0.3 for “yes” and 0.7 for “no.” Thus, each figure in this model shows the relationships between the variables that affect the subject's health.

We perform the triangulation of the graph in Figure 3. According to the theory, during triangulation, we first eliminate vertices with one neighbor, A and X. Then, we eliminate vertices with two neighbors, T and D. These vertices are part of cliques, so there is no need to add edges. The remaining subgraph consists of the vertices S, L, E, B, and the edges {SL, LE, EB, BS}, forming a cycle of length 4.

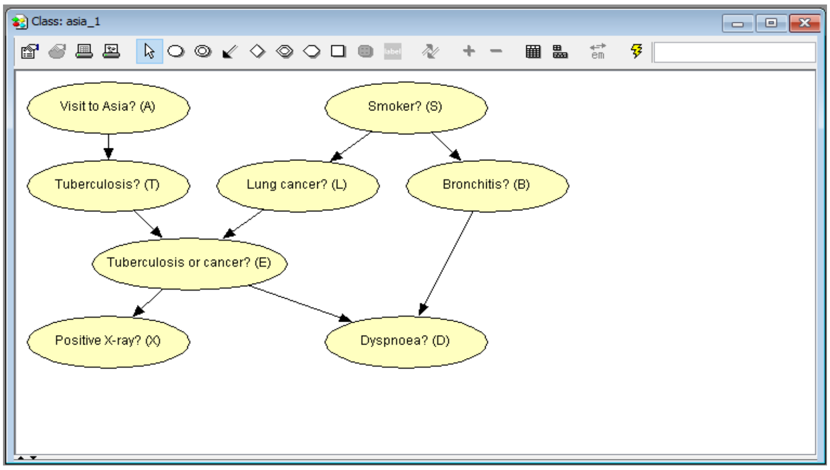

Figure 1. Educational example ASIA

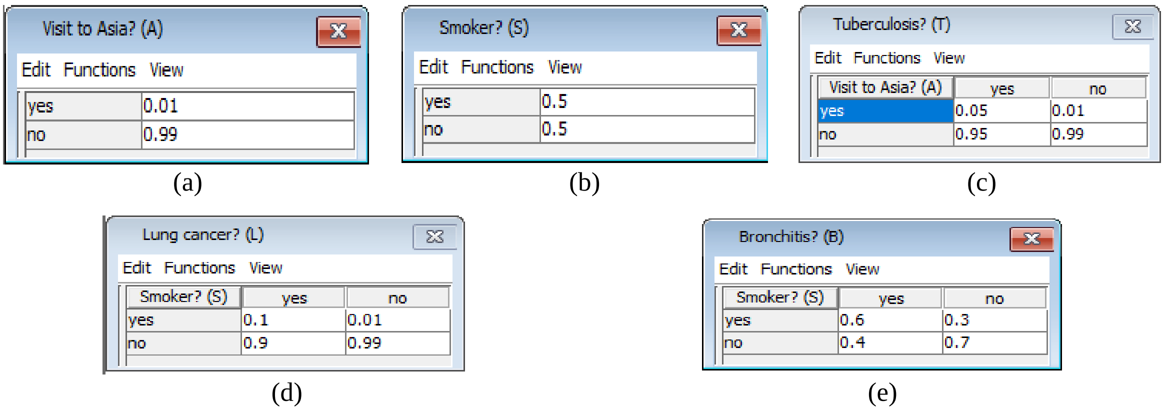

Figure 2. The fact of the subjects' visits to Asia

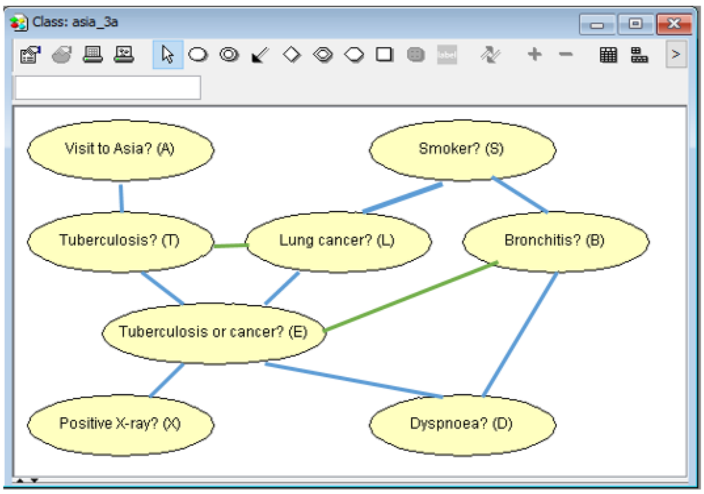

Figure 3. Graph triangulation

For the graph in Figure 1, we construct the moral graph. Here, the directions of the edges are removed, and two edges are added to connect the nodes T and L, as well as the nodes E and B. An edge L LB (or edge SE) must be added to complete the triangulation. The triangulated graph is shown in Figure 4. To construct the secondary graph, we select two minimal cliques AT and XE from the graph in Figure 4. The neighboring clique to these cliques will be the clique TLE. The neighboring clique to TLE will be the clique ELB. The neighboring cliques to ELB will be LSB and EBD.

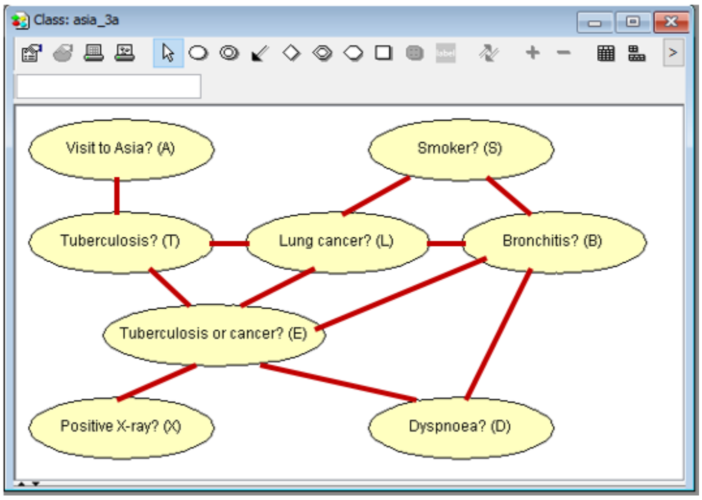

Figure 4. The join tree corresponding to the Bayesian network (BN)

The analysis of the ASIA educational example shows that, despite significant opportunities for simplifying calculations, the computational load for large Bayesian networks (BNs) can be very substantial, and transformations into a secondary graph can be complex and not always transparent to researchers. This raises the question: how can we reduce and simplify computations, making them more transparent? To achieve these goals, it is important to ensure that the accuracy of the calculations remains within reasonable limits and that the theoretical foundation of the new approach is based on previously developed theories and methods. It is necessary for the algorithms, their software implementation, and the speed of calculations to be significantly simpler and faster, allowing for the efficient processing of any reasonable BN topologies. A key aspect is also the ability to identify certain subgraphs of the BN for integrated computations, which will help optimize the process. At the same time, potential errors that may arise during the simplification of computations should be considered, and their impact on the final results should be assessed. As part of the development of a new approach to computations in Bayesian networks, it was decided to refer to a well-known paradigm from the theory of Markov chains, according to which information in the node under study depends only on the information in neighboring nodes, regardless of how that information was obtained. This work assumes that the information in the node under study in a Bayesian network depends only on specific neighboring nodes, including parents and children. Various implementations of this approach are possible: in one, the information in the node depends only on the parents, children, and other parents of the children, which corresponds to the concept of the node's Markov blanket Figure 5. In another variant, the information in the node depends on neighboring nodes, designated as set A, with the addition of all neighboring nodes for each node in set A.

When considering options, we will consider the d-separation of nodes in the Bayesian network (BN). Next, we will consider the choice of the Bayesian blanket. Without evidence, computations in the BN do not present difficulties. Let's systematize the order of computations in the absence of proof. We will divide the vertices of the BN into several levels. We will include vertices where no computations are required at the zero level. These are nodes without parents—A and S. A and S are at the zero level. T, L, and B are at the first level, with parents belonging only to the zero level. E is at the second level, with parents belonging only to the first level. X and D are at the third level, with parents belonging only to the first and second levels. A and S are at the zero level. T, L, and B are at the first level, with parents belonging only to the zero level. E is at the second level, with parents belonging only to the first level. X and D are at the third level, with parents belonging only to the first and second levels. X is at the zero level, and the vertex with evidence does not need to be calculated. E is at the first level, with vertex X connected only to vertex E. T, L, and D are at the second level; these vertices are not d-separated from vertex X through vertex E, and thus, they are influenced by vertex X through vertex E. A and S are at the third level; these vertices are not d-separated from vertex X and are therefore influenced by vertex X through vertices E, T, and L. B is at the fourth level; this vertex is not d-separated from vertex X. Vertex B is influenced by vertex S at the third level; however, vertices E and B are independent concerning the converging connection at vertex D. Thus, all vertices in the BN end up depending on vertex X. The following order of calculating the vertices in the BN can be proposed (11):

$$X -> E -> T -> L -> D -> A -> S -> B \tag{11}$$

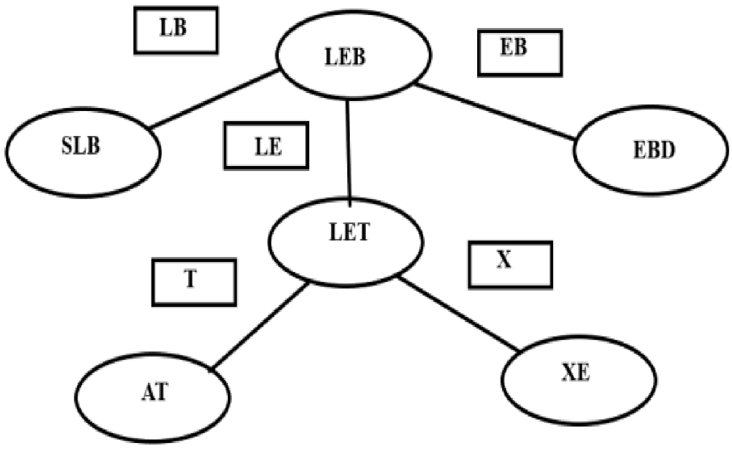

Figure 5. Moral graph

Bayes' theorem and the prior probabilities of specific vertices may be used in the calculations. The vertices involved in the calculations are determined by the Markov Bayesian blanket of the vertex being calculated or by neighboring vertices. The Bayesian blankets for the vertices will be as follows: the blanket for vertex E will be {T, L, X, D, B}; the blanket for vertex T will be {A, E, L}; the blanket for vertex L will be {S, E, T}; the blanket for vertex D will be {B, E}; the blanket for vertex A will be {T}; the blanket for vertex S will be {L, B}; and the blanket for vertex B will be {S, D, E}. We considered a relatively simple example. Evidence in the single vertex X influences all the remaining vertices. Vertex X is not separated from any of the other vertices. Let's complicate the task a bit. Suppose the evidence is in vertex A. Vertex E contains a converging connection. This means that vertices A and T will be independent (d-separated) from vertices L and S. Similarly, vertices E and B are independent concerning the converging connection at vertex D. The partitioning of vertices by levels in this case will be as follows: A and S are at the zero level, with vertex A containing evidence and vertex S being d-separated from vertex A, meaning they are independent. T, L, and B are at the first level, with vertex T depending only on vertex A, and vertices L and B depending only on vertex S. E is at the second level, depending on vertices T and L. X and D are at the third level. The vertices involved in the calculations are determined either by the Markov Bayesian blanket of the calculated vertex or by neighboring vertices. The Bayesian blankets for the vertices will be as follows: the blanket for vertex T will be {A, E, L}; the blanket for vertex L will be {S, E, T}; the blanket for vertex B will be {S, D, E}; the blanket for vertex E will be {T, L, X, D, B}; the blanket for vertex X will be {E}; and the blanket for vertex D will be {B, E}.

In the calculations, we will use the probabilities of the nodes calculated in the previous step. In the initial step, we calculate the nodes without evidence. The data obtained is then used in the calculations for the subsequent steps. In the first step, we estimate the probabilities of one of the neighboring nodes of the first piece of evidence. If there are more pieces of evidence, we estimate the probabilities of one of the neighboring nodes of the second piece of evidence, and so on. After processing all the evidence, we replace the proof with the nodes that we have just processed. We then construct their Markov blankets. In the second step, we calculate the probabilities of the new nodes. This BN example is too simple to provide examples with multiple pieces of evidence. Let's consider an example of a second BN. Figure 6 shows a network containing 20 vertices.

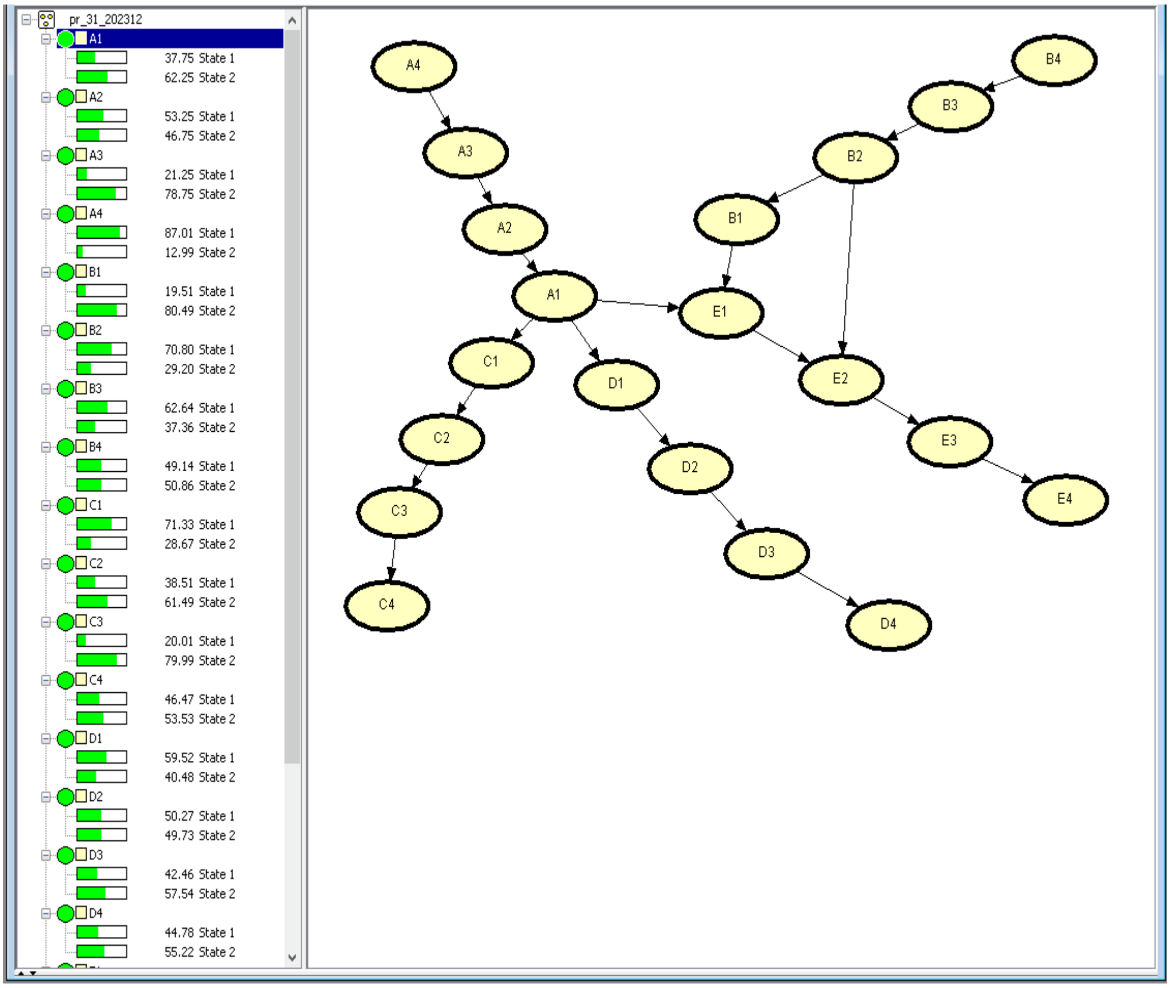

Figure 6. Level partitioning considering d-separation

Of particular interest are vertices E1 and E2, which possess a converging connection, determining the existence of d-separation in certain cases. When considering a scenario where vertices A4 and C3 have received evidence, the level partitioning will be as follows: At the zero level are vertices A4, C3, and B4, where the first two vertices have received evidence, and B4 is not connected to the vertices generated by vertices A4 and C3 through converging connections in E1 and E2. The first level includes vertices A3, C4, C2, and B3. The second level consists of vertices A2, C1, and B2. The third level includes vertices A1 and B1. Vertices D1 and E1 represent the fourth level, the fifth level by vertices D2 and E2, the sixth level includes D3 and E3, and the seventh level consists of vertices D4 and E4. The following order of calculating the vertices in the BN can be proposed: A3 -> C4 -> C2 -> B3 -> A2 -> C1 -> B2 -> A1 -> B1 -> D1 -> E1 -> D2 -> E2 -> D3 -> E3 -> D4 -> E4. The Bayesian blankets for the vertices will be as follows: the blanket for vertex A3 will be {A4, A2}; the blanket for vertex C4 will be {C3}; the blanket for vertex C2 will be {C3, C1}; the blanket for vertex B3 will be {B4, B2}; the blanket for vertex A2 will be {A3, A1}; the blanket for vertex C1 will be {C2, A1}; the blanket for vertex B2 will be {B3, B1, E2, E1}; the blanket for vertex A1 will be {A2, C1, D1, E1, B1}; the blanket for vertex B1 will be {B2, E1, A1}; the blanket for vertex D1 will be {A1, D2}; the blanket for vertex E1 will be {A1, B1, E2, B2}; the blanket for vertex D2 will be {D1, D3}; the blanket for vertex E2 will be {B2, E1, E3}; the blanket for vertex D3 will be {D2, D4}; the blanket for vertex E3 will be {E2, E4}; the blanket for vertex D4 will be {D3}; and the blanket for vertex E4 will be {E3}. The article explores the possibilities of using the Markov property to simplify computations in Bayesian networks, which introduces new challenges while solving a number of existing problems. Specifically, it requires justification for applying the Markov property, adaptation of existing probability tables despite their inconvenient structures, and proper justification for using Bayes' theorem. It is essential to prove that in the presence of multiple pieces of evidence and changes in the direction of some edges, cycles do not arise and to investigate potential contradictions in Bayesian networks that may occur due to incorrect or incompatible receipt of multiple pieces of evidence, offering solutions to these issues. Despite the challenges, the approach under consideration appears to be quite promising and warrants further investigation. The effectiveness of using Bayesian networks largely depends on selecting the correct concept for querying the network and efficient propagation methods, especially in the presence of multiple pieces of evidence. The complex and extensive topology of Bayesian networks can significantly increase computation time, and many existing algorithms currently use an approach based on representing the original network as a junction tree. This method has several advantages, such as universality, obtaining accurate probability values, and high-speed implementation of algorithms in code. However, this approach also has limitations, including difficulties in computations for specific topologies and complexity in understanding and developing algorithms. As a result, many developers are seeking ways to simplify algorithms and improve their efficiency, highlighting the need for new, correct paradigms for working with Bayesian networks. One such direction is a return to the idea of Markov chains, where the information in a node depends only on the state of neighboring nodes, which simplifies algorithms and increases their speed. It is essential to consider the order of node computations and the need for iterations in the presence of evidence. Controlling the d-separation of nodes and using the Markov blanket also play a crucial role in improving the accuracy of calculations, although they may complicate the process. Thus, the proposed approach to developing algorithms for Bayesian networks requires further investigation and practical testing to assess its effectiveness in real-world conditions.

4. CONCLUSION

This article examines the potential of using Markov processes to simplify Bayesian networks (BNs) computations. Despite the obvious advantages, such as speeding up calculations and simplifying algorithms, introducing the Markov property into Bayesian networks also introduces several new challenges. These include the need to justify the application of Markov properties, adapt existing probability tables to new structures, and correctly apply Bayes' theorem. Additionally, it is important to prove that no cycles arise in the presence of multiple pieces of evidence and changes in the direction of certain edges and to investigate possible contradictions that may occur when multiple pieces of evidence are received incompatible.

Nevertheless, the approach under consideration is promising and requires further research, especially in complex and large BN topologies. The practical application of these methods could significantly improve the efficiency of working with Bayesian networks, particularly in processing data with multiple uncertainties. The proposed algorithms and approaches need detailed testing and optimization for use in real-world conditions, which will allow their total practical value to be realized.

ACKNOWLEDGEMENTS

This paper was written under grant AP19679142 “Search for optimal solutions in Bayesian networks in models with linear constraints and linear functionals. Development of algorithms and programs” (2023-2025) of the Ministry of Education and Science of the Republic of Kazakhstan.

REFERENCES

[1] J. K. Ghosh, “Probabilistic networks and expert systems: exact computational methods for Bayesian networks,” *International Statistical Review*, vol. 76, no. 2, pp. 306–307, Jul. 2008, doi: 10.1111/j.1751-5823.2008.00054_8.x.

[2] F. J. Radermacher, “Probabilistic reasoning in intelligent systems: networks of plausible inference,” *SIAM Review*, vol. 32, no. 4, pp. 704–707, Dec. 1990, doi: 10.1137/1032155.

[3] L. E. Sucar, “Bayesian networks: representation and inference,” in *Probabilistic Graphical Models*, Springer International Publishing, 2020, pp. 111–151.

[4] J. Baz, P. Alonso, J. M. Peña, and R. Pérez-Fernández, “Gaussian Markov random fields and totally positive matrices,” *Journal of Computational and Applied Mathematics*, vol. 430, Oct. 2023, doi: 10.1016/j.cam.2023.115098.

[5] F. Leonardi, R. Carvalho, and I. Frondana, “Structure recovery for partially observed discrete Markov random fields on graphs under not necessarily positive distributions,” *Scandinavian Journal of Statistics*, vol. 51, no. 1, pp. 64–88, Aug. 2023, doi: 10.1111/sjos.12674.

[6] B. Panić, M. Borovinšek, M. Vesenjak, S. Oman, and M. Nagode, “A guide to unsupervised image segmentation of mCT-scanned cellular metals with mixture modelling and Markov random fields,” *Materials and Design*, vol. 239, Mar. 2024, doi: 10.1016/j.matdes.2024.112750.

[7] M. Soleimani, F. Campean, and D. Neagu, “Integration of hidden Markov modelling and Bayesian network for fault detection and prediction of complex engineered systems,” *Reliability Engineering & System Safety*, vol. 215, Nov. 2021, doi: 10.1016/j.ress.2021.107808.

[8] A. El-Awady and K. Ponnambalam, “Integration of simulation and Markov chains to support Bayesian networks for probabilistic failure analysis of complex systems,” *Reliability Engineering & System Safety*, vol. 211, 2021, doi: 10.1016/j.ress.2021.107511.

[9] D. Kaur *et al.*, “Application of Bayesian networks to generate synthetic health data,” *Journal of the American Medical Informatics Association*, vol. 28, no. 4, pp. 801–811, Dec. 2020, doi: 10.1093/jamia/ocaa303.

[10] A. V. Makkuva *et al.*, “Attention with Markov: a framework for principled analysis of transformers via Markov chains,” *arXiv preprint arXiv:2402.04161*, 2024.

[11] G. Wolfer and S. Watanabe, “Geometric aspects of data-processing of Markov chains,” *Transactions of Mathematics and Its Applications*, vol. 8, no. 1, May 2024, doi: 10.1093/imatrm/tnae001.

[12] W. Wang and T. Osaragi, “Generating and understanding human daily activity sequences using time-varying Markov chain models,” *Travel Behaviour and Society*, vol. 34, Jan. 2024, doi: 10.1016/j.tbs.2023.100711.

[13] A. Waqar, “Intelligent decision support systems in construction engineering: an artificial intelligence and machine learning approaches,” *Expert Systems with Applications*, vol. 249, Sep. 2024, doi: 10.1016/j.eswa.2024.123503.

[14] A. Patnaik and K. P. K., “Intelligent decision support system in healthcare using machine learning models,” *Recent Patents on Engineering*, vol. 18, no. 5, Jul. 2024, doi: 10.2174/1872212118666230606145738.

[15] T. Xu, Y. Xu, D. Wang, S. Chen, W. Zhang, and L. Feng, “Path planning for autonomous articulated vehicle based on improved goal-directed rapid-exploring random tree,” *Mathematical Problems in Engineering*, vol. 2020, pp. 1–14, May 2020, doi: 10.1155/2020/7123164.

[16] V. Muneeswaran, P. Nagaraj, and M. F. Ijaz, “An articulated learning method based on optimization approach for gallbladder segmentation from MRCP images and an effective IoT-based recommendation framework,” in *Studies in Computational Intelligence*, vol. 1021, S. Mishra, A. González-Briones, A. K. Bhoi, P. K. Mallick, and J. M. Corchado, Eds. Cham: Springer International Publishing, 2022, pp. 165–179.

[17] A. Jain *et al.*, “An enhanced self-learning-based clustering scheme for real-time traffic data distribution in wireless networks,” *Heliyon*, vol. 9, no. 7, Jul. 2023, doi: 10.1016/j.heliyon.2023.e17530.

[18] X. Hou, W. Hu, and M. Luo, “Short-term wind farm cluster power point-interval prediction based on graph spatio-temporal features and S-Stacking combined reconstruction,” *Heliyon*, vol. 10, no. 14, Jul. 2024, doi: 10.1016/j.heliyon.2024.e33945.

[19] D. W. Wajgi and J. V Tembhurne, “Localization in wireless sensor networks and wireless multimedia sensor networks using clustering techniques,” *Multimedia Tools and Applications*, vol. 83, no. 3, pp. 6829–6879, Jun. 2023, doi: 10.1007/s11042-023-15956-z.

[20] S.-K. Han, “Use of fibrin glue to support wound healing,” in *Innovations and Advances in Wound Healing*, Springer Nature Singapore, 2023, pp. 127–139.

[21] A. Shayakhmetova, N. Litvinenko, O. Mamyrbayev, W. Wójcik, and D. Zhamangarin, “Cycles in Bayesian networks,” *International Journal of Electronics and Telecommunications*, pp. 181–186, Dec. 2020, doi: 10.24425/ijet.2021.135962.

[22] A. Akhmetova, A. Shayakhmetova, A. Khabduali, M. Shurenov, and A. Sultangaziyeva, “The system of finding people in an emergency situation,” *Journal of Theoretical and Applied Information Technology*, vol. 100, no. 9, pp. 2924–2934, 2022.

[23] K. Alimhan, N. Tasbolatuly, and A. Yerdenova, “Global output tracking control for high-order non-linear systems with time-varying delays,” *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 13, pp. 3337–3352, 2021.

[24] K. Alimhan, N. Otsuka, M. N. Kalimoldayev, and N. Tasbolat, “Practical output tracking for a class of uncertain nonlinear time-delay systems via state feedback,” *MATEC Web of Conferences*, vol. 189, 2018, doi: 10.1051/matecconf/201818910027.

[25] K. Alimhan, M. N. Kalimoldayev, A. A. Adamov, O. Mamyrbayev, N. Tasbolatuly, and A. Smolarz, “Further results on output tracking for a class of uncertain high-order nonlinear time-delay systems,” *Przeglad Elektrotechniczny*, vol. 95, no. 5, pp. 88–91, May 2019, doi: 10.15199/48.2019.05.22.

[26] N. K. Kitson, A. C. Constantinou, Z. Guo, Y. Liu, and K. Chobtham, “A survey of Bayesian network structure learning,” *Artificial Intelligence Review*, vol. 56, no. 8, pp. 8721–8814, Jan. 2023, doi: 10.1007/s10462-022-10351-w.

[27] M. Scutari and J.-B. Denis, *Bayesian networks: with examples in R*, Chapman and Hall/CRC, 2021.

[28] D. Heckerman and D. Geiger, “Likelihoods and parameter priors for Bayesian networks,” *arXiv preprint arXiv:2105.0624*, 2021.

[29] A. Orynbayeva, N. Shyndaliyev, and A. Aripbayeva, “Improving statistical methods of data processing in medical universities using machine learning,” *World Transactions on Engineering and Technology Education*, vol. 21, no. 1, pp. 58–63, 2023.

[30] G. Abdikerimova *et al.*, “Detection of lung pathology using the fractal method,” *International Journal of Electrical and Computer Engineering*, vol. 13, no. 6, pp. 6778–6786, Dec. 2023, doi: 10.11591/ijece.v13i6.pp6778-6786.

[31] A. Shekerbek, S. Serikbayeva, M. Tulenbayev, G. Bakanov, S. Beglerova, and A. Makovetskaya, “Application of mathematical methods and machine learning algorithms for classification of X-ray images,” *Eastern-European Journal of Enterprise Technologies*, vol. 3, no. 2, pp. 6–17, Jun. 2022, doi: 10.15587/1729-4061.2022.259710.

BIOGRAPHIES OF AUTHORS

Assem Shayakhmetova in 2004, she graduated from the Abai Kazakh National pedagogical university with a degree in computer science. In 2006, she received a master's degree in computer science. In 2016, she graduated from the doctoral program Kazakh National Technical University named after K.I. Satbayev, specialty 6D07300 - Information systems. In 2017 she defended her doctoral dissertation and received a Ph.D. degree. She is the author of more than 50 works and participation with reports in international conferences in countries far and near abroad and in Kazakhstan. In 2018-2020 was the scientific director of the GF project No. AP05131293 “Development and software implementation of a package of application programs for solving applied problems on Bayesian networks”. Currently, she is the scientific director of the AP 9679142 project “Search for optimal solutions in Bayesian networks in models with linear constraints and linear functionals. Development of algorithms and programs” (2023-2025). Her research interests include developments in computer science, data mining, artificial intelligence, and working with Bayesian networks. She is a Ph.D., associate professor at the Department of Artificial Intelligence and Big Data at the Al-Farabi Kazakh National University. She can be contacted at email: asemshayakhmetova@mail.ru.

Nurbolat Tasbolatuly in 2020, Nurbolat Tasbolatuly defended his doctoral dissertation on the topic "Global practical tracking for high-order uncertain nonlinear systems and creating program complex for them" in the specialty 6D070400 - Computer engineering and software. He currently works at Astana International University as docent of the Higher School of Information and Computing Technologies. His field of research includes control theory, numerical methods for solving tracking problems, mathematical and computer modeling. In the field of research, he has more than 50 scientific papers, 8 of them in highly rated journals. He can be contacted at email: tasbolatuly@gmail.com.

Ardak Akhmetova received a specialist degree in 1996, and a master's degree in Computer Science from Kainar University in 2009. In 2015-2018 studied for a doctorate at Al-Farabi Kazakh National University, majoring in information systems. In 2019, she defended her doctoral dissertation and received a Ph.D. degree. More than 20 scientific papers have been published. works in republican and high-ranking foreign scientific journals, participation with reports in international conferences in countries far and near abroad and in Kazakhstan. Akhmetova A.M.-Ph.D., acting associate professor of the Department of Artificial Intelligence and Big Data, Al-Farabi Kazakh National University, Almaty, Kazakhstan. She can be contacted at email: ardak_66@mail.ru.

Assel Abdildayeva received a bachelor's degree in 2002, and a master's degree in computer science in 2005 from the Kazakh National Pedagogical University. Abay. In 2013-2016 studied for a doctorate at the International University of Information Technologies, majoring in information systems. In 2016 she defended her doctoral dissertation and received a Ph.D. degree. More than 30 scientific papers have been published. works in republican and high-ranking foreign scientific journals, participation with reports in international conferences in countries far and near abroad and in Kazakhstan. In 2018-2020 was the scientific director of the GF project No. AP05132020 "Development of an intelligent information complex and mathematical models for modern electric power systems." Currently acting as scientific director of the GF project No. AP 9679142 "Search for optimal solutions in Bayesian networks in models with linear constraints and linear functionals. Development of algorithms and programs" (2023-2025). Abdildayeva Assel is a Ph.D., associate professor at the Department of Artificial Intelligence and Big Data at the Al-Farabi Kazakh National University. She can be contacted at email: asselabdildayeva5@gmail.com.

Marat Shurenov graduated from the Kazakh Academy of Transport and Communications named after M. Tynyshpayev in 2008 in the specialty organization of transportation, traffic and operation of transport with the qualification of Bachelor of Transport, in 2010 this academy with this specialty with a master's degree. In 2014, he graduated and successfully defended his doctorate at the National Technical University named after K.I. Satpayev in the specialty 6D071300-Transport, transport equipment and technology. By authority of the Committee for Control of Education and Science of the Ministry of Education and Science of the Republic of Kazakhstan dated March 18, 2015, order number 365, Shurenov Marat was awarded the degree of Doctor of Philosophy 6D071300-Transport, Transport Engineering and Technology. He has more than 70 works published in local and foreign publications with an impact factor. He can be contacted at email: iskendirmk@mail.ru.

Anar Sultangaziyeva in 2004, she graduated from the Abai Kazakh National pedagogical university with a degree in computer science. In 2008, she received a master's degree in computer science. From 2020 to the present, she has been a lecturer at the Astana International University. She is the author of more than 10 works. Her research interests include developments in computer science, cloud technologies, and internet of things. She can be contacted at email: anara77777@mail.ru.