UDC 662.753.3+547.915
DOI https://doi.org/10.32782/2663-5941/2025.5.2/02

***Bugaieva L.M.***
National Technical University of Ukraine
«Igor Sikorsky Kyiv Polytechnic Institute»

***Abramova A.O.***
National Technical University of Ukraine
«Igor Sikorsky Kyiv Polytechnic Institute»

# DEVELOPMENT OF AN INTELLIGENT SYSTEM USING A BAYESIAN NETWORK FOR EVALUATING AIR FILTRATION EFFICIENCY

*The article presents the results of a study aimed at developing an intelligent system in the form of a Bayesian network for evaluating the efficiency of the air filtration process in gas turbine installations. There is a need to create systems capable of operating in conditions of uncertainty and limited data in the context of intensive development of information technologies. Bayesian belief networks combine statistical observations with expert knowledge, making them an effective tool for modeling complex technical and environmental processes. Particular attention is paid to the use of the HUGIN Lite software environment, which allows automating the construction of the network structure, determining the conditional probability tables, and conducting computational experiments. Nine key factors have been identified for modeling the air purification process, including dust particle diameter, air temperature and velocity, pressure drop across the filter partition, dust penetration coefficient, and filter performance. The probabilistic model was constructed based on experimental data, it reflects the cause – and – effect relationships between these parameters and allows conclusions to be drawn about the condition of the filter element. The simulation results confirmed the system's ability not only to reproduce known patterns, but also to generate recommendations for operational decisions, including the need to replace filters. It has been established that the use of a Bayesian belief network allows both a priori expert knowledge and empirical data to be taken into account, which increases the accuracy of the system status assessment and the effectiveness of management decisions. The proposed approach can be considered a promising direction for the further development of intelligent systems for monitoring and predicting the efficiency of filtration processes in industrial and environmental applications.*

***Key words:*** *Bayesian network, intelligent system, air filtration, gas turbine plant, modeling, probabilistic models.*

**Formulation of the problem.** In addition to traditional statistical data processing, a large amount of experimental or monitoring data makes it possible to perform intelligent data analysis, which will provide new knowledge about the process. Nowadays, intelligent tools such as neural networks, fuzzy inference systems, and many others, including Bayesian networks (BN), are widely used. These networks are graphical models of probabilistic and causal relationships between variables in statistical information modeling. Bayesian networks can organically combine empirical frequencies of occurrence of different variable values, subjective assessments of "expectations," and theoretical ideas about the mathematical probability of certain consequences from a priori information. This is an important practical advantage and distinguishes Bayesian networks from other intelligent modeling techniques. The use of Bayesian networks to create an intelligent system for evaluating the air filtration process in a gas turbine installation is a relevant area of research.

**Analysis of recent research and publications.** The development of modern information technologies is increasingly linked to the need to create intelligent systems capable of functioning effectively in conditions of uncertainty and incomplete data. One of the key devices that has proven its effectiveness in such tasks is Bayesian belief networks (BBN) [1]. They allow the creation of probabilistic models of complex processes, combining statistical data with expert knowledge, and providing decision support in various fields [2].

In Ukraine, this topic has been actively developed over the past two decades. Domestic scientists are researching both the theoretical foundations of and training BBN and the applied aspects of their use in intelligent systems for various purposes, from transportation to cybernetic and economic – ecological. However, despite the existence of a fundamental basis, there are significant challenges related to the limited number of large – scale empirical studies and the need to standardize model validation methods.

Contemporary scientific works in Ukraine confirm the active development of learning methods and improvement of BN structures. Thus, in [3], a new approach to learning network structures using genetic algorithms was proposed. It is emphasized that classical algorithms often prove to be ineffective for large systems, therefore the use of evolutionary methods allows optimizing the process of BN development. The advantage of the study is the combination of theoretical foundations with practical examples, although the lack of testing on large real data sets can be considered a limitation.

BN are quite actively used in applied fields. Thus, the authors [4] proposed the use of BN for forecasting nonlinear non – stationary processes. A positive aspect of the work is the introduction of structural time series, which allows modeling complex dynamic dependencies. At the same time, a certain limitation of the study is the need for significant computational resources.

In environmental research, the author [5] used BN to evaluate the accuracy of atmospheric pollution modeling results. The value of the work is the attempt to integrate dispersion models with decision support systems. However, a certain weakness is the dependence on the quality of the initial measurements, which are often inaccurate.

The authors [6] use a BN to assess the impact of air pollution. The model takes into account both external monitoring data and biological indicators, which increases its interdisciplinary value. However, the proposed model is limited to one city and focuses on impact assessment rather than air quality forecasting or management.

Despite the availability of thorough theoretical and applied research, the Ukrainian scientific school on BN faces a number of challenges.

1. Limited number of large – scale empirical studies. The majority of studies are limited to laboratory experiments or narrow applied tasks.

2. Insufficient standardization of evaluation methods. Many publications lack clear metrics for validating models and comparing their effectiveness.

3. Problems with data openness. A significant part of applied research is based on proprietary data sets that are not publicly available, which complicates the reproducibility of results.

4. Lack of interdisciplinary projects. Although BBN are universal in nature, there are still few examples of their integration into complex intelligent systems in various industries in Ukraine.

Given these problems, several promising areas for development can be identified.

1. Expansion of the empirical base. Large– scale research involving large data sets is needed, particularly in the fields of transport, medicine, and energy. Standardization of methods. Development of uniform criteria for assessing the quality of models and procedures for their validation.

2. Development of software tools. Creation of local libraries and platforms for developing open – source BBN tailored to Ukrainian users.

3. Multidisciplinary. Involvement of specialists from various fields to create complex systems based on BBN.

4. Educational integration. Further implementation of BBN methods in educational programs in computer science, artificial intellect, and systems analysis.

**Task statement.** The purpose of the article is to develop an intelligent system in the form of a Bayesian network to evaluate the efficiency of the air filtration process in a gas turbine installation.

In accordance with the stated goal, the authors solved the following tasks.

1. Assessment of the state of the process and directly the state of the filter element for developing a Bayesian network.

2. Implementation of decision support based on the results obtained from the system on the state of the air purification process for a gas turbine installation.

**Outline of the main material of the study.** BBN are used in areas characterized by uncertainty that may arise due to: incomplete understanding of the subject area; incomplete knowledge; when the task is characterized by randomness.

BBNs are typically used to model situations in which random events are linked by cause – and – effect relationships. BBN are directed acyclic graphs with the following properties: each vertex is an event described by a random variable that can have several states; all vertices connected to "parent" vertices are determined by a conditional probability table (CPT) or a conditional probability function (CPF); for vertices without "parents," the probabilities of their states are unconditional (marginal).

In other words, in BBN, vertices are random variables, and arcs are probabilistic dependencies defined by conditional probability tables. The conditional probability table for each vertex contains the probability of the states of that vertex given the states of its "parent vertices." In BN, there are three possible types of relationships between variables: sequential connections; divergent connections; convergent connections.

The HUGIN Lite system allows to create decision support systems in conditions of uncertainty based on models of the problem domain. The system is focused on developing models based on BN theory and influence diagrams. The system can be used to create expert systems in a wide variety of problem domains, including the construction of environmental decision support systems. Modern software tools such as HUGIN provide the tools for developing such networks, as well as the ability to use BBN to enter new evidence and obtain a decision (conclusion) by recalculating the probability at all vertices corresponding to this evidence.

The design of a BBN consists of the following stages:

– adding new vertices to the designed BBN;
– establishing cause – and – effect relationships between the vertices of the designed BBN;
– determining all possible states of each of the BBN vertices;
– assigning values to the conditional probability tables for each of the vertices of the BBN;
– compilation of the designed BBN;
– change in probability in the BBN when new knowledge is introduced.

Based on the analysis of the air filtration process using a cartridge filter for a gas turbine engine, the main factors affecting the efficiency of the process can be identified [7]:

– the diameter of dust particles in the filtered air;
– air temperature at the inlet to the cleaning system;
– input air velocity;
– pressure drop across the filter partition.

Of course, there are many more factors that affect the final result of air purification. However, this study will only consider the main factors. The indicators describing the filtration process also include such indicators as the dust particle transmission coefficient through the filter partition, the reduction in engine power (caused by the pressure drop across the partition), and the filter performance.

Based on the above, the following nodes (vertices) of the BN can be identified (Fig. 1). There will be nine of these nodes.

The data used to construct conditional probability tables are experimental values of process factors accumulated over a certain period of time. They fully reflect the nature of the process. In total, there are 820 statistical data points. Table 1 shows a part of the statistical data.

When the network structure is defined and the data for its construction is determined, it is necessary to define all possible states for the variables.

Table 2 shows that the conditional probability table for the network should consist of 64,800 values (4*3*6*5*5*2*6*3). This means that, for example, to construct a conditional probability table for each variable, it is necessary to have 64,800 values for it. However, the HUGIN software allows to automatically generate a conditional probability table based on

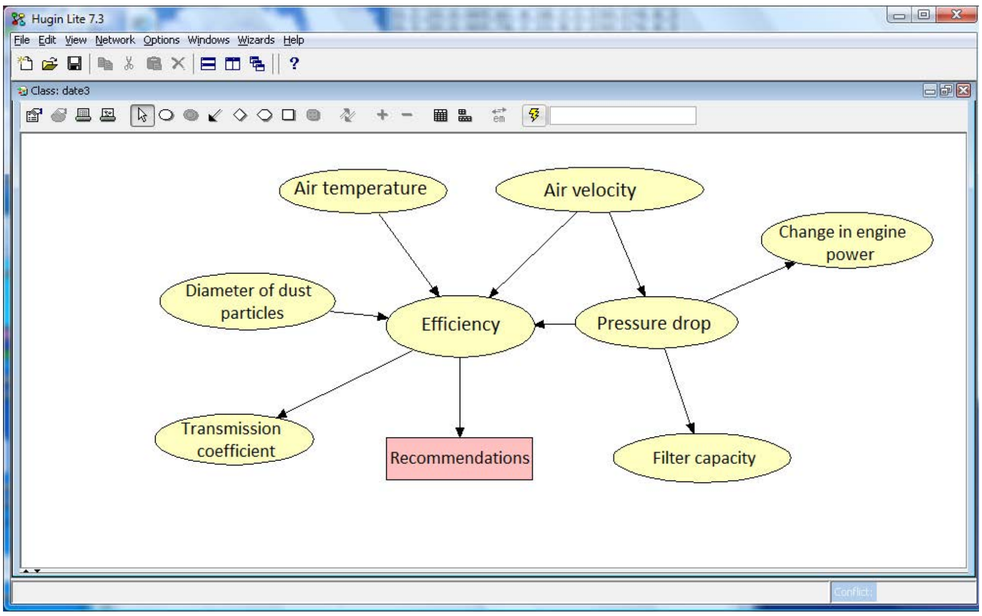



**Fig. 1. Nodes of the BN for describing the air filtration process**

existing data. To do this, it uses a built – in algorithm called *Estimation – Maximization* (EM) learning. The procedure for constructing a BN is as follows.

*Step 1.* The first step in developing a network is to create and load a file with data in one of the possible formats (for example, .txt or .csv) into the program, as shown in Fig. 2.

*Step 2.* Launch the built – in network preparation wizard and specify the list of network variables.

*Step 3*. The next window (Fig. 3) Data preprocessing is for preparing the data. At this step, the program allows to specify which variables to include for developing the network. Here, all variables are selected, i.e., they are included.

At this step, it is important to specify the possible states of each variable. In our case, Table 2 was used for this purpose. To assign a discrete value to a variable, select the "*Discretize*" type and enter the range of values for the variable. Fig. 3 shows the range for variable E (efficiency, %).

*Step 4*. When all possible states for all variables have been entered, you can view the table (Fig. 4).

*Step 5.* Here you can specify the relationships between the variables in the BN. Even if the user is

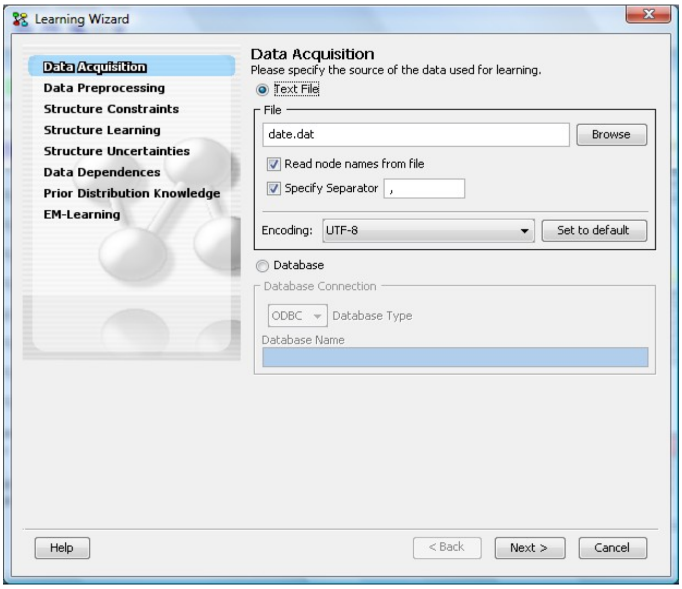



**Fig. 2. Program window at the data loading stage**

Table 1

**Part of the statistical data for developing the network**

| **Particle diameter, cm** | **0,3** | **0,5** | **0,6** | **1,5** | **1,7** | **1,9** |
|---|---|---|---|---|---|---|
| Air temperature, °C | 20 | 32 | 32 | 21 | 22 | 32 |
| Filtration, m/s | 0,005 | 0,005 | 0,01 | 0,005 | 0,005 | 0,01 |
| Filter efficiency, % | 74,2 | 88,9 | 92,9 | 99,88 | 99,95 | 92,9 |
| Transmission coefficient, % | 58,3 | 11,1 | 7,2 | 0,1 | 0,05 | 7,2 |
| Filter productivity, $m^3$/h | 176,8 | 186,7 | 173,5 | 213,8 | 215,1 | 173,5 |
| Pressure drop, Pa | 155 | 168 | 158,9 | 212 | 298,9 | 158,9 |
| Reduction engine power, % | 3 | 3,4 | 3,2 | 4,2 | 6,2 | 3,2 |

Table 2

**Values of possible states for network variables**

| **Particle diameter, cm** | **0–0,5** | **0,5–1,0** | **1,0–1,5** | **1,5–2,0** |
|---|---|---|---|---|
| Air temperature, °C | 0 | 15 | 25 | 30 |
| Filtration, m/s | 0,001–0,005 | 0,005–0,01 | 0,01–0,015 | 0,015–0,02 |
| Filter efficiency, % | 0 | 20 | 40 | 60 |
| Transmission coefficient, % | 0 | 20 | 40 | 60 |
| Filter productivity, $m^3$/h | 0 | 100 | | |
| Pressure drop, Pa | 0 | 100–150 | 150–200 | 200–250 |
| Reduction engine power, % | 0 | 3 | 6 | |

unsure about the presence or absence of a relationship between certain vertices, the program will find the necessary relationships based on the specified data.

*Step 6.* At this step, the program prompts to select an algorithm that will be used to determine the data structure. The most commonly used algorithm is *Necessary Path Condition* (NPC). The default significance level is set to 0.05 (Fig. 5).

*Step 7.* At this stage, the strength of the connection between the vertices is set using the slider (Fig. 6). Fig. 6 shows that the connection between some variables is stronger.

*Step 8.* At this step, the program prompts you to set probability values for the variables. Since we do not have such information in our case, we leave the value "1", which indicates the absence of any information. Fig. 7 shows the conditional probability table for variable E (efficiency, %).

*Step 9.* The last step before calculating the conditional probability table is to set the number of iterations required for the process. In Fig. 8, the value "0" is specified in the *Number of iterations field*. This value indicates that the number of iterations is not limited. Also, at this stage, the convergence threshold

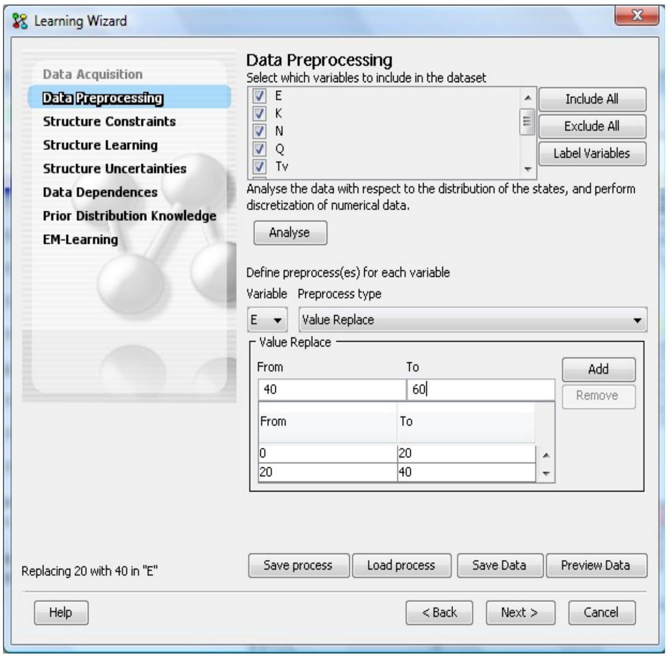
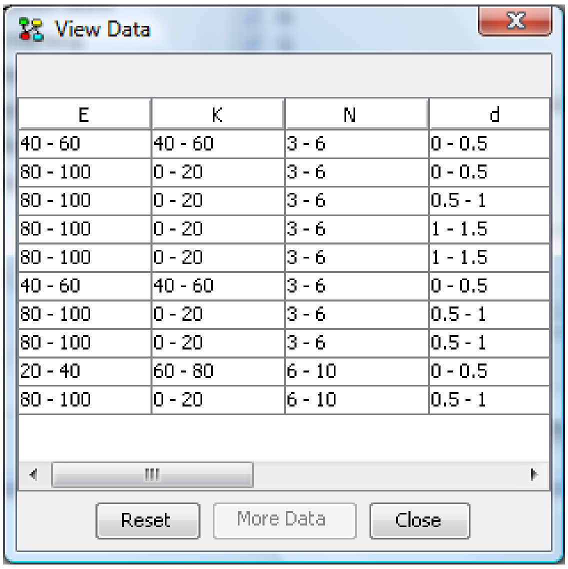



**Fig. 3. Program window at the data preparation stage**




| E | K | N | d |
|---|---|---|---|
| 40 - 60 | 40 - 60 | 3 - 6 | 0 - 0.5 |
| 80 - 100 | 0 - 20 | 3 - 6 | 0 - 0.5 |
| 80 - 100 | 0 - 20 | 3 - 6 | 0.5 - 1 |
| 80 - 100 | 0 - 20 | 3 - 6 | 1 - 1.5 |
| 80 - 100 | 0 - 20 | 3 - 6 | 1 - 1.5 |
| 40 - 60 | 40 - 60 | 3 - 6 | 0 - 0.5 |
| 80 - 100 | 0 - 20 | 3 - 6 | 0.5 - 1 |
| 80 - 100 | 0 - 20 | 3 - 6 | 0.5 - 1 |
| 20 - 40 | 60 - 80 | 6 - 10 | 0 - 0.5 |
| 80 - 100 | 0 - 20 | 6 - 10 | 0.5 - 1 |

**Fig. 4. Program window at the stage of reviewing the entered possible states of variables**

value is set. It must be greater than zero, but usually takes the value $10^{-4}$ (Fig. 8).

For a more detailed description of the process of creating a Bayesian network using the HUGIN Lite software product, you can use the built – in help.

*Step 9*. The last step is to generate a table of conditional probabilities. For variable *E,* it is shown in Fig. 9.

Fig. 10 shows the final appearance of the constructed BN after compilation. The result can be described as follows. In our case, at an air temperature of 25–35 $^{O}$C, air velocity within 0,01– 0,02 m/s, and dust particle size in the air within 1 cm, the probability of high– quality cleaning was low (efficiency value is 20–40%, a pass rate within 40– 60%, and a very large pressure drop of 300350 Pa, which indicates that the filters are already contaminated and need to be replaced. The top of the "Recommendations" section offers three possible courses of action, one of which is to replace the filters.

Further in the process, to obtain new probabilities of network values when new knowledge appears (i.e., when the fact of an increase (decrease) in the value of one of the variables appears), it is necessary to

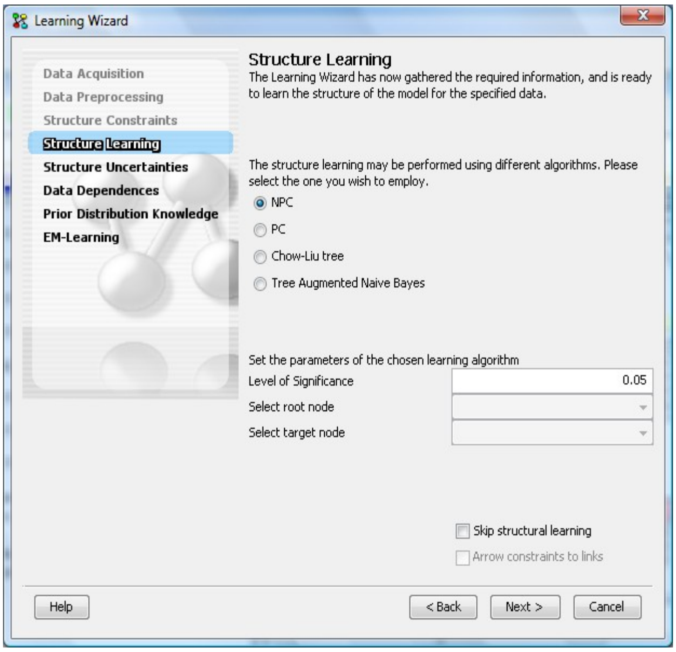



**Fig. 5. Program window for selecting the network construction algorithm and significance level**

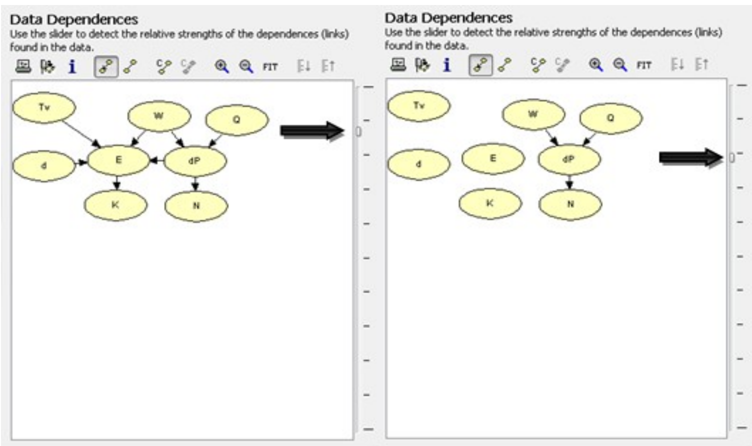



**Fig. 6. Program window for setting the strength of the connection between vertices**

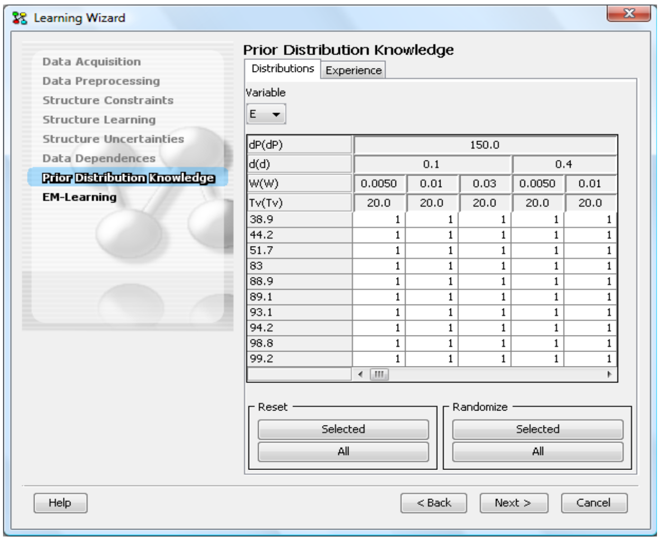



**Fig. 7. Program window at the stage of specifying probability information for variables**

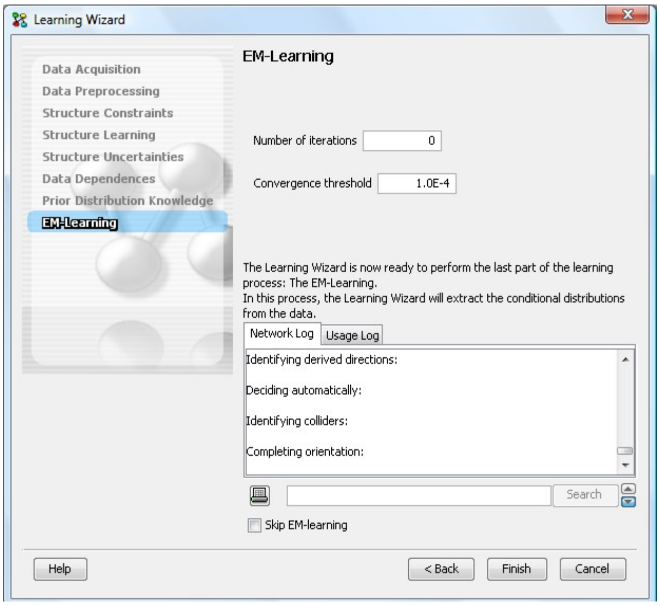



**Fig. 8. Program window at the stage of setting the number of iterations**

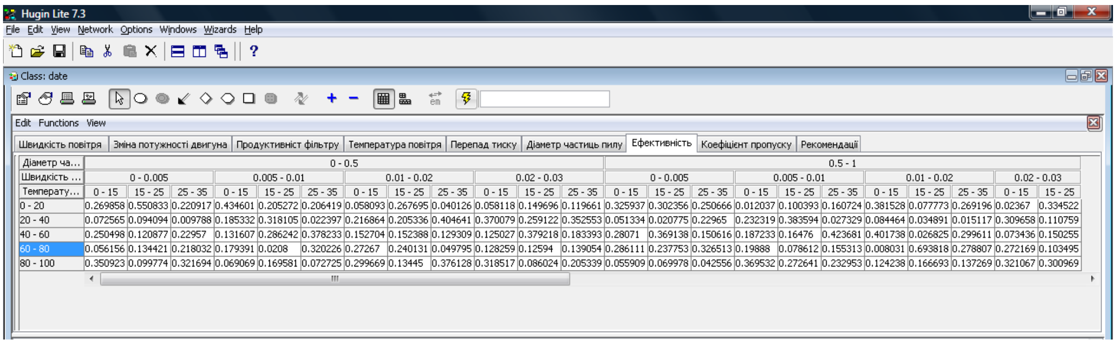



**Fig. 9. Program window with a table of conditional probabilities for variable *E* (efficiency, %)**

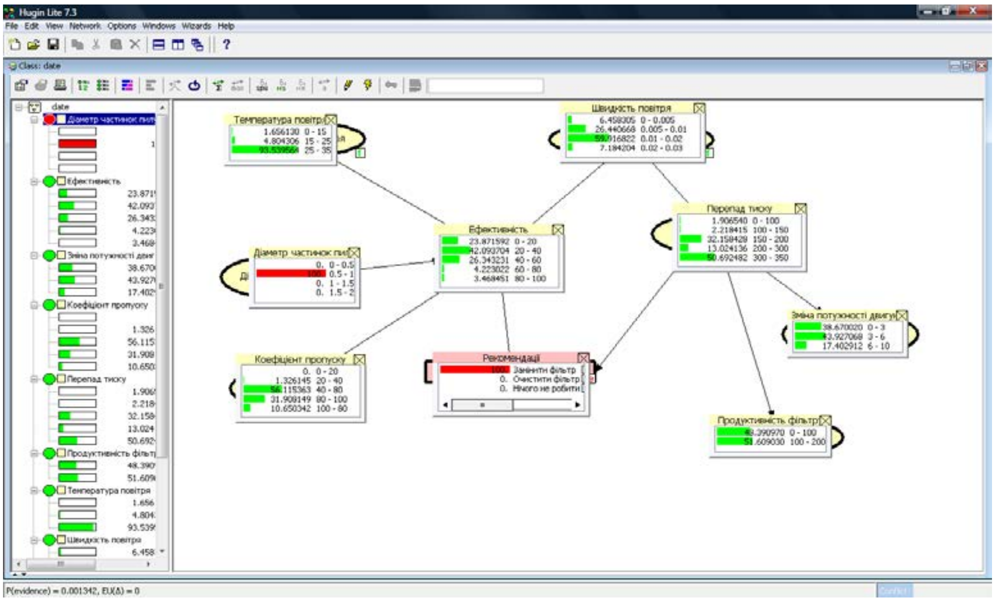



**Fig. 10. View of the program window with the projected and compiled BN**

specify this value for the top and click on the icon to spread the new fact throughout the network.

Once the results have been obtained, HUGIN allows you to generate an HTML document with the results.

Comparing the results obtained using the traditional method, it can be seen that there is still a discrepancy in the results obtained. This discrepancy exists because at the stage of developing the system in the HUGIN, information about the probability of a particular event occurring is unknown, and only statistical data on the process of purifying atmospheric air using filter cartridges accumulated over the last year is known. Using this information, the program generated a probability table for all variables. However, in general, it can be assumed that the developed intelligent system in the form of a BN fully reflects the investigated process.

**Conclusions.** Bayesian probability methods are a significant step forward compared to popular "black box" models. They provide a clear explanation of their conclusions, allow for logical interpretation and modification of the structure of relationships between the variables of the problem, and also allow for explicit consideration of the a priori experience of experts regarding the behavior of the object under study. Thanks to their clear representation in the form of graphs, BN are quite user – friendly. BN are based on the fundamental principles and results of probability theory, which ensures their successful use in practical tasks.  The Bayesian methodology is, in fact, broader than the possibility of using conditional probability operations in directed graphs. It also includes models with symmetric connections (random fields and lattices), models of dynamic processes (Markov chains), as well as a broad class of models with hidden variables that allow solving probabilistic problems of classification, pattern recognition, and prediction. The developed intelligent system using a BN in the HUGIN Lite environment allows evaluating the state of the filter element under specific operating conditions and the efficiency of the entire process as a whole.

**Bibliography:**

1. Бідюк П. І., Коновалюк М. М., Терентьєв О. М. Байєсівські мережі в технологіях інтелектуального аналізу даних. *Штучний інтелект*, 2010. 15 (2). С. 104–113.

2. Згуровський М. З., Бідюк П. І., Терентьєв О. М., Просянкіна– Жарова Т. І. Байєсівські мережі в системах підтримки прийняття рішень. Київ : ТОВ «Видавниче Підприємство «Едельвейс», 2015. 300 с.

3. Салій А. С. Методи навчання структури баєсівської мережі. *Наукові записки НаУКМА. Комп'ютерні науки*, 2021. №4. С. 56–59.

4. Калініна І.О., Гожий О.П. Моделювання і прогнозування нелінійних нестаціонарних процесів на основі Байєсівських структурних часових рядів. *Прикладні аспекти інформаційних технологій*, 2022. Том 5, №3. С. 240–255.

5. Kryvakovska R.V. Application of Bayesian networks for accuracy estimation of modeling results of the air pollution dispersion given inaccurate input data. *System Research and Information Technologies*, 2020. №2. С. 82–87.

6. Lytvynenko V., Voronenko M., Sitalo S., Boskin O., Lurie I., Savina N., Tanasiichuk Y., Krugla N. Using a Bayesian Network to Assess the Atmospheric Pollution Influence on Immunological Parameters. *CEUR Workshop Proceedings*, 2019. С. 1–12.

7. Бугаєва Л. М., Полякова К. М., Статюха Г. О. Моделювання та аналіз процесу очищення атмосферного повітря для газотурбінних установок на патронних фільтрах. *Східно-Європейський журнал передових технологій*. 2010. № 6/6(48). С. 47–51.

**Бугаєва Л.М., Абрамова А.О. РОЗРОБЛЕННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ З ВИКОРИСТАННЯМ БАЙЄСІВСЬКОЇ МЕРЕЖІ ДЛЯ ОЦІНКИ ЕФЕКТИВНОСТІ ФІЛЬТРУВАННЯ ПОВІТРЯ**

*У статті представлено результати дослідження, спрямованого на розроблення інтелектуальної системи у вигляді байєсівської мережі для оцінювання ефективності процесу фільтрації повітря в газотурбінних установках. В умовах інтенсивного розвитку інформаційних технологій виникає потреба у створенні систем, здатних працювати в умовах невизначеності та обмеженості даних. Байєсівські мережі довіри забезпечують поєднання статистичних спостережень із експертними знаннями, що робить їх ефективним інструментом моделювання складних технічних та екологічних процесів. Особливу увагу приділено застосуванню програмного середовища HUGIN Lite, що дозволяє автоматизувати побудову структури мережі, визначити таблиці умовних імовірностей та проведення обчислювальних експериментів. Для моделювання процесу очищення повітря виділено дев'ять ключових факторів, серед яких діаметр частинок пилу, температура та швидкість повітря, перепад тиску на фільтрувальній перегородці, коефіцієнт пропуску пилу та продуктивність фільтра. На основі експериментальних даних побудовано ймовірнісну модель, що відображає причинно-наслідкові зв'язки між цими параметрами та дозволяє робити висновки щодо стану фільтрувального елемента. Результати моделювання підтвердили здатність системи не лише відтворювати відомі закономірності, а й формувати рекомендації щодо експлуатаційних рішень, зокрема необхідності заміни фільтрів. Встановлено, що застосування байєсівської мережі довіри дає змогу враховувати як апріорні знання експертів, так і емпіричні дані, що підвищує точність оцінки стану системи та ефективність управлінських рішень. Запропонований підхід можна розглядати як перспективний напрям для подальшого розвитку інтелектуальних систем моніторингу та прогнозування ефективності фільтраційних процесів у промислових і екологічних застосуваннях.*

***Ключові слова:*** *байєсівська мережа, інтелектуальна система, фільтрація повітря, газотурбінна установка, моделювання, ймовірнісні моделі.*

Дата надходження статті: 27.09.2025
Дата прийняття статті: 14.10.2025
Опубліковано: 16.12.2025