The Mathematical Approach to the Identification of Trouble-Free Functioning of Mining Facilities

Fares Abu-Abed¹,*

¹Tver State Technical University, A. Nikitin Street, 22, 170026, Tver, Russia

Abstract. Modern drilling rigs are complexes of high-tech equipment operated in difficult climatic and technological conditions, characterized by sudden spasmodic changes in the process of drilling a well, which contributes to increased wear of drilling components and increases the likelihood of pre-emergency and emergency situations. Drilling equipment has a wide range of characteristics and technological parameters, the values of which are available during drilling due to the use of modern software and hardware systems for processing geological and technological information. In order to single out the most frequent pre-emergency situations in practice and to preliminarily determine the set of signs necessary for their recognition, a corresponding analysis of the complications arising during well drilling has been carried out.

1 Introduction

The main causes of accidents at industrial facilities include:
• Failures of technological and electrical equipment, automatic control, control and protection devices;
• Non-observance of operating rules and erroneous actions of staff;
• Violations of the technological regime;
• Errors in the design;
• Non-observance of safety regulations when performing work associated with increased danger;
• The occurrence of discharges caused by the accumulation of static electricity.

To solve the problems of ensuring safety and trouble-free operation, it is necessary to carry out a set of theoretical and practical studies, which, first of all, include:
• Further development of the theoretical foundations for solving trouble-free problems, the creation of new models and algorithms for the design of technological processes and their management, taking into account the accident-free requirements; moreover, the models should reflect the causal relationships between the technological processes, changes in the state of equipment operability, the impact of maintenance personnel, environmental influences and other influences leading to an accident;

* Corresponding author: aafares@mail.ru
Development of new diagnostic, identifying and monitoring devices, emergency protection systems based on microprocessor technology, ensuring timely adoption of measures to prevent accidents;

Improving technological processes to eliminate or reduce the number of hazardous stages, operations, modes;

Modernization, improving the quality and reliability of the main technological equipment, failures of which can lead to accidents [1].

Analysis of the development of severe accidents shows that eliminating the causes of accidents (if known) is much easier than preventing their development [2]. In this regard, the creation of methods and means for identifying emergency situations at industrial facilities is relevant.

Previously, issues of identification of emergency situations were considered in the framework of the problem of early recognition of accidents, statistical theory of reliability, problems of technical diagnostics, ergonomics. In particular, in [2], consideration of this issue involves the identification of the so-called “harbingers of the accident”, defined as a violation of normal operation, characterized by the values of a specially calculated rating. In [3], the solution to the problem is considered in terms of technical diagnostics and involves the formation of forecasts of equipment status at certain points in time.

A number of methods and algorithms used to ensure trouble-free production can actually be extended to the problem of identifying pre-emergency situations. However, this task is characterized by certain specificity, as a result of which a number of definitions and provisions require additional refinement.

The definition of a pre-emergency situation given in [4] describes it as a situation at an industrial facility, characterized by the following features:

- Deviation from the nominal operating mode, which under adverse conditions will lead to an accident;
- Each process parameter is still within acceptable ranges, although some may reach limit values;
- In general, a combination of parameter values can characterize the situation at the facility as close to emergency.

Timeliness of returning the process to normal mode allows you to avoid stopping it, which increases the economic efficiency of production by reducing equipment downtime, and timeliness of stopping the process avoids the occurrence of an emergency itself, which provides the following advantages:

- Technological equipment is not subjected to the stresses characteristic of emergency conditions, thus, its repair and maintenance require lower costs, which means the actual physical resource of the equipment increases;
- The number of direct industrial accidents is reduced, which contributes to the reduction of the total annual cost of eliminating the consequences of accidents;
- The emergency identification system can perform the functions of automatic protection in the event of a failure of the latter.

Currently, production safety-assurance systems rely, as a rule, on monitoring process state parameters in an acceptable range. This approach does not allow to take into account situations determined by the combination of permissible values of several parameters. As a result of this, the operation of an industrial facility is characterized by insufficient accident-freeness and is associated with violations of environmental safety. Therefore, for complex industrial facilities, it is necessary to recognize pre-emergency situations, which allows predicting the occurrence of an emergency mode, and then preventing the process from stopping or reducing losses from the occurrence of an accident.

Existing methods and means of identifying and predicting pre-emergency situations do not provide the required effectiveness in solving these problems, therefore, their solution is...
carried out with the mandatory participation of plant personnel. However, the effectiveness of recognition of emergency situations by personnel also depends on the complexity of the installation, the characteristics of the equipment, and the qualifications of the operator and may be insufficient, which leads to accidents at an industrial facility [5,6,7].

In many works aimed at improving the safety of industrial facilities, the problem of identifying emergency situations is considered from the following positions:

• The use of statistical data on the reliability of typical blocks and the construction of transition graphs between their states;
• Use of technical diagnostic methods;
• Recognition of the image of the pre-emergency situation;
• Construction and consideration of a process model in the space of phase coordinates.

The methods of the statistical group can be applied in isolation or in addition to each other, and the methods of qualitative analysis can include quantitative risk criteria (mainly, according to expert estimates using, for example, the matrix “probability - severity of consequences” of hazard ranking).

Methods of technical diagnostics [8,9] involve constant monitoring of the condition of the equipment, its diagnosis and prediction of the condition.

In the process of diagnosing the technical condition of an object, it is first necessary to solve the problem of studying the physical properties of the object, as well as constructing mathematical models of malfunctions.

Identification of pre-emergency situations using pattern recognition methods [10, 11] relies on various mathematical methods used in the field of artificial intelligence.

The issue of observing and identifying the phase coordinate vector of the process is currently developed only under the assumption that the object is in a state of normal functioning, and the measurement errors of the input and output are small [12]. In the analysis of trouble-free, when it is necessary to take into account the change in the state variables of the object at arbitrary points in time, adaptive observation and identification methods are used. In this case, both the values of the phase coordinate vector and the model parameters are determined.

In the case of a high level of interference at the input or output of the object, the value of the phase coordinate vector at the current time (with the known model of the object and the probabilistic characteristics of the random component of the input and the error in measuring the output) is estimated by optimal filtering methods [12].

Since the actions of the operator of production facilities are largely based on the principles of intelligent control and, in particular, the rules of fuzzy logic, it is necessary to consider in more detail the options for using artificial intelligence methods for identifying pre-emergency situations.

2 Method used

The scientific literature widely presents the results of building systems and means for identifying accidents and emergency situations.

In [13], a number of systems are described, such as PILOTE, RTIME, PICON, etc. The special value of this work is determined by the fact that it considers a process control system using fuzzy logic elements.

In [14], emergency situations of the technological process are identified by using an expert system based on a frame-production two-level model.

The work [15] describes an approach to building expert systems for diagnosing electrical equipment using a knowledge base of a hierarchical structure, a mechanism for accumulating statistics during work, and a relational database for storing knowledge. The described system is self-learning.
In [16], plant safety management using new information technologies and computer decision-making tools is considered. The above-described methods for identifying pre-emergency conditions are characterized by a number of disadvantages.

Statistical methods of hazard analysis and methods of technical diagnostics, although they are necessary [17], but allow predicting emergency situations only from the number of previously occurred emergency and pre-emergency situations. This disadvantage can be explained by the existence of improbable impacts on the system, incorrectly introduced controls, etc. To eliminate such situations, it is necessary to supplement the security management system with a subsystem that allows real-time timely identification of the occurrence of an emergency situation, identifies the cause of its occurrence, determines the set of acceptable and unacceptable scenarios for resolving the situation, fixes the violations in the cards of the corresponding units, etc.

When using the methods of technical diagnostics, there is no assessment of the general situation at the facility. In addition, the determination of the states of all diagnostic elements is time-consuming and equipment-intensive. It should also be noted the fact that technical diagnostics relies mainly on the study of the physical characteristics of the diagnostic object. In some cases, these characteristics are characterized by a significant time drift, which reduces the applicability of this approach to identifying pre-emergency situations.

The methods associated with pattern recognition are characterized by the following disadvantages:
• Lack of a single concept for pattern recognition;
• Orientation of many methods to a specific task.

A mathematical study in the space of phase coordinates is feasible only for cases with a small number of process state variables. An increase in the number of variables or especially pronounced process features require the development of a special system for identifying emergency situations for each process.

The identification of pre-emergency situations using the knowledge of experts and the apparatus of fuzzy sets allows you to save knowledge about the identifier of the pre-emergency situations in the form of production rules, with which it is possible to automatically identify the pre-emergency situations during the process. However, this method does not provide for the formation of quantitative assessments, with the help of which it would be possible to accurately assess emerging pre-emergency situations. At the same time, such assessments could reduce the probable damage from an incorrect decision by the operator in the event of pre-emergency situations. A comparative analysis of pre-emergency situations identification methods showed that there is no universal method that can be used to solve any class of pre-emergency situations identification problems. With regard to solving the problem of identifying pre-emergency situations in drilling rigs during well drilling, a promising direction is the use of pattern recognition methods.

In a general sense, recognition is the task of converting input information, which is appropriate to consider some parameters into output, which is a conclusion about which class the recognizable image (object) belongs to [2]. The simplified structure of the recognition system is shown in Fig. 1.

![Fig. 1. Simplified recognition system structure.](image-url)
The signs of recognizable objects are used as input information, and they try to make the number of signs is minimal, and the information contained in them is sufficient to obtain a result with high reliability.

The result of the classifier is a decision on whether the object belongs to one of the classes. It is very important that the classification accuracy is as large as possible, and the time to achieve the result as short as possible.

Classifier development stages.

The classification process can be divided into two stages [18]:
1. The stage of development of the classifier;
2. The stage of determining the state of an object using this classifier - stage of functioning.

The stage of development of the classifier is schematically presented in the form of two blocks (Fig. 2) [19]

![Classifier development phase.](image)

At the input of the source information processing unit, there are many signs of the object, \( \overline{P} = \{ p_1, p_2, \ldots, p_m \} \), \( |P| = m \) which may be, for example, sensor readings, trends, analysis results, etc. At the output, we have a sign vector \( \overline{X} = \{ x_1, x_2, \ldots, x_n \} \), \( |X| = n \), and \( \overline{X} \subseteq \overline{P}, n \leq m \).

In addition, the number of features of the object can be quite large, therefore, their processing will require large expenditures of computer time, which will certainly affect the efficiency of the classifier. Therefore, a software tool that implements the functions of this unit should also correctly reduce the dimension of the measurement space from several thousand to hundreds or even tens of signs. Consequently, this block is entrusted with the functions, firstly, of determining the informational significance of signs, secondly, of identifying the interdependence of signs, and thirdly, of discarding insignificant signs. To do this correctly is not so simple due to the fact that all the signs are interconnected, and the value of some can change very significantly when discarding other, even non-essential, signs.

Thus, the main tasks of the source information processing unit are:
1. Determination of the informational significance of the signs.
2. Identification of the interdependence of signs.
3. Correct decrease in the dimension of the attribute space.

The so-called feature vector of the object \( \overline{X} = \{ x_1, x_2, \ldots, x_n \} \), \( |X| = n \), which contains only those features without which the organization of the recognition process with the required accuracy is impossible, is received at the input of the classifier construction block. The output of this block should be a decision on assigning the recognizable object to a certain class, or, in other words, the class of the recognizable object \( K_j, j = \frac{1}{l} \).

Thus, the main task of the classifier building block is to develop a structure that allows you to determine the state of an object by its characteristics \( \overline{X} = \{ x_1, x_2, \ldots, x_n \} \). This
structure can be a set of possible states of an object in the form of classes, separating functions, criteria, etc. depending on the availability of a priori information, types of features used and development methods. It is important to note that all deterministic features applied to the input of this block must be formalized.

Thus, the main functions of the first stage are:
1. Input of initial information;
2. Processing of source information;
3. Formalization of signs;
4. Development of a classification algorithm;
5. Determination of the possible states of the object (classes, criteria, etc.);
6. Analysis of the quality of the classifier.

The stage of determining the state of an object using the classifier is schematically presented in the form of two blocks (Fig. 3).

Fig. 3. The stage of determining the state of the object.

Selection of informative signs.

We assume that objects are described by a set of features \( F = \{f_1, \ldots, f_n\} \). Each attribute \( f_j \) is a mapping from \( X \) to some set \( D_j \) of admissible values of the attribute, in the general case it is not necessarily numerical.

Vector \( f_1(x), \ldots, f_n(x) \in D_1 \times \ldots \times D_n \) is called the characteristic description of the object \( x \).

Let \( G \subseteq F \) be an arbitrary subset of features. We will denote by \( \mu G \) the training method that constructs algorithms using only features from the subset \( G \). We assume that the \( \mu G \) method selects an algorithm from the algorithm model \( A(G) \) that uses only features from \( G \).

The number of features used \( |G| \) will be called the complexity of the model \( A(G) \).

Consider the most common ways of selecting features that are compatible with arbitrary training methods and with arbitrary quality functionals, and therefore equally applicable to various classification algorithms.

In all methods, the lower envelope of the set of points \( (|G|, Q_{\text{opt}}(G)) \) is constructed. In the case of complete enumeration, its minimum corresponds to the optimal set \( G^* \) of optimal complexity \( j^* = |G^*| \). Other algorithms solve the problem of finding the optimal set of features only approximately.

3 Results and Discussion

Definition of a set of signs for recognition of emergency situations

To determine the reliability of the relationship between the two features \((X, Y)\) we will use the Pearson parametric correlation coefficient. The value of this correlation indicator is determined by the formula:

\[
 r^\rho_{xy} = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}}
\]
The number of features used and development methods. It is important to note that all deterministic functions, criteria, etc. depending on the availability of a priori information, types of structure can be a set of possible states of an object in the form of classes, separating complexity features only approximately.

To determine the reliability of the relationship between the two features (X, Y) we will use the Pearson parametric correlation coefficient. The value of this correlation indicator is determined by the formula:

\[
\rho = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum (x_i - \mu_x)^2 \sum (y_i - \mu_y)^2}}
\]

where: \( x_i \) - statistics of the characteristic \( x \), \( y_i \) - statistical data of the characteristic \( y \).

The list of basic parameters (features) measured during the operation of drilling rigs is set in the technical standards for their operation and is shown in Table 1.

To assess the mutual influence of features, the correlation coefficients were calculated for all features available in the database of geological and technological information.

Table 1. shows the values of the correlation coefficients of the signs used to recognize pre-emergency situations.

| No | Sign | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|----|------|----|----|----|----|----|----|----|----|----|----|
| 1  | S100 | -  | 0.11| -  | -0.03| 0.01| -0.02| -0.11| 0.03| 0.12| -0.05| 0.67 |
| 2  | S101 | 0.11| -  | 0.02| 0.09| -0.04| 0.06| 0.01| -0.04| 0.03| 0.45|    |
| 3  | S102 | -0.03| 0.02| -  | 0.35| 0.09| 0.04| -0.03| -0.01| 0.02| 0.05|    |
| 4  | S103 | 0.01| 0.09| 0.35| -  | 0.12| -0.05| 0.01| 0.02| -0.03| 0.02|    |
| 5  | S105 | -0.02| -0.04| 0.09| 0.12| -  | 0.03| -0.04| -0.06| 0.11| -0.03|    |
| 6  | S106 | -0.11| 0.06| 0.04| -0.05| 0.03| -  | 0.05| 0.03| 1.00| -0.05|    |
| 7  | S111 | 0.03| 0.01| -0.03| 0.01| -0.04| 0.05| -  | 0.86| 0.02| -0.03|    |
| 8  | S112 | 0.12| -0.04| -0.01| 0.02| -0.06| 0.03| 0.86| -  | 0.03| -0.05|    |
| 9  | S116 | -0.05| 0.03| 0.02| -0.03| 0.11| 1.00| 0.02| -0.01| -  | 0.03| 0.04 |
| 10 | S120 | 0.67| 0.45| 0.05| 0.02| -0.03| -0.05| -0.03| -0.05| -0.02| -  |    |
| 11 | S300 | 0.03| -0.02| 0.04| 0.02| 0.05| -0.04| 0.08| 0.04| 0.01| 0.03|    |
| 12 | S600 | 0.01| 0.05| 0.08| -0.03| 0.08| 0.05| -0.05| -0.01| 0.01| 0.04| -0.01|
| 13 | S605 | -0.02| 0.11| -0.06| 0.06| -0.04| 0.03| 0.03| 0.02| 0.01| 0.06|    |
| 14 | S719 | -0.04| -0.06| 0.04| 0.03| 0.06| 0.02| 0.04| 0.27| 0.02| 0.02|    |
| 15 | S720 | 0.09| 0.01| -0.13| -0.01| 0.03| -0.04| -0.01| 0.01| 0.026| -0.05|    |
| 16 | S900 | 0.13| -0.04| -0.07| 0.08| 0.09| -0.01| 0.06| 0.02| 0.012| -0.02|    |
| 17 | S910 | 0.04| 0.02| 0.05| 0.04| -0.08| -0.07| -0.03| 0.024| 0.019| 0.021|    |
| 18 | S1001| -0.02| -0.08| 0.10| -0.03| 0.04| 0.08| 0.04| 0.03| 0.003| 0.05|    |
| 19 | S1002| -0.03| 0.07| 0.06| 0.07| 0.01| 0.02| 0.01| 0.013| 0.009| 0.01|    |
| 20 | S1005| -0.02| 0.01| 0.05| 0.01| 0.02| -0.03| 0.04| -0.04| 0.01| 0.055|    |
| 21 | S1300| 0.05| 0.03| -0.04| 0.02| -0.07| 0.05| 0.02| 0.15| 0.02| 0.21|    |

Table 1. The values of the correlation coefficients of signs

To assess the mutual influence of features, the correlation coefficients were calculated for all features available in the database of geological and technological information.

Table 1. shows the values of the correlation coefficients of the signs used to recognize pre-emergency situations.
Selection of features for pre-emergency situations of a drilling rig using the branch and bound method

The selection of signs for recognition of emergency situations was carried out on the basis of the use as an internal criterion of the coefficient of correlation of signs with the value of the target attribute (identifier of the emergency situation).

To assess the relationship of signs, Pearson's parametric correlation coefficient was used.

The main types of emergency situations are given in [20]. For the study, real data of geological technology information containing information on the occurrence of the following pre-emergency situations were selected:
- Gas and oil occurrences;
- Absorption of drilling or grouting fluid;
- Grips of the drilling tool;
- Destruction of the walls of the well.

These situations are quite common in the practice of drilling coal wells, and according to the data given in [21], they make up from 50 to 60 percent of all the pre-emergencies that occur on the wells during the period of drilling.

The values of the correlation coefficient of signs with the value of the target sign for the types of pre-emergency situations under consideration are given in Table 2.

| No | Sign | Gas and oil water manifestations | Destruction of the walls of the well | Mud or cement slurry absorption | Drill tool grips |
|----|------|---------------------------------|--------------------------------------|---------------------------------|-----------------|
| 1  | S100 | -0.02                           | 0.04                                 | 0.09                            | 0.04            |
| 2  | S101 | 0.03                            | -0.03                                | -0.07                           | 0.03            |
| 3  | S102 | 0.04                            | 0.01                                 | 0.01                            | -0.04           |
| 4  | S103 | -0.03                           | 0.02                                 | -0.02                           | -0.05           |
| 5  | S105 | 0.11                            | 0.04                                 | 0.05                            | 0.01            |
| 6  | S106 | -0.06                           | 0.07                                 | 0.06                            | -0.03           |
| 7  | S111 | 0.04                            | -0.03                                | -0.02                           | 0.04            |
| 8  | S112 | 0.72                            | 0.46                                 | 0.04                            | 0.55            |
| 9  | S116 | 0.08                            | -0.06                                | -0.05                           | 0.11            |
| 10 | S200 | -0.12                           | 0.07                                 | 0.03                            | 0.66            |
| 11 | S300 | 0.81                            | 0.61                                 | 0.64                            | 0.44            |
| 12 | S600 | 0.52                            | 0.09                                 | 0.03                            | 0.05            |
| 13 | S605 | -0.02                           | 0.44                                 | -0.01                           | 0.03            |
| 14 | S719 | -0.06                           | -0.06                                | 0.02                            | 0.06            |
| 15 | S720 | -0.01                           | 0.34                                 | 0.54                            | -0.12           |
| 16 | S900 | 0.32                            | -0.12                                | 0.43                            | -0.02           |
| 17 | S910 | 0.05                            | 0.11                                 | 0.03                            | -0.07           |
| 18 | S1001| 0.12                            | -0.09                                | 0.37                            | 0.01            |
| 19 | S1002| 0.04                            | -0.05                                | 0.04                            | -0.08           |
| 20 | S1005| -0.03                           | 0.06                                 | 0.02                            | 0.02            |
| 21 | S1300| 0.11                            | 0.11                                 | 0.13                            | 0.53            |

The set of attributes for each situation was determined by the branch and bound method.

To increase the efficiency of the search, the algorithm added the ability to limit the width of the search, which is characteristic of the width search algorithm.

To do this, at each step of adding features to the current set, no more than $B$ of previously not selected features are selected that have the highest correlation coefficients with the target feature.
The criterion for the selection of signs is the expression:

\[ Q_N = \sum_{i \in T_N} K_i^C - \sum_{i \in T_N} \sum_{j \notin T_N} |K_{ij}| \]  \hspace{1cm} (2)

where: \( Q_N \) - is the value of the criterion for choosing a feature in step \( N \); 
\( T_N \) - is the set of already selected features in step \( N \); 
\( K_i^C \) - is the correlation coefficient of attribute \( i \) with the target attribute; 
\( K_{ij} \) - is the value of the correlation coefficient of attribute \( i \) with attribute \( j \).

The results of the algorithm for the pre-emergency situation "gas and oil occurrences" are given in [20]

4 Conclusions

The results allow us to conclude that it is advisable to continue the study of the search parameters of the presented algorithm with the aim of further improvement:
1. The general structure of the recognition system, refinement of the identified main stages of the development of the classifier.
2. Definitions of an a priori vocabulary of features used for recognition of emergency situations during the operation of mining machines and installations.

References

1. F. Abu-Abed, N. Borisov, E3S Web Conf., 21, 01019 (2017)
2. R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification* (Wiley, New York, 2001)
3. S. I. Gallant, *Neural Network Learning and Expert Systems* (MIT Press, Boston, 1993)
4. Rotary Foundation Drilling Rig Safety (OAFS, Paris, 2016)
5. F. N. Abu Abed, Bulletin of Tula State University. Earth sciences, 1, 121-131(2018)
6. F. N. Abu-Abed, A. R. Khabarov, Dilemas contemporáneos: Educación, Política y Valores, 6:S1:P, 67 (2019)
7. European Standards References: En 16228-1 Drilling and Foundation Equipment-Safety (ESA, Brussels, 2016)
8. Prudent Development: Realizing the Potential of North America’s Abundant Natural Gas and Oil Resources (OPC, Oslo, 2016)
9. Predictive asset analytics to improve mining operations (Schneider Electric, Berlin, 2015)
10. M. D. Hammer, V. N. Syzrantsev, S. L. Golofast, *Simulators Based on Software and Hardware Platform for Technical Education* (Oil and Gas University, Novosibirsk, 2011)
11. I. P. Pustovoitenko, *Prevention and Methods for Eliminating Accidents and Complications in Drilling* (Nedra, Moscow, 1987)
12. D. E. Rumelhart, G. E. Hinton, R. J. Williams, *Learning Internal Representations by Error Propagation. Parallel Distributed Processing* (MIT Press, Massachusetts, 1986)
13. V. Gel'Man, O. Marganiya, *Resource Curse and Post-Soviet Eurasia: Oil, Gas, and Modernization* (Drilling Pub., Trenton, 2010)
14. F. Abu-Abed, A. Khabarov, E3S Web Conf., 41, 01028 (2018)
15. Drilling Technologies (Universita Campus Bio-Medico, Roma, 2017)
16. F. Abu-Abed, E3S Web Conf., 41, 01025 (2018)
17. F. N. Abu-Abed, L. G. Naumova, J. Min. Geot. Eng., 3:6, 21-35 (2019)
18. F. Abu-Abed E3S Web Conf., 105, 03006 (2019)
19. M. Alrumah, T. Ertekin, J. Pet. Gas Eng., 10:2, 14-22 (2019)
20. F. Abu-Abed, A. Khabarov, JFAS, 9:2S, 1171-1181 (2017)
21. F. Abu-Abed, E3S Web Conf., 105, 03017 (2019)