Use of sensor networking technology to build a power transformer monitoring system

A Smerdin1,*, G Ermachkov1, V Nezevak1, O Sidorov1, and A Golubkov1

1Omsk State Transport University, 644046, Marx av., 35, Omsk, Russia

Abstract. The paper discusses the technology of designing a monitoring system for power transformers and power supply devices, based on the use of sensor networks, describes the creation of a Bayesian network to predict the probability of failure. Conditions for achieving the maximum diagnostic efficiency with the required accuracy are formulated. A diagnostic system is proposed that allows assessing the influence of each parameter on the accuracy of predicting failures. Using the method of expert assessments, the possibility to calculate the probabilities of failures with a lack of information is shown.

1 Introduction

From 2011 to 2019, electricity consumption in transformer systems is growing at an outstripping pace. Most of the electricity is processed by traction and transformer substations and fed to the electric traction network of electric railway transport. For example, in the Russian Federation, more than 80 billion kWh per year are spent on train traction, or 8% of the annual consumption of the entire country [1]. A similar situation is observed in the energy systems of other countries with a developed railway network, such as China, Japan, and the EEC countries.

From 2011 to 2019, electricity consumption in transformer systems is growing at an outstripping pace. Most of the electricity is processed by traction and transformer substations and fed to the electric traction network of electric railway transport. For example, in the Russian Federation, more than 80 billion kWh per year are spent on train traction, or 8% of the annual consumption of the entire country [1]. A similar situation is observed in the energy systems of other countries with a developed railway network, such as China, Japan, and the EEC countries.

Modern approaches to monitoring the state of power transformers are based on measurement processing methods that allow monitoring the dynamics of change and target values of parameters characterizing the state of the active part of the transformer oil, switches in the tank of the voltage control device under load (CUL), transformer oil, high-voltage inputs and other auxiliary equipment.

* Corresponding author: alexandr.smerdin@omgups.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
Currently ongoing research in the field of monitoring the state of power transformers is associated with determining the optimal number of diagnosed parameters, choosing the structure of the diagnostic system, sensors, system of collecting, processing, analyzing information from installed sensors and choosing a mathematical apparatus for predicting the state of power equipment.

2 State of the issue

In most works, the main diagnostic parameters processed in real time are: load, temperature, gas concentration, partial discharges, vibration and acoustic signals. The development of “digital substations” has led to an interest in processing diagnostic data of power equipment using systems with artificial intelligence [2]. One of the directions of obtaining a forecast of the power equipment state is the development of probabilistic models based on measurements of the load and temperature of the equipment [3]. A general trend in the development of condition monitoring systems is the increasingly widespread use and analysis of equipment diagnostic data, which can be obtained online [4, 5]. One of the most promising methods is the acoustic method for diagnosing the state [6, 7], the method for determining partial discharges [8]. Since the power equipment contains auxiliary systems, it is also desirable to determine their state online [9]. The difference in monitoring systems has given rise to the problem of choosing the complexity of systems that correspond to the cost of power equipment [10], the choice of diagnostic methods [11].

3 Formulation of the problem

The development of the “digital substation” project allows us to take a fresh look at the structure of the system for monitoring the state of power transformers of traction substations. Existing monitoring systems do not provide for online analysis and forecasting of the state of the main equipment. Diagnostic results are obtained on the basis of measurements and tests performed during maintenance by operating personnel (inspection, current repair, overhaul tests, thermal imaging diagnostics, transformer oil sampling). This system has a number of disadvantages associated with the rate of detection of developing defects, since the rate of change is fixed in accordance with the measurement and test schedules, the typical frequency of which is from six months to several years.

Modern monitoring systems make it possible to reduce the frequency of measurements of diagnostic parameters and, by determining the rate of change of parameters, to identify developing defects at an early state.

The rate of change of diagnostic parameters in general is determined by the formula, %:

\[ V_i = \frac{P_{\text{cur}i} - P_{\text{prev}i}}{P_{\text{cur}i} \cdot \Delta t}, \]  

where \( P_{\text{cur}i} \), \( P_{\text{prev}i} \) are diagnostic parameters obtained during measurements or tests in the current or previous periods, respectively;

\( \Delta t \) is the time period separating the current and previous periods.

The state of power transformers is determined using a system of diagnosed parameters that depend on influencing factors. Control over changes in the latter allows predicting the state of transformers (Fig. 1).
Thus, in relation to railway traction substations, the construction of a system for monitoring the state of power transformers must satisfy the following conditions:
- measurement of diagnostic parameters should be carried out online;
- the analysis of the diagnostic parameters must be carried out taking into account the test results of the power equipment obtained during the scheduled maintenance;
- the forecast of the state should be based on the results of measurements, their comparison with the boundary values and the rate of change together;
- the results of monitoring should be presented in a form that meets the requirements of the "digital substation" project.

4 Formation of a monitoring system

Depending on the design options, monitoring systems allow measuring and analyzing data using the following sensors: vibration; pressure; current; voltage; humidity; temperature; gas analyzers, etc.

The influence of factors on the diagnosed parameters allows, when organizing systems for monitoring the state of power equipment with artificial intelligence, to consider sets of parameters and their rate of change to determine and predict the state.

Setting up an algorithm for detecting developing defects is based on controlling the rate of change of the diagnosed parameters and their compliance with influencing factors. An example is the change in temperature, vibration displacement and gas concentration when considered together with a change in the load of a power transformer (Fig. 2).
Fig. 2. Typical changes in the diagnosed parameters: a) "current – gas concentration – time"; b) "current – vibration displacement – time"; c) "current – temperature – time"; d) "temperature – gas concentration – time".

Consideration of the measurement results in the coordinates "current – gas concentration – time" makes it possible to determine the discrepancy between the increase in concentration and the decrease in the load level, which may indicate the occurrence of partial discharges, turn-to-turn closures and an increase in the resistance of contact connections (Fig. 2, a). A similar consideration in the coordinates "current – vibration displacement – time" makes it possible to reveal the moment of vibration growth with a decrease in the load (Fig. 2, b). An increase in temperature with a decrease in the load may indicate a malfunction in the cooling system, a hot spot or the appearance of discharges (Fig. 2, c), and an increase in the gas concentration with increasing temperature may indicate discharges or overheating (Fig. 2, d).

To obtain quantitative and qualitative relationships between influencing factors and current collection rates, the authors compiled a computer network based on the theorem about Bayesian posteriori probabilities.

For the correct operation of such a network, probabilistic relationships between factors are required, as well as an array of data on the change in the values of these factors over time.

Currently, correlation multivariate analysis, the method of expert estimates and laboratory experiments are used to detect relationships. Regression models, radial-basis neural networks, Bayesian networks, etc. [12-15] can be used as a mathematical apparatus.

Research in the application of Bayesian networks for predicting failures of railway infrastructure [16] and power line devices [17] have shown the promise of this method in diagnostics. In these models, various combinations of weather factors, failure statistics, and operational features are used as input parameters. The results of the model operation are most
often forecasts or probabilities of failure of individual elements or subsystems in the short and long term.

Since the results of measurements of diagnostic parameters are random values, and failures of technical objects are random events, the Bayesian network is a suitable type of mathematical model for predicting failures.

Bayesian network is a graphical probabilistic model, where each node is a random variable, and the connections between the nodes show the dependencies between them.

The main reasons for choosing Bayesian networks for modeling failures are: the possibility of a simple causal interpretation of the network structure and an explicit representation of influencing factors, as well as the ability to work with random events. Bayesian networks are widely used in the diagnosis of failures and prediction of damage in technical systems [18].

The initial data for constructing the proposed probabilistic model for predicting failures based on a Bayesian network is a matrix of variable values obtained during the interpretation of previous observations, changes in parameters and operation indicators of transformers. The matrix contains various types of data: integers and real numbers, named values. Random variables in a Bayesian network can have different types and distribution laws.

To form the input matrix of the model, the authors use an array of initial data, including the results of diagnostics, cases of failures. Some of the necessary information is available upon request from existing databases on equipment failures and maintenance, while the other part cannot be obtained without additional costs.

The work used the results of studying the statistics of transformer failures taking into account weather factors. It has been found that there is a relationship between different combinations of weather conditions and the number of transformer failures. We used a database of transformer failures for three years (from 2016 to 2019).

When building a Bayesian network using regression analysis, the most significant factors affecting the reliability of transformer operation were identified.

Comparison of weather factors with failures was carried out with reference to time and coordinates on the ground.

The relational table obtained as a result of summarizing all the available data (Table of the initial data "T") was used to determine the structure of the network.

The assessment of the statistical significance of each possible relationship between factors and the determination of the network structure was carried out using the Hill-Climbing method [19].

The graphic structure of the developed Bayesian network is an acyclic directed graph $G = (V, A)$, where $V$ is the set of graph vertices and $A$ is the set of edges. The graph determines the factorization of the joint probability distribution $V = \{X_1, X_2, \ldots, X_v\}$ of the events taken into account by the model. The form of factorization was determined in accordance with the Markov property of the Bayesian network, which states that the random variable $X_i$ depends only on its direct parent variables:

$$P(X_1, \ldots, X_v) = \prod_{i=1}^{v} P(X_i | \Pi_{X_i}),$$

The first step of the created algorithm is to determine the network structure, after which local probability distributions are calculated based on the obtained structure.

The structure determination algorithm is based on the evaluation method, in which each variant of the network structure is assigned a certain score, after which heuristic optimization is performed in order to obtain the structure with the maximum score.

As an evaluation function for the network structure, various criteria were applied (logarithmic similarity function, information criteria of Akaike and Bayes, logarithm of the
equivalent evaluation of Bayes-Dirichlet) [20]. The best results were obtained using the logarithmic similarity function, which was chosen as the main one:

$$L(G|T) = \sum_{i=1}^{v} \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} N_{ijk} \log \left( \frac{N_{ijk}}{N_{ij}} \right),$$  \tag{3}$$

where $N_{ijk}$ is the number of cases in the table of initial data "T", in which the variable $X_i$ takes its $k$-th value $x_{ik}$ and the variables $\Pi_{X_i}$ take their $j$ configuration $w_{ij}$; $N_{ij}$ is the number of cases in the table "T" in which the variable $X_i$ takes its $k$-th value $x_{ik}$; $r_i$ is the number of possible states of the variable $X_i$; $q_i$ is the number of possible configurations of the parent variables $\Pi_{X_i}$ for the $X_i$ variable.

This evaluation function tends to its maximum for a fully connected network, which negatively affects the display of links for possible independent variables, therefore, for cases with a large number of links in the network structure, the Bayes-Dirichlet evaluation function was used:

$$L(G|T) = \log(P(G)) + \sum_{i=1}^{v} \sum_{j=1}^{q_i} \left( \log \left( \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \right) + \sum_{k=1}^{r_i} \log(N_{ijk}) \right).$$  \tag{4}$$

When adding new initial data, the network structure can change, so the algorithm provides for updating the structure every time the table "T" changes.

A probabilistic model for predicting the operability of transformers based on a Bayesian network for the electric traction network of the West Siberian region has been created at the Omsk State University of Railways.

The sensor network technology is designed to solve the widest range of industrial monitoring and control tasks and has the following undeniable advantages over other existing wireless and wired systems:

- the ability to install sensors on an existing and operating facility without additional work on laying a wired network;
- low cost of a separate control element;
- low cost of installation, commissioning and maintenance of the system;
- minimum restrictions on the placement of wireless devices;
- high resiliency of the sensor network as a whole.

The database accumulated by information from sensors has an object-oriented structure and is based on a post-relational model, which makes it possible to bind by key fields, provided that the identifier values match.

To obtain the initial data on the power transformers, devices are placed that form a sensor network by sensors: atmospheric, internal (current, voltage, temperature, gas concentration, etc.), external (vibration).

Tables 1 – 3 describe the structures of the database tables for storing data received from diagnostic devices.

Registration of atmospheric factors is carried out by the device for monitoring the parameters of the air by Meteometer MES-200A (Table 1).

**Table 1. Structure of the data table of atmospheric factors.**

| Description                  | Unit of measurements | Variable name |
|------------------------------|----------------------|---------------|
| Weather station number       | –                    | –             |
| Date of registration of observations | dd/mm/yy            | data         |
Average temperature | °C | temp_mean
Maximum temperature | °C | temp_max
Minimum temperature | °C | temp_min
Air humidity indicator (average dew point temperature) | °C | dewp_mean
Precipitation | mm | rain
Snow cover thickness | mm | snow
Atmospheric visibility | mm | visible
Average atmospheric pressure | mmHg | pressure
Average wind speed | m/s | wind_mean
Maximum wind speed | m/s | wind_max
Maximum speed of wind gusts | m/s | wind_gust
Weather phenomena (fog, rain, snow or freezing rain, hail, thunderstorm, tornado) | – | weather_event

Load measurement for monitoring systems in real time at traction substations is carried out by connecting the measuring devices of the system to the current circuits of the measuring transformers.

To measure the temperature of the upper layers of oil at traction substations, various brands of manometric condensation temperature alarms are used to ensure the automatic operation of the cooling system and alarm. The specified devices are not intended for transmitting measurement data, they are used when inspecting transformers and in control circuits of the cooling and alarm system. To ensure the operation of the predictor module, it is necessary to use additional temperature sensors, which make it possible to measure the temperature of the most heated point by indirect or direct methods. The outputs of the temperature sensors are connected to the input of the predictor module.

The existing methods for monitoring the gas concentration dissolved in oil can be divided into three groups: analysis of oil samples by a stationary laboratory at a frequency set by the maintenance system; analysis of oil or gas samples from the Buchholz relay at the location of the power transformer; on-site gas concentration analysis with remote transmission of measurement results.

Diagnostic indicators associated with sampling from the transformer tank are determined periodically at intervals of 10-90 days.

**Table 2.** Structure of the data table of internal diagnostic indicators.

| Description                  | Unit of measurements | Variable name |
|------------------------------|----------------------|---------------|
| Phase current from HV side   | A                    | current_hv    |
| Phase current from MV side   | A                    | current_mv    |
| Phase current from LV side   | A                    | current_lv    |
| HV voltage                   | B                    | volt_hv       |
| MV voltage                   | B                    | volt_mv       |
| LV voltage                   | B                    | volt_lv       |
To measure vibration, a portable vibration analyzer is used in the mode of measuring vibration accelerations, vibration velocities, or root-mean-square values of vibration displacements (Table 3). Indicators of mechanical vibrations are recorded automatically in the form of 30 second files with a periodicity of 24 times a day.

Table 3. The structure of the data table of external diagnostic indicators.

| Description                                      | Unit of measurements | Variable name    |
|--------------------------------------------------|----------------------|------------------|
| Oil temperature                                  | °C                   | oil_temp         |
| Winding temperature                              | °C                   | winding_temp     |
| Insulation resistance of windings                | MOm                  | insulation_resist|
| Breakdown voltage of oil                         | kV                   | oil_break_volt   |
| Acid number (KOH) of oil                         | mg/g                 | koh_oil          |
| Flash point of oil                               | °C                   | temp_flash_oil   |
| Oil moisture content                             | %                    | oil_moisture_cont|
| Mechanical impurities in oil                     | %                    | impurities_oil   |
| Oil dielectric loss tangent                      | degree (%)           | tangent_oil      |
| Content of water-soluble acids and alkalis       | %                    | acids_alkalis_cont|
| Content of antioxidant additive                  | %                    | antioxid_cont    |
| Pour point                                       | °C                   | freez_temp       |
| Gas content                                      | %                    | gas_cont         |
| Soluble sludge content                           | %                    | sludge_cont      |
| Content of furan derivatives                     | %                    | furan_cont       |
| Sulfur content                                   | %                    | sulf_cont        |
| Viscosity                                        | мм²/с                | viscosity        |
| Tangent of dielectric loss angle of winding insulation | degree (%)         | tangent_isolation|
| DC resistance of windings                        | mΩ                   | dc_resis_wind    |
| On-load tap-changer position                     | №                    | rpn_position     |
| CO2 content                                      | ppm                  | co2              |
| CH4 content                                      | ppm                  | c4               |
| C2H6 content                                     | ppm                  | c2h6             |
| C2H4 content                                     | ppm                  | c2h4             |
| C2H2 content                                     | ppm                  | c2h2             |
| CO content                                       | ppm                  | co               |
| H2 content                                       | ppm                  | h2               |
| O2 content                                       | ppm                  | o2               |
| N2 content                                       | ppm                  | n2               |
| Re-analysis                                      | -                    | repeat_test      |
| Vibration displacement | mm | vib_displ |
|------------------------|----|----------|
| Vibration velocity     | mm/s | vib_veloc |
| Vibration acceleration | mm²/s | vib_accel |

Figure 3 shows the network topology automatically generated using the developed algorithm and software.

The degree of importance of factors (line color)

- Atmospheric factors
- Internal diagnostic indicators
- External diagnostic indicators

**Fig. 3.** Bayesian network topology for predicting transformer failures.

The resulting network consists of nodes denoting discrete variables and connections between them, showing the influence of nodes on each other. For independent variables, the probability of each of the possible outcomes is determined, the sum of the probabilities of which is equal to one. For the dependent variables, a table of the probabilities of possible outcomes is compiled, provided that certain outcomes of the influencing variables occur. An
example of the relationship of several diagnostic predictors of the transformer with the tables of conditional probabilities of the TDRUNG-20000 transformer failure is shown in Fig. 4.

![Calculated relationship of diagnostic predictors of TDRUNG-20000 transformer.](image)

**Fig. 4.** Calculated relationship of diagnostic predictors of TDRUNG-20000 transformer.

As a result of modeling, matrices of conditional probabilities are generated for all nodes included in the model. The choice of influencing factors is carried out in automatic mode (by rank), or manually.

The developed software allows, based on the generated network structure and the table of conditional probabilities for various factors, determining the missing diagnostic factors and predicting the probability of failures in the short and medium term.

In addition to direct dependencies between variables, the Bayesian network allows obtaining tables of conditional probabilities between any nodes, while the table of conditional probabilities is supplemented with attributes, the influence of which is planned to be investigated.

## 5 Conclusion

The software allows modeling on a prepared network. By introducing artificial values of influencing variables into the model, it is possible to obtain the conditional probability of failure or non-failure with a combination of these factors.

The model allows, relying on the weather forecast, to calculate the probability of failure-free operation of transformers, to make a decision on the need to increase the availability of service personnel, on the need to reduce the load, etc., to prevent possible failures.

Further improvement of the predictive probabilistic model is possible with a significant expansion of the set of data on weather factors, such as solar radiation, which has a serious impact on the change in the electrical capacity and load resistance of the electric traction network [21], as well as the number of lightning during a thunderstorm.

It should be noted that for the effective operation of the proposed system, algorithms for obtaining, processing and interpreting information, as well as software that allows one to effectively solve problems of predicting failures and preparing operational decisions and control actions related to the power supply subsystem of railways, are of key importance [22].

At Omsk State Transport University, work is underway to improve the methodology for predicting transformer failures based on regression analysis and Bayesian networks.
References

1. 2020 Report on the functioning of the UES of Russia in 2019 (SO UES) p 36
2. Al Mhdawi A K, Al-Raweshidy H S 2019 IEEE Systems Journal 1–11 doi:10.1109/jsyst.2019.2921867
3. Bracale A, Carpinelli G, Pagano M, De Falco P 2018 IEEE Transactions on Power Delivery 33(4) 1825–1834 doi:10.1109/tpwrdr.2018.2791181.
4. Sun C, Ohodnicki P R, Stewart E M 2017 IEEE Sensors Journal 17(18) 5786–5806 doi:10.1109/jsen.2017.2735193
5. Karandaeva O I, Yakimov I A, Filimonova A A, Gartlib E A, Yachikov I M 2019 Machines 7(4) 77 doi:10.3390/machines7040077
6. Hekmati A, Hekmati R 2017 IET Science, Measurement & Technology 11(5) 581–589 doi:10.1049/iet-smt.2016.0417
7. Tenbohlen S, Coenen S, Djamali M, Müller A, Samimi M, Siegel M 2016 Energies 9(5) 347 doi:10.3390/en9050347
8. Rahman M S A, Lewin P L, Rapisarda P 2016 IEEE Transactions on Dielectrics and Electrical Insulation 23(2) 1088–1098 doi:10.1109/tdei.2015.005070
9. Djamali M, Tenbohlen S 2017 International Journal of Thermal Sciences 116 224–233 doi:10.1016/j.ijthermalsci.2017.02.012
10. Hernandez M D P C, Labib A 2017 Journal of Quality in Maintenance Engineering 23(4) 400–414 doi:10.1108/jqme-07-2015-0027
11. Koltunowicz W, Badicu L-V, Broniecki U, Belkov A 2016 IEEE Transactions on Dielectrics and Electrical Insulation 23(3) 1347–1354 doi:10.1109/tdei.2015.005579.
12. Golubkov A S 2017 Fundamental foundations, theory, methods and means of measurement, control and diagnostics: Materials of the international. Scientific-practical conf. (Novocherkassk: Lik) p 173
13. Smerdin A N 2018 Transsib news. Omsk state un-t of ways of communication Omsk 1(33) 69 – 79
14. Burkov A T 2018 Transsib news. Omsk state un-t of ways of communication Omsk 3(35) 91 – 100
15. Golubkov A S 2018 Transsib news. Omsk state un-t of ways of communication Omsk 4(36) 2 – 9
16. Wang G, Xu T, Tang T, Yuan T, Wang H 2017 Expert Systems With Applications 69 247–256
17. Zhou Y, Pahwa A, Yang S S 2006 IEEE Transactions on Power Systems 21(4) 1683Neapolitan R E 1990 Probabilistic reasoning in expert systems: theory and algorithms (John Wiley & Sons)
18. 2014 Bayesian Networks with Examples in R M Scutari and J-B Denis (Statistical Science, Chapman & Hall/CRC, US)
19. Cooper G and Herskovits E 1992 Machine Learning 9 309 – 347
20. Girshin S S, Sorokin I A, Smerdin A N, et al 2020 Przeglad Elektrotechniczny 96(6) 59Mescheryakov V A, Smerdin A N, Golubkov A S 2013 Bulletin of the All-Russian Research and Design Institute of Electric Locomotive Engineering 2(66) 104-121