Study Committee D2

Information Systems and Telecommunication

Paper D2_315_2018

DATA ANALYTICS PLATFORM FOR POWER EQUIPMENT INTELLIGENT LIFECYCLE MANAGEMENT

Alexandra I. Khalyasmaa, Stanislav A. Eroshenko

Ural Federal University, Ekaterinburg, Russian Federation

Motivation

- **Low reliability of power equipment** operation under the conditions of its high wear and tear.
- **Vast multidimensional and dissimilar data**: sensors, monitoring and technical diagnostic systems, testing protocols of power network equipment, etc.
- **The need to optimize the repair cycles** of the power equipment both to minimize technical risks and to manage production assets and investment programs of enterprises.
- **Production asset management platforms** are often based on technical state assessment, so the accuracy and reliability of the first depend greatly on the performance of the latter.

Objects of investigation

- **A model of the data analytics platform** of the functional state assessment of the power network equipment based on technologies of Knowledge Discovery in Databases.
- **The application of Gradient Boosting Trees** Machine Learning approach as a Data Mining tool for a problem of power equipment functional state identification.
- **The development of the training sample** with a minimum required set of training data in the form of retrospective states of the power equipment under consideration.
- **Functional state assessment** of 110 kV oil-filled power transformers in a large power hub of a regional GridCo.

Method

The algorithm of **Gradient Boosting Trees** in the given system is implemented in such a way, that, firstly, the basic $b_{n=0}$ is initialized. For n=1,...,N the following steps are performed:

- shift vector S , which characterizes how to correct the predictions of the given composition in order to reduce the training error, is calculated:

$$S_n = (-2(a_{n-1}(x_1) - y_1),\dots,-2(a_{n-1}(x_l) - y_l))$$

- a basic algorithm b_n is constructed by approaching its responses on the training sample to the shift vector s_i:

$$b_N(x)=\operatorname{argmin}\tfrac{1}{l}\textstyle\sum_{j=1}^{l}(b(x_i) - s_i)^2 = \sum_{j=1}^{l}[x \in R_{Nj}]b_{Nj}$$

- after the algorithm is found, it is added to the composition:

$$a_n(x) = a_{n-1} + \eta \sum_{j=1}^{J}[x \in R_{Nj}]b_{Nj}$$

The average error Q in percentage [%], which was used to evaluate the testing error of the algorithm a on the sample X^l is calculated as follows:

$$Q(a, X^l) = \frac{1}{l}\sum_{i=1}^{l}|a(x) \neq y^*(x)| \cdot 100\%$$

where $|a(x) \neq y^*(x)|$ – error indicator; $y^*(x)$ – true value of the parameter; l –number of observations.

Approach

The integrated assessment of the functional state of 35-220 kV substations is carried out based on a set of state estimates of low-level objects (sub-objects), such as:

- power transformers;
- power transmission lines (overhead and cable lines);
- auxiliary transformers;
- reactors;
- relay protection and automation systems;
- switches; disconnectors;
- busbars;
- instrument transformers (current and voltage);
- surge protection devices, etc.

In accordance with the developed approach, **risk is determined by a universal fuzzy set** $R\{r_1, r_2, \dots, r_N\}$, where N is the number of risk states corresponding to linguistic characteristics ***{"Very low"; "Low"; "Average"; "High"; "Critical"}***, described by membership functions and fuzzy production rules.

Generalized risk assessment is calculated on the basis of average weighted score of the results of risk distribution density, which is calculated using Weibull distribution function.

Study Committee D2

Information Systems and Telecommunication

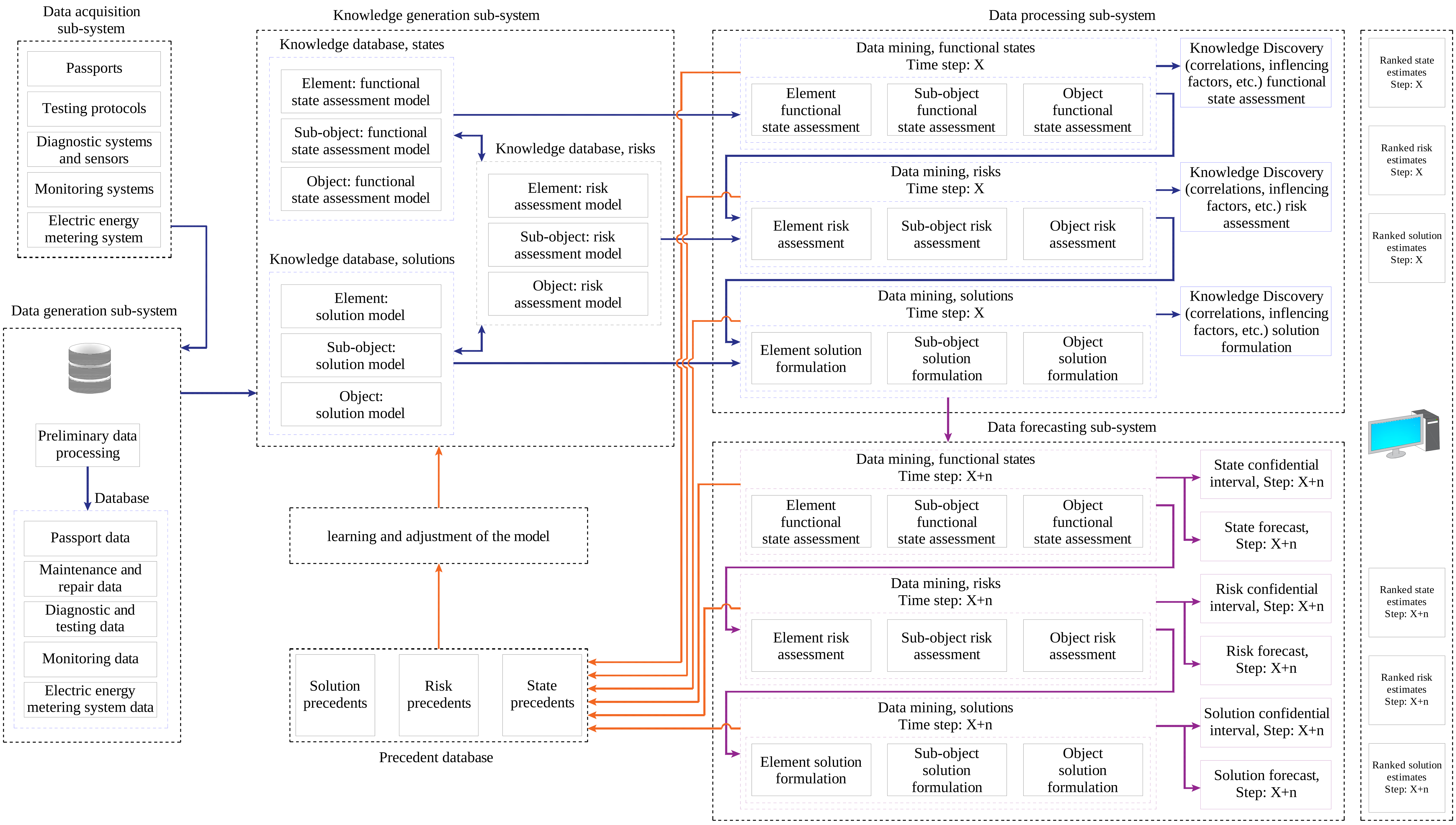
Paper D2_315_2018

DATA ANALYTICS PLATFORM FOR POWER EQUIPMENT INTELLIGENT LIFECYCLE MANAGEMENT

Alexandra I. Khalyasmaa, Stanislav A. Eroshenko

Ural Federal University, Ekaterinburg, Russian Federation

Intelligent data analytics platform structure



Experimental setup & Technical state assessment results

Retrospective								
Data type	**Number of pairs in the training sample, pcs.**	**Number of pairs in the testing sample, pcs.**	**Mean training error, %**	**Mean testing error, % Step: X+1**	**Errors of the 1st / 2nd kind, pcs. Step: X+1**	**Mean testing error, % Step: X+2**	**Mean testing error, % Step: X+3**	**Mean testing error, % Step: X+4**
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Power transformer								
Commissioning year, rated data, capital repairs, oil type, tank volume, d.c. resistance, etc.	411	407	5,3	11,8	36/12	14,0	16,9	19,3
Trans. tank oil chromatographic analysis of dissolved gases	562	541	2,1	7,3	32/7	8,3	9,5	11,4
Transformer oil physical and chemical testing	382	376	2,8	8,3	21/11	9,7	11,3	13,5
No load losses	367	341	5,1	11,5	24/15	13,9	16,7	19,0
Insulation resistance	379	331	3,4	9,2	22/8	10,6	12,5	14,8
Trans. OLTC oil chromatographic analysis of dissolved gases	342	350	3,1	8,1	23/5	9,5	11,2	13,3
Trans. inputs oil chromatographic analysis of dissolved gases	478	432	2,5	7,8	26/8	9,0	10,5	12,1
Thermal-imaging control	361	367	4,8	10,2	24/14	11,4	13,1	15,2

Study Committee D2

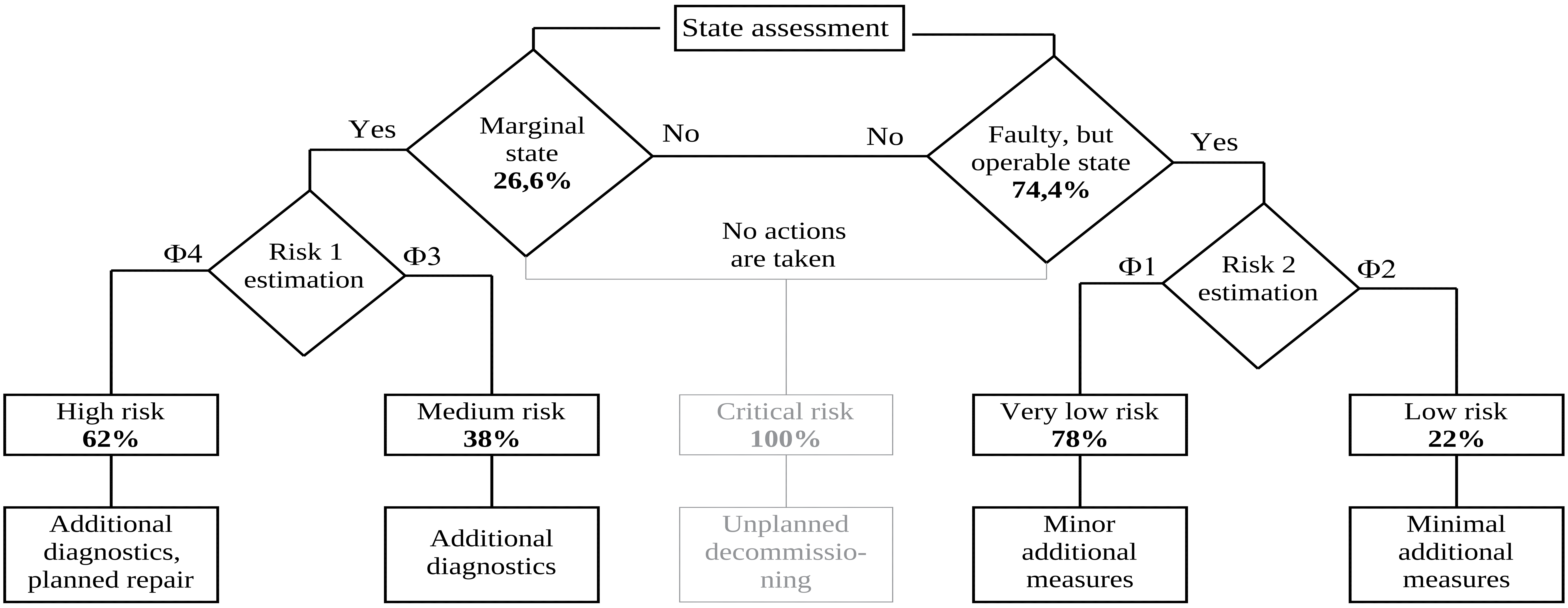
Information Systems and Telecommunication

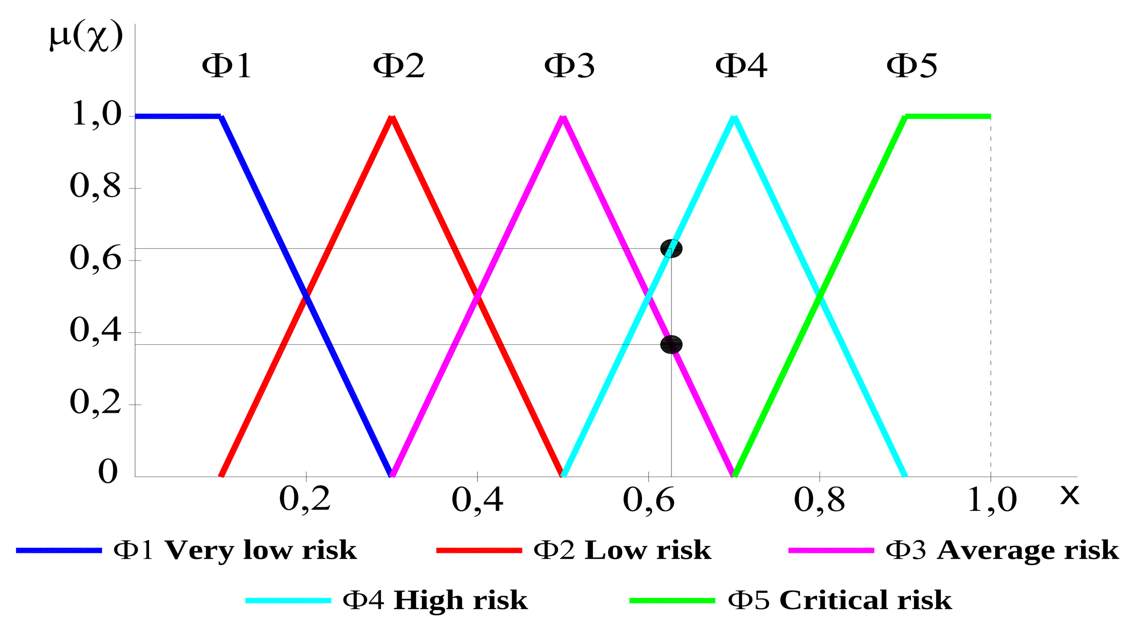
Paper D2_315_2018

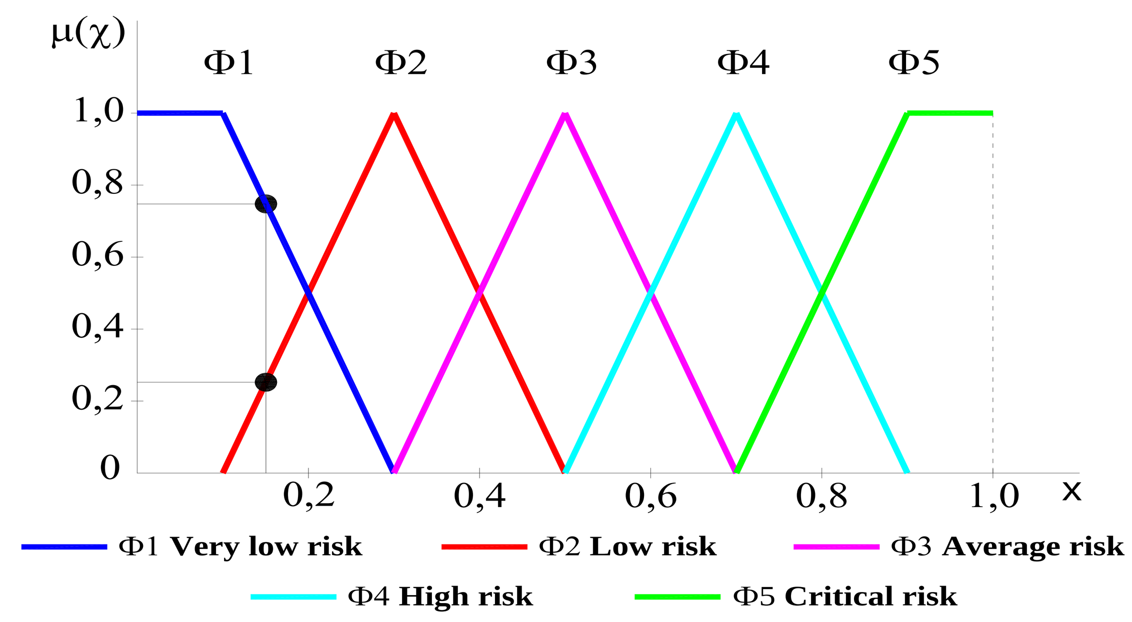
DATA ANALYTICS PLATFORM FOR POWER EQUIPMENT INTELLIGENT LIFECYCLE MANAGEMENT

Alexandra I. Khalyasmaa, Stanislav A. Eroshenko

Ural Federal University, Ekaterinburg, Russian Federation







Discussion

- **Current operation conditions were not taken into account** due to the absence of real-time monitoring data for the majority of cases under consideration.
- The **average testing error of power transformer functional state identification for *X+1* time period equals to 9%** (identification accuracy 91%).
- **The first kind errors prevail in the total number**, which illustrates the effectiveness of the proposed system, aimed at reducing the risk of defect missing.
- Power transformer **functional state forecasting for the time periods *X+2, X+3* and *X+4* correspond to 18% mean identification error** for the power transformers.
- **Proposed multiparameter approach gives the possibility to avoid incorrect state estimates** by taking into account implicit internal and external effects and their mutual correlation.

Conclusion

- **The proposed model can be used as a tool for integrated assessment of power network equipment, providing:**
  - ✓ effective investment planning of the utilities by shifting from preventive to condition-based repairs;
  - ✓ optimization of energy and resource saving strategies as well as gradual reliability improvement.
- **The Gradient Boosting Trees algorithm used as a tool for implicit dependencies identification provided:**
  - realization of the possibility to identify and predict gradual failures and developing defects;
  - taking into account the individual conditions of the power equipment, rather than by benchmarking with the marginal values provided by industrial regulations.
- **Future work:** practical implementation at existing power plant switchgear.