PAPER • **OPEN ACCESS**

Intelligent analysis and early warning model of grid meteorological disaster risk

To cite this article: Jiahui Xu *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **242** 022030

View the article online for updates and enhancements.



IOP ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

This content was downloaded from IP address 218.240.53.53 on 24/09/2019 at 11:49

Intelligent analysis and early warning model of grid meteorological disaster risk

Jiahui Xu[12], Shouqin Wang [12], Haoming Hu[3]， Xiaojie Wang [4], Wenbin Wu [4], Jianhui Ye[5], Tianyue Lou[12], Tian Xiu[12]

[1]NARI GROUP CORPORATION/STATE GRID ELECTRIC POWER RESEARCH INSTITUTE，NanJing, China 211000

[2]BeiJing KeDong Electric Power Control System Co.,Ltd.,Beijing China 100192

[3]Jiangsu Electric Power Company Research Institute, Nanjing, Jiangsu 210096 China

[4] Fujian Electric Power Company Research Institute, Fuzhou, Fujian 35007 China

[5] State Grid Hunan Electric Power Company, Changsha, Hunan 410000 China

e-mail: yuxujiahui@126.com

Abstract. Meteorological disaster is an important factor that affecting the operation of power system stability. Through data mining the relationship between the electric power equipment failure, the limit of water, and the meteorological disaster, we can find the main meteorological factors that cause failure or the limit. In order to effectively reduce the security risk of power grid operation, it is necessary to alarm the fault or the risk of exceeding the limit in advance.

1.Introduction

Natural disasters have caused great harm to domestic and international power grids, leading to huge economic losses to the society. Based on China's energy and load distribution, it is necessary to conduct long-distance, large-capacity transmission. With the continuous development and expansion of the grid scale in recent years, the threat of meteorological disasters [1] to transmission systems and equipment has also increased. Many measures have been taken within the power system to reduce its impact on the safe operation of the grid.

In real-time meteorological monitoring, the grid company implements high-precision weather monitoring for local areas by installing monitoring devices. However, monitoring points are only deployed in a relatively small area with a low coverage, making installation and maintenance extremely difficult and equipment reliability low. In addition, the constraints of natural conditions and facility costs make it difficult to increase micrometeorological coverage. In terms of meteorological forecasting, the meteorological department is unable to predict disaster weather with long-term and unconventional means (such as encrypted observations), and the local meteorological department can only provide qualitative weather forecasting due to technical level and calculation limitations, therefore unable to meet the power company's requirements for disaster weather forecasting in terms of precision (meteorological, geographical, etc.). Although there are already some applications and systems in the aspects of meteorology and power grid, and the risk warning for grid equipment affected by meteorological disasters is initially realized [2] [3], the result set of early warning contains a large number of equipment, affecting the accuracy of warning for faulty equipment. Therefore, it is

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

urgent to obtain timely, comprehensive and accurate early warning of grid equipment from existing and coarse-grained meteorological data so as to build a grid meteorological disaster early warning system with high integration, comprehensive functions and tight integration with the power grid.

This paper comprehensively analyzes the phenomenon of transmission equipment failure caused by meteorological disasters and water level changes caused by meteorological disasters based on previous equipment and plant data, and previous meteorological data by using clustering algorithm. Combined the above analysis with statistical data of other effects of grid equipment failure, this paper puts forward a grid meteorological disaster risk of estimate and smart warning model in order to make early warning more accurate, and provide technical support and basis for reducing the adverse effects of disaster weather on the safe, stable, and reliable operation of power grid and power supply.

2. Data Analysis

2.1 Definition of Terms

(1)Threshold rule: a rule for defining the value of meteorological elements, such as "24-hour cumulative precipitation: XXXX mm/h". The purpose of the threshold rule definition is to generate an initial set of disaster warning results.

(2)Streamlined rule: a more detailed disaster warning rule that includes not only the values of meteorological elements, but also fields such as regions, time zones, grid equipment, plant stations, equipment faults [4] or water level limits [5].

(3) Intelligent decision-making expert database: a set of rules for auxiliary decision-making including the threshold rule and the streamlined rule.

2.2 Data Processing Process

As shown in Figure 2-1, the meteorological disaster data processing process consists of two parts: the generation of intelligent decision-making expert database based on historical data and real-time meteorological disaster warning.

(1)The intelligent decision-making expert database explores the association rule between meteorological, fault and auxiliary factors based on historical meteorological data, grid management data and equipment operation data, using data analysis algorithms [6].

(2) Meteorological disaster warning uses real-time meteorological data as input, and compares and calculates the threshold rule and the streamlined rule in the intelligent decision-making expert database to produce the final result of equipment failure or water level over-limit warning.

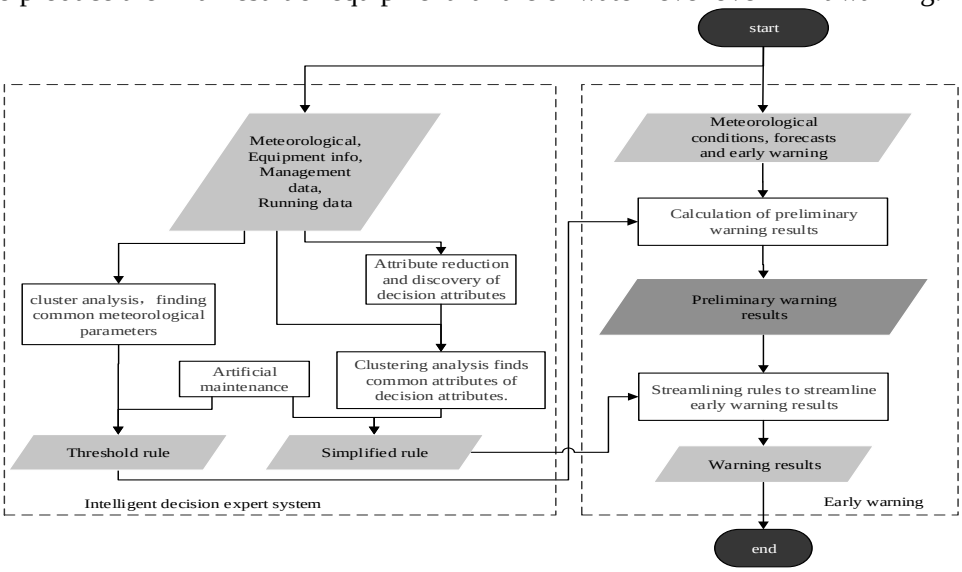

Figure 2-1. Disaster warning data processing flow.

2.2.1 Intelligent Decision Expert Database

The processing flow generated by the intelligent decision-making expert database mainly includes four parts: grid data integration, meteorological data integration, data modeling, and intelligent decision-making expert database generation.

(1) Grid data integration

Access data includes grid management data and device operational data. The management data is accessed from the grid OMS system with the content including equipment account, schedule log, fault information, maintenance information, mode adjustment, power generation plan, and disposal plan. The operation data is accessed from a real-time monitoring system such as D5000 with the content including operation data and device warning data.

(2) Integration of meteorological data

Meteorological data includes meteorological elements such as precipitation, wind, temperature, humidity, air pressure, lightning, and light. The data comes from multiple meteorological systems, including provincial and municipal meteorological stations, China Electric Power Research Institute Meteorological Forecast Center, and public weather information on the Internet.

(3) Data modeling

Meteorological data and grid data are based on regions, meteorological stations, plant stations, and equipment ledgers. The latitude and longitude, plant station ID, line ID, and equipment ID are used as reference values to establish the relationship between meteorological data, management data, and operational data. Subsequent data analysis lays the data foundation, as shown in Figure 2-2.

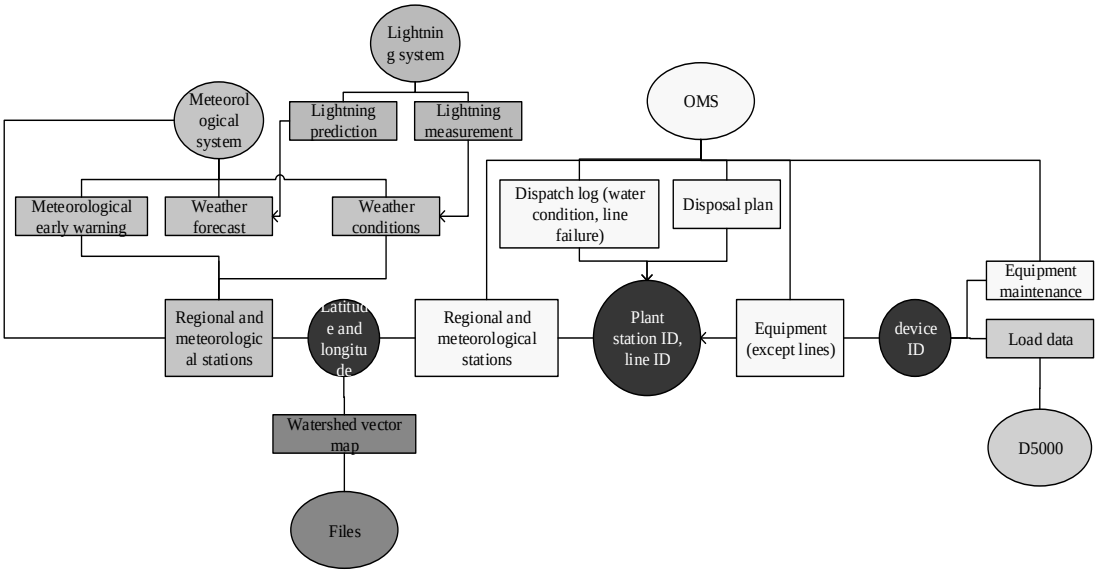

Figure 2-2. Data Modeling.

(4) Generation of intelligent decision expert database

1 Perform automatic text parsing on the dispatch log to obtain equipment faults and water level limit data caused by meteorological factors;

2 By clustering the meteorological elements at the time of failure or threshold-crossing, the commonality of the meteorological elements is used as the threshold rule for the disaster warning;

3 Introduce data of maintenance, technical modification, and operation mode adjustment to select the decision attributes of meteorological aids that cause failure or over-limit by using attribute reduction algorithms such as rough set [7];

4 By using the clustering algorithm for the decision attributes of the fault or the time limit of occurrence, the commonality of time, region, meteorological elements, and other decision attributes in each cluster is found as a streamlined rule for equipment failure or water over-limit warning.

2.2.2 Meteorological Disaster Warning

The meteorological disaster early warning process processes real-time data and provides early warning of future equipment failures and water over-limit risks.

(1) Meteorological disaster warning is based on meteorological conditions, forecasts, and early warning data;

(2) Using the threshold rule generated by the expert database to adjust the model parameters in order to obtain preliminary warning results;

(3) Simplify the rules using the expert database and produce a more accurate warning result after a series of comparisons and calculations using the current operating state of the grid involved in the rules as auxiliary decision attributes.

2.3 Rainstorm Warning Model Data Processing Flow

Take the rainstorm warning model as an example. The model is based on general data processing flow while focusing on various parts of the process, as shown in Figure 2-3.

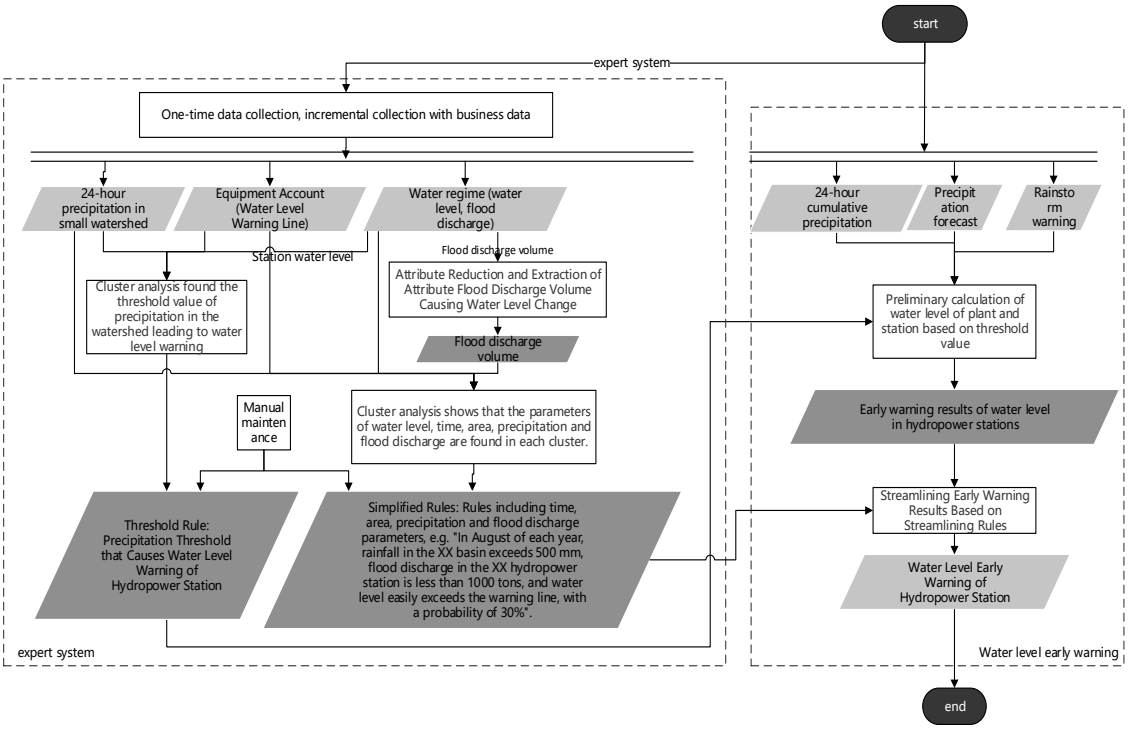

Figure 2-3 Rainstorm warning data processing flow.

2.3.1 Intelligent Decision Expert database

In the rainstorm warning model, the intelligent decision expert database takes historical precipitation, hydropower station warning water level, water condition and operation mode as input, takes historical water level warning data of water stations as the entrance during the analysis process, queries the 24-hour cumulative precipitation at the alarm time, and carries out precipitation cluster analysis to find the main area of precipitation and the precipitation threshold caused by the water level warning of the hydropower station. At the same time, the decision database takes water level alarm data of the

historical hydropower station as the entrance to query and calculate accumulated precipitation and flood discharge at the alarm time by clustering the combination of warning period, area, and flood discharge so as to produce the streamlined rule.

2.3.2 Meteorological Disaster Warning

The meteorological disaster warning takes 24-hour cumulative precipitation, precipitation forecast, and rainstorm warning data as input, adjusts the threshold of the early warning model according to the precipitation threshold rule, and preliminarily calculates the hydropower station where the water level alarm may occur; then uses the streamlined rule with such parameters as time, region, precipitation, and flood discharge quantity in order to further streamline the list of hydropower stations generated in the previous step and make water level warning of the hydropower station more accurate.

3. Key Technologies

3.1Data Preprocessing

3.1.1Vacancy Field Value Supplement

Inside the power grid, there inevitably exists missing and misdiagnosed equipment fault data (including time, equipment, fault causes, etc.), which may affect the precision of data. Therefore, before data analysis and mining, the field vacant values shall be filled according to certain rules. Some common filling methods are listed as follows:

(1) Global constant filling

Fill in all field vacancies with a constant such as "unknown". This approach saves resources and is simple and easy. However, this filling method also has great defects, which may mislead the data calculation process, resulting in a big gap between the training results and the actual situation, and even completely wrong conclusions. This method may also mislead data mining, which may cause bias or even erroneous conclusions depending on filled values.

(2) Average sample filling of similar samples

Fill with the average of all samples of the same class of a given tuple. This method is more reasonable than the global constant filling and can reduce the error caused by the filled value.

(3) Filling of the most likely value

The most likely value can be determined using regression analysis and decision tree induction. The values obtained through statistical processing are in some cases more accurate than the sample mean.

3.1.2Error Field Value Elimination

Samples of data analysis should be as clean and non-polluted as possible, otherwise excessive error data will lead to biased analysis. Grid data used for analysis is generated using both machine acquisition and manual reporting. There is a certain amount of error in the data generated by these two methods, especially manual reporting. Therefore, some techniques of data mining (regression, clustering, binning, etc.) can be used to process the data, thereby eliminating noise and improving data quality.

3.1.3 Data Consistency

The inconsistency of the data comes from the integration process of the data. The data set of training sample may come from different acquisition systems, and the data maintained by each system differs in terms of units, resulting in integration data inconsistency. Under this circumstance, manual intervention, such as reference to the source data of the data dictionary table or data table can be applied.

3.2 Clustering Algorithm

This paper takes K-Means [8] algorithm to cluster the historical data to obtain expert database rules. The basic idea of the K-Means algorithm is to perform clustering calculation based on k points in the multi-dimensional space so as to realize the k-class division of the training object. By continuously iterating the coordinates of the k center points and updating the values of the various center points one by one until the iterative result no longer changes, the final clustering result can be obtained.

In the rainstorm warning model, the 24-hour cumulative precipitation at the time when the water level exceeds the limit is taken as a sample. Assume the sample set is classified as c-class, and the calculation steps are as follows:

(1) Initialize the initial center of the c category to ensure that c centers are not equal to each other;

(2) In the kth iteration, find the distance of any sample to c centers, and classify the sample into the class with the shortest center;

(3) Update the central values of the c class by means like averaging;

(4) For all c cluster centers, repeat steps (2) and (3); stop iteration until the center value remains unchanged; otherwise, iteration continues.

After processing by the K-Means algorithm, the data is divided into c clusters according to the size and dispersion of the data set. After comparing the amount of data and the number of clusters, the centroid of the cluster with the largest amount of data will be used for the new threshold of the early warning model.

The expert database rules generated by the clustering algorithm are more statistically persuasive than the manually defined empirical values and more accurate in early warning model to some extent.

3.3 Attribute Reduction

Attribute reduction [9] is also called dimension specification or feature selection. Mathematically speaking, there is p-dimensional data x=(x1,x2...xp), and new data x=(x' 1,x'2...x'k), k≤p is obtained by adopting a certain kind of measure. New data retains the characteristics of the original data to a maximum extent under certain criteria. The goal of attribute reduction is to eliminate the effects of redundant and unrelated attributes on the calculation process and the final result. In this paper, attribute reduction is mainly used to solve the decision attribute of extracting multi-dimensional grid management and operation data as the auxiliary factor of meteorological factors, resulting in equipment failure or water over-limit.

This paper uses rough set theory as a method of attribute reduction. Rough set theory [10], based on the classification mechanism, considers categorization as the equivalence relation in a specific space and the equivalence relation constitutes as division of the space. The most striking difference between this theory and other theories dealing with uncertain and inaccurate problems is that it does not need to provide any priori information beyond the data required by the problem, so the description or treatment of the uncertainty of the problem can be considered as more objective .

4. Research Results and Conclusion

This paper, based on the relationship between grid equipment and meteorological disasters, obtains the rules of early warning experts for meteorological disasters through data mining training including the threshold of meteorological elements for each early warning model (Figure 4-1), and description of associated streamlined rules for meteorological elements, other decision attributes, and equipment failures (Figure 4-2).

According to research results, the intelligent analysis and early warning system of meteorological disaster risk based on geographic information platform [11] is constructed, which makes possible not only the query of basic data and data display based on geographic information platform, but also more comprehensive access data, more comprehensive early warning object, more accurate results, and improved visual display, thus providing more comprehensive support for equipment failure warning, overhaul construction, hydropower generation, and load analysis, especially for hydropower station water level warning and formulation of the plan of hydropower plant.

no.	model name	threshold factor	threshold	unit	risk grade	update time
1	rainstorm	Total rainfall	25	mm/h	1	2015.12.21
2	gale	wind speed	10.8	m/s	1	2015.12.21
3	frozen	(temperature, relative humidity, wind speed)	(-5, 80, 10)	(℃,%,m/s)	1	2015.12.21
4	fire	(temperature, humidity)	(35,30,)	(℃,%)	1	2015.12.21
5	thunder	(lightning current amplitude, side distance)	(35,250)	(A, m)	1	2015.12.21

Figure 4-1. Threshold rule.

no.	streamline rules	target	failure or overshoot	time	repair	run type
1	The precipitation in Wuqiangxi River Basin is large from July to August every year. The water level of Wuqiangxi Power Plant is easy to reach the warning line with a probability of 30%.	Wuqiangxi	over-limit	July,august	no	unrelated
2	In August, the partial operation mode of XX was A, the number of long I lines was more frequent, and the probability of failure was 30%.	LinePaichangl	fault	august	no	A

Figure 4-2. Streamlined rule.

5. Conclusion

With ever deeper understanding of the relationship between grid business and meteorology and continuous improvement of various types of data, it is expected to learn from big data storage and analysis technology to further improve the accuracy of the model and the precision of early warning results.

With the accumulation of data and the continuous improvement of the early warning model, the early warning model will be applied not only to the field of fault and water level warning, but also to all aspects of grid operation that are susceptible to meteorological disasters, such as forecasting of regional load change caused by temperature and prediction of generating capacity so as to provide better assistance and support for grid operation safety.

REFERENCES

[1] LI, Y. B., HAN, Y., ZHANG, R., LI, Y..(2013). Research on impact model of meteorological factors on the power accidents. Power System Technology.

[2] LU, Y. L., TAO, F. B., ZHOU, Z. C. .(2015). Analysis of influence and prevention of meteorological disasters on jiangsu's power grid equipment. Journal of Nanjing University of Information Science & Technology.

[3] Lu, Y. L. , Tao, F. B. , Hu, C. B. , Zhou, Z. C. , & Sa, L. . (2017). Statistical analysis of meteorological disasters in the jiangsu power grid and prevention methods. IOP Conference Series Earth and Environmental Science, 86(1), 012021.

[4] MA, Y .S., (2005). Analysis of the equipment faults in beijing power grid in 2004. Electric Power.

[5] Hartmann, & Holly, C. . (1988). Historical basis for limits on lake superior water level regulations. Journal of Great Lakes Research, 14(3), 316-324.

[6] Agrawal, R. , Srikant, R. , Bocca, J. B. , Jarke, M. , Zaniolo, C. , & Kaufmann, M. . (2014). Fast Algorithms for Mining Association Rules in Large Databases. International Conference on Very Large Data Bases. IEEE.

[7] Dou, H., Yang, X., Song, X., Yu, H., Wu, W. Z., & Yang, J. (2016). Decision-theoretic rough set. Knowledge-Based Systems, 91(C), 71-83.
[8] Likas A, Vlassis N, Verbeek J J.(2003,)The global k -means clustering algorithm[J]. Pattern Recognition, 36(2):451-461.
[9] Xin, Y. U. , Da, D. W. , & Sheng, P. J. . (2005). A rough set approach to incomplete information systems. Computer Engineering & Science, volume 112(4), 39-49(11).
[10] Pawlak, Z., & Skowron, A. (2006). Rudiments of rough sets. Information Sciences, 177(1), 3-27.
[11] Wilhelmi, O. V., & Brunskill, J. C. (2003). Geographic information systems in weather, climate, and impacts. Bulletin of the American Meteorological Society, 84(10), 1409-1414.