A Data-Driven Self-Learning Evaluation Method of Voltage Sag Severity

Shuming LIU, Chen ZHENG, Bo ZHANG, Shuangyin DAI, Yuzheng TANG, and Yi WANG

***Abstract*—Lightning is one of the main causes of voltage sags, and it is of great significance for subsequent analysis and governance to accurately evaluate the severity of the sag events caused by lightning. There are many uncertain factors between lightning fault events and voltage sag events. To evaluate the severity of voltage sag events caused by lightning, a data-driven self-learning evaluation method for voltage sag severity is proposed. According to a large number of online monitoring data of lightning positioning system and power quality monitoring system, the association rule mining algorithm based on incremental learning is used. And the rules are kept updating through the accumulation of historical data, which may give it the abilities of self-learning. The empirical analysis is carried out based on the monitoring data of a regional power grid. The results show that the method in this paper can accurately mine more valuable rules in reality and solve the problem of low efficiency of mining algorithm when the database changes dynamically.**

***Index Terms*—Association rule, lightning, lightning location system, power quality monitoring system, self-learning, voltage sag severity.**

I. Introduction

VOLTAGE sag has become one of the most concerned problems due to the huge economic losses. The common cause of voltage sag is the sudden appearance of a large current in the system, such as a short circuit caused by a lightning strike [1], [2]. It is of great significance for subsequent governance to accurately assess the severity of sag events caused by lightning.

The traditional voltage sag severity evaluation methods [3]–[6] are mainly divided into three categories: the measured statistical method, the stochastic prediction method and the state estimation method. The measured statistical method refers to online monitoring at selected sites in the system, and the monitoring data is analyzed and calculated to determine the sag level of the measured points. But this method is only applicable to the situation where the system topology and parameters are determined and cannot assess the sag level of the whole system. The accuracy of the evaluation results depends on the length of monitoring time. The stochastic prediction method is based on stochastic modeling. The probability model is used to simulate the voltage sag randomly, and the sag level evaluation of the whole system can be achieved without long-term online monitoring in the system. However, the historical statistical data used by the random prediction method is affected by factors such as weather and maintenance conditions, and changes greatly every year; on the other hand, if the system is rebuilt or expanded, the historical statistical data will no longer be applicable. The state estimation method is a combination of the above two evaluation methods, which can accurately evaluate the sag level of the system over time using limited monitoring data. However, the accuracy of this method is based on the configuration method of the monitoring device.

In recent years, the massive monitoring data stored in the power grid monitoring system provides a new way for voltage sag evaluation. The basic idea of the data-driven voltage sag evaluation method is to obtain the correlation between the parameters and the sag indexes to realize the system voltage sag evaluation. In [7], based on the data of the power quality monitoring system, data mining technology is used to predict power quality disturbance levels. In [8], the improved correlation rule algorithm is used to evaluate the fault level of sensitive equipment caused by voltage sag from the perspective of the power grid. Reference [9] proposes a data-driven prediction method for the characteristics of voltage sags based on fuzzy time series. Reference [10] studies the propagation characteristics of voltage sag based on data driven method. Reference [11] proposes a residual voltage data-driven pridiction method for voltage sag. References [7]–[11] do not depend on the network topology and have high practical applications in certain historical databases.

However, with the accumulation of data in the power grid, the timeliness of association rules is constantly changing. The above algorithm is difficult to apply to a dynamic database, and the database needs to be rebuilt and scanned when the database changes, and the algorithm efficiency is too low. To solve the above problems, this paper proposes a data-driven self-learning evaluation method of voltage sag severity. Where self-learning means that the rule base can be updated based on the association rules that have been mined in the face of the dynamic changes of the database, thereby improving the efficiency of the algorithm and the accuracy of

Manuscript received April 30, 2022; revised July 4, 2022; accepted July 29, 2022. Date of publication September 30, 2022; date of current version September 13, 2022. This work was supported by the State Grid Corporation Headquarters Science and Technology Project under Grant 5400-202124153A-0-0-00. *(Corresponding author: Chen Zheng.)*

All authors are with the Electric Power Scientific Research Institute of State Grid Henan Province Electric Power Company, Zhengzhou, Henan 450052, China (e-mail: losuwing@126.com; zhengchen725@163.com; zhangbo@163.com; daishuangying@163.com; tayuzhneg@163.com; wangyi@163.com).

Digital Object Identifier 10.24295/CPSSTPEA.2022.00030

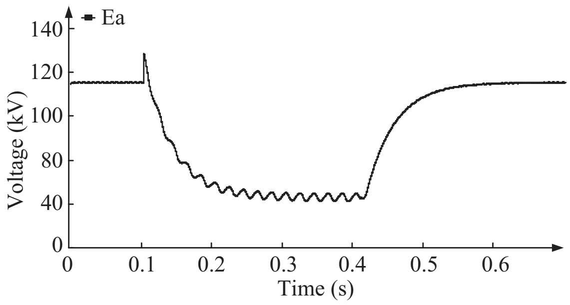

Fig. 1. Characteristics of voltage sags caused by lightning.

the rules. The novelties of this paper are as follows. First, a mapping relationship between known inputs and the severity of voltage transients is constructed based on a large amount of historical data, which is more applicable when dealing with different types of lightning activities. Second, the proposed method achieves continuous correction of the rules through the accumulation of historical data, thus continuously improving its decision-making accuracy or efficiency in practical applications. In addition, based on the actual data in the lightning location system and power quality monitoring system, the correctness and validity of the proposed method were verified.

II. Mining Parameter Selection

When lightning strikes tower top, lightning conductor, or phase line, the potential for the struck points will rise suddenly due to lightning current. If the voltage difference exceeds the flashover voltage of the insulator string, the insulator string will flashover and the transmission line may occur short circuit fault. Then, the voltage of other lines, which are closing to the faulty line, will also be correspondingly reduced.

Since there is a strong correlation between lightning data and voltage sag data, but it is not a one-to-one mapping or a simple causal relationship, it is necessary to analyze and mine the data of multiple monitoring platforms in the power grid. Lightning Location System (LLS) can provide real-time and historical lightning data [12], including lightning strike location, time, peak current, polarity, and number of strikes back. Power Quality Monitoring System (PQMS) stores sag information [13], including sag start and end time, location, voltage level, etc. In China, each lightning trip accident has been analyzed on a case-by-case basis. By observing the lightning strike point or arc trajectory of the struck tower, engineers determine the lightning location. Therefore, according to the line trip record and the sag event record, the lightning parameters causing sag events can be determined, and a single lightning-sag event can be obtained.

The power quality monitoring system can obtain the magnitude and duration of sag. Combined with the voltage tolerance curve (VTC) corresponding to the equipment type of the node, this paper uses the IEEE 1564 standard to calculate the severity of a lightning-sag event.

TABLE I
Selected Parameters

Parameters	identifier	Parameter Meaning
date	$C1$	date of a lightning-sag event
time	$C2$	time of a lightning-sag event
lightning location	$C3$	location of lightning
lightning peak current	$C4$	maximum current during lightning
sag distance	$C5$	distance between the lightning location and the concerned node
sag severity	D	event Indicators S_e

$$S_e = \frac{1-V}{1-V_{\text{curve}}(d)} \tag{1}$$

Where $V_{\text{curve}}(d)$ is the corresponding voltage magnitude on the VTC when the duration is d, and V is the actual voltage sag magnitude. To obtain the voltage sag severity at the unmonitored nodes of the system, the electrical distance is selected as one of the mining parameters in this paper. The electrical distance is the geometric distance between the lightning location and the concerned node. In summary, the parameters selected in this paper are shown in Table I.

III. Self-learning Evaluation Method of Voltage Sag Severity

A. Association Rule Mining Algorithm

As a classical association rule mining algorithm, the basic idea of AprioriTid algorithm [14] is to obtain the set of frequent items and mine the association rules based on the threshold value set by the user. Let $DA = \{t_1, t_2 \ldots, t_n\}$ be an event database, where each event t_i is a set of items. Items are the values taken for a single parameter in a single event, and the set of these items is IT. For the set of items T, $T \in IT$, if there are k items in T, then T is also called a k-item set. For a k-item set T, if its computed support is higher than the support threshold (min_sup), then it can be called a frequent k-item set.

The frequent item set represents an association rule, and the association rule is an implication of $X \Rightarrow Y$, where $X \subseteq IT$, $Y \subseteq IT$ and $X \cap Y = \Phi$. To reflect the correlation between X and Y, the user sets a confidence threshold (min_conf) to determine whether the rule is a strong association rule, and the support and confidence of the rule are calculated as shown in (2) and (3), respectively.

$$\sup(X \Rightarrow Y) = \frac{|X \cap Y|}{|U|} \tag{2}$$

$$\operatorname{conf}(X \Rightarrow Y) = \frac{|X \cap Y|}{|X|} \tag{3}$$

Where $|X|$ and $|X \cap Y|$ represent the number of events of the set X and $|X \cap Y|$ in the original database DA, respectively. The support represents how frequently the event appears in the database, and the confidence represents the confidence of the rule.

The AprioriTid algorithm improves the computational efficiency by using the k-level TID table instead of the original database when generating frequent k-item sets, which reduces the amount of scanning. Since the k-level TID table stores all the generated frequent k-item sets, the number of events stored in the table will be higher than the original database when k is small. To further improve the efficiency of the algorithm, this paper uses the decision table instead of the TID table, and each scan only needs to delete the rows in the decision table where the events without frequent k-item sets are located to get a new decision table.

$$J=\begin{bmatrix} \text{No.} & V_{C1} & V_{C2} & V_{C3} & V_{C4} & V_{C5} & V_D \\ 1 & & & & & & \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ n & & & & & & \end{bmatrix} \tag{4}$$

Where V_{ci} is the identifier of the i-th parameter Ci, the first column of the decision table is the number of the event, and the i-th row indicates the value of each parameter in the i-th event.

B. Incremental Update Algorithm

The data in the actual power grid is dynamically changing and the timeliness of the association rules is constantly changing. As time changes, new events are added to the lightning-sag event database. The incremental update algorithm is to mine new association rules and delete old association rules that do not satisfy the conditions based on the rules already acquired in the case of database changes, which avoids repeated scanning of decision tables and realizes the update of the association rule base.

Assuming that the original database is DA and the added event database is da, the updated database is $DA+da$, and the new support is calculated as

$$\sup(DA+da)=\frac{\sup(DA)\times|DA|+\sup(da)\times|da|}{|DA|+|da|} \tag{5}$$

where $|DA|$ denotes the number of events in the original database DA, and $|da|$ denotes the number of events in the added database da. When updating the database, let the set of frequent items of the database be $L(DA)$ and $L(da)$, respectively, then there are four cases as follows.

1) If the itemset $T\in L(DA)$, and $T\in L(da)$, then the itemset T must be a frequent itemset.
2) If the itemset $T\in L(DA)$, and $T\notin L(da)$, then you need to scan DA to calculate the support of T.
3) If the itemset $T\notin L(DA)$, and $T\in L(da)$, then you need to scan DA to calculate the support of T.
4) If the itemset $T\notin L(DA)$, and $T\notin L(da)$, then the itemset T is not a frequent itemset.

When calculating the support of each set T in the newly added database da, sup(T) is calculated based on the size of the database da. Let T be the frequent itemset in DA, if sup(T) $\geqslant$ min sup, then T is the frequent itemset after the database change, else, it is calculated according to (5) to determine whether it is the frequent itemset. Let T is not a frequent itemset in DA, if sup(T) $\geqslant$ min sup, then it is calculated according to (5) to determine whether it is a frequent itemset; else, T is not a frequent itemset after database changes. After obtaining the frequent itemset by the association rule mining algorithm based on incremental learning, let the voltage transient severity be Y, the remaining parameters be the set X, and the rule form be $X\Rightarrow Y$, whose confidence is calculated by the following formula.

$$\mathrm{conf}(X\Rightarrow Y)=\frac{\sup(X\cap Y)\times(|DA|+|da|)}{\sup_{DA}(X)\times|DA|+\sup_{da}(X)\times|da|} \tag{6}$$

where $\sup_{DA}(X)$ and $\sup_{da}(X)$ denote the support of the item set X in the databases DA and da, respectively.

C. Association Rule Matching

The X of the actual scenario cannot be same as the filtered strong correlation rules, so the similarity between the actual scenario and the rules in the rule base needs to be calculated. The rule with the highest similarity is used as the matching result for the voltage sag severity. Y is affected by each parameter in X to a different degree, so the parameters in X are given weights before matching. Combining the advantages of the subjective and objective assignment methods, this paper uses a comprehensive assignment method combining the Analytic Hierarchy Process (AHP) [15] and the entropy weight method [16].

The hierarchical analysis method is based on the importance of the parameters according to a certain scale to obtain the weight. According to the requirements of experts or users of the importance of the parameters for two comparisons, the scale matrix R is built. After consistency testing for R, the weight Cj of the parameters β_j is calculated.

$$\beta_j=\frac{\sqrt[5]{\prod_{i=1}^{5} r_{ji}}}{\sum_{k=1}^{5}\sqrt[5]{\prod_{i=1}^{5} r_{ki}}} \tag{7}$$

where r_{ji} denotes the element i of the jth row of the matrix R.

The entropy weight method is based on the fuzzy transformation theory to obtain the fuzzy evaluation matrix, then the entropy value e_j of the parameter Cj is calculated. The higher the parameter entropy value represents the lower the importance, so $1-e_j$ is used to calculate the weight w_j of the parameter.

$$w_j=\frac{(1-e_j)}{\sum_{k=1}^{5}(1-e_k)} \tag{8}$$

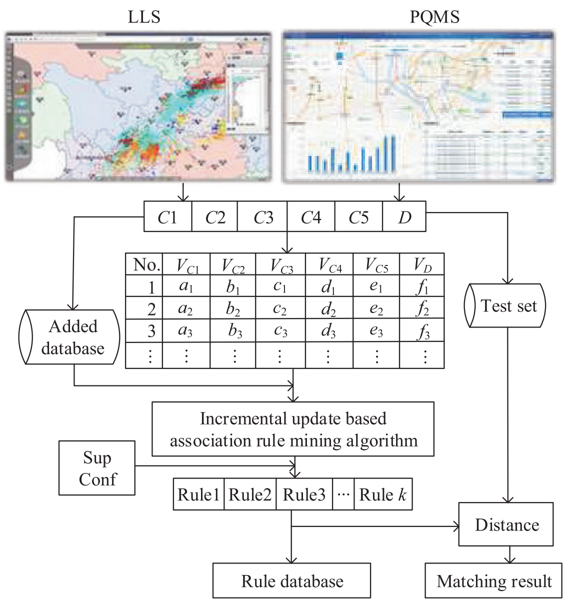

Fig. 2. Association rule mining flow chart.

Then combined weight is calculated by the following.

$$\alpha_j = \frac{\beta_j \times w_j}{\sum_{j=1}^{5} \beta_j \times w_j} \tag{9}$$

After the comprehensive weights are obtained, the distance between the actual scene x and the association rule r is calculated by the weighted Euclidean distance. A smaller distance indicates a higher similarity, and the calculation formula is as follows.

$$d(x,r) = \sqrt{\sum_{j=1}^{5} \alpha_j (Vc_j^x - Vc_j^r)^2} \tag{10}$$

where α_j is the combined weight of parameter Cj, and Vc_j^x and Vc_j^r are the values of parameter Cj in the actual scenario and association rules, respectively.

The flowchart of this paper is shown in Fig. 2. First, the monitoring data of LLS and PQMS are matched by spatio-temporal correlation to construct the database. Second, discrete the continuous data in the database by clustering algorithm. Subsequently, the association rules are mined by the AprioriTid algorithm based on the incremental update algorithm. Finally, the actual scenarios and rules are matched based on weighted Euclidean distance to obtain the voltage sag evaluation result.

IV. Case Study

Since lightning activities are seasonal, the measured data of LLS and PQMS from June to September 2016−2020 of a regional power grid are used as an example, and 127, 141, 139, 103, and 98 lightning-strike-temporary fall events are matched year by year, respectively. The events in 2016 are used as the original database, the events in 2017−2019 are used as the incremental training set, and the 98 events in 2020 are used as the test set for validation.

A. Data Processing

The AprioriTid algorithm can only handle discrete attributes, so it is necessary to discretize the continuous attributes. Among the six mining parameters selected in this paper, the lightning location is expressed in latitude and longitude, and a lightning – sag event record in the order of lightning date, time, location, peak lightning current, electrical distance and voltage sag severity is shown as 2019-06-06, 00:16:28.779, [102.5204, 28.3845], −67.9 kA, 35.78 km, 1.129. The lightning date and time need to be converted into numerical data. In this paper, the lightning date is converted into a variable value of 1−122 with a period of years, and the time is converted into a continuous value of 1−144 with a unit of every 10 min. The date of 2019-06-06/00:16:28.779 is 6 and the time is 1.6. The data types of lightning location, peak lightning current, electrical distance and voltage sag severity are continuous numerical data.

The K-means algorithm is a classical discrete algorithm [17], [18], but its number of discrete intervals relies on subjective experience. To avoid the artificial selection of the number of discrete intervals affecting the subsequent association rule mining, Silhouette Coefficient (SC) [19] is selected in this paper to determine the optimal number of clustering intervals.

The calculation steps of SC are as follows: 1) Calculate the average distance $m(a)$ between the data value a and the rest of the data in the interval, reflecting the degree of closeness within the interval; 2) Calculate the minimum value $n(a)$ of the distance from the data value a to the other intervals, reflecting the degree of dispersion between the intervals; 3) Calculate the contour coefficient of the data point a with the following formula.

$$s(a) = \frac{n(a) - m(a)}{\max\left(m(a), n(a)\right)} \tag{11}$$

The SC reflects the discrete validity of individual data, and the larger $s(a)$ is, the more reasonable the discrete result is. The average value of the contour coefficients of all data is used as the dispersion evaluation index of the algorithm. With different numbers of dispersion intervals set, the changes of SC values of the six parameters are shown in Fig. 3.

According to Fig. 3, the number of discrete intervals for parameters $C1$, $C2$, $C3$, $C4$, $C5$, and D are 3, 3, 2, 2, 3, and 4, respectively. The discrete results of the parameters are shown in Table II.

B. Rule Mining

The initial decision table is constructed based on the original database and discrete results, and the TID table is replaced using the decision table. Set the minimum support to 0.01 and the minimum confidence to 0.5, then 13 association rules are

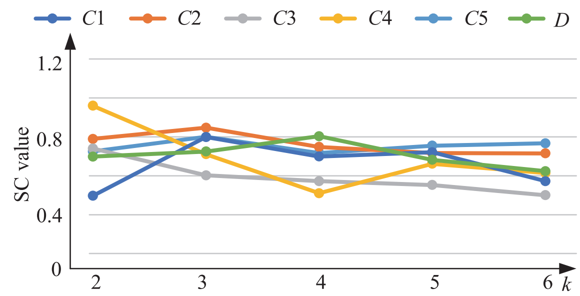

Fig. 3. The SC values of the 6 parameters under different intervals.

TABLE II
THE DISCRETE RESULTS OF THE PARAMETERS

Parameter	Interval			
$C1$	6.1–6.30 ($V_{C1}=1$)	67.1–7.31 ($V_{C1}=2$)	8.1–8.31 ($V_{C1}=3$)	
$C2$	0:00–8:00 ($V_{C2}=1$)	8:00–16:00 ($V_{C2}=2$)	16:00–24:00 ($V_{C2}=3$)	
$C3$ [Latitide, Longitude]	[(260, 330), (99, 105)] ($V_{C3}=1$, South)	[(300, 330), (105, 108)] ($V_{C3}=2$, North)		
$C4$ (kA)	(−150)–(−100) ($V_{C4}=1$)	(−100)–150 ($V_{C4}=2$)		
$C5$ (km)	0–35 ($V_{C5}=1$)	35–80 ($V_{C5}=2$)	>80 ($V_{C5}=3$)	
D	0–0.6 ($V_D=1$, excellent)	0.6–1.0 ($V_D=2$, good)	1.0–1.6 ($V_D=3$, Medium)	1.6–2.0 ($V_D=4$, poor)

TABLE III
THE DISCRETE RESULTS OF THE PARAMETERS

No.	Rule	Conf
1	V_{C1}=1, V_{C2}=1, V_{C3}=1, V_{C4}=1, V_{C5}=3 => V_D=1	1
2	V_{C1}=2, V_{C2}=1, V_{C3}=1, V_{C4}=1, V_{C5}=3 => V_D=2	0.5
3	V_{C1}=3, V_{C2}=2, V_{C3}=1, V_{C4}=1, V_{C5}=3 => V_D=1	1
4	V_{C1}=1, V_{C2}=1, V_{C3}=1, V_{C4}=2, V_{C5}=3 => V_D=1	0.5
5	V_{C1}=1, V_{C2}=1, V_{C3}=1, V_{C4}=1, V_{C5}=1 => V_D=4	0.5
.	......	.

mined. Considering the space factor, some of the rules are given, as shown in Table III.

Analysis of the rules can yield valuable information. For example, rule 1 indicates that when lightning occurs in June, the time is between [0:00, 8:00], the location is in the south, the peak lightning current is between (−150)–(−100) kA, and the electrical distance is between 0–35 km, the voltage sag severity of the concerned node is low; the confidence of 1 indicates that the rule is completely plausible. The incremental training set is processed by discrete results and added to the original database. When the minimum support and minimum confidence are set to 0.01 and 0.5, respectively, the rules are mined using the Apriori algorithm, the AprioriTid algorithm and the algorithm in this paper, respectively. The consumption time of algorithms are shown in Table IV.

As shown in Table IV, the efficiency of the algorithm used in this paper is significantly better than the traditional algorithms. Based on the database of 2016–2019, the computation time of the algorithm in this paper is 0.4684 of the Apriori algorithm and 0.6216 of the computation time of the AprioriTid algorithm.

TABLE IV
THE CONSUMPTION TIME OF ALGORITHMS

Database	Apriori algorithm	AprioriTid algorithm	Algorithm in this paper
2016	738 ms	568 ms	265 ms
2016 – 2017	1176 ms	682 ms	354 ms
2016 – 2018	1456 ms	1010 ms	558 ms
2016 – 2019	1834 ms	1382 ms	859 ms

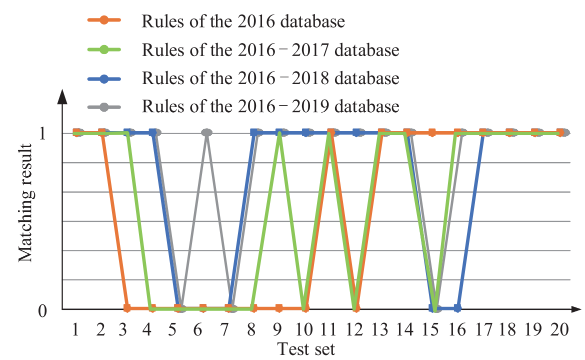

Fig. 4. Matching results based on four rule bases.

It is because this paper uses decision tables instead of TID tables, which avoids the problem of generating too many item sets. Moreover, when the database is updated, the association rule update is implemented based on the already mined rules.

C. Rule Matching

The minimum support and minimum confidence are set to 0.01 and 0.5, respectively. 13, 22, 32, and 40 strong association rules are mined using the algorithm of this paper based on four databases (i.e. 2016, 2016–2017, 2016–2018, and 2016–2019). In addition, the rules are mined using the Apriori algorithm and the AprioriTid algorithm under the same minimum support and minimum confidence. The rules mined based on the Apriori algorithm and the AprioriTid algorithm are the same as the rules mined by the algorithm in this paper. It indicates that the algorithm in this paper can improve the mining efficiency without reducing the quality of the mined rules. Therefore, the algorithm in this paper is more suitable for dynamic databases.

The weighting results of each parameter using the integrated assignment method are as follows: the weights of electrical distance, peak lightning flow, geographic location, date, and moment are 0.31, 0.24, 0.18, 0.15, and 0.12, respectively. the accuracy and validity of the mined rules are verified using the test set. Due to space limitation, this paper shows 20 actual scenarios with rule matching results, as shown in Fig. 4.

The 1 in Fig. 4 represents a correct matching result. In all test sets, the rule matching accuracy based on the database obtained for 2016, 2016–2017, 2016–2018, and 2016–2019 is 55%, 60%, 75%, and 87%, respectively. Thus, the rule matching accuracy has improved with the increase of rules. In practice by continuously expanding the database in order to improve the accuracy of the evaluation.

V. Conclusion

Making full use of the lightning and voltage sag data stored in the existing monitoring system, this paper proposes a self-learning evaluation method of voltage sag severity based on the improved association rule mining algorithm, which can obtain the voltage sag severity of the concerned nodes without the network topology. The following conclusions can be drawn.

(1) The proposed method has the ability of self-learning, which overcomes the disadvantage of the inefficiency of the AprioriTid algorithm when the dynamics of the database changes. Compared with traditional methods, it can improve the computational efficiency of the algorithm through continuous self-learning during long-term utilization. Therefore, it is more suitable for dynamic databases.

(2) The matching accuracy will improve with the accumulation of the number of rules. The method proposed in this paper can be trained by continuously accumulating voltage sag events and continuously updating the association rule base, and its matching accuracy will improve. The rule matching accuracy based on the database obtained for 2016, 2016−2017, 2016−2018, and 2016−2019 is 55%, 60%, 75%, and 87%, respectively.

References

[1] D. Raghvendraprasad and M. S. Raviprakasha, "Analysis of power quality variations in distributed generation systems," in *Renewable Energy Research*, vol. 9, pp. 281–289, 2019.

[2] M. Marzinotto, F. Fiamingo, C. Mazzetti, and G. B. L. Piparo, "Quality of the supplied electric service: A tool to evaluate the need of protection against lightning surges," in *Electric Power Systems Research*, vol. 85, pp. 75–81, Apr. 2012.

[3] Y. Wang, S. Li, and X. -Y. Xiao, "Estimation method of voltage sag frequency considering transformer energization," in *IEEE Transactions on Power Delivery*, vol. 36, no. 6, pp. 3404–3413, Dec. 2021.

[4] X. Zambrano, A. Hernández, M. Izzeddine, and R. M. de Castro, "Estimation of voltage sags from a limited set of monitors in power systems," in *IEEE Transactions on Power Delivery*, vol. 32, no. 2, pp. 656–665, Apr. 2017.

[5] A. Hernandez, E. Espinosa-Juarez, R. M. de Castro, and M. Izzeddine, "SVD applied to voltage sag state estimation," in *IEEE Transactions on Power Delivery*, vol. 28, no. 2, pp. 866–874, Apr. 2013.

[6] H. Liao, J. V. Milanović, M. Rodrigues, and A. Shenfield, "Voltage sag estimation in sparsely monitored power systems based on deep learning and system area mapping," in *IEEE Transactions on Power Delivery*, vol. 33, no. 6, pp. 3162–3172, Dec. 2018.

[7] F. Xiao and Q. Ai, "Data-driven multi-hidden markov model-based power quality disturbance prediction that incorporates weather conditions," in *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 402–412, Jan. 2019.

[8] F. Xu, K. Guo, H. Li, Y. Lin, L. Xu, C. Wang, J. Zhao, and D. Song, "Evaluation of fault level of sensitive equipment caused by voltage sag via data mining," in *IEEE Transactions on Power Delivery*, vol. 36, no. 5, pp. 2625–2633, Oct. 2021.

[9] Y. Wang, M. -H. Yang, H. -Y. Zhang, X. Wu, and W. -X. Hu, "Data-driven prediction method for characteristics of voltage sag based on fuzzy time series," in *International Journal of Electrical Power & Energy Systems*, vol. 134, pp. 107394, Jan. 2022.

[10] J. Jia, L. Qi, and H. Wang, "Research on propagation characteristics of voltage sag based on data driven and stream computing," in *Electric Power Information and Communication Technology*, vol. 19, no. 6, pp. 22–31, Jun. 2021.

[11] C. Zheng, S. Dai, B. Zhang, Q. Li, S. Liu, Y. Tang, Y. Wang, Y. Wu, and Y. Zhang, "A residual voltage data-driven prediction method for voltage sag based on data fusion," in *Symmetry*, vol. 14, no. 6, pp. 1272, Jun. 2022.

[12] L. Cai, Q. Hu, M. Zhou, Y. Ke, Q. Li, J. Wang, and Y. Fan, "Difference in lightning current estimated by lightning location system between triggered lightning to ground and distribution line," in *IEEE Transactions on Electromagnetic Compatibility*, vol. 63, no. 6, pp. 2041–2049, Dec. 2021.

[13] A. Sur and S. L. Dhar, "Power quality monitoring using virtual instrumentation & LABVIEW," in *International Journal of Engineering and Advanced Technology*, vol. 10, no. 5, pp. 197–205, Jun. 2021.

[14] A. Saxena and V. Rajpoot, "A comparative analysis of association rule mining algorithms," in *IOP Conference Series: Materials Science and Engineering*, vol. 1099, Dec. 2021.

[15] Y. -G. Lee, J. -K. Kim, and C. -H. Lee, "Analytic hierarchy process analysis for industrial application of LNG bunkering: A comparison of Japan and South Korea," in *Energies*, vol. 14, no. 10, pp. 2965, May 2021.

[16] N. Singhal, P. Joshi, and B. Mazumdar, "Entropy reduction model for pinpointing differential fault analysis on SIMON and SIMECK ciphers," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 40, no. 6, pp. 1090–1101, Jun. 2021.

[17] X. Li, Q. Han, and B. Qiu, "A clustering algorithm with affine space-based boundary detection," in *Applied Intelligence*, vol. 48, pp. 432–444, Jul. 2017.

[18] A. Al-Wakeel, J. Wu, and N. Jenkins, "*k*-means based load estimation of domestic smart meter measurements," in *Applied Energy*, vol. 194, pp. 333–342, May 2017.

[19] Z. Wu, Y. Zhang, J. Z. Zhang, K. Xia, and F. Xia, "Determining optimal coarse-grained representation for biomolecules using internal cluster validation indexes," in *Journal of Computational Chemistry*, vol. 41, no. 1, pp. 14–20, Jan. 2020.

Shuming Liu received his B.S. degree from North China University of Electric Power in 2007 and his M.S. degree from North China University of Electric Power in 2010.

Since 2011, he has been working in the power quality assessment position at the Electricity Science Research Institute of State Grid Henan Province Electric Power Company. His research interests include power quality analysis and control.

Mr. Liu's awards and honors include the First Prize of China Electric Power Science and Technology Progress, and the Second Prize of Henan Province Science and Technology Progress.

Chen Zheng received his B.S. degree from Henan University of Technology in 2013 and Ph.D. degree from Chongqing University in 2018.

He has been working in the position of power quality assessment in Electric Power Research Institute of State Grid Henan Province Electric Power Company since 2018. His research interests are power quality testing analysis and evaluation.

Mr. Zheng received the Second Prize of Science and Technology Progress of State Grid Corporation in 2021.

Bo Zhang received his bachelor's degree from Zhengzhou University in 2008 and master's degree in engineering from Zhengzhou University in 2012.

Since 2012 , he has been working as a power quality specialist in the Institute of Electric Power Science of State Grid Henan Province Electric Power Company. He has been working in the field of power quality analysis and control for a long time.

Mr. Zhang has received three Science and Technology Progress Awards from State Grid Corporation and five Science and Technology Progress Awards from State Grid Henan Province Electric Power Company.

Shuangyin Dai received his bachelor's degree from North China University of Electric Power in 2008 and his master's degree from North China University of Electric Power in 2011.

Since 2011, he has been working as a power quality specialist at the Electricity Science Research Institute of State Grid Henan Province Electric Power Company. His research interests include power quality analysis and control.

Mr. Dai's awards and honors include the First Prize of China Electric Power Science and Technology Progress, and the Second Prize of Henan Province Science and Technology Progress.

Yuzheng Tang received his bachelor's degree from Southeast University in 2010 and master's degree in engineering from Southeast University in 2013.

Since 2012, he has been working as a power quality specialist in the Institute of Electric Power Science of State Grid Henan Province Electric Power Company. He has been working in the field of power quality analysis and control for a long time.

The awards and honors received by Mr. Tang include the First Prize of China Electric Power Science and Technology Progress, and the Second Prize of Henan Province Science and Technology Progress, etc.

Yi Wang received his B.S. degree from Wuhan University in 2015 and M.S. degree from Wuhan University in 2017.

He has been working in the position of power quality assessment in Electric Power Research Institute of State Grid Henan Province Electric Power Company since 2017. His research interests are power quality testing analysis and evaluation.

Mr. Wang received the Second Prize of Science and Technology Progress of State Grid Corporation in 2021.