



Intelligent Filling Method of Power Grid Working Ticket Based on Historical Ticket Knowledge Base

Zhiguo An¹, Mancheng Yi¹, Jing Liu¹, Ying Li¹, Zheng Peng¹, Sifan Yu¹, Jianxin Liu¹, Weirong Huang¹ and Chunhua Fang^{2*}

¹Guangzhou Power Supply Bureau, Guangzhou, China, ²College of Electrical Engineering and New Energy, Three Gorges University, Yichang, China

OPEN ACCESS

Edited by:

Zhenhao Tang,
Northeast Electric Power University,
China

Reviewed by:

Zhibin Qiu,
Nanchang University, China
Xiaoyue Chen,
Wuhan University, China

*Correspondence:

Chunhua Fang
45946866@qq.com

Specialty section:

This article was submitted to
Smart Grids,
a section of the journal
Frontiers in Energy Research

Received: 12 November 2021

Accepted: 06 December 2021

Published: 24 December 2021

Citation:

An Z, Yi M, Liu J, Li Y, Peng Z, Yu S,
Liu J, Huang W and Fang C (2021)
Intelligent Filling Method of Power Grid
Working Ticket Based on Historical
Ticket Knowledge Base.
Front. Energy Res. 9:813855.
doi: 10.3389/fenrg.2021.813855

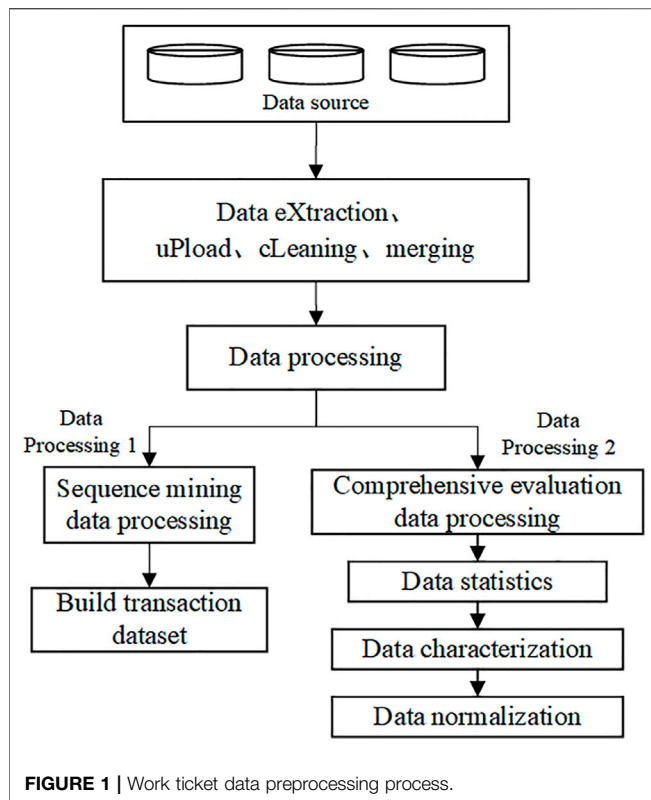
The traditional power grid ticket filling method has a large workload, low efficiency, and cannot achieve comprehensive and effective reference of historical tickets. This paper proposes a method of intelligent filling in a power grid working ticket based on a historical ticket knowledge base. Firstly, the historical ticket data are preprocessed, then the historical ticket information is mined by the association rule algorithm, and the method of establishing the historical ticket knowledge base is proposed. Based on the improved word bag model, an intelligent grid work ticket filling model is established based on the historical ticket knowledge base, and the correctness of the method is verified by an example. The results show that the accuracy of the proposed method is at least 18% higher than that of the traditional model, and the matching efficiency is 50% higher than the evaluation results of the three models based on semantic expressions. The method enables the identification and extraction of similar and associated work tickets, improves the efficiency of filling work tickets for power grids, and promotes the intelligence of the safety procedures for power grid operations.

Keywords: work ticket, intelligent filling, the knowledge base, historical ticket, data mining

1 INTRODUCTION

The historical work ticket of a power grid contains a lot of valuable information. The retrieval and utilization of the historical work ticket can effectively assist the filling of the power grid work ticket. The traditional power grid work ticket is mainly filled in by staff according to their tasks, who compare the grid wiring diagram, reference the maintenance equipment historical tickets, and combined them with their own experience. Among them, the retrieval process of historical tickets includes obtaining historical tickets with similar names by inputting the name of the maintenance equipment, and then the employees screen some work tickets based on their experience for reference. This retrieval method has a large workload, poor reliability, and low automation level. Therefore, it is of great significance to study the intelligent filling method of power grid work tickets (Liu et al., 2020; Gui et al., 2021).

Text mining, as a branch of data mining (Wang et al., 2016; Lin et al., 2017; Gao et al., 2019), can fully exploit the potential value of information and has been gradually applied to the field of electric power in recent years. An earlier study (Jiang et al., 2019) mined transformer operation and maintenance texts based on deep semantic learning, achieved text classification through text mining, and considered semantic factors in the text vectorization process to improve classification accuracy and achieve the assessment of transformer operation status. Other studies (Cao et al., 2017; Pang



et al., 2017) established a semantic framework-based text mining model to achieve classification and statistics of grid defect text. The study by Lynnette and David (2015) proposed a bag-of-words model based on FCM clustering to identify aerial targets, which improved the original “either/or” hard classification feature. The paper by Yuan and Zhou (2018) proposed a supervised bag-of-words model for multimedia information of different objects, and the text vector and retrieval results obtained by tagging the training samples were highly accurate compared with the traditional bag-of-words model. In this paper, according to the characteristics of the working ticket text, the original bag-of-words model was improved for problems such as multiple meanings of words and reversed order of words caused by irregular expressions, and an improved bag-of-words model containing both main words and auxiliary words was proposed.

Therefore, a method of intelligent filling of grid work tickets based on text mining of the historical ticket knowledge base is proposed. By analyzing the preliminary processing of historical ticket data and the establishment of a knowledge base, an improved two-layer bag-of-words model for intelligent filling of grid work tickets is constructed. The method uses Term Frequency-Inverse Document Frequency (TF-IDF) for feature extraction, and then uses cosine similarity to achieve multivariate matching between work tickets, after which the results are sorted in descending order, and finally the proposed method is proved to be more suitable for historical work tickets than other methods based on practical cases. The proposed method is finally proved to be more suitable for historical work ticket mining and work ticket filling than other methods. Therefore, the historical work

tickets that have been stored in the long-term operation of the power grid can be fully utilized to effectively reduce the error rate of power system work and enhance the efficiency of filling work tickets in the power grid.

2 BIG DATA ANALYSIS OF WORK TICKETS

2.1 Data Sources of Work Tickets

Work ticket big data are mainly distributed in three information systems, including an Equipment Asset Management System, Production Scheduling and Managing System, and Distribution Automation System. Finally, the data are extracted and summarized in the intelligent verification system of the work ticket. When acquiring the working data, the historical data are firstly extracted from the Equipment Asset Management System, and then uploaded to the intelligent verification system of the working ticket. The data tables in the historical database are sorted out, the analysis objectives are clarified, and the corresponding data files are finally selected according to the objectives (Wu and Yu, 2021).

2.2 Work Ticket Data Preprocessing

In the data mining of a work ticket, the data sample of the work ticket should be preprocessed first. The workflow of preprocessing is shown in **Figure 1**, including data cleaning, data merging, comprehensive evaluation modeling data processing, and sequence mining models data processing.

2.2.1 Data cleaning

Work ticket data generally come from different information systems, so there are many differences in the integrity and format of data (Wu et al., 2015). To comprehensively investigate the basic situation of data, data cleaning occurs from the following four aspects:

2.2.1.1 Name normalization

The format of the work ticket type and the company field is unified. The format of “affiliated company” is “**power supply company”. For example, after the format of the first kind of work ticket of B power supply company is unified, the fields are “B power supply company of the first kind of work ticket” and “B power supply company”, respectively.

2.2.1.2 Time field specification and logical judgment

The time format is unified into “year/month/day hour/minute”. Work ticket data contain the start time and end time. Generally, the end time is later than the start time, but the end time of some data is earlier than the start time, resulting in logical errors (Li, 2021). For this part of the data, the end time and start time are exchanged. If the system data only contain “operation time”, the field is changed to “start time” and the “end time” is added.

2.2.1.3 Type and quantity of violation

Data processing is classified according to the type of violation and the name of the company, and each type of violation contains statistical quantity. Therefore, it is necessary to input all kinds of

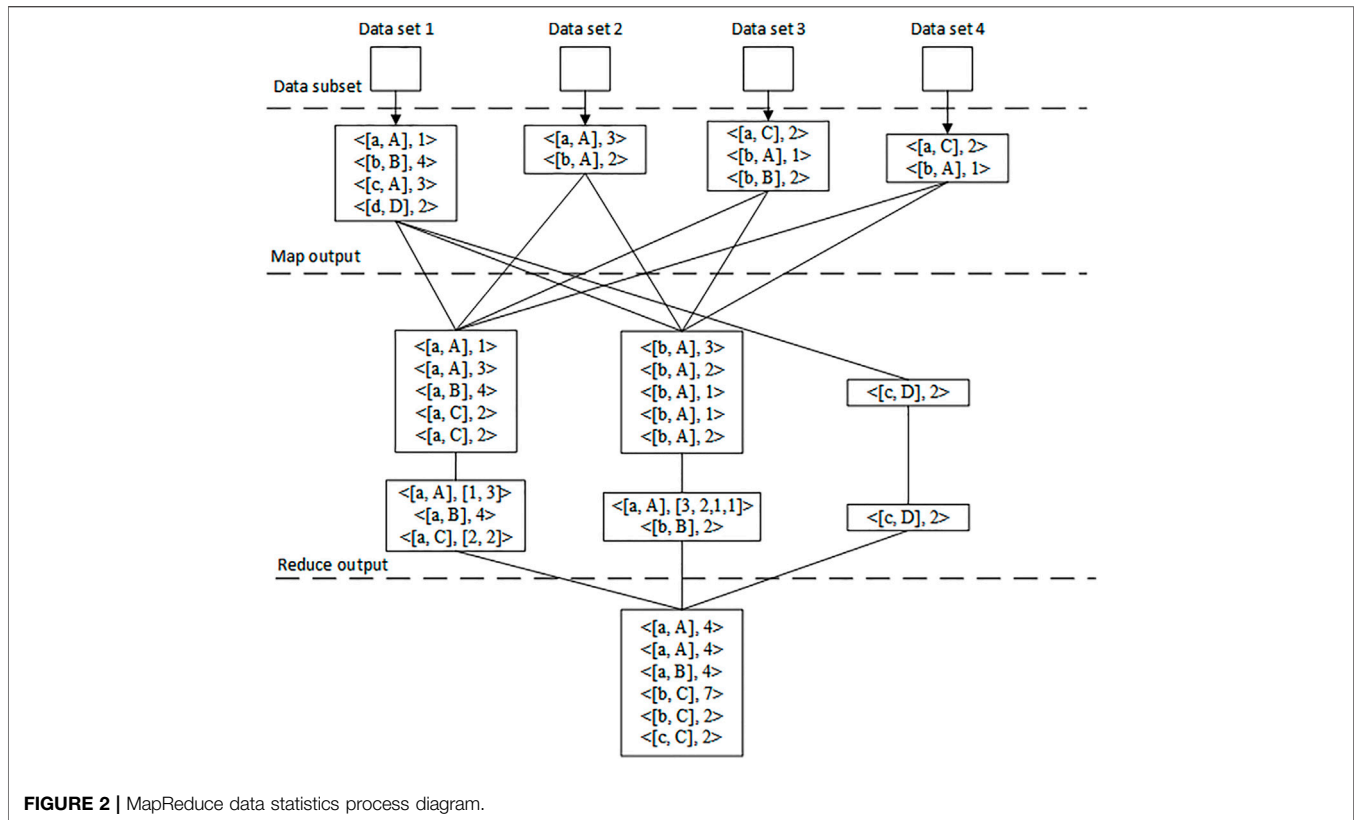


FIGURE 2 | MapReduce data statistics process diagram.

violation types into the system in a unified format to achieve accurate and rapid data classification.

2.2.1.4 Null data processing method

Because some important fields of data are null values, such as equipment name, line name, and sampling value, these null values will make it difficult for the system to carry out subsequent data mining, so these kinds of data will be screened out.

2.2.2 Data merging

After data cleaning, the data are combined to make data mining more convenient. Firstly, the fields of work ticket type, power supply company name, and planned time are extracted from the data (Zhu et al., 2003), then the fields of violation type and operating equipment are added, and the data are further divided. After data merging, the new data table contains the following five fields: work ticket type, name of power supply company, operating equipment, scheduled time, and violation type.

2.2.3 Comprehensive evaluation modeling data processing

Before the comprehensive evaluation of the data, the data need to be processed. The processing process includes: data statistics stage, data normalization, and data characterization.

2.2.3.1 Data statistics stage

In the map stage, we first define the input key value pair as < name, type >, and the intermediate key value pair generated after the map

processing task as < (power supply company, violation type), quantity >, such as < [B, b] 2 >, indicating that there are two pieces of data for the violation type b of power supply company B. In the reduce phase, all key value pairs containing the same key are sent to the same reduce phase, and then the values corresponding to the same key are added to generate new key value pairs. Finally, the results are written into HDFS. The data statistics process based on MapReduce is shown in Figure 2. A, B, and C represent three different power supply companies, respectively, and A, B, and C represent different data indicators, respectively.

2.2.3.2 Data characterization

Due to the complexity of violation types of the work ticket, when evaluating the work ticket data, if the data preprocessing is carried out in a simple statistical detail way, it is easy to ignore the objective differences existing in the work ticket, which will affect the final evaluation results (Wang et al., 2016). Therefore, it is necessary to preprocess the data according to the characteristics of different data.

2.2.3.3 Data normalization

In order to avoid the impact of data-level differences on evaluation results, Eq. 1 is used to normalize the data.

$$y_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

In this formula, y_i is the normalized data; x_i is an attribute value of the original data; and x_{\min} and x_{\max} represent the minimum and maximum values for this property.

TABLE 1 | Serial transaction database.

Affair	Sequence
S1	[(a, c), b, e, f]
S2	(g, a, e)
S3	(g, a, b, f)
S4	(d, a)
S5	(a, b)

2.2.3.4 Sequence mining models data processing

In sequence mining, the corresponding transaction database should be established first. Then we use MapReduce to set up the transaction database. The setup steps are as follows:

- 1) Map input definition key value pair < name, type >, map function output key value pair < line, (begintime, name) >, the specific format is < name of power supply company, (operation equipment, violation type) >.
- 2) The reduce phase obtains key value pairs with the same key from the map task, and sorts the corresponding values in chronological order. For example, < Line1, list < (begintime1, Name2), (begintime2, name3), (begintime3, name1) >.
- 3) We use the map function to input key = name of power supply company, value = list, and output key = name of power supply company, value = violation type. Finally, we store the results. The above examples are: < Line1, (name2, name3, name, 1) >. The final sequence transaction database is shown in **Table 1**.

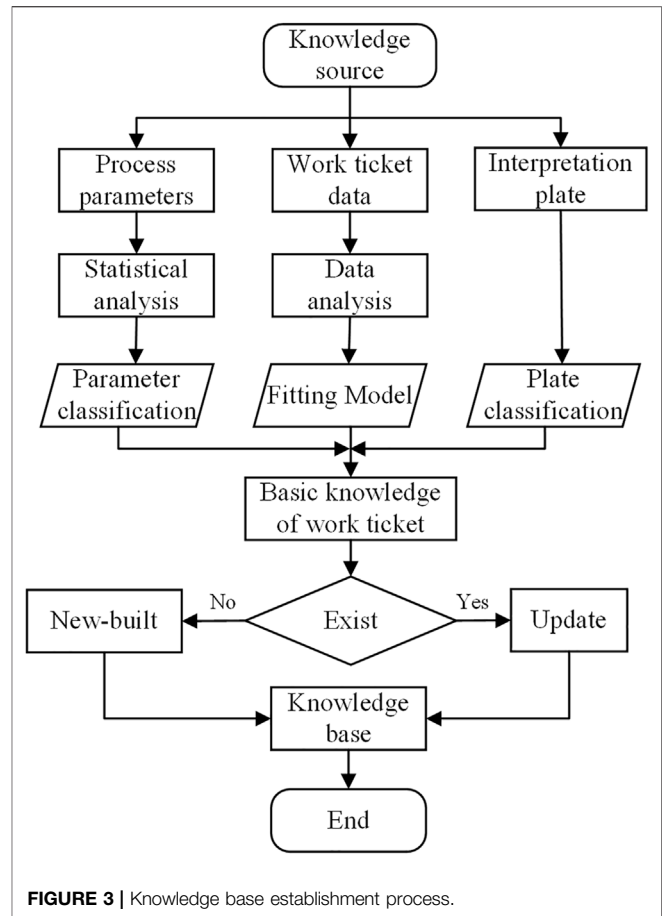


FIGURE 3 | Knowledge base establishment process.

3 INTELLIGENT FILLING MODEL OF WORK TICKETS

3.1 The Establishment of the Historical Ticket Knowledge Base

The establishment of the historical ticket knowledge base includes extracting historical ticket processing and interpretation information such as work ticket data, processing parameters, and interpretation templates from the knowledge source, converting them into specific computer code, and finally obtaining the preferred parameters for work ticket processing and interpretation (Jiang et al., 2019). Knowledge acquisition includes passive and active methods: 1) Active knowledge acquisition automatically obtains historical ticket processing interpretation rules based on accumulated ticket processing interpretation information and stores them in the knowledge base. For example, data fitting is carried out on the parameter frequency of the ticket, and the confidence interval method is used to statistically analyze the fitting results. 2) Passive knowledge acquisition uses the information base editor to store the logic rules of historical tickets in the knowledge base. When the working ticket processing and interpretation parameters are the knowledge source, the work ticket knowledge base conducts statistical analysis and classification of the parameters, and integrates the priority of the parameters. When the work ticket data are the knowledge source, the work

ticket fitting model is established according to the work ticket data, and the classical formula parameters are obtained by using the work ticket data.

The establishment process of the knowledge base is shown in **Figure 3**. After preprocessing the knowledge source data, the data items are composed of a set $I = (I_1, I_2, \dots, I_n)$. The logical rules of the data can be expressed as follows:

$$\begin{cases} A \geq B \\ A \subset I, B \subset I, A \cap B = \Phi \end{cases} \quad (2)$$

In this formula, A is the antecedent of the logical rule, B is the consequent of the logical rule, and Φ is an empty set.

Support S is used to measure the applicability of logical rules, and reliability C is used to measure the accuracy of logical rules. P is the probability of logical rules, then:

$$\begin{cases} S(A \geq B) = P(A \cup B) \\ C(A \geq B) = P(A \cup B) / P(A) = P(B|A) \end{cases} \quad (3)$$

A frequent set is a set of items whose specified threshold is less than the support count. The information acquisition of the historical ticket includes two steps: 1) Finding all frequent sets of history tickets whose frequency is greater than or equal to the minimum support threshold. 2) The corresponding history ticket knowledge is generated by using frequent sets, and this

knowledge must meet the minimum trust threshold. The parameter model of the work ticket knowledge base is established by using the association rule algorithm, and the frequent set is calculated. For instance, in the process of obtaining information on historical tickets, the Archie formula can be chosen when the frequent set is less than 20%. Each item of the frequent set must be arranged according to the generated time order, and the specified time interval must be greater than the generated time difference (Zhu et al., 2003). In addition, minimum trust and minimum support must be set. Setting the minimum trust C_{min} to 70% and the minimum support S_{min} to 35%, the association rules of the model must meet the above conditions. If the logical association rule obtained during mining meets the following constraints, the association rule is acceptable.

$$\begin{cases} S(V_{sh} < 20\%, Archie) \geq 35\% \\ C(V_{sh} < 20\%, Archie) \geq 70\% \end{cases} \quad (4)$$

The significance of S in Eq. 4 is that at least 35% of all model preferences exhibit behavior where both the frequent set of less than 20% and the Archie formula are used together, and the significance of C in this example is that at least 70% of the frequent set of less than 20% model preferences use the Archie formula together. Therefore, if there is a preference for an explanatory parameter with a frequency set of less than 20%, the knowledge base will be able to recommend that the user prefers the Archie formula.

3.2 Establishment of Intelligent Filling Model

Matching between texts is performed by relying on the similarity calculation of text features. Text vector feature extraction using the TF-IDF method, which has a wide range of applications in finding text, text classification, and other similar fields, has been called the most meaningful creation in information retrieval. In this paper, the multivariate similarity calculation between the feature vectors of query text and historical text is performed using the cosine function after deactivating words, splitting words, vectorization, and feature extraction of historical text and query text (Shang et al., 2015). The n-dimensional vector cosine similarity calculation formula is shown in Eq. 5.

$$S_{\cos}(a, b) = \frac{ab}{|a||b|} = \frac{\sum_{i=1}^n a_i b_i}{\sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{i=1}^n b_i^2}} \quad (5)$$

$a = (a_1 a_2 \dots a_n)$, $b = (b_1 b_2 \dots b_n)$ are the query text and historical text vectors, respectively, and a_i , b_i are the TF-IDF values corresponding to the i -th phrase in the text. The larger $\cos\theta$ is, the higher the similarity is. Since each variable in the vector is positive, the similarity takes values between 0 and 1.

In order to allow the computer to process text information, each word of the text is mapped into a number set. A word bag model is a common method of text processing. Firstly, the document set S is merged by Eq. 6 to generate word bag T_0 ; then, the text vocabulary in the document is counted to obtain the text vector D_V composed of a number.

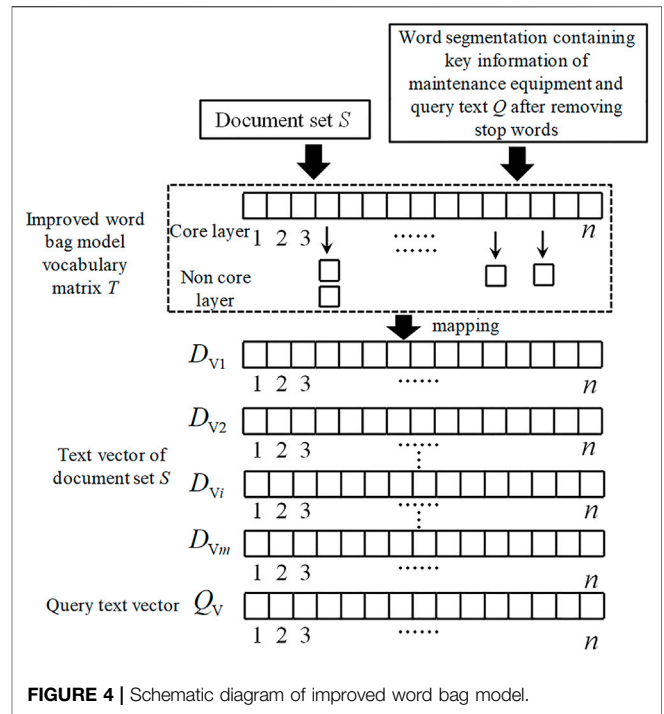


FIGURE 4 | Schematic diagram of improved word bag model.

$$T_0 = [t_{01} t_{02} t_{03} \dots t_{0n}] \quad (6)$$

In this formula, $t_{01}, t_{02} \dots t_{0n}$ are the words in D , respectively.

In the process of filling in work tickets, word order is often confused due to non-standard filling, which leads to the decrease of similarity (Han et al., 2016; Du et al., 2018). Therefore, this paper improves the original word bag model T_0 and adds a vocabulary relation expansion table on the basis of the original word bag model. The text is divided into a subject core layer and auxiliary non-core layer, and the vocabulary matrix T is expressed as:

$$T = \begin{bmatrix} t_{01} & t_{02} & \dots & t_{0n} \\ t_{11} & t_{12} & \dots & t_{1n} \\ \dots & \dots & \dots & \dots \\ t_{m1} & t_{m2} & \dots & t_{mn} \end{bmatrix} \quad (7)$$

In this formula, the first line is the core layer of the subject, and the second line to the m line is the non-core layer of the auxiliary word.

In addition, on the basis of the original bag model, each text in S is successively mapped to n-dimensional vectors, as shown in Eq. 8:

$$\begin{cases} D_{V1} = [d_{11} d_{12} \dots d_{1j} \dots d_{1n}] \\ D_{V2} = [d_{21} d_{22} \dots d_{2j} \dots d_{2n}] \\ \dots \\ D_{Vi} = [d_{i1} d_{i2} \dots d_{ij} \dots d_{in}] \\ \dots \\ D_{Vk} = [d_{k1} d_{k2} \dots d_{kj} \dots d_{kn}] \end{cases} \quad (8)$$

In this formula, k is the number of power grid safety specification texts in the historical ticket knowledge base; d_{ij} is

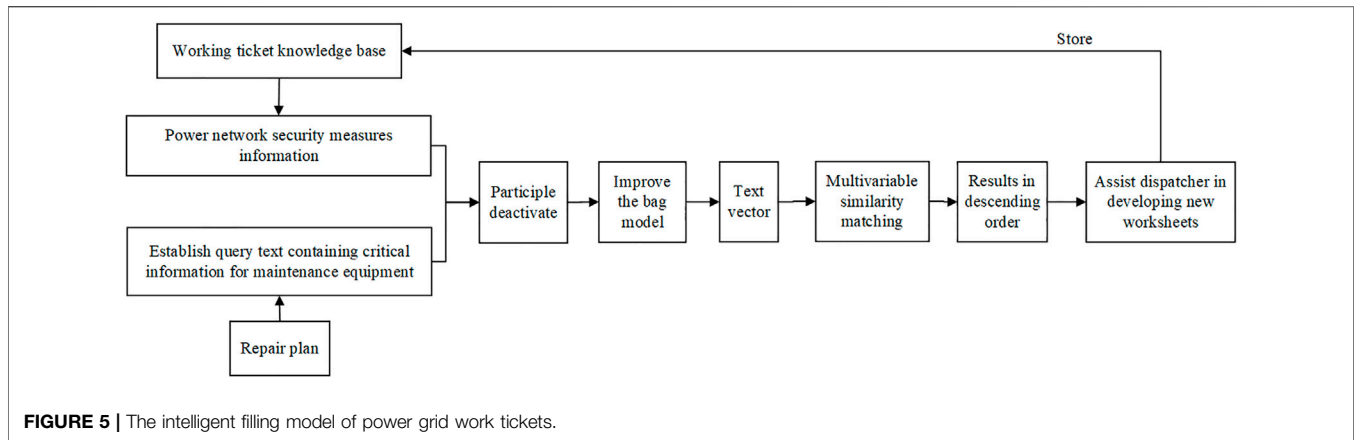


FIGURE 5 | The intelligent filling model of power grid work tickets.

the word frequency of the j -th word in the word bag model i ; when the text contains an auxiliary word, we put the corresponding word frequency of the auxiliary word into the position of the subject word frequency, and add the word frequency. The working principle diagram is shown in **Figure 4**.

In order to improve the efficiency of power grid work ticket filling and promote the intelligent development of the power grid, an intelligent power grid work ticket filling model based on the historical ticket knowledge base is built on the basis of the improved word bag model. The process of the model is as follows: 1) Analyze and obtain the text features of the work ticket, and mine the potential power grid security measures; 2) preprocess the unstructured text and segment the text according to the knowledge of the power field; 3) the improved model is applied to the vectorization of words after word segmentation; 4) extract the features of the text vector; and 5) cosine similarity is used to realize variable matching and descending arrangement of text. Employees can refer to the historical work ticket text at the top of the order. The intelligent filling model of the power grid work ticket is shown in **Figure 5**.

4 CASE ANALYSIS

4.1 Case Analysis 1

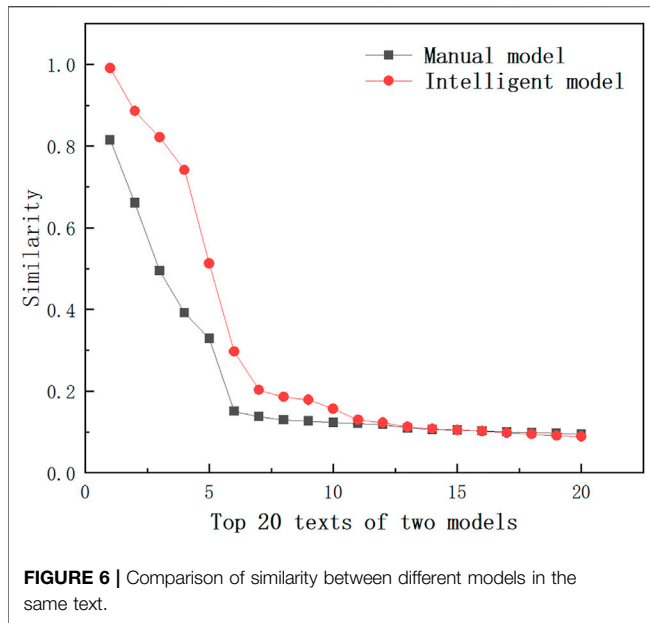
In order to demonstrate the accuracy of the proposed method for filling grid work tickets intelligently based on text mining of the historical ticket knowledge base. Taking the historical working tickets of power transmission and transformation equipment maintenance of a Power Supply Bureau in 2020 as an example, the work tickets are 2,542 in total. We take the key information of the equipment to be repaired as the query content, such as: “CT current rise of G02 cabinet of group junction no. 1 outdoor switch station of 10 kVA line, voltage withstand test of cable from group junction no. 1 outdoor switch station to no. 10 European style box transformer”. After being segmented, it can be expressed as “10 kVA line, group junction, no. 1 outdoor switch station, G02 cabinet, CT upflow, group junction, no. 1 outdoor switch station, no. 10 European style box transformer, cable, voltage withstand test”. The combination of historical ticket keyword association and

TABLE 2 | Top six historical texts in similarity ranking.

Text serial number	Similarity	Similarity ordering
36	0.956	1
53	0.901	2
18	0.874	3
79	0.793	4
3	0.705	5
48	0.608	6

plant station topology is used for analysis, the optimal parameters are obtained and returned to the work ticket application. The features of each text are extracted through the text similarity TF-IDF model, and the work ticket knowledge base is called to calculate the text similarity of the work task. The descending order of historical text similarity is obtained through intelligent matching, the top six texts are shown in the following table.

In **Table 2**, the text with serial number 36 has the highest similarity, and the corresponding historical text is “voltage withstand test for cable from outdoor switch station no. 1 to switch room no. 14 of 10 kVA line”. The historical text of serial number 53 is “replace the transformer of no. 6 complex room of 10 kVA line”, it is the same as the power failure line in the query content, so the power failure scope, power failure layout, and power failure time can be used as reference information to form a certain rule. The text no. 18 is “cable head made for cabinet G02 of Outdoor Switch Station no. 1 of 10 kVA Line Cluster”. “Power cut line”, “switch cabinet”, and “power cut equipment” are similar to the query content, so the reference content can be filled in the corresponding column of the work ticket. The text of no. 79 corresponds to “voltage withstand test of newly laid cable from G02 cabinet in no. 2 switch room of 10 kVA line to newly installed G13 cabinet in this room”. The work content in the historical text is also the cable withstand voltage test, which is similar to the work content in the text, and can be used as a reference when the staff arranges safety measures and precautions, and can be combined with the text no. 18 content, making the work ticket content more manageable. The text of no. 3 corresponds to “CT current rise test of G04



traditional model, and the accuracy is at least 18% higher than the traditional model.

4.2 Case Analysis 2

To prove that the proposed intelligent filling method of power grid work tickets based on historical ticket knowledge base text mining can improve the filling efficiency, the historical work tickets of transmission and substation equipment overhaul in a Power Supply Bureau in 2020 are analyzed as an example, with a total of 2,542 work tickets. The key information of five groups of overhaul equipment is selected as query items, where the average length of query items is 10 and the average length of documents is 500. Each query item will index the top five texts with similarity ranking, and after that, MAP (Mean Average Precision) and NDCG (Normalized Discounted cumulative gain), which are commonly used in the field of information retrieval, are used to match the text with the similarity ranking. Cumulative gain indexes commonly used in the field of information retrieval are used to evaluate the matching results, and the average of the five evaluation results is used as the final evaluation index, expressed as η . Where, the larger

TABLE 3 | Results of the model in the matching of similar and correlated work tickets.

Category	Models	η (MAP)	η (NDCG)
Traditional features	BM25 model	0.23	0.4
Based on single semantics	DSMM model	0.08	0.2
Text expression	CDSSM model	0.06	0.14
Multi-semantic text-based representation	MV-LSTM model	0.12	0.18
Improving traditional features	The model proposed in the paper	0.28	0.47

standby cabinet in 10 kV 5 switchgear room”, content similarity is “switchgear related tests”, and the precautions for power failure can be used as an appropriate reference. The text of serial no. 48 is “switch room no. 11 of 10 kVA line is newly installed with 5 M busbar and four sides of newly installed open pressure cabinets (cabinets G11 to G14)”, which has low similarity with the query content and is of no value.

According to the historical work ticket data of transmission and substation equipment maintenance in a Power Supply Bureau in 2020, 2,542 tickets were issued. Using the model proposed in the previous paper, the similarity comparison graph shown in **Figure 6** is drawn; the red curve is the similarity size under the improved model and the black curve is the similarity size under the traditional model. The similarity between the first four texts and the query text under the improved model is high and has a certain degree of connectivity; the similarity of the fifth text decreases sharply and continues to decrease in the latter. The similarity of the first two texts under the traditional model is higher.

It can be seen that the results under the traditional model have no obvious transition phase and are less accurate, while the text similarity results under the improved model have distinct boundaries. Dispatchers have clear targets when extracting work tickets, the similarity size is generally higher than the

the average value, the more accurate the matching result is (Wang et al., 2014; Umagandhi and Senthil Kumar, 2015).

The proposed model is compared with the traditional BM25 model, Deep Semantic Structured Model (DSSM) based on single-semantic text representation, and Convolutional Deep Semantic Structured Model (CDSSM). The Multi View-Long Short Term Memory (MV-LSTM) model based on multi-semantic representation is compared. The results of the evaluation metrics are shown in **Table 3**.

From **Table 3**, the MAP and NDCG index values of the model in this article are the largest among the average of the five evaluation results, the matching results are the best, and the effectiveness is the highest; the evaluation results of the three models based on semantic expression are very low, among which the MV-LSTM model has better evaluation results. Both MAP and NDCG values are lower than 50% of the models mentioned in the article.

5 CONCLUSION

By analyzing the preprocessing of work ticket data and the establishment of a knowledge base, this paper constructs an intelligent filling model of power grid work tickets based on

the historical ticket knowledge base, and verifies the specificity and effectiveness of the proposed method combined with an actual power grid case. The main conclusions are as follows:

- 1) Using the association rule algorithm to mine historical ticket knowledge, we can obtain the key words of historical ticket information effectively, and then assist the staff to carry out comprehensive referencing.
- 2) According to the non-standard expression of the work ticket text, the original word bag model is improved to reduce the similarity error caused by environmental noise and improve the matching accuracy.
- 3) The case study shows that the accuracy of the proposed model in the paper is at least 18% higher than that of the traditional model, and the matching efficiency is 50% higher than that of the evaluation results of the three models based on semantic expressions. Therefore, the work ticket management

application module can be automated, saving time for filling and verifying tickets and improving work efficiency.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

ZA, MY, and JL (3rd author) conceived the idea and designed the experiments. YL and ZP led the experiments. SY, JL (3rd author), and WH contributed to data analysis and interpretation. CF and ZA wrote the paper. All authors read and approved the final manuscript.

REFERENCES

- Cao, J., Chen, L., Qiu, J., Wang, H., Ying, G., and Zhang, B. (2017). Semantic Framework-Based Text Mining Technique for Grid Defects and its Application. *Power Grid Technol.* 41 (02), 637–643. doi:10.13335/j.1000-3673.pst.2016.1044
- Du, X., Qin, J., Guo, S., and Yan, D. (2018). Text Mining of Typical Fault Cases of Power Equipment. *High Voltage Technol.* 44 (04), 1078–1084. doi:10.13335/j.1003-6520.hve.20180329005
- Gao, X., Chen, G., and Zhao, H. (2019). Data Mining-Based Cyber Security Posture Assessment of Electric Power Information Systems. *Electr. Meas. Instrumentation* 56 (19), 102–106. doi:10.19753/j.issn1001-1390.2019.019.017
- Gui, Q., Yao, X., Wang, L., Shao, Z., Yang, H., Hao, Z., et al. (2021). Intelligent Error Prevention Technology of Maintenance Work Ticket Based on Layered Finite State Machine. *J. Hefei Univ. Technol. (Natural Sci. Edition)* 44 (04), 458–463.
- Han, W., Fan, S., Wei, Z., and Wang, W. (2016). “A Method for Similarity Analysis of Power System State,” in IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), 2033–2036. doi:10.1109/appeec.2016.7779842
- Jiang, Y., Li, L., Li, Z., Su, C., Wang, G., and Peng, S. (2019). A Deep Semantic Learning-Based Text Information Mining Method for Power Transformer Operation and Maintenance. *Chin. J. Electr. Eng.* 14, 4162–4172. doi:10.13334/j.0258-8013.pcsee.181457
- Li, S. (2021). Intelligent Operation Ticket System for Power Grid Dispatch Based on Hierarchical Structure Knowledge Base and Reasoning Machine Technology. *Electric Age* 23 (04), 52–54.
- Lin, S., Xie, C., Tang, B., Pan, A., and Zhou, J. (2017). Application of Data Mining in Power Quality Monitoring Data Analysis. *Electr. Meas. Instrumentation* 43 (09), 46–51.
- Liu, T., Li, S., Gu, X., Wang, T., Wu, Y., and Wang, W. (2020). Research on Text Mining Based Power Grid Maintenance Work Ticket Assisted Decision Method. *Electr. Meas. Instrumentation* 57 (02), 39–45. doi:10.19753/j.issn1001-1390.2020.002.006
- Lynnette, P., and David, S. (2015). Accounting Variables, Deception, and a Bag of Words: Assessing the Tools of Fraud Detection. *Contemp. Account. Res.* 32 (3), 1193–1223. doi:10.1111/1911-3846.12089
- Pang, L., Lan, Y., Xu, J., Guo, J., Wan, S., and Cheng, X. (2017). A Review of Deep Text Matching. *J. Comput. Sci.* 40 (04), 985–1003. doi:10.27312/d.cnki.gshsu.2020.001431
- Shang, F., Yuan, Y., Wang, C., Cao, M., and Feng, Z. (2015). A Knowledge-Base Based Method for Intelligent Preferential Logging Data Processing of Interpretation Models. *J. Pet.* 36 (11), 1449–1456.
- Umagandhi, R., and Senthil Kumar, A. V. (2015). Evaluation of Reranked Recommended Queries in Web Information Retrieval Using NDCG and CV. *Int. J. Inf. Technol. Comput. Science(IJITCS)* 7 (8), 23–30. doi:10.5815/ijitcs.2015.08.04
- Wang, S., Bi, S., Xu, Y., and Sun, Y. (2016). Short-term Load Forecasting Based on Data Mining Techniques and Support Vector Machines. *Electr. Meas. Instrumentation* 15 (10), 62–67.
- Wang, Y., Huang, Y., Lu, M., Pang, X., Xie, M., and Liu, J. (2014). A Multi-Sorting Model Fusion Approach for Direct Optimization of Performance Metrics. *J. Comput. Sci.* 42 (08), 1658–1668.
- Wu, J., and Yu, W. (2021). A Study of Text Clustering with Fused Knowledge Base Semantics. *J. Intelligence* 35 (05), 156–164. doi:10.3778/j.issn.1002-8331.2104-0286
- Wu, K., Liu, W., Li, Y., Su, Y., Xiao, Z., Pei, X., et al. (2015). Cloud-based Power Big Data Analysis Technology and Application. *China Electric Power* 24 (02), 111–116+127.
- Yuan, G., and Zhou, X. (2018). A Supervised Bag-Of-Words Model Based on Multimedia Information Retrieval. *Comput. Eng. Des.* 32 (09), 2873–2878. doi:10.16208/j.issn1000-7024.2018.09.031
- Zhu, Y., Zhang, J., Yang, Z., Ji, X., and Yang, Y. (2003). An Object-Oriented General Expert System for Work Tickets and Operation Tickets in Agricultural Network Substations. *Power Grid Technol.* 24 (04), 27–30.

Conflict of Interest: ZA, MY, JL, YL, ZP, SY, JL, and WH were employed by the company Guangzhou Power Supply Bureau.

The remaining author declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 An, Yi, Liu, Li, Peng, Yu, Liu, Huang and Fang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.