

Research on Short-term Prediction Method of Substation Bus-bar's Voltage Trend Based on Multidimensional Time Series Data Mining

Fei Xiao¹, Xiwu Leng², Kang Ye¹, Youlin Hu¹, Xiongli Li³, Licheng Zhu³

¹ State Grid Shanghai Electric Power Company Power Dispatch Control Centre, Shanghai, P.R. China

² National Power Dispatch Communication Centre, Beijing, P.R. China

³ Tellhow Software Co. Ltd., Nanchang, P.R. China

E-mail: xiaofei@sh.sgcc.com.cn; xiwu-leng@sgcc.com.cn;
yekang_yk@sh.sgcc.com.cn; 3311481@qq.com; 66653752@qq.com;
zhulc@tellhow.com

Abstract. The short-term prediction of voltage trend is the important technical foundation of the voltage exceeding intelligent alarm system in power substation. According to correlation between the adjacent lines in the same substation, the time series data of different bus-bars in same section, was constructed into a multi-dimensional time series data matrix model for prediction of voltage trend. Moreover, based on above matrix model, a novel prediction method based on multidimensional time series data was proposed, which transformed the multidimensional time series data matrix into a classical two-dimensional decision information system in the first stage via preprocessing and clustering. In the second stage, various classical machine learning algorithms are ensembled to forecast the short-term future trend of the specific bus-bars' voltages. The efficiency of the prediction method based on multi-dimensional time series data mining was validated by the implement in a 500 KV bus-bar's prediction in a Substation of the State Grid Shanghai Company. The results represented that the prediction method in this paper had sound precision in practice and can improve the functionality of voltage exceeding intelligent alarm system via assistance to filter plenty of fake alarms.

1. Introduction

The short-term prediction of bus-bars' voltage trend in substation is basic technology of the voltage exceeding intelligent alarm system in power substation. At present, the voltage exceeding intelligent alarm system in our country mainly depended on the real-time alarm data from the SCADA system of power grid D5000 and the operation state data of voltage control system from AVC system. However, in practice, the voltage value of bus-bars' voltage exceeding alarm preset in D5000 system and the busbars' voltage automatic regulating preset in AVC system are always inconsistent. D5000 system can deliver a large number of alarms in a short time, which is mixed with the true and fake alarms. Thus, the monitor in control centre almost cannot identify the effective alarm information quickly, and the plenty of fake alarms are seriously affecting the efficiency and functionality of the voltage exceeding alarm system.

Based on the traditional voltage exceeding alarm technology, this paper proposed an effective intelligent alarm method for busbars' voltage exceeding alarm with busbars' voltage short-time prediction technology. Furthermore, the principle of this method is to forecast the busbar's voltage



trend in short-term future when the D5000 system delivered alarms and the AVC reactive power regulation capacity has been exhausted. If the trend is upward in short-term future, the alarm is a real alarm, otherwise it is a false alarm. The practical results in a substation in Shanghai showed that this method can effectively remove a large number of fake alarms and improve the efficiency of busbars' voltage exceeding alarm system.

The traditional methods of voltage prediction include least squares estimation [1], power flow estimation [2], sensitivity matrix estimation, and public coupled point voltage estimation [3], which is suitable for small and medium-sized electricity grid with only partial information of grid. In the era of big data, the data acquisition of the power grid is increasing rapidly, and the information has been improved. However, these information with multi-source, time series, heterogeneous structure and diverse characteristics are difficult to be implemented into the existing traditional method of voltage prediction at present. Y. Wei proposed a voltage prediction algorithm based on the dynamic pattern matching of voltage time-series data [4]. B. Wei proposed a method to predict the average output voltage of the system based on BP Neural Network [5]. the methods based on non-parametric estimation and adaptive neural network is proposed to realize the voltage rapid estimation [6], [7]. However, in practice, the voltage prediction methods, which based on a single algorithm or one-dimensional data of grid, cannot figure out the dynamic characteristics of the data from large power grid. The stability and accuracy of their prediction are seriously limited.

Given the defects of traditional voltage prediction, this paper constructed a multi-dimensional time series data model and utilized ensemble learning strategy with various classical data mining algorithms to establish a new framework of busbars' voltage prediction. The research results have been successfully implemented in busbars' voltage exceeding intelligent alarm system of a substation in Shanghai.

2. The features of voltage trend prediction data

The data of voltage trend prediction involving working conditions of related power lines in substation were shown in Fig. 1, which included the historical voltage time series data of busbar, the history voltage time series data of relevant power lines, history remote signal displacement data of related switch, the operation records of relevant AVC system, the historical data of environmental temperature and so on. Firstly, all multiple historical time series data of working conditions were segmented by sliding windows with certain time span in the paper. Moreover, the voltage trend in short-term future was taken as a predictive attribute, which have 3 decision values as upward trend (recorded as P), downward trend (recorded as N) and stable trend (recorded as B).

After data preprocessing, a data model was established as following. A specified relevant L data set were represented to $D = \{D_1, D_2, \dots, D_L\}$, $D_i \in D$, Among them, D was the time series

information system shown in Fig.1, and can be describe as $D_i = (A_i, V_i, N, f, g)$.

$A_i = \{a_{i1}, a_{i2}, \dots, a_{im}\}$ means the attribute i time series dataset of busbar working condition with m segments. $V_i = \{TS_{i1}, TS_{i2}, TS_{i3}, \dots, TS_{im}\}$ stands for the attribute i time series dataset of

transformer working condition with m segments. f is a mapping relationship: $f: a_{ij} \rightarrow TS_{ij}$.

$N_i = \{N_1^k, N_2^k, \dots, N_m^k\}$ represents the dataset of busbar voltage trend values for a continuous K -minute after each historical working condition V_i . Its value is determined by the mapping $g: TS_{ij} \rightarrow N_j^k$, where $j \in \{1, 2, \dots, m\}$. In this paper, finally, the future k -minute voltage trend

predictive value $N = \{N_1, N_2, \dots, N_m\}$ synthesize into the corresponding time period, and the voltage trend data matrix model DB as shown in following equation is the training dataset of predictive model.

$$DB = \begin{bmatrix} TS_{11} & TS_{21} & L & TS_{L1} & N_1^k \\ TS_{12} & TS_{22} & L & TS_{L2} & N_2^k \\ M & M & M & M & \\ TS_{1m} & TS_{2m} & L & TS_{Lm} & N_m^k \end{bmatrix}, j \in \{1, 2, L, m\}, i \in \{1, 2, L, L\}$$

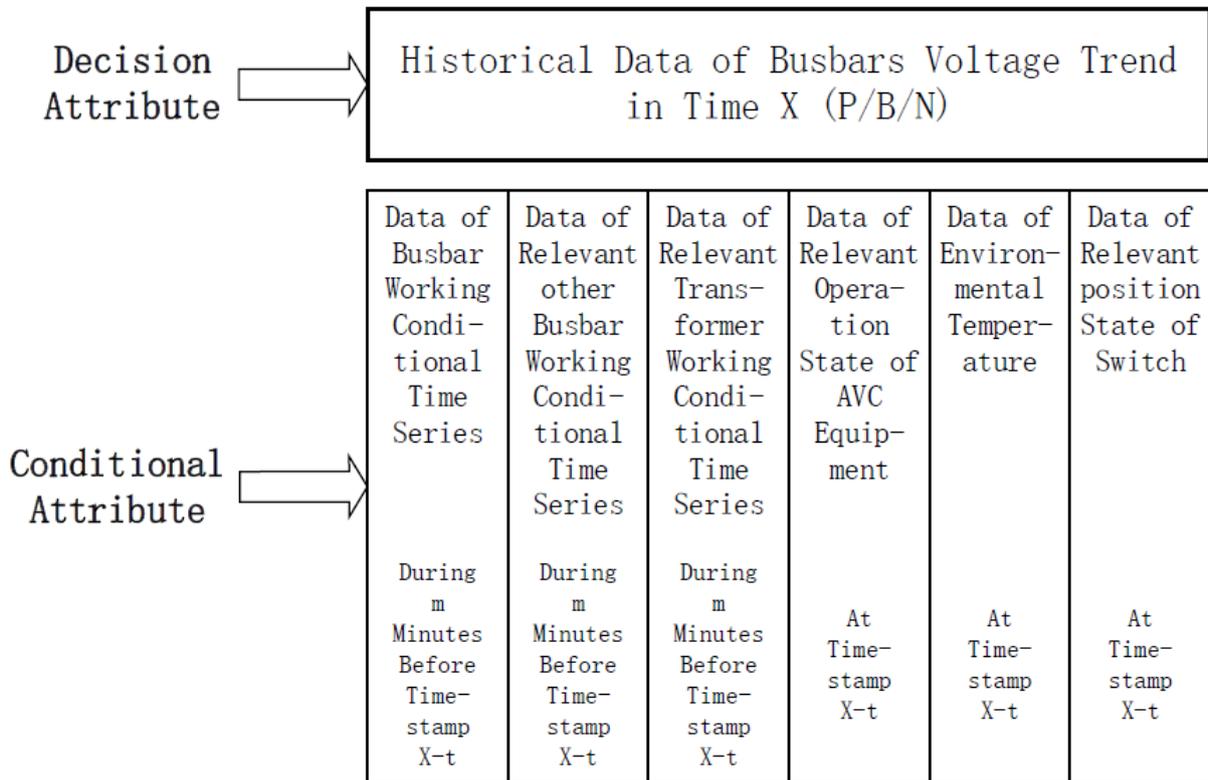


Figure 1. Diagram of data model of relevant busbar's voltage prediction

3. The method for short-term prediction of busbars' voltage trend based on multi-dimensional time series data mining

The short-term prediction method of busbars' voltage trend based on multidimensional time series data mining is shown in Figure 2 as following. Firstly, the multi-dimensional time series dataset of all relevant historical dataset for busbars' voltage trend prediction were modelled and transformed into multi-dimensional time series matrix model. After that, the time series clustering method was used to transform and reduce the multi-dimensional time series matrix model into the classical two-dimensional information table. Then, the classical two-dimensional information system was imported into multi-machine learning group. Finally, an optimal classifier was generated via competition of the multiple algorithm models (ANN, C4.5, SVM, etc.) based on test dataset.

As shown in Fig. 2 and Fig. 3, the short-term prediction model for the busbar voltage trend prediction in this paper was based on 2 critical algorithms, dimensional reduction algorithm based on DTW [6]-[9] (Dynamic Time Wrapping), and ensemble learning [10]-[14] algorithm.

The steps of short-term prediction model of busbar voltage trend are described as follows:

Input: a set D of historical data of L -Lines related to a given busbar,

Output: optimal classifier.

Step 1: The Set D of historical data was preprocessed and transformed into relevant time series dataset of decision information system for busbar voltage trend prediction.

Step 2: The time series dataset was clustered for each column of time series data (except the decision column) in the DB matrix.

Step 3: Based on the clusters of Step 2, DB matrix was transformed into a two-dimensional decision information table IS_{DB} , in which any C_{xy} is a TS_{xy} cluster type.

Step 4: Based on the two-dimensional decision information table IS_{DB} of step 3, the classical machine learning classification algorithm (SVM, RNN, C4.5 decision tree, etc.) was training to model classifier.

Step 5: Based on the decision knowledge and classifier models of step 4, the bagging strategy of ensemble learning [12]-[17] algorithm was utilized to choose the optimal classifier model

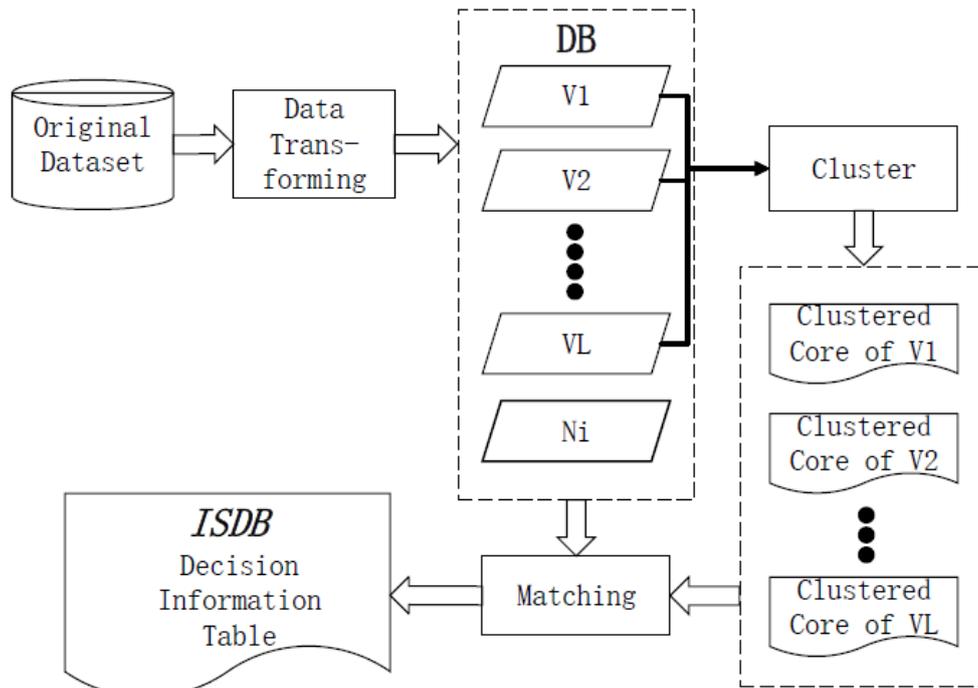


Figure 2. Flowchart of DTW dimensional reduction algorithm

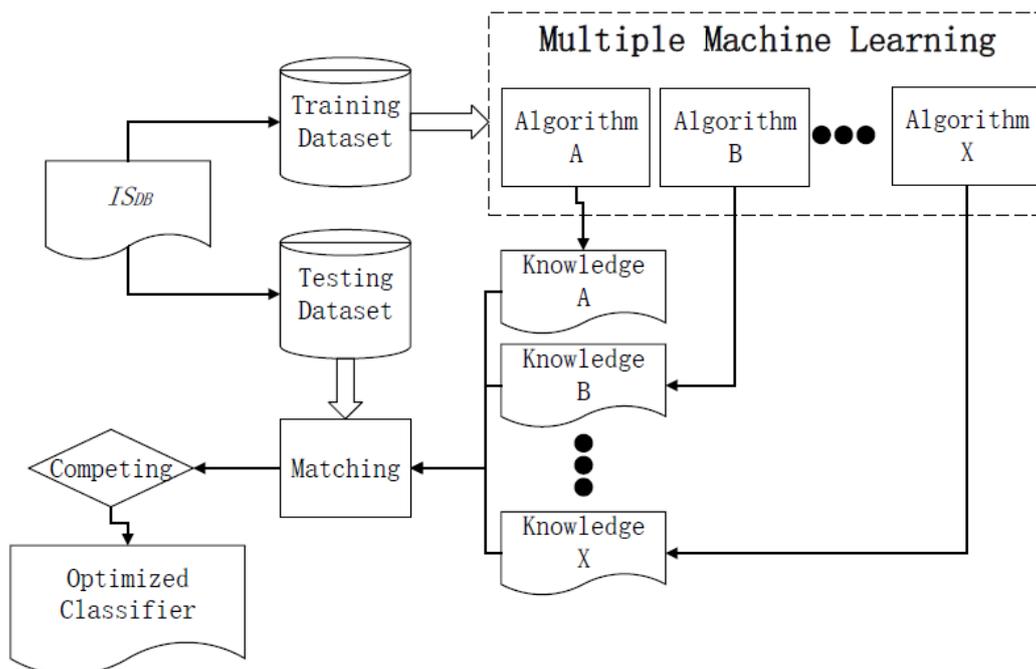


Figure 3. Flowchart of ensemble learning algorithm

$$DB = \begin{bmatrix} TS_{11} & TS_{21} & L & TS_{L1} & N_1^k \\ TS_{12} & TS_{22} & L & TS_{L2} & N_2^k \\ M & M & M & M & \\ TS_{1m} & TS_{2m} & L & TS_{Lm} & N_m^k \end{bmatrix} \xrightarrow{\text{Cluster}} IS_{DB} = \begin{bmatrix} C_{11} & C_{21} & L & C_{L1} & N_1^k \\ C_{12} & C_{22} & L & C_{L2} & N_2^k \\ M & M & M & M & \\ C_{1m} & C_{2m} & L & C_{Lm} & N_m^k \end{bmatrix}$$

4. Experiments and analysis

The data of experiment was abstracted from D5000 system in a substation of the State Grid Shanghai company. The goal of busbar voltage trend prediction was to forecast No. I 500kV busbar historical Uac line voltage sequence from 2015 to 2017, and the relevant data included 4 related historical 220kV busbar line Uac line voltage sequence, the relevant transformer time sequence, switching log records and status records of related AVC equipment, related historical remote signal of switch, and environmental temperature sequence. Support Vector Machines, BP Artificial Neural Network, C4.5 Decision Tree were chosen as the algorithms of multiple machine learning for ensemble learning.

Figure 4-6 demonstrated three algorithms (C4.5 decision tree, BP neural network, SVM) with different h values, which represent the time window granularity of predictive model, with the dataset of year 2015-2017. The results of experiment validated that the efficiency of predictive model was depended on the granularity of time windows and the feature of algorithms. The predictive accuracy of the optimal classifier from ensemble learning was the highest, which also indicated that ensemble learning method can effectively resist the conceptual drift of the short-term predictive data of voltage trend.

In addition, three experiments showed that the effect of bus voltage trend prediction with optimal classifier model window granularity $h = 8$ was better than the others with granularity $h = 4$ or $h = 12$. The predictive accuracy of model in year 2015 was 91.3% when the granularity was $h=8$. The accuracy was 92.5% when the granularity was $h=4$, and the accuracy was only 88.5% when the granularity was $h=12$. Furthermore, in 2016 and 2017, the predictive accuracy of model was also highest when granularity was $h=8$. The experimental results discovered that the potential knowledge in the experiment cannot be completely extracted if the timing window granularity of model was too small. On the other side, if the timing window granularity of model was too large, it was possible to overfit the real pattern of knowledge. Therefore, in the short-term prediction method of busbars' voltage trend, it is an important factor to choose the appropriate timing window granularity of model.

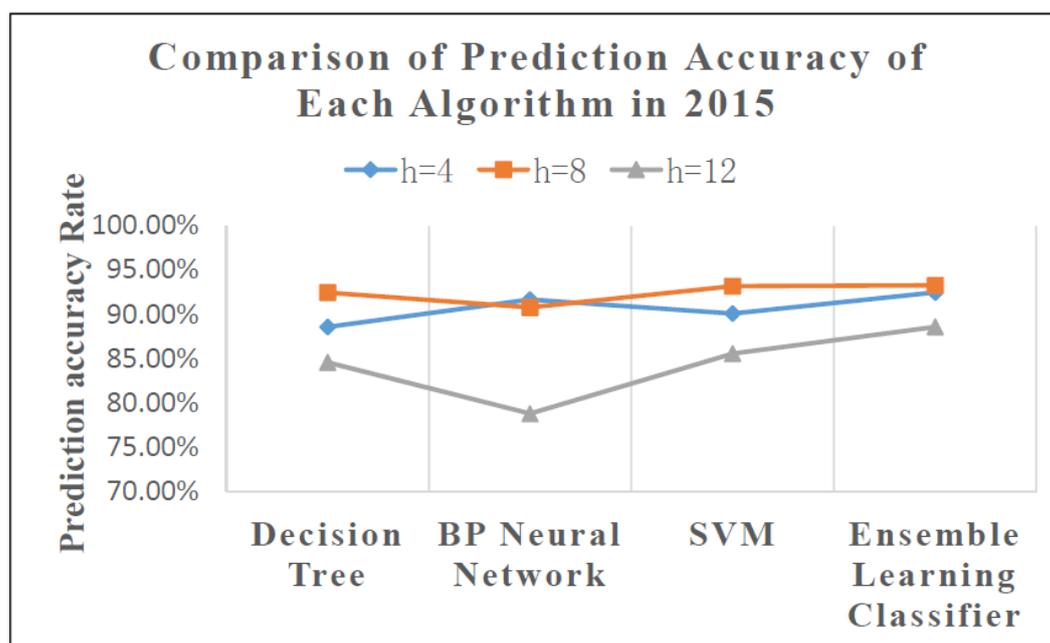


Figure 4. Comparison of predictive accuracy between each predictive models in year 2015.

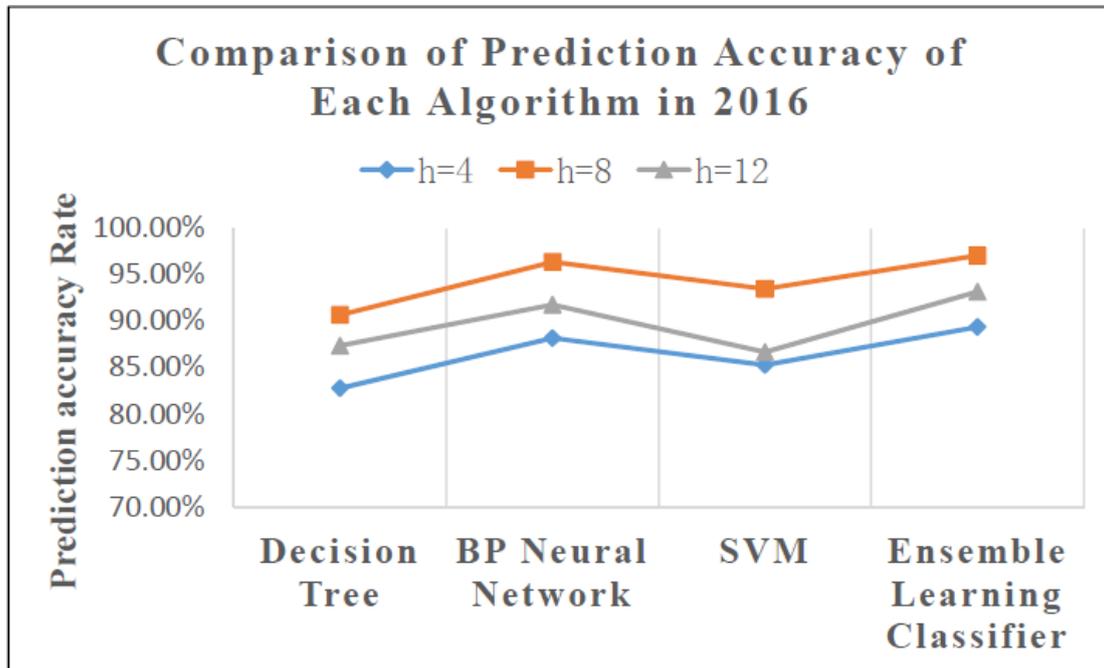


Figure 5. Comparison of predictive accuracy between each predictive models in year 2016.

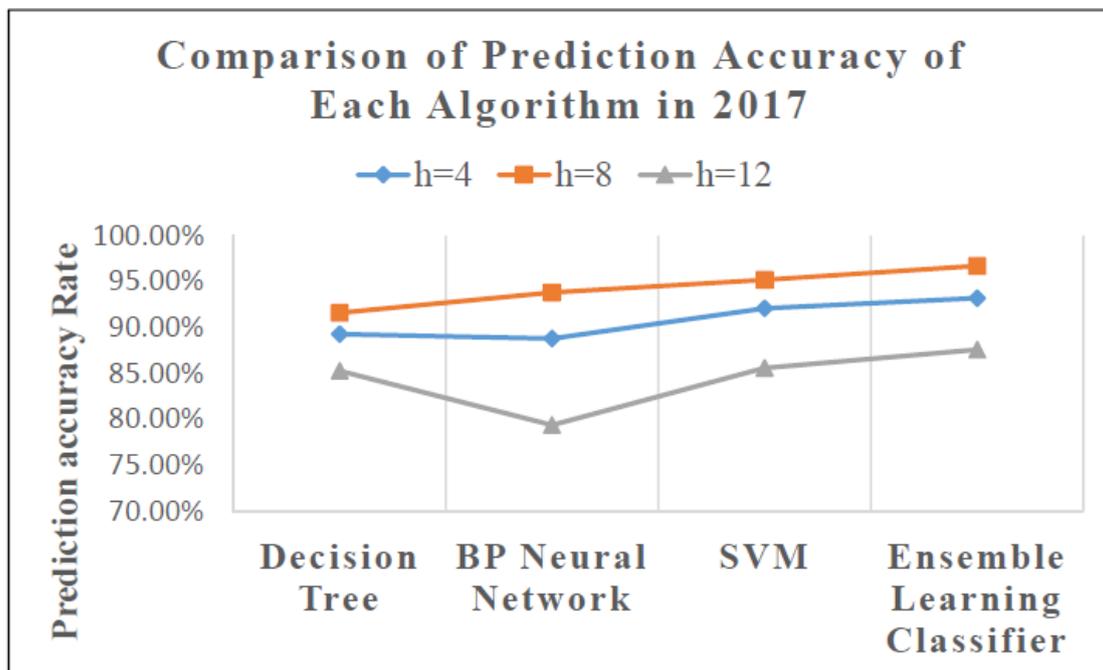


Figure 6. Comparison of predictive accuracy between each predictive models in year 2017.

5. Conclusion

This paper proposed a short-term prediction method based on multi-dimensional time series data mining and had been succeeded to implement in practical scenarios of State Grid Shanghai Company. Compared with the traditional voltage prediction algorithm based on one dimensional data, the ensemble learning strategy with various classic machine learning algorithm was proposed in this paper and validated its efficiency for short-term prediction of voltage trend. The results of prediction can support the function of fake alarms filtering in the voltage exceeding intelligent alarm system, and

improve the efficiency of whole alarm system. However, the prediction method was still possible to improve, for example, via Dynamic Time Wrapping (DTW) clustering algorithm used in the decision table. On account of the high computational complexity of classical DTW algorithm, efficiency of clustering is the critical tasks in the next stage. In addition, the strategy of ensemble learning classifier in this paper is Bagging strategy based on voting mechanism, and whether Boosting strategy [14] is more efficient than Bagging strategy or not in short-term prediction for busbar's voltage trend, also need further research.

6. Acknowledgments

I really appreciate the support and assistant of every expert and colleague from State Grid Shanghai Company for this project. Their advices and knowledge made this method discussed in the paper to be feasible and implement. However, I also thank for the help of Dr. Yu, Dr. Xu and Dr. Zhao from Nanchang University and Tellhow Company.

7. References

- [1] Wang Chen, Dong Xiaoyang, Wang Ruihan. 2015. Wide area back-up protection algorithm for least squares estimation under finite phasor measurement unit. *Power system automation*, 39 (20), p124-129.
- [2] Shao Li, Xie Kaigui, Wang Jin et al. 2008. Distribution network Reliability evaluation algorithm based on power flow estimation and block load reduction. *Power grid Technology*, 32 (24), p34-37.
- [3] Zhang Jihong, He Zhiyong, Xing Jun. 2015. Droop control strategy of micro-power grid inverter using public coupled point voltage estimation. *Large motor Technology*, 1 (1) p57-61.
- [4] Zeng Yi, Zhu Zhenmin, Liu Jinla, Chen Yuanfei. 2010. Dynamic voltage regulation prediction algorithm based on periodic pattern matching. *Computer Engineering*, 36 (13) p265-268
- [5] Li Xiaomin, Yin Ming. 2016. Research on the state prediction method of BP neural network Electronic system based on genetic algorithm. *Electronic measurement technology*, 39 (9) p182-186
- [6] You Jin, Liu Junyong, Qiu Gao. 2017. Voltage fast estimation based on adaptive neural network. *Power Exploration Design*, 02 (1) p44-49
- [7] You Jin. 2017. On-line voltage prediction on the basis of nonparametric estimation. *Sichuan Electric Power Technology*, 40 (5) p1-5
- [8] Phan, T., Caillaud, E. 2017. Dynamic time warping-based imputation for univariate time series data. *Pattern Recognition Letters*.
- [9] Górecki, Tomasz, Łuczak, Maciej. 2014. Non-isometric transforms in time series classification using DTW. *Knowledge-Based Systems*, 61 p98-108.
- [10] Łuczak, Maciej. 2016. Hierarchical clustering of time series data with parametric derivative dynamic time warping. *Expert Systems with Applications*, 62 p116-130.
- [11] Yang Yanlin, Ye Feng. 2016. a method for mining hydrological time series similarity based on DTW clustering. *Computer science*, 42 (2) p245-249
- [12] Krawczyk, B. Minku, L. 2017. Ensemble learning for data stream analysis: A survey. *Information Fusion*
- [13] Li, S. Goel, L. 2016. An ensemble approach for short-term load forecasting by extreme learning machine. *Applied Energy*, p22-29
- [14] Chen, K., Jiang, J. 2018. A novel data-driven approach for residential electricity consumption prediction based on ensemble learning. *Energy*, 150 p49-60.
- [15] Heinermann, J., Kramer, O. 2016. Machine learning ensembles for wind power prediction. *Renewable Energy*, 89 p671-679.
- [16] Peimankar, A., Weddell, S. J. 2018. Multi-Objective Ensemble Forecasting with an Application to Power Transformers. *Applied Soft Computing*
- [17] Collell, G., Prelec, D. 2018. A simple plug-in bagging ensemble based on threshold-moving for classifying binary and multiclass imbalanced data. *Neurocomputing*, 275 p330-340