Research into voltage sag online detection technology based on wavelet tree

Xianfeng Zheng[1, 2*], Zheng Fan[2]

[1]*Faculty of Electrical Engineering Henan Mechanical and Electrical Engineering College, No.699 Ping Yuan Road, Xinxiang, Henan, China*

[2]*Faculty of Automatic Control Engineering Henan Mechanical and Electrical Engineering College, No. 699 Ping Yuan Road, Xinxiang, Henan, China*

Received 6 February 2014, www.tsi.lv

Abstract

A general process model is established using the real-time requirements of data stream processing, and the data is constantly processed with a sliding window. This paper selects the recursion-based complex wavelet as the detecting algorithm for voltage sag, and tries to detect when the voltage sag occurs and ends with amplitude and phrase information contained in the wavelet analysis results. Meanwhile, this paper seeks to improve the precision of detection by looking for optimal wavelet scales with information entropy. The shifted wavelet tree-based data flow anomaly detection algorithm and data update method of shifted wavelet tree have been improved to make rapid detection possible. Finally, this paper reports the experimental simulation, which proved the instantaneity and accuracy of this method.

Keywords: Data Stream, Voltage Sag, Recursive Complex Wavelet Transform, Shifted Wavelet Tree

1 Introduction

Power quality is a common concern in certain parts of the world. Voltage sag - a major quality problems affecting the normal and safe operation of electrical equipment - is a fast short-term decline of voltage effective value caused by a system short circuit fault, overload or a large motor starting [1-2].

According to one survey on power supply: with the exception of outages, voltage sag is the greatest power quality issue, accounting for more than 80% of the problems. Voltage sag causes widespread equipment failure, causing economic losses, such as damage to or malfunctioning of electronic, computer and control equipment and other sensitive devices.

In response to market demand, accurate real-time detection of magnitude, duration and occurrence of voltage sag is required. At present, the most commonly used method is, according to the definition of continuous periodic signal effective value, to obtain sag amplitude by calculating voltage effective values, among which the voltage effective value can be obtained with digital RMS operation in one cycle of time domain [3-4]. As the voltage sag problem becomes increasingly evident, an effective method is needed for a real-time detection. Accordingly, real-time detection of occurrence time/end time, duration, frequency of voltage sag is a primary subject of the anomalous voltage detection research.

2 Data stream processing model

A windowed process with a sliding window was developed in light of the very obvious temporal characteristics of voltage sag information detection, the special technical requirements of one access, limited storage, continuous processing, and rapid response during the real-time processing of the data stream formed by detection-related data, and the higher demands for instantaneity [5]. The deadline time of data processing is defined to ensure real-time constraints on data stream processing. The created process system model is shown in Figure 1.

Figure 1 shows that the system core is the data stream processing engine. Different processing algorithms should be adopted according to function. In this system, data flow anomaly detection is the main system function.

* ***Corresponding author*** e-mail: hnxxzxf@126.com

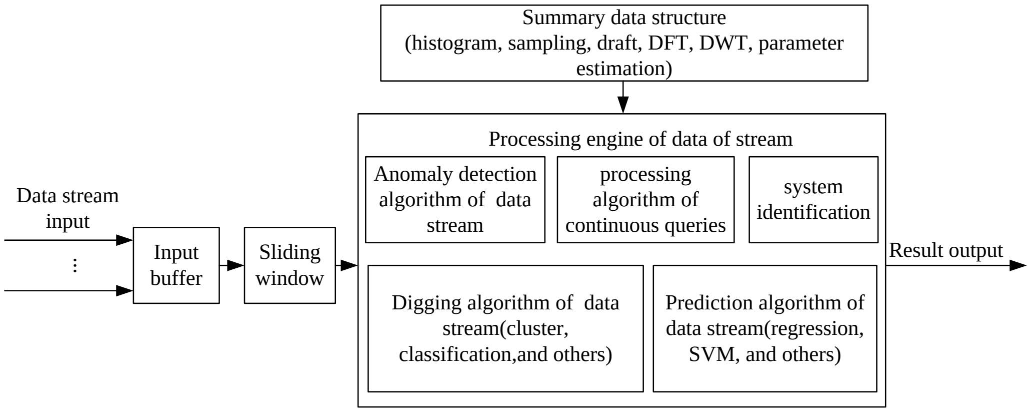


FIGURE 1 Data stream processing model

3 Detection of data flow anomaly

The essence of anomaly detection is to find significantly different data from a large number of data, and the definitions of anomalous data vary in different fields of application. The detection method we used for data stream anomaly with mobile wavelet tree data structure uses the SWT structure to do the anomaly detection, and is capable of detecting anomalies of different length.

In the anomaly detection process, firstly, the data stream is x_i (i=1... n), the aggregation function H (including data maximum, minimum, average, sum, standard deviation, and etc.) can be obtained by calculation, and according to different wavelet resolution, several sliding windows w_j and corresponding threshold H (w_j) are set, and then compared with the threshold value, we screen out the anomalous data so that the aggregation function is greater than the corresponding threshold. The algorithm, based on the anomalous duration, detects corresponding wavelet levels. Once the value of a window in this level is found to exceed the predefined threshold, then the corresponding lower window search is continued, until the anomalous position is found.

In the above process, the data in different levels should use a different sliding window length and different wavelet scale factor, so that the time series is decomposed into a multi-level wavelet tree structure, in which the original data sequence consisting of zeroth layer wavelet decomposition tree (Level0), wherein the average and difference of adjacent data constitutes the first layer (Level1) wavelet coefficients. And then the process repeats itself, the wavelet coefficients of the next layer will be obtained through successive recursion, until only one average value and difference are left in the top layer. Therefore, the wavelet coefficients in wavelet tree contain the average value and difference of each layer, and the original signal can be reconstructed from these coefficients without distortion [6].

For a specific data aggregation, each definite anomalous length can be detected at a corresponding level. If it is over the threshold, it can be assumed that each data aggregation contained in this window exceeds the threshold, and concluded that the data includes an anomaly. The anomaly detection algorithm filters windows that are found with no anomaly, and does not detect it. Instead, this detection algorithm only detects a few windows aggregations that are over the threshold, which thus greatly reduces the detection range and improves detection efficiency, thereby making the detection omission rate relatively low. And by changing the length of elastic windows, we can carry out anomaly detection of an arbitrary length of data flow.

The summary data of the above aggregation function is obtained through a wavelet tree and a real-time complex wavelet algorithm; an improved recursive wavelet is used here to process voltage sag information [7].

First, the selected base wavelet is equation (1)

$$\varphi(t)=\left(-\frac{\sigma^3t^3}{3}-\frac{\sigma^4t^4}{6}-\frac{\sigma^5t^5}{15}\right)e^{(\sigma+j\omega_0)}u(-t)\,. \tag{1}$$

Here, $\sigma=2\pi/\sqrt{3}$; $\omega_0=2\pi$; at this point, $\varphi(0)=0$, which ensures that the selected base wavelet satisfies the admissibility condition. The base wavelet can be represented in the frequency domain

$$\psi(\omega)=\left[\frac{6\sigma^5-2\sigma^3(\omega-\omega_0)^2}{\sigma+i(\omega-\omega_0)}\right]^*. \tag{2}$$

Discretizing the above expressions, we can get equation (3)

$$\varphi(fnT)=\left(-\frac{\sigma^3(fnT)^3}{3}-\frac{\sigma^4(fnT)^4}{6}-\frac{\sigma^5(fnT)^5}{15}\right)e^{(\sigma-\omega_0)fnT}u(-fnT)\,. \tag{3}$$

Among them, f=1/a, a is the wavelet scale factor, T is sampling period. After transformation and sorting out via Z, the following equation (4) can be obtained

$$\Psi(Z)=\frac{\delta_1Z^{-1}+\delta_2Z^{-2}+\delta_3Z^{-3}+\delta_4Z^{-4}+\delta_5Z^{-5}}{1+\lambda_1Z^{-1}+\lambda_2Z^{-2}+\lambda_3Z^{-3}+\lambda_4Z^{-4}+\lambda_5Z^{-5}+\lambda_6Z^{-6}}\,. \tag{4}$$

Among them, $\delta_n = \left[a_n(\sigma fT)^3 + b_n(\sigma fT)^4 + c_n(\sigma fT)^5 \right] A^n$, a_n, b_n, c_n is the coefficient; after calculation, the following can be obtained a_1=1/3; a_2=2/3; a_3=-2; a_4=2/3; a_5=1/3; b_1=-1/6; b_2=-5/3; b_3=0; b_4=5/3; b_5=1/6; c_1=1/15; c_2=26/15; c_3=22/5; c_4=26/15; c_5=1/15. $A = e^{-fT(\sigma - j\omega_0)}$, λ_1=-6A; λ_2=15A^2; λ_3=-20A^3; λ_4=15A^4; λ_5=-6A^5; λ_6=A^6.

And available expressions

$$\begin{aligned} W_{s,\phi} &= \sqrt{f}T[\delta_1 s((k-1)T, f) + \delta_2 s((k-2)T, f) + \delta_3 s \\ &((k-3)T, f) + \delta_4 s((k-4)T, f) + \delta_5 s((k-5)T, f)] - \\ &\lambda_1 W_{s,\phi}((k-1)T, f) - \lambda_2 W_{s,\phi}((k-2)T, f) - \lambda_3 W_{s,\phi} \\ &((k-3)T, f) - \lambda_4 W_{s,\phi}((k-4)T, f) - \lambda_5 W_{s,\phi}((k-5)T, f) \\ &-\lambda_6 W_{s,\phi}((k-6)T, f) \end{aligned} \quad (5)$$

According to different application requirements, the signal, by decomposing wavelet packet subspaces, can be decomposed into a high frequency part and a low frequency part as the summary data signal with the above wavelet decomposition.

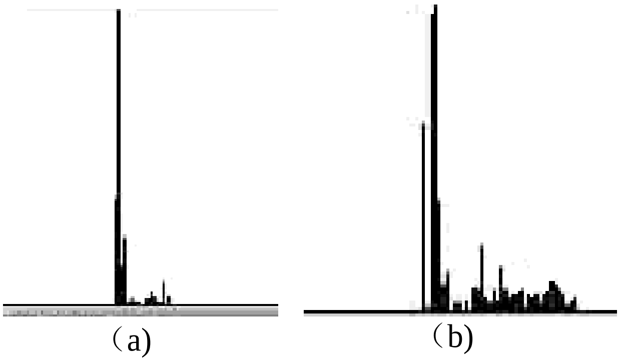


FIGURE 2 The wavelet analysis results of the end point of voltage sag (a) f=3000 (b) f=5000

4 Establishment and search of wavelet error tree

In certain data stream applications (e.g., correlation analysis) only a low frequency coefficient is needed for synoptic data of the data stream and analysis. In other applications (such as power quality disturbance identification), detailed coefficients are required for summary purposes [8]. The algorithm must be able to handle the real-time, continuous, unlimited data stream; accordingly, the wavelet tree is constructed as shown in Figure 3.

Based on the wavelet tree shown above, the source data can be reconstructed, and the range, threshold value and other parameters will also be rebuilt; thereby aggregation query will be achieved.

In the actual detection process, because the anomaly detection algorithm of the data stream based on the shifted wavelet tree needs to simultaneously detect the entire length of the sliding windows, which takes much time and space, causing deterioration of system performance; therefore this search method need to be optimized [9]. The binary search method adopted here can greatly reduce system overhead.

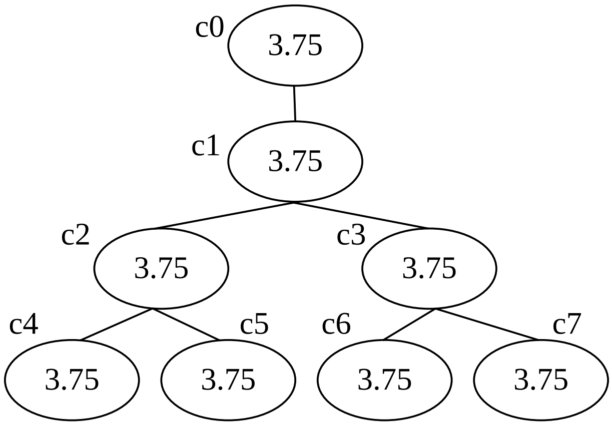


FIGURE 3 Wavelet Tree Structure Figure

First, the monotonic search space will be established, and all the sliding windows, which need to be detected are sorted according to length. Based on this, a binary search is carried out to find the largest anomaly window length.

According to the methods above, and considering the fact that the given data may include anomaly data, whose length is W_i (where in, i=1,2,... m, m is various window length of abnormal data to be detected), the monotonic search space will be constructed first, that is to sort space (W_1,…,W_m) according to detection window size. And then, the binary search is carried out on the basis of this space sequence. First, the detection begins from the intermediate position window of entire search space. If the anomaly is found, then the latter half part will be searched. If no anomaly is found, then the former will be searched. Then the process is repeated ceaselessly, until the anomaly window is found. And then, the anomaly starting position will be located in the window.

5 Data pre-processing

In order to eliminate the jolts data and improve detection accuracy, the original data stream should be properly pre-processed first [10].

The aim of data pre-processing is to eliminate the jolts data and improve detection accuracy. The jolts data are data in the data stream that reach certain intensity, which, however, is not enough to become anomaly data. The jolts data of different window length has a varied effect on detection results. Some is not considered anomaly data in a small window. When an anomaly detected in a larger window, these cumulative small changes are likely to be mistaken as an anomaly.

In order to improve the accuracy of anomaly detection, the ratio threshold anomaly detection method is adopted here to remove the interference of jolts data. The ratio threshold, using the recent two equal length window aggregation ratio in the data stream, determines whether the anomaly has occurred. It is defined as follows

$$(x_{w+1}+\cdots+x_{2w})>\beta(x_1+\cdots+x_w)\ \beta>1, \tag{6}$$

$$(x_{w+1}+\cdots+x_{2w})<\beta(x_1+\cdots+x_w)\ 0<\beta<1. \tag{7}$$

In the above formula, β, corresponding to the upper limit threshold value and lower limit threshold value, is determined by the ratio of the two same windows before the current time point.

In practical application, if no anomaly occurs in the data stream, it is necessary to consider the possibility that the size of the time window might lead to a false anomaly and try to prevent it from happening. Since the large window anomaly caused by jolts data does not appear in a small window, therefore, the window whose length is 1 should be detected first with the above ratio threshold method. If no anomaly occurs in this window, then it will not be considered as an anomaly in the large window, which reduces the probability of misinformation.

After processing the jolts data, no anomaly caused by jolts data appears in the large window, so the maximum length of the anomaly sliding window is the actual length of the detected anomaly window.

6 Update of data of shifted wavelet tree

With the continuous updating of the data stream in the sliding window, calculation of the DWT coefficient of the data stream in the sliding window is done in two ways: one is direct convolution calculation of all the data in the sliding window, which is called direct update; the other is incremental update, that is, incremental calculation on the entire sliding DWT coefficient with the DWT coefficients before window sliding and the newly arrived data. The incremental update algorithm can avoid the disadvantages of the traditional algorithm; the wavelet summary data structure in the sliding window need not be rebuilt when new data comes into the sliding window, thus improving time cost.

The direct update algorithm can ensure accurate completion of each node update when new data arrive. But taking into account the window sliding, and as long as the update data length is less than the sliding window size, data redundancy exists before and after sliding, and the direct update of reconstruction of the wavelet summary reduces processing efficiency.

When receiving new data in the sliding window, it is only necessary to carry out wavelet decomposition on new arrival data with the incremental update method, and the other part can be obtained by shifting the wavelet decomposition results before window sliding. This avoids time overhead of re-computing. Compared with the direct update algorithm, the short calculation time of the SWAT algorithm is particularly prominent in the analysis application of a large data stream. One limitation is when only half of the new data of the sliding window arrives; the summary has to be completely updated again. And when the amount of new arrival data is less than half of the sliding window, the summary data structure cannot accurately represent the window data, thus precision cannot be guaranteed in some applications. Another limitation is that because an update operation is carried out during one unit of time, the large data stream and frequent update operation need a certain time overhead in practical application, which affects process efficiency.

Considering the limitation of the above two kinds of incremental updating algorithm, an improved algorithm is proposed in this paper, whose main idea is as follows: the sliding window is divided into several basic windows (sub windows), wavelet approximation trees are constructed respectively in each sub window and transition wavelet approximation trees (similar to the intermediate nodes) are built between the adjacent two sub windows. When new data arrives, it is not only updated in each tree node, but also updated between trees. Therefore, the complete update of the entire wavelet summary only needs the data of half the amount of the basic window length.

After analysis, the update process of the improved algorithm can be divided into two cases.

(1) Update of internal nodes in the first tree. It can be further divided into two types: node update in the same layer. After one update cycle, some node value is shifted to the left node, and the right node value is shifted to this node; to achieve the node update between different layers, as for the lower node, the value of the upper node located in the left and right sides of this node can be calculated, the result is stored in this node.

(2) Update between different trees. Like the node update inside the tree, the entire tree achieves dynamic incremental update through shifting.

The specific approach is that we add more redundant data windows to the wavelet tree, and store one summary data set for 2^i number data. w-2^i+1 summary data is saved in each layer, and w is the length of the selected sliding window. According to the real time incremental updating algorithm, incremental update involves carrying out a small amount of data summary update in each wavelet level when one data set arrives. It is in the *i* layer of the wavelet tree where one summary data set for 2^i number data is saved. Whenever new data arrives, the summary corresponding to the wavelet level of the oldest data as the starting point is expired and, at the same time, the upper data summary of the new data as the end point is generated. So when one data set arrives, a small amount of the summary corresponding to wavelet level is updated, which reduces the time complexity of algorithm detection. The real time continuous updating algorithm can detect the anomaly more quickly and accurately, meeting the requirement of real-time detection of data stream and improving the efficiency of anomaly detection.

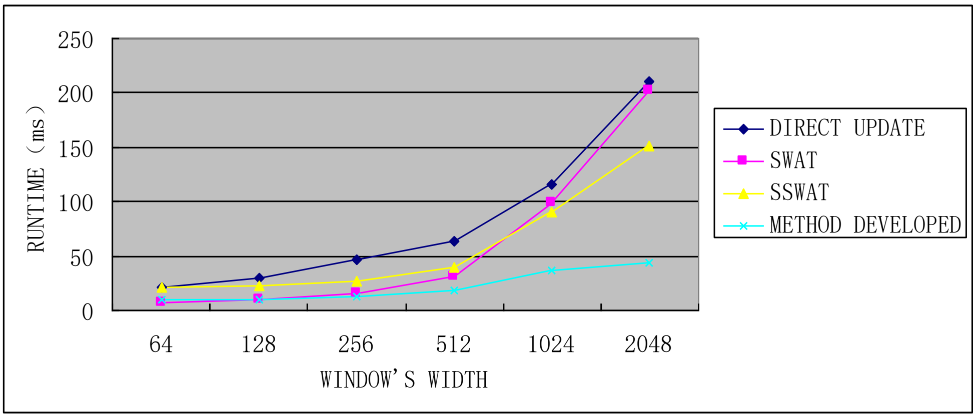


FIGURE 4 Running Time Ratio of updating algorithm for All Kinds of Data

The full update can be completed when two new data sets (half-length of base window) arrive in the improved algorithm. Compared with the foregoing two traditional algorithms, the necessary data length of completing one full update is shorter, and suitable for higher requirements on the accuracy and speed of data stream processing system.

The specific data is shown in Figure 4. It can be seen that the required time of the SWAT algorithm to complete one update is much shorter than a direct update, but the time of completing one update is equal to half the sliding window size, which is not suitable for the case of: (i) an update interval of less than one-half of the sliding window size, or (ii) an accuracy requirement.

When the sliding window is small, the time it takes for SWAT to be completely updated should be less than SSWAT. But when the sliding window length is greater than 1024, the time required for SSWAT is less than for SWAT, and real-time data streams are relatively large. Therefore, SSWAT is more suitable for high-speed data stream processing.

Compared to SSWAT, the improved algorithm only needs one half of the base window to complete the full update, which is suitable for the application of a small time interval of output result, with better accuracy than SSWAT.

Given that the implementation of wavelet decomposition is extraction from convolution of the input sequence and wavelet decomposition filter, so when the length of wavelet decomposition filter coefficients is too long and the input sequence is limited, it will cause boundary distortion, and the above discussed wavelet decomposition incremental update algorithm will also sometimes cause boundary distortion. The solution is to adopt extension, such as zero extension, periodic extension, symmetric extension and so on. In this case, the wavelet decomposition that results after window shifting cannot be simply obtained by the above method of shifting. As to the edge part, we will obtain it by direct convolution calculation, while the middle part can be obtained by the wavelet decomposition results before window shifting.

7 Example analysis

The anomaly detection algorithm of the data stream with improved shifting wavelet tree carries out detection on simulation signal which involves voltage sag, and compares it to the common error tree algorithm to show the relative effectiveness of this method.

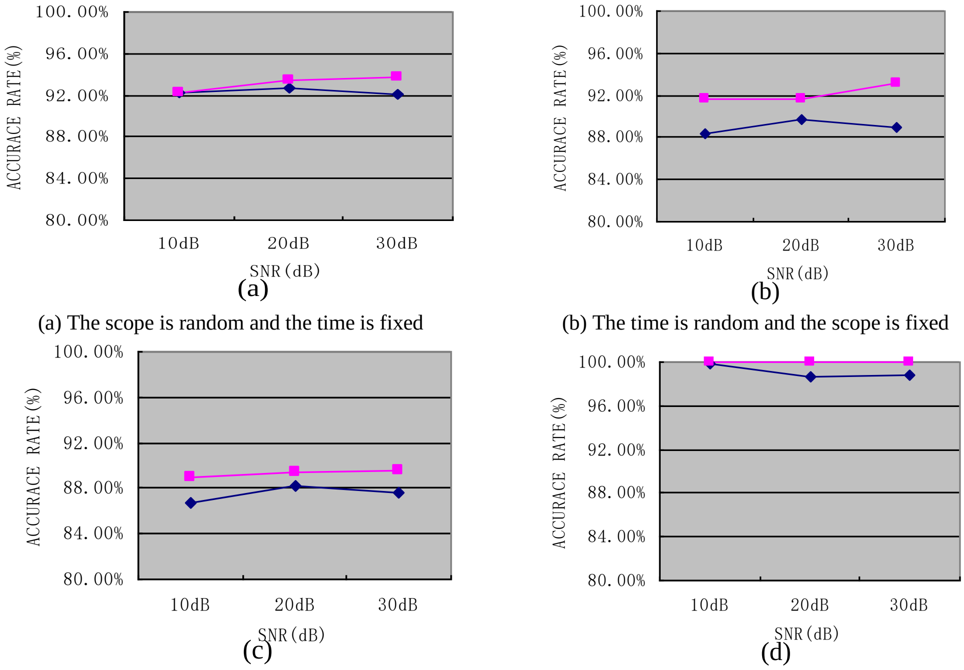
First, the typical power quality disturbance signal model is established, and the signal is generated in Matlab, which includes four common kinds of transient signal in an electric power system (voltage swell, voltage sag, and voltage interruption). In order to simulate an actual situation, the parameters of voltage sag, voltage swell and interruption are allowed to vary randomly within the permitted range (the parameters that characterize the disturbance signal fluctuate randomly in a certain range), and random white noise of 10-30dB noise ratio is added to them [11-12].

TABLE 1 Disturbance signal model

Voltage swell	$0.1\leq\alpha\leq0.8$, $T\leq t_1-t_2\leq9T$	$v(t)=A(1+\alpha(u(t_2)-u(t_1)))\sin(\omega t)$
Voltage sag	$0.1\leq\alpha\leq0.8$, $T\leq t_1-t_2\leq9T$	$v(t)=A(1-\alpha(u(t_2)-u(t_1)))\sin(\omega t)$
Voltage interruption	$0.9\leq\alpha\leq1$, $T\leq t_1-t_2\leq9T$	$v(t)=A(1-\alpha(u(t_2)-u(t_1)))\sin(\omega t)$
Harmonic	$0.05\leq\alpha_3(\alpha_5,\alpha_7)\leq0.15$, $\sum\alpha_i^2=1$	$v(t)=A(a_1\sin(\omega t)+a_3\sin(3\omega t)+a_5\sin(5\omega t)+a_7\sin(7\omega t)$

In the analysis process, the input signal analysis time length adopts 10 sine wave cycles (0.2S), 6.4kHz sampling rate, 50Hz voltage frequency. Given the noise condition, the accuracy of detection result is shown in Figure 5.

From the results of the table above, the accuracy of the shifting wavelet tree algorithm is much better than that of the error tree algorithm. Lower accuracy of detection results when the detection is affected by noise.



(a) The scope is random and the time is fixed

(b) The time is random and the scope is fixed

(c) Both time and scope are random

(d) Both time and scope are fixed

FIGURE 5 Data table of detection result under different disturbances

As for the required time of detection, under the same condition, one time of detection needs 0.8ms with the algorithm of the shifting wavelet tree, while the error tree algorithm needs about 1.2ms. The shifting wavelet tree algorithm is obviously better than the error tree algorithm.

8 Conclusion

By using an improved detecting method of data stream based on the shifted wavelet tree data structure, we constructed a wavelet tree data model and reduced the time complexity with binary search error of the wavelet tree. We designed a voltage sag detecting algorithm based on the adapted recursive wavelet, which is capable of rapid and precise detection of the start-stop of voltage sag. Because it is simple and effective - not at the cost of precision - this method guarantees instantaneity of the system.

In addition, with this real-time incremental updating algorithm to detect wavelet tree data, we can meet the requirement of anomaly detection, and improve detecting efficiency and instantaneity, only through updating a small amount of data summary of the relevant wavelet levels. Analysis of the actual example proved that this algorithm, with advantages like low time consuming, high real-time and so on, provides accurate results for the range, duration and frequency of voltage sag, offering a new method for detecting voltage sag.

References

[1] Babu S, Widom J 2001 Continuous queries over data streams *ACM SIGMOD Record* **30**(3) 109-20

[2] Chakrabarti K, Garofalakis M, Rastogi R, et al 2001 Approximate query processing using wavelets *VLDB Journal* **10**(2-3) 199-223

[3] Gerbec D, Gasperic S, Smon I,Gubina F 2005 *IEEE transactions on power systems* **20**(2) 548-55

[4] Guha S, Meyerson A, Mishra N, Motwani R, O'Callaghan L 2003 *IEEE Transactions on Knowledge and Data Engineering* **15**(3) 515-28

[5] Barbara D 2003 Requirements for clustering data streams *In Proceedings of ACM SIGKDD Explorations Newsletter* **3**(2) 23-27

[6] Abdel-Galil T K, Kamel M, Youssef A M, et al 2004 Power quality disturbance classification using the inductive inference approach *IEEE Transactions on Power Delivery* **19**(4) 1812-7 *(In chinese)*

[7] Fan Zheng, Tian Xiaowu, et al 2008 Design of virtual instrument for voltage sag detection based on recursive complex wavelet transform *Electric power automation equipment* **28**(6) 100-2 *(In chinese)*

[8] Fan Zheng, Tian Xiaowu, et al 2008 Design of virtual instrument for voltage sag detection based on recursive complex wavelet transform *Electric power automation equipment* **28**(6) 100-2 *(In chinese)*

[9] Figueiredo V, Rodrigues F, Vale Z, Gouveia J B 2005 *IEEE Transactions on Power Systems* **20**(2) 596-602

[10] Wang Yongli, Zhou Jinghua, Xu Hongbing, Dong Yisheng, Liu Xuejun 2007 Adaptive prediction of time series data stream *Journal of automation* **2**(33) 197- 201 *(In chinese)*

[11] Monedero I, Leon C, Ropero J, Garcia A, Elena J M, Montano J C 2007 *IEEE Transactions on Power Delivery* **22**(3) 1288-96

[12] Wu Shanshan, Gu Yu, Lv Yanfei, Yu Ge 2007 Sliding window processing strategy of sensitive deadline in data stream *Computer science* **34**(7) 99-102 *(In chinese)*

[13] Liu Lianguang 2008 Achieving accurate measurement of voltage sag with wavelet transform and RMS algorithm *Automation of electric power systems* **27**(11) 30-3 *(In chinese)*

Author

Zheng Xianfeng, born in March, 1972, Xinxiang County, Henan Province, P.R. China

Current position, grades: the Associate Professor of Henan Mechanical and Electrical Engineering College, China.
University studies: received his B.Sc. in Industrial Electrical Automation from Shanxi Institute of Mining Technology in China. He received his M.Sc. from Xi'an Jiaotong University in China.
Scientific interest: His research interest fields include measurement and control technology and electrical insulation testing.
Publications: more than 31 papers published in various journals.
Experience: He has teaching experience of 18 years, has completed ten scientific research projects.

Fan Zheng, born in June, 1973, Xinxiang County, Henan Province, P.R. China

Current position, grades: the Associate Professor of Henan Mechanical and Electrical Engineering College, China.
University studies: received his B.Sc. in Electrical Engineering and Automation from Yan Shan University in China. He received his M.Sc. from Xi'an Jiaotong University in China.
Scientific interest: His research interest fields include computer measurement and control technology, electrical testing and analysis of power quality.
Publications: more than 28 papers published in various journals.
Experience: He has teaching experience of 19 years, has completed five scientific research projects.