Journal of Engineering and Applied Sciences

MEDWELL PUBLICATIONS SINCE 2000

Application of Adaptive Threshold Algorithm in Data Processing of Optical Fiber Time Ratio

[1]Xu Lu, [2]Danhe Chen, [1]Yahong Wan and [1]Pan Wang
[1]Chinese Academy of Science National Time Service Center, Xi'an 710600, China
[2]Nanjing University of Science and Technology, Nanjing, China

Key words: Fiber optical time transfer, segmented adaptive threshold algorithm, data processing, clock difference correction

Corresponding Author:
Xu Lu
Chinese Academy of Science National Time Service Center, Xi'an 710600, China

Page No.: 3074-3079
Volume: 15, Issue 16, 2020
ISSN: 1816-949x
Journal of Engineering and Applied Sciences

Abstract: In this study, a promotion of quantum frequency standard and precision measurement and other professional development are proposed in order to obtain a unified time benchmark, higher precision requirements have been put forward on the high-precision time frequency ratio. On the basis of 124KM optical fiber time transfer system platform, combined with the experimental data in the recent 2 years, analysis of various factors that affect the time comparison of optical fiber are presented and the adaptive threshold method is proposed to deal with the time difference of fiber optics in the system. In addition, simulation results show that the application of proposed method can obviously improve the accuracy of comparison which also prove the validity of the algorithm.

INTRODUCTION

With the development of worldwide economy, more researches have been focused on topics of navigation, aerospace and electric power engineering. Time as a physical quantity which can ensure the effectivity and reliability in the field of state construction and data processing. For instance, the standard time in China is defined as Beijing time which was established and maintained by the National Timing Service Center of the Chinese Academy of Sciences. Meanwhile, the requirement of time frequency datum is increasing. At present, the means of high precision time comparison mainly is based on GPS common-view and satellite bidirectional frequency comparison method, however, these two approaches still cannot meet the demand of high precision time frequency transmission.

In order to obtain high precision of time and frequency, the user needs to transmit the high precision time reference signal over long distance. Fiber as an excellent material has become an important transmitter for high precision time frequency ratio and in recent years, in view of optical fiber has the advantages of large capacity, long distance, anti-jamming, low loss, small noise and etc., the time frequency ratio of optical fiber has become a hot topic in the field of time-frequency. Compared with the other time and frequency ratio approaches, optical fiber method has a superiority in both precision and stability. The optical fiber networks can stay stable during the special time against the factors of destruction and replace the satellite timing system when for example, the Beidou system is unavailable.

In the process of optical fiber time transmission, the signal received contains a series of interference signals, especially, the line between the optical fiber exposed in the surface section. Considering of the external temperature, humidity and other environment caused by the interference, the amplitude of signal is significant and the error data is mixed with the measured optical fiber time frequency transmission data. In other researches, method of removing the gross error as the measurement in order to improve the reliability and credibility of the collected data is applied. At present, there are four commonly used methods to eliminate abnormal data: Raida guidelines, Grubb's guidelines, Chauville guidelines, Dixon Guidelines. These statistical methods are applied to large sample events which normal distribution and a lot of data points are removed. Based on the above methods, combined with the clock data obtained by the optical fiber time frequency transfer system in the latest 2 years, a segmented adaptive threshold algorithm in this paper is presented and the experimental results show that the proposed method can effectively detect the outliers.

COMPOSITION OF FIBER OPTICAL TIME FREQUENCY TRANSFER LINK

Structure of optical fiber transfer system is presented in Fig. 1. The standard time signal which is from precise main clock, through delay device, optical fiber time transmission signal transmitter. The time and frequency are transferred from place A to place B (which requires a C repeater) by a fiber, the receiver is in place B and the time and frequency is output from place B. The 1PPS signal received on the fiber time and frequency receiver in place B is synchronized with the 1PPS of place A. In this fiber, a bidirectional comparison method is used to compensate the transmission delay of the fiber link and the change of transmission delay. Of course, the 1PPS signal of place A and place B is kept in basic synchronization.

The 1PPS signal input interval counter received by the receiver from place B can measure the deviation value of the working clock in place B and main clock in place A (Standard time), that is, the clock difference between the two places. The clock difference processing module collects the clock difference in real time and enters the clock difference adjustment of place B into the HROG-5 by using the clock difference and deducting the clock rate, thus keeping the time of the work clock in place B and main clock in place A is synchronization. Therefore, the integrity and usability of the clock difference data directly affect the correction of the atomic clock in place B.

Adaptive threshold algorithm: The measurement error is mainly divided into: stochastic error, system error and gross error[1]. In the case of repeated measurements of the same measurement, the absolute value and the symbol change unpredictably but the error with certain statistical law is called stochastic error. The error is caused by random factors which is the measuring personnel and measuring equipment during the measurement process. In the case of repeated measurements of the same measurement, if the error according to a certain rule, the error is called system error such as the inaccuracy of the standard measurement value and the inaccuracy of the instrument scale caused by the error[2]. If the cause is mainly due to the measurement implementation or the imperfection of the measuring instrument system which can reduce the system error by improving the measurement or improving the measuring instrument system. The gross error refers to the error that obviously deviates from the measurement result which is mainly caused by unstable and unpredictable reasons such as sudden change of measurement environment or error in measuring the implementation process which is generally eliminated or ignored in the process of measurement result analysis[3, 4]. Different elimination methods are selected for large error in different fields and the segmented adaptive threshold algorithm is studied to eliminate the clock difference data measured by optical fiber time and frequency transfer system.

Segmented threshold algorithm: According to the analysis of optical fiber time and frequency transfer data in the past 2 years, it can be found that the approximate normal distribution is obeyed while in some extreme cases the interference is large and the data distribution pattern is often irregular. Further research on the measurement data shows that after rearranging the data over a period of time from small to large, the standard deviation of the data points in the upper and lower parts is larger and the number in the middle (or the average of the two digits in the middle) is approximately the mean value. There are three kinds of rearranged data according to linear fitting: the rapid growth rate is slow after the first growth rate is fast, the slope growth rate is slow, the growth rate is fast and the slope growth rate is slow. Therefore, it needs to be judged from the side where the slope is growing fast. Eliminate points where the slope is growing at a greater rate and iterate continuously to bring it closer to the mean point. According to the description of the literature[5], due to the external temperature changes, the digital signal drift when the drift to a certain amount of error generation is easy to cause delay,

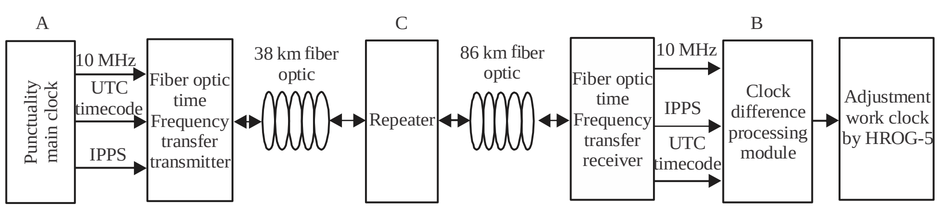

Fig. 1: Link composition of fiber optic time frequency system

although, the system uses a bidirectional link on a fiber, there is still the uncertainty of the measurement value caused by temperature change. On the basis of studying a large number of experimental data, this method selects different standard deviation as the threshold value of discrimination, combines the clock difference data characteristics of optical fiber time transfer system, carries on a lot of experiments and gives the segmented adaptive threshold algorithm.

The principle of algorithm and the determination of adaptive threshold: The general idea of the adaptive threshold algorithm given in this study is as follows: first, sort the measured data, bounded by the midpoint data and divide the data into two parts if the total number of data is even and repeat the median by two times if the total number is odd, so that, it becomes an even number of data points. The standard deviation of the front and rear parts is calculated, respectively, the standard deviation of the first half is recorded as Front_var, the mean variance of the latter part is recorded as Back_var and the comparison between the two mean variance is Max_var and the Max_var and threshold Th_var are compared, if the former is larger than the latter, The data values corresponding to the maximum mean variance are deleted and the order of the original data remains unchanged while the total number of samples is reduced by 1 and if the generous difference Max_var and Front_var or Max_var and Back_var are equal or the remaining data <3, the operation ends. The algorithm flowchart is shown in Fig. 2.

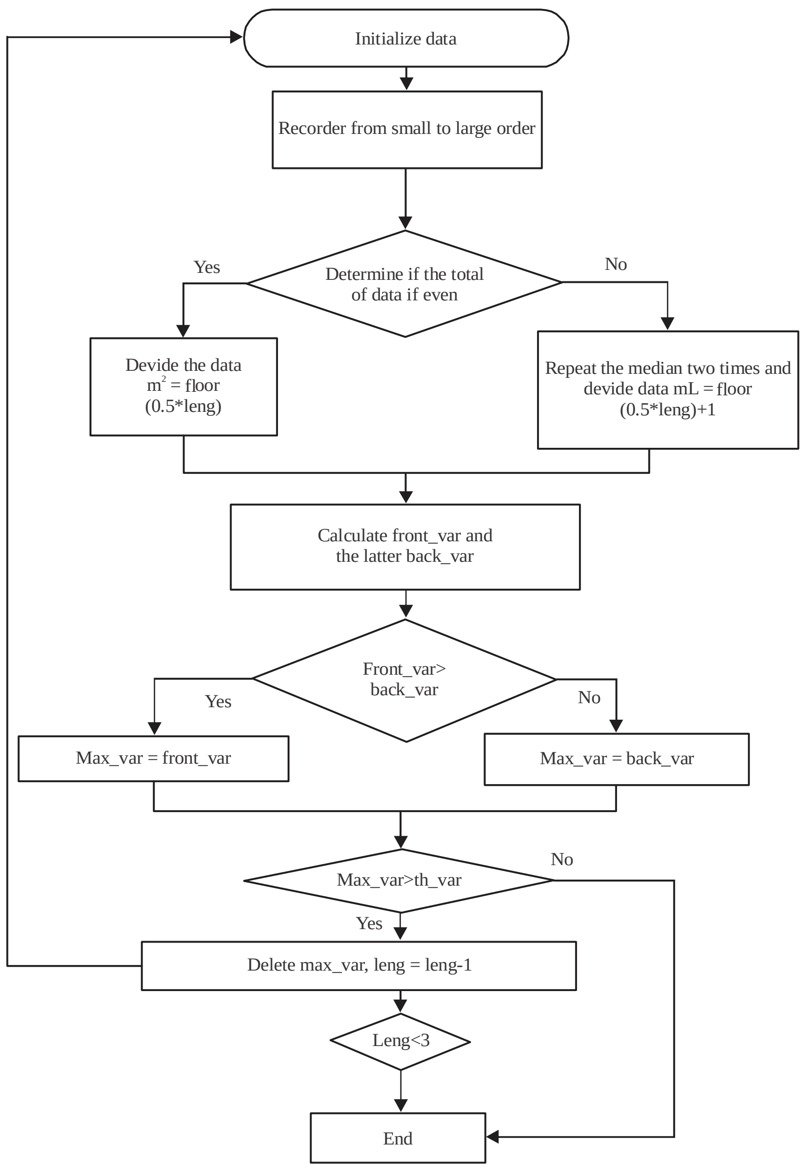

Fig. 2: The flowchart of algorithm

Obviously, the clock difference data obtained by the optical fiber time frequency transfer system has its own characteristics in the expression of the literature[5], because the fiber link has its own instability factors, the external environment has a certain impact on the optical fiber transmission, the paper is based on different original data segmentation to give different thresholds for data filtering. For each threshold determination, select a set of data with a sample length of n based on demand, the last value for that group of data, both the data collection $S = \{y_i/y_{m-n}, y_{m-n+1}, \ldots, y_{m-1}\}$. Since, the original data is a large sample event, it is calculated by the standard deviation formula modified by Bessel when calculating the dispersion degree of the research object:

$$s=\sqrt{\frac{\sum_{m=m-n+1}^{m}(y_m-\bar{y})^2}{n-1}} \quad (1)$$

$$\bar{y}=\frac{\sum_{m=m-n+1}^{m}y_n}{n} \quad (2)$$

Therefore, the scope of the algorithm setting threshold is:

$$\bar{y}_t-\gamma\sqrt{\frac{\sum_{m=m-n+1}^{m}(y_m-\bar{y})^2}{n-1}} \leq th\_var \leq \bar{y}_t+\gamma\sqrt{\frac{\sum_{m=m-n+1}^{m}(y_m-\bar{y})^2}{n-1}} \quad (3)$$

γ is the coefficient of variation. It is found that the clock difference between AB and two obtained by optical fiber time frequency transfer system can be fitted into two equation model in segments $y = f(x_1, x_2, \ldots, x_n) = 1/2ax^2_i+bx_i+c+\varepsilon(x_i)$. Divide the total amount of data N into three sub-parts, n_1, n_2, n_3. According to Eq. 1, the standard deviation calculated separately is the standard difference synthesis calculation of the three parts:

$$th\_var=\sqrt{\sum_{i-1}^{m}(\frac{\partial f}{\partial x_i})s^2_{x_i}+2\sum_{i\neq j}^{m}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}\rho_{ij}s_{x_i}s_{x_j}} \quad (4)$$

Among them $\partial f/\partial x_i$ is the error transfer coefficient of each part which is the correlation coefficient. The correlation coefficient is calculated according to the formula based on the corresponding values (ε_i, η_i) of the two sets of data measured:

$$\rho_{ij}=\frac{\sum(\varepsilon_i-\bar{\varepsilon})(\eta_i-\bar{\eta})}{\sqrt{\sum(\varepsilon_i-\bar{\varepsilon})^2\sum(\eta_i-\bar{\eta})^2}} \quad (5)$$

For different sampling lengths, the data corresponds to different thresholds, so, the algorithm has the characteristics of self-adaptation.

Advantages of the algorithm: At present, Raida guidelines, Gebras guidelines, Chauville guidelines and Dixon guidelines are commonly used to remove coarse data. Taking 2000 measurement data at the same time using the above four algorithms and comparing with the algorithm, it is found that the adaptive threshold algorithm has obvious advantages over the four algorithms in the clock difference calculation in the optical fiber time transfer system which eliminates the large data caused by the line fault and retains a lot of raw data to participate in the clock difference calculation. It lays a foundation for the subsequent work of clock difference correction.

Adaptive threshold algorithm has obvious advantages in processing speed because this algorithm takes all the data as the boundary of midpoint data which is divided into two parts corresponding to the processing calculation standard deviation, the algorithm is relatively easy, the calculation cost is small, the processing speed is fast.

The application scope of the algorithm is wide, for the unpredictable factors such as the drift of atomic clocks caused by the proportion of the error data ratio of the sample, the commonly used large error elimination algorithm can not be calculated normally and this algorithm is also applicable and this algorithm can reduce the probability of larger normal measurement value is eliminated.

EXPERIMENTAL VALIDATION

Using the optical fiber time transfer system to obtain the effectiveness of the clock difference data verification algorithm, the experimental environment is the computer, in MATLAB R2010b on the implementation of programming. Combined with the measured data, the adaptive threshold algorithm is validated by selecting a month data. In order to reduce the computation of the algorithm, first of all, the system integrity should be pre-determined, for the continuous N>60 data is greater than the threshold, it should be judged as the system link interrupt, this piece of data directly removed and alerted do not participate in the algorithm calculation as shown in Fig. 3, after the system link to return to normal. After removing the abnormal data, the temperature adaptive threshold algorithm is used to calculate the re-clock difference.

Figure 4a is a month measured by the fiber time frequency transmission system clock difference raw data, choose a data length of 2000 piece of data, in the calculation of thresholds using Eq. 4, the data into the I = 3 group calculation threshold, Fig. 4b for the use of the clock difference threshold algorithm after the remaining data as can be seen from the diagram, After applying the fixed clock difference threshold algorithm, the clock difference data is more concentrated, the original 2000

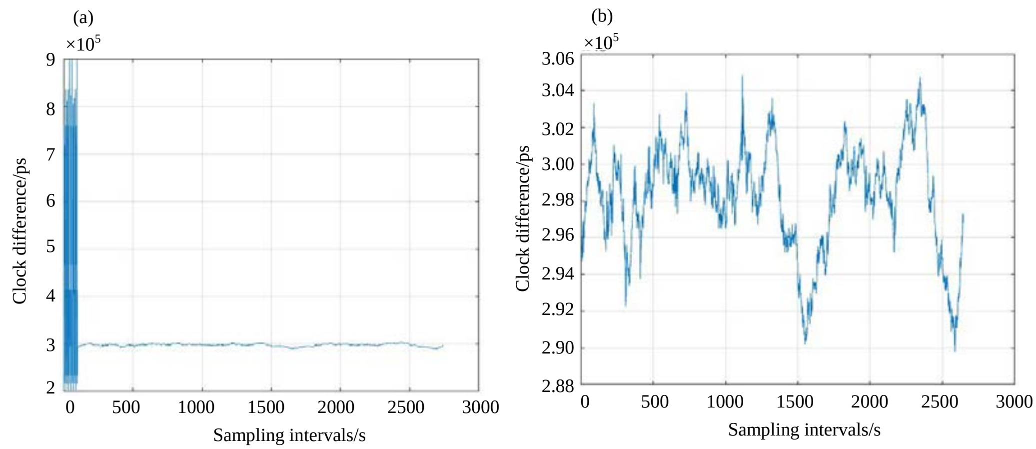

Fig. 3(a, b): Link integrity preview

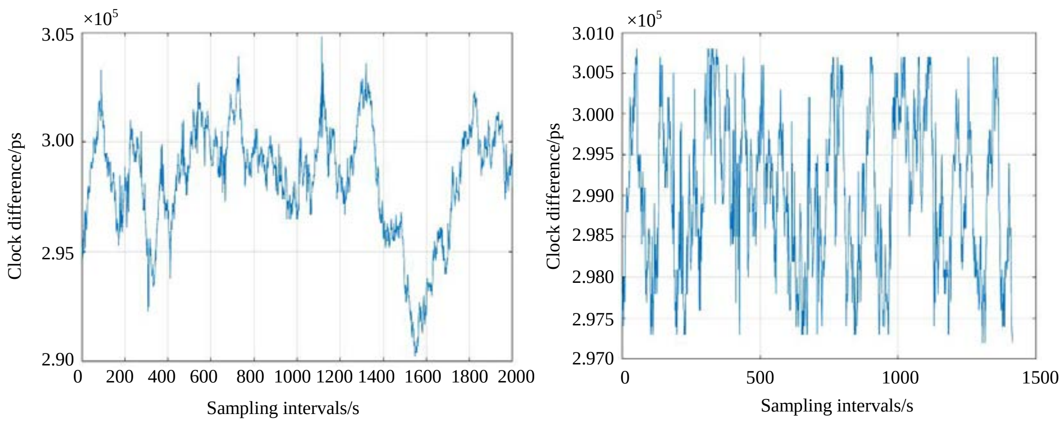

Fig. 4(a, b): Adaptive threshold algorithm effect

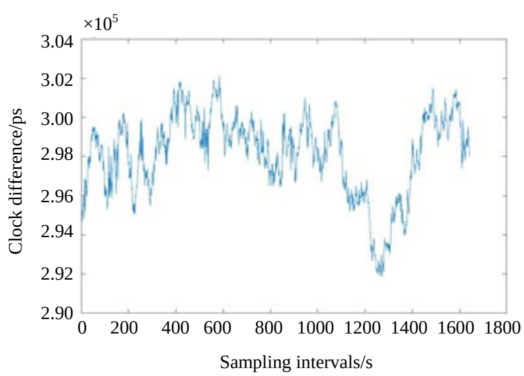

Fig. 5: The effect of segmented adaptive threshold algorithm i = 10

data becomes 1432 and nearly 70% data is retained to participate in the subsequent clock difference correction calculation but because the amount of data is too large, the grouping is not detailed enough and the original data needs to be reordered after removing the data greater than the threshold value, This results in a number that should have been removed in the I = 1 group during the sorting process which may actually exclude the same data from the I = 3 group but the recalculation threshold in the I = 3 group is likely not to be excluded, resulting in a change in the order of the data in the event that the total amount of remaining data remains the same. This will then affect the next calculation.

In order to solve this problem, the same data is divided into several intervals, each interval using the above algorithm, respectively, the data into 10 groups or 20 groups, namely i = 10 and i = 20, respectively, each set of data using this algorithm to deal with, the test results as shown in Fig. 5 and 6, from which can be seen that: take i = 10 o'clock Data surplus is 1623, accounting for 81% of the total data. When taking i = 20, the amount of remaining data is 1781, accounting for 89% of the total data. The adaptive threshold algorithm of i = 20 basically retains the order of the original data, the more the

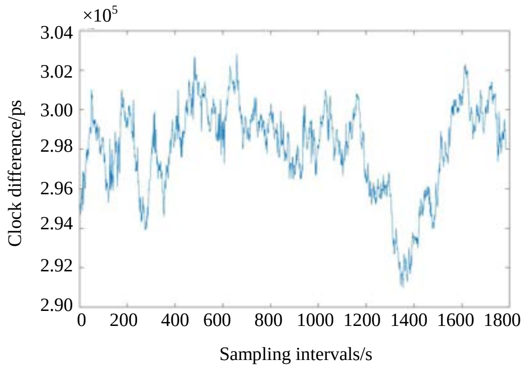

Fig. 6: The effect of segmented adaptive threshold algorithm i = 20

grouping in theory, the stronger the self-adaptability but the high cost of too much packet computation. It is found that it is better to select a window with a length of 100 to carry out the calculation of "moving window".

CONCLUSION

This study analyzes the clock difference data obtained by the optical fiber time frequency transfer system and puts forward the use of segmented threshold algorithm to eliminate the coarse error in the extreme environment of the system without link interruption, because the data is distributed locally in two times, so, the algorithm of segmented synthesis standard deviation is designed in the selection of threshold value and it is proved by experiments that the algorithm can keep as much useful data as possible while removing the coarse error. Because this algorithm reordered the original data in advance, resulting in the algorithm in the case of a large amount of data, the elimination of the original data location is uncertain, affecting the subsequent calculation results. It is proved by experiments that by grouping the original data first, the influence of the uncertainty of the original data position can be greatly reduced by the method of this algorithm. Therefore, the algorithm removes the coarse error and retains rawer data which lays a foundation for the subsequent research work and has practical significance[6].

REFERENCES

01. Changlian, X., 1992. Prestack mid-value filtering. Oil Geophys. Prospect., 27: 387-396.
02. Liu, Z., S. Zhou, J. Tan and H. Tang, 2013. Multiple parameters correlation analysis and prediction method of product performance based on multi-criteria modification. J. Mech. Eng., 49: 105-114.
03. Liu, Z., 2009. Treatment analysis of rough error about measurement instrument. Electron. Meas. Technol., 32: 55-57.
04. Zhonghao, S. and M. Suofang, 1996. The testing of temperature characteristics of delay in optical fiber. J. Nanjing Instit. Posts Telecommun., 16: 82-85.
05. Zongbao, L., G. Shiqiao and D. Jingqing, 2012. An algorithm for measurement data eliminating gross error and smoothing. Proceedings of the 31st Chinese Control Conference, July 25-27, 2012, IEEE, Hefei, China, pp: 7582-7587.
06. Yu, J., 2003. Gross error procesing technology in intelligence measuring system with high precision. J. Test Meas. Technol., 17: 258-261.