

Application of Sliding Nest Window Control Chart in Data Stream Anomaly Detection

Guang Li ^{1,2,*}, Jie Wang ^{1,*}, Jing Liang ¹ and Caitong Yue ¹

¹ School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China; liangjing@zzu.edu.cn (J.L.); zzyuecaitong@163.com (C.Y.)

² The 22th Research Institute of China Electronics Technology Group Corporation, Xinxiang 453003, China

* Correspondence: goldsun18@126.com (G.L.); wj@zzu.edu.cn (J.W.); Tel.: +86-138-3710-6273 (J.W.)

Received: 9 April 2018; Accepted: 12 April 2018; Published: 17 April 2018



Abstract: Since data stream anomaly detection algorithms based on sliding windows are sensitive to the abnormal deviation of individual interference data, this paper presents a sliding nest window chart anomaly detection based on the data stream (SNWCAD-DS) by employing the concept of the sliding window and control chart. By nesting a small sliding window in a large sliding window and analyzing the deviation distance between the small window and the large sliding window, the algorithm increases the out-of-bounds detection ratio and classifies the conceptual drift data stream online. The designed algorithm is simulated on the industrial data stream of drilling engineering. The proposed algorithm SNWCAD is compared with Automatic Outlier Detection for Data Streams (A-ODDS) and Distance-Based Outline Detection for Data Stream (DBOD-DS). The experimental results show that the new algorithm can obtain higher detection accuracy than the compared algorithms. Furthermore, it can shield the influence of individual interference data and satisfy actual engineering needs.

Keywords: data stream; nested sliding window; anomaly detection; machine learning; concept drift; control chart

1. Introduction

The data stream is ubiquitous in real-life and industrial data [1] such as the industrial sensor data stream [2,3], the Internet data stream [4–8], cloud computing [9], Internet of Things [10], the stock data stream [11], and the traffic data stream [12]. The data stream has the characteristics of changing with time, being endless, real-time processing, being high-dimensional, irregular processing, and more. There are many anomaly detection models such as the landmark model [13], the sliding window model [14], and the snapshot model [15]. The sliding window model only considers the latest M (M is the sliding window size) data in the data stream. As the data continues to arrive, the data in the window is updated over time. Because the sliding window model is simple and effective in the real application, it is widely studied. This paper discusses the data flow anomalies in many practical engineering applications as well as the degree of the abnormal size. Then, the size of the abnormal degree is put into the model for prediction or decision-making. One of the practical application problems is the industrial warning system.

The data stream anomaly detection includes concept drift [16] and drift level [17] detection, which is an online classification problem. According to the papers published in the international authoritative journals in machine learning and data mining and top international conferences in recent years, the research of data stream anomaly detection [18] is becoming a popular research topic. The characteristics of the data stream make data stream anomaly detection a challenging task. This is because most machine learning algorithms are based on static data environment.

Therefore, they are not suitable for data stream anomaly detection. There are two tricky problems. One is the high computational complexity of the model and the other is the concept of drift that will change the model. Therefore, researchers are trying to solve data stream anomaly detection using the sliding window model and the control chart [19], evolutionary computation [20], transfer learning [21], clustering [22], feature selection [23], integrated learning [24], and the combination of various model classification theories [25]. Several combination models have been proposed such as the combination of the sliding window and control chart for Internet data stream [26] anomaly detection, the combination of sliding window, and partial differential equation sorting (PDEs) algorithm to detect an abnormal data stream [27], data stream anomaly detection combining the sliding window, and integrated learning [28]. There have been several control chart methods for data stream anomaly detection such as non-parametric accumulation and control charts electrocardiographic (ECG) anomaly detection [29], weighted calculation cumulative sum (CUSUM) study [30], CUSUM-based satellite power supply system anomaly detection [31], and an online consumption forecast based on CUSUM [32]. The contributions of this paper include adopting the sliding window nested way to increase trend prediction accuracy, increasing the out-of-bounds detection rate, weakening the influence of the current point, and using the CUSUM algorithm to combine the above two points and then test the real-time data stream.

As shown in Figure 1, due to the influence of the sensor's position, working environment, and the quality of the sensor, the data stream usually has a small cluster of points (rectangle point data in Figure 1), which deviates from the overall trend. The rectangle point data are not really abnormal data but a kind of interference data. If this kind of data are also considered abnormal data, it will lead to the frequent triggering of the prediction model and increase the false alarm frequency of abnormal detection. Late-period data deviations from the normal trend (trapezoidal points in Figure 1) are the real underlying causes, which lead to data rising. Such data is often caused by the internal environment change and are the focus of our attention. This kind of internal environmental change and breakdown are what we want to analyze, which is consistent with our prediction goal. These data need to be correctly labeled. Therefore, there is a research direction for data stream anomaly detection to reduce the influence of individual point randomness, reduce false positives, increase trend judgment, and improve prediction accuracy.

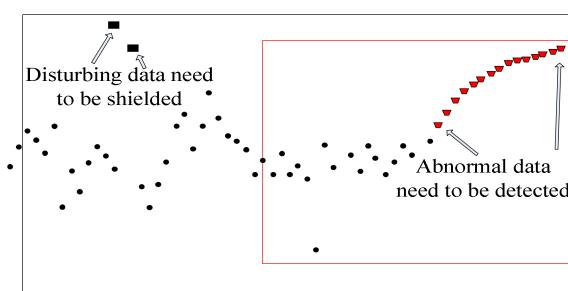


Figure 1. Schematic Diagram.

The long window is a piece of data that extracts the fixed length of the distance from the infinite data stream to analyze the trend of the data stream. The short window is built to suppress the randomness of the current data of the data stream and displays the immediate changes of the current data. Comparing and analyzing the changes in the data in the long and short windows, we can analyze the changes of the current data and can also shield some misclassification caused by the randomness of the current data.

As shown in Figure 2, data stream anomaly detection is a necessary part of many decisions in reality, such as network attack defense, industrial security accident early warning, traffic congestion warning, and stock decisions. First, the data stream anomaly and the degree of anomaly must be

detected, and the abnormal data stream of several parameters is combined to train the decision-making model. With enough data trained models to monitor incoming data streams, the model is triggered when several data stream exceptions are combined. Afterwards, automatic decisions are made.

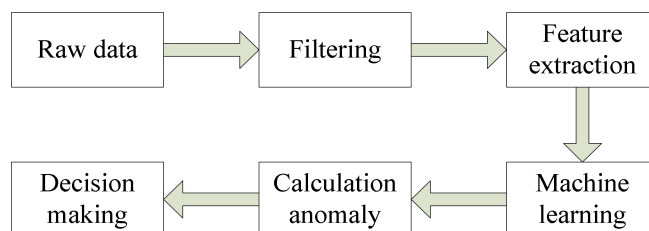


Figure 2. Data stream decision process.

The difficulty of online data stream classification lies in the unknown distribution of the data stream. It is difficult to judge whether the arriving data are real anomalies or interference and it is not easy to distinguish only by several incoming adjacent data. If the concept data drift of adjacent data occurs, follow-up data continues this trend in accordance with the trend of drift, which can be considered true anomalous data. This kind of data are caused by some inherent mechanism, which require special attention. However, if less data return to the overall trend over a shorter period, it may be due to erratic sensor operation or some changes in the working environment. A traditional machine learning classification model is based on historical data training. Due to the characteristics of real-time data flow, distinguishing between disturbing data and abnormal data is difficult in data stream anomaly detection. To solve this kind of problem, the Piecewise Linear Approximation (PLA) [33] was proposed for adopting the best fitting of the distance error. It adopts the linear regression method with piecewise linear approximation. The discrete wavelet transform (DWT) method and complex discrete wavelet transform (cdwt) method employed fragment similarity of the time data stream [34]. Even though the above method analyzes the data stream anomaly detection, they do not analyze whether the current point has abnormal interference or is a normal point. This paper proposes a sliding nest window method. It can weaken the influence of the sampling data of the concept drift and it enhances the trend analysis ability. In addition, it can suppress the interference of noise.

This paper first introduces the background of the algorithm. Then the nested sliding window anomaly detection algorithm is described in detail. The proposed algorithm is compared with other relevant algorithms. Lastly, the paper is summarized.

2. A Nested Sliding Window Data Stream Anomaly Detection Algorithm

The object of machine learning mining is the property of commonality in the data. The data stream is affected by the conceptual drift and the feature needs to be extracted from the original value to explore the common attributes in the data stream. Although data streams drift in concept as the environment changes, the degree of drift is generally constant, or the degree of drift varies proportionately with an attribute value. Therefore, extracting this constant attribute value from the original value is the key for data stream machine learning.

In this paper, we use the box diagram control chart to detect the abnormal data stream and its size. The data stream traverses close to the median including from the lower part to the upper part. The area from the lower limit to the upper limit is normal. The data above the upper limit and the lower limit are key monitoring objects. After a few data points of a certain percentage continuously exceed the limit, the data stream can be regarded as an anomaly according to a certain trend. The distance between the current point and the upper limit or the lower limit needs to be calculated as the magnitude of the anomaly of the point. Since the data stream drifts with the time concept, the median value in the box graph changes constantly over time and can adaptively track the data stream trend. However,

the relative value is relatively fixed. The upper limit and lower limit values will be relatively fixed or take fixed adjustment with certain parameters.

The mean value in Figure 3 is obtained by truncating the data stream in the sliding window. Lower and upper hinges are the normal data boundary lines. The data between the two lines are normal fluctuations without monitoring. The data between the lower hinge and the lower limit and those between the upper hinge and the upper limit are suspicious abnormal data, which need to be targeted, but no abnormality calculations are carried out on them. Data out of the lower limit and the upper limit are identified as anomalies and need to be processed.

$$\text{RootMeanSquare}(RMS) = \left[\frac{1}{N} \sum_{i=1}^N [X(i)]^2 \right]^{\frac{1}{2}} \quad (1)$$

$$\text{Mean} = \frac{1}{N} \sum_{i=1}^N X(i) \quad (2)$$

$$\text{Variance}(\delta^2) = \frac{1}{N} \sum_{i=1}^N (X(i) - \mu)^2 \quad (3)$$

$$\text{Skewness} = \frac{1}{N} \sum_{i=1}^N \left(\frac{X(i) - \mu}{\delta} \right)^3 \quad (4)$$

$$\text{Kurtosis} = \frac{1}{N} \sum_{i=1}^N \left(\frac{X(i) - \mu}{\delta} \right)^4 \quad (5)$$

$$\text{GrestFactor} = \frac{\max(|X|)}{RMS} \quad (6)$$

$$\text{ImpluseFactor} = IF = \frac{\max(|X|)}{\frac{1}{N} \sum_{i=1}^N |X(i)|} \quad (7)$$

$$\text{ShapeFactor} = SF = \frac{RMS}{\frac{1}{N} \sum_{i=1}^N |X(i)|} \quad (8)$$

$$\text{Median} = \text{magnitude}\left(\frac{N+1}{2}\right) \quad (9)$$

$$\text{Range} = \text{Max}(X(i)) - \text{Min}(X(i)) \quad (10)$$

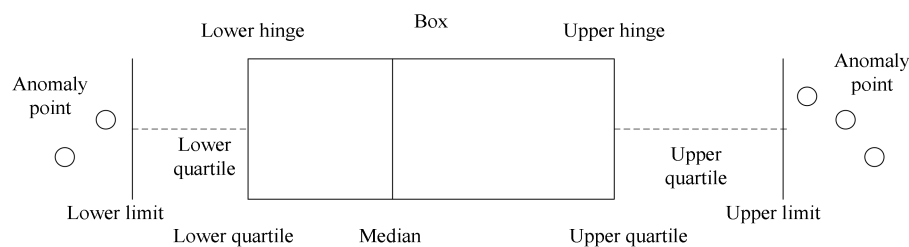


Figure 3. Parameters of Box plot.

This paper analyzed the data stream anomaly detection. The concepts and definitions are presented as follows:

Definition 1. Data stream DS. The data stream is a sequence of data generated continuously in chronological order, which is represented mathematically as $DS = \{(x_0, t_0), (x_1, t_1), \dots, (x_i, t_i), \dots\}$. Among them, x_i is the data arriving at t_i .

Definition 2. Short window distance (number of data) SWM. Number between the current point and the closer fixed interval data. The length of short window is m .

Definition 3. Long window distance (number of data) LWM. Number between the current point and the far fixed interval data. The length of the long window is M .

Definition 4. Sliding window. Given a data stream sequence T with length n and a user-defined subsequence with length M . The length of the sliding window is M . All subsets of all subsequences taken from the data stream sequence T by the sliding window constitute s .

Definition 5. Threshold $\mu\sigma$. μ is a constant. If the average value exceeds μ , it will be considered abnormal.

Definition 6. The number of out-of-bounds data is h . The out-of-bounds rate is h/M .

The nested sliding window data stream anomaly detection algorithm is outlined below.

- (1) Calculate the short window mean (SWM).

$$SWM = \frac{1}{m} \sum_{i=1}^N x_i \quad (11)$$

- (2) Calculate long window mean (LWM).

$$LWM = \frac{1}{M} \sum_{i=1}^M x_i \quad (12)$$

- (3) Calculate the mean difference.

$$mdv = SWM - LWM \quad (13)$$

- (4) Use CUSUM principle to calculate the cumulative sum.

$$cusum = \sum_{i=1}^M mdv_i \quad (14)$$

- (5) Calculate the cumulative sum.

$$mcusum = \frac{1}{M} \sum_{i=1}^M cusum_i \quad (15)$$

- (6) Calculate the cumulative sum and variance.

$$slc = std(cusum - mcusum) \quad (16)$$

- (7) Calculate the upper critical line.

$$UpperCriticalLine(UCL) = mcusum + \mu \times slc \quad (17)$$

- (8) Calculate the lower critical line.

$$DownCriticalLine(DCL) = mcusum - \mu \times slc \quad (18)$$

- (9) Analyze whether the current value is abnormal. If $cusum > UCL$, the current point is considered abnormally increased and the point flag is marked as 1. The abnormal accumulated number $mcusum = 1$. If $DCL > cusum$, the current point is considered to be abnormally decreased and

the point flag is marked as 2. The abnormal accumulated number $mcusum = 1$. If there are no abnormal points detected, the flag is marked as 0 and $mcusum = 0$.

- (10) Calculate the current point flag. If the current point is the same as the previous moment mark, then the abnormal point accumulates as 1. Otherwise, the abnormal point is 1.
- (11) Calculate the abnormal cumulative number. If the abnormal cumulative number $> h$, the current point is the abnormal point.

The traditional Shewhart control chart confirms that the 3σ principle is used for UCL and DCL. The 3σ principle is based on the data that are normally distributed or approximate to normal distribution and the sampling number is large enough. The data stream sliding window analysis method truncates the amount of sampling data. The concept deviation is common in the dynamic data stream. They are not normal distribution or approximate normal distribution, which means the traditional 3σ mode is not appropriate. As shown in Figure 4, This paper treats 3σ as a parameter to be determined. This is obtained through data training.

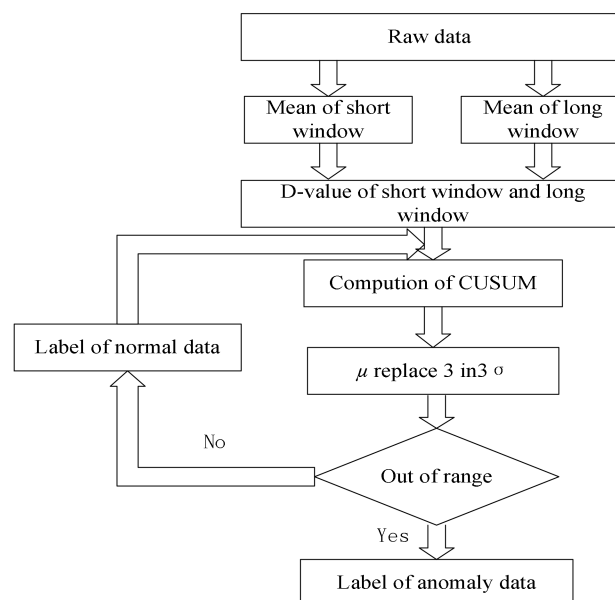


Figure 4. Algorithm stream chart.

SNWCAD-DS adopt nested windows to cut data streams. It only analyzes the data in long and short windows and classifies them by comparing data deviations in long and short windows. As time goes on, the data in the window is constantly updated, the new data constantly enters the window, the old data is constantly removed from the window, and the data in the window is updated each time. The SNWCAD-DS algorithm is performed, and computation is the eleventh step such as the nested sliding window data stream anomaly detection algorithm. The real-time classification of the data flow values is realized. The algorithm can not only satisfy the classification difficulty brought by the concept drift, but also shield the randomness of the current data and improve the accuracy of data stream anomaly detection.

3. Simulation and Comparison

There are two key points in data stream anomaly detection including the accuracy of data stream classification and the computational complexity of the algorithm. If the computational complexity is high, the algorithm is not suitable for online data stream detection even though its accuracy is high. In addition, if the algorithm classification accuracy is too low, it is still meaningless even though its computational complexity is low. Therefore, the online classification accuracy should be improved as

much as possible, which is the principle of data stream anomaly detection. The proposed algorithm is compared with DBOD-DS [35] and A-ODDS [36]. The AUC (Area under Curve) index and the Jaccard coefficient in the receiver operating characteristic curve (ROC) are employed as a performance indicator. A satisfactory outlier detection technique is one that maximizes true positive (TP) values and minimizes false negative (FN) and false positive (FP) values. ROC detection accuracy is calculated according to the equations below.

$$\text{TruePositiveRate}(TPR) = TP / (TP + FN) \quad (19)$$

false alarm rate:

$$\text{FalsePositiveRate}(FPR) = FP / (FP + TN) \quad (20)$$

Jaccard coefficient

$$\text{JaccardCoefficient}(JC) = TP / (TP + FP + FN) \quad (21)$$

where TP (True Positive) is the correct number, FN (False Negative) is the number of missed reports and the unmatched one, FP (False Positive) is mis-declaration, and TN (True Negative) is correct rejective Non-matching logarithms. The traditional method to calculate AUC is by changing the threshold, but this algorithm adopts the method of online learning and the coefficient of variance, which considers the whole $\mu \times slc$ as one threshold. In addition, ROC is a dichotomy, but this article solves the problem of three classes. Therefore, this paper regards the abnormal rise and abnormal decline as one category and no abnormalities as another category. In the anomalous ascent and the decline class, we still distinguish the anomalous rise and the abnormal decline. Therefore, this article focuses on three categories. In this paper, the AUC area of the ROC is used as the classification accuracy and the test time is used as algorithm computational complexity.

The data are from a certain well in the Tarim Oilfield and they are tagged by expert experience. Machine learning was performed on slowly changing data, which were based on the data stream anomaly detection of the segment distance. The obtained graph is shown in Figure 5.

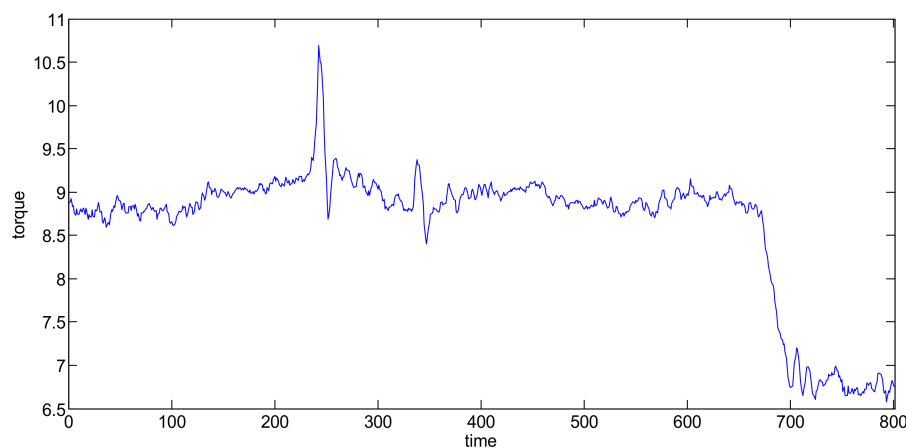


Figure 5. Drilling engineering torque data stream.

As can be seen from Figure 5, the point near 230 clearly deviates from the normal trend range. According to the characteristics of the data stream analysis, when only a few data points arrive, it cannot be analyzed whether the data are really abnormal or caused by interference. So, the analysis of the latter data is required. The purpose of data stream online machine learning is to correctly classify the data stream.

The parameters of each algorithm are set according to Table 1. The algorithms are tested on the data shown in Figure 5. The normal data are labeled 0, the abnormal increase data are labeled 1, and the abnormal decline are labeled 2. The simulation results are shown in Figure 6. The parameters that are not in the algorithm are replaced by oblique lines.

Table 1. Parameter setting table.

Item	Short Window	Long Window	Threshold	Out Rate
SNWCAD-DS	20	50	0.3	8
DBOD-DS	/	50	0.3	/
A-ODDS	20	50	0.3	/

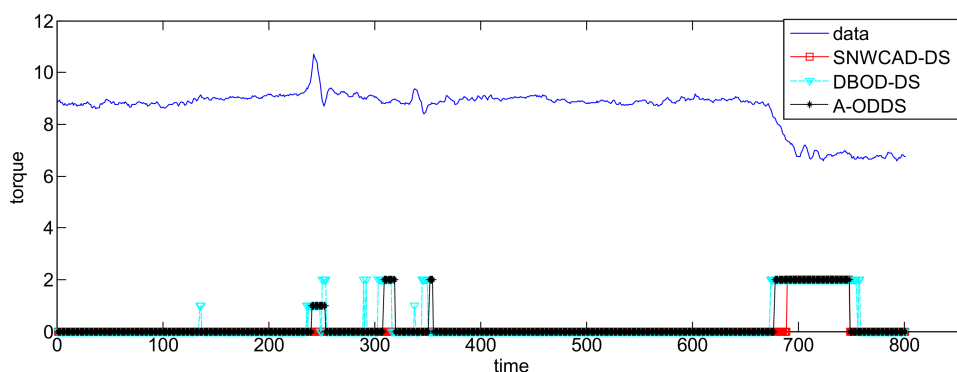


Figure 6. Comparison of classification results.

As can be clearly seen from Figure 6, SNWCAD-DS masks interferences in the range of 200–400 and can determine the actual anomalous drop in the 600–800 range. The original algorithms, DBOD-DS and A-ODDS cannot shield the interference.

Figure 7 is the distribution diagram of each algorithm of the Jaccard coefficient. As can be seen from Figure 7, the Jaccard coefficient of the proposed algorithm is higher than DBOD-DS and A-ODDS for different settings of long window length.

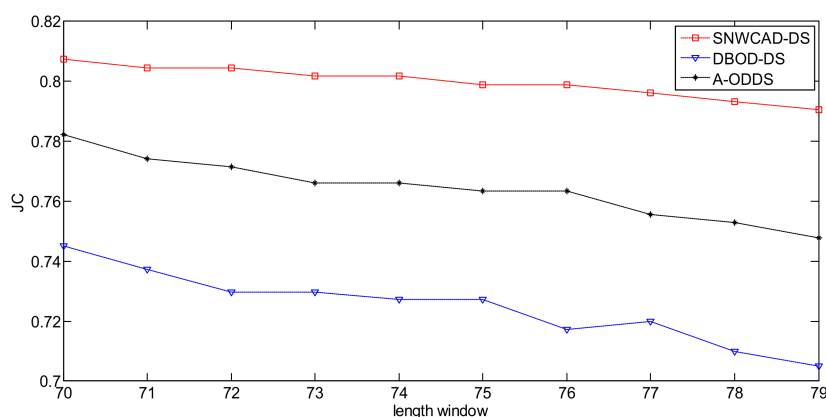


Figure 7. JC of each algorithm.

As shown in Table 2, The running time of the proposed algorithm does not increase too much and can fully meet the requirements of online real-time running. The above analysis is based on online machine learning for slowly changing parameters. The following analysis simulates a three-category data stream with dramatic changes and a rising trending to verify the effectiveness of the proposed algorithm.

Table 2. Running time of different length of long window.

Length of Long Window	SNWCAD-DS	DBOD-DS	A-ODDS
70	0.0312	0.0312	0.0312
71	0.0312	0.0156	0.0312
72	0.0312	0.0312	0.0156
73	0.0312	0.0156	0.0312
74	0.0468	0.0156	0.0312
75	0.0312	0.0156	0.0315
76	0.0312	0.0312	0.0312
77	0.0312	0.0157	0.0312
78	0.0313	0.0156	0.0314
79	0.0313	0.0156	0.0313
average	0.0328	0.0203	0.0297

As shown in Figure 8, the (+) points are the normal data, the (o) points are rising data, and the (trapezoid) points are the falling data. Non-abnormal data flag is “-”, abnormal decline flag is “o”, and abnormal rise flag is “+”. The purpose of the anomaly detection algorithm is to mark the entire curve in different positions correctly. To compare the performance of different algorithms, the data detection is performed under the same setting for the four parameters. The settings of the four parameters include the long window 500, the short window 30, threshold 0.3, and the abnormal number 8.

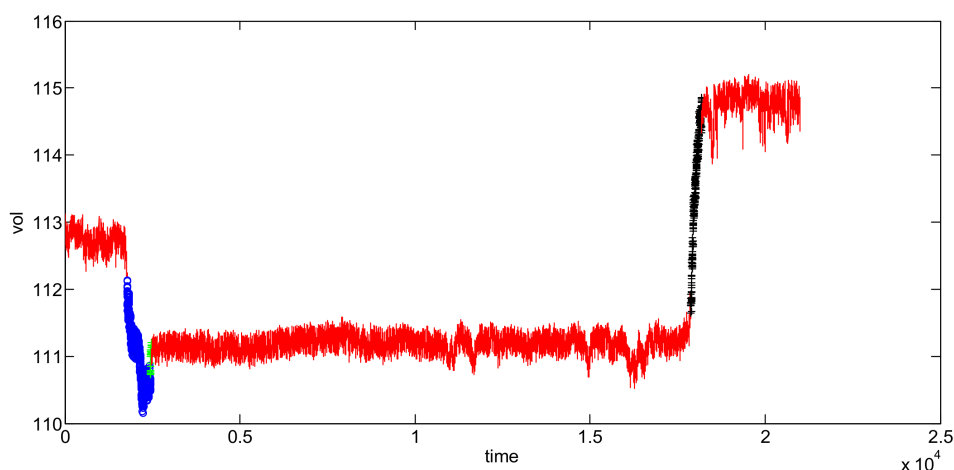
**Figure 8.** Raw data and abnormal signs.

Figure 9 is a schematic diagram of the algorithm abnormal judgments. The top curve is the upper threshold line. The bottom curve is the lower threshold line. The middle line is considered normal data. Data which exceed the upper and lower threshold line are false outliers. If the number of false outliers in the long window is larger than h , then the current sampling point will be the abnormal point. The whole figure implies window sliding operation.

Figure 10 maps the original value to the SNWCAD-DS space to determine whether the parameters are abnormal through the Shewhart control chart. ‘A’ is the original data graph, ‘B’ is the converted data graph, and ‘C’ is the data graph mapped to the SNWCAD-DS space. As shown from the bottom graph in Figure 10, the original data have clear abnormalities in two places. The Shewhart control chart can quickly determine the abnormal point and its location.

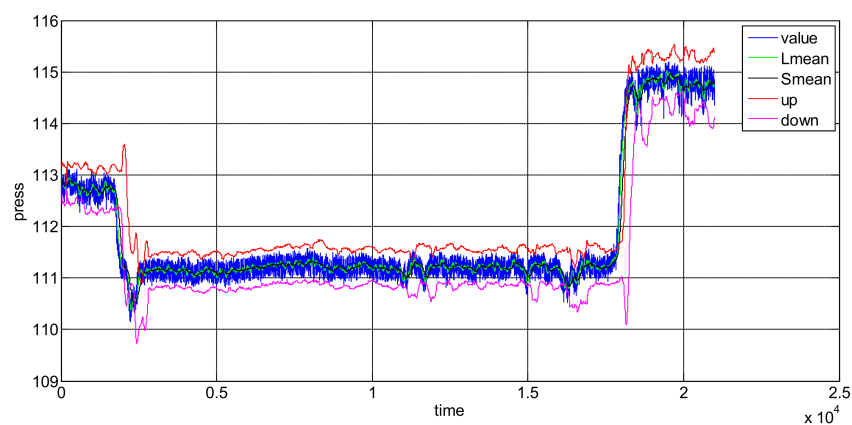


Figure 9. Data analysis diagram.

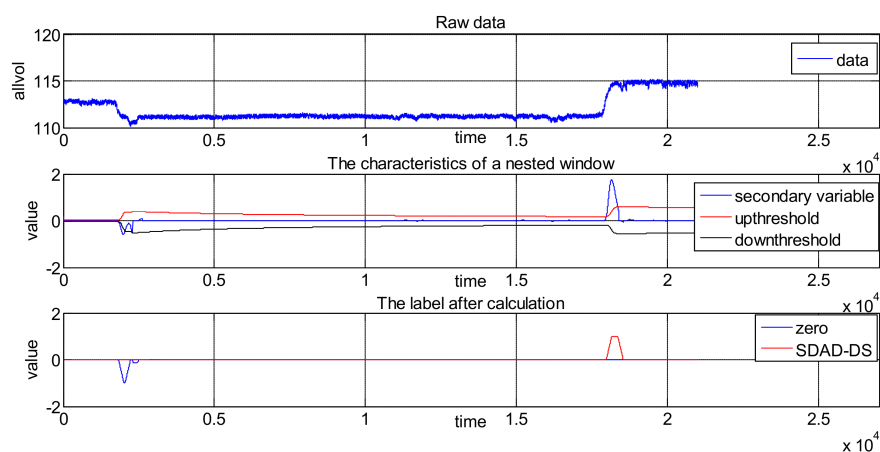


Figure 10. Schematic diagram of data decomposition.

As shown in Figure 11, the proposed algorithm SNWCAD-DS has clear advantages than DBOD-DS and A-ODDS, which can significantly improve TPR and reduce FPR.

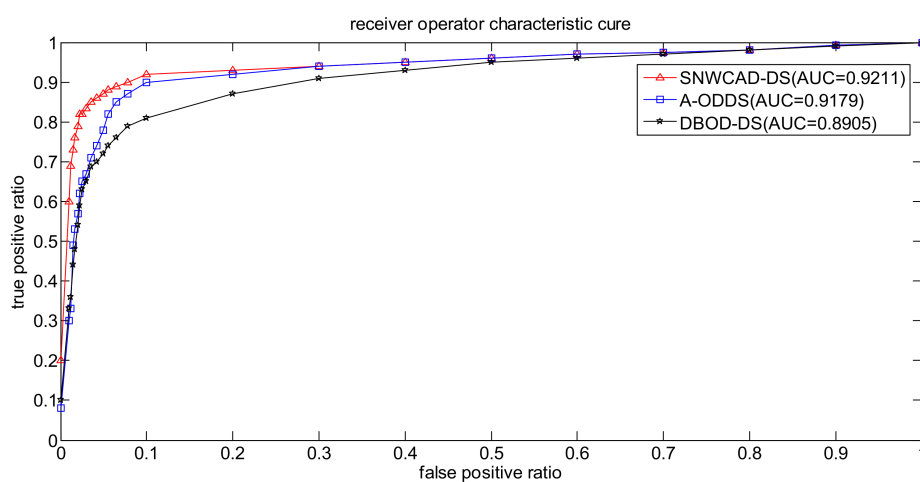


Figure 11. ROC curve and AUC area.

The running time of different algorithms is compared. This experiment adopts 500 labeled data points. The operating environment involves the CPU dual-core 2.1 GHz, Win7 Sp1 x86, and memory 2 G. The running time is shown in Table 3.

Table 3. Algorithm running time (s) comparison table.

Algorithm	SNWCAD-DS	DBOD-DS	A-ODDS
Running time	0.049660	0.0456750	0.0487450

It can be seen from Table 3 that the complexity of the SNWCAD-DS algorithm is slightly larger, but the computational accuracy is higher than those of DBOD-DS and A-ODDS. In the practical application, the accuracy of SNWCAD-DS can be improved without affecting the online operation.

DBOD-DS adopts the adaptive probability density function to detect anomaly data. This method executes in a single window and does not analyze the influence of current data randomness on the classification error rate. A-ODDS detects anomaly data by calculating the deviations of global data and local data changes. In fact, the method adds one window because it is the calculation of all the data. With the accumulation of data, more and more historical data are obtained, which is equivalent to the increasing weight of the history data in the method of judgment and the existence of the history data weight. For the same problem of the DBOD-DS algorithm, the algorithm does not shield the randomness of the current data.

Conclusion 1: SNWCAD-DS algorithm can effectively improve the detection accuracy.

Conclusion 2: SNWCAD-DS algorithm can effectively reduce the false alarm rate.

Conclusion 3: SNWCAD-DS algorithm complexity meets the needs of practical application.

4. Summary

In this paper, the SNWCAD-DS is proposed to detect data stream anomaly. It is based on the segmented distance data stream and simulated in data-driven drilling fault diagnosis. Compared with the DBOD-DS algorithm and the A-ODDS algorithm, the proposed algorithm can not only meet the online operation complexity, but also significantly improve the accuracy of online data classification and reduce the false alarm rate. It provides a new distance-based algorithm for online machine classification and promotes the improvement of data stream anomaly detection technology.

Acknowledgments: National Natural Science Foundation of China (61473266 and 61673404), Project supported by the Research Award Fund for Outstanding Young Teachers in Henan Provincial Institutions of Higher Education of China (2014GGJS-004) and Program for Science & Technology Innovation Talents in Universities of Henan Province in China (16HASTIT041).

Author Contributions: Guang Li and Caitong Yue conceived and designed the experiments; Guang Li performed the experiments and analyzed the data; Jie Wang and Jing Liang provided guidance and recommendations for this research. Guang Li contributed to the contents and writing of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Paul, G.V.; Brauker, J.H.; Kamath, A.U.; Thrower, J.P.; Carr-Brendel, V. Systems and Methods for Replacing Signal Artifacts in a Glucose Sensor Data Stream. U.S. Patent 8,790,260B2, 29 July 2014.
2. Ge, M.; Wang, J.; Ren, X. Fault Diagnosis of Rolling Bearings Based on EWT and KDEC. *Entropy* **2017**, *19*, 633. [CrossRef]
3. Al-Janabi, S.; Rawat, S.; Patel, A.; Al-Shourbaji, I. Design and evaluation of a hybrid system for detection and prediction of faults in electrical transformers. *Int. J. Elect. Power Energy Syst.* **2015**, *67*, 324–335. [CrossRef]
4. Baccarelli, E.; Cordeschi, N.; Mei, A.; Panella, M.; Shojafar, M.; Stefa, J. Energy-efficient dynamic traffic offloading and reconfiguration of networked data centers for big data stream mobile computing: Review, challenges, and a case study. *IEEE Netw.* **2016**, *30*, 54–61. [CrossRef]

5. Siddique, K.; Akhtar, Z.; Lee, H.G.; Kim, W.; Kim, Y. Toward Bulk Synchronous Parallel-Based Machine Learning Techniques for Anomaly Detection in High-Speed Big Data Networks. *Symmetry* **2017**, *9*, 197. [\[CrossRef\]](#)
6. Javanmardi, S.; Shojafar, M.; Shariatmadari, S.; Ahrabi, S.S. Fr trust: A fuzzy reputation-based model for trust management in semantic p2p grids. *Int. J. Grid Util. Comput.* **2014**, *6*, 57–66. [\[CrossRef\]](#)
7. Shojafar, M.; Pooranian, Z.; Naranjo, P.G.V.; Baccarelli, E. FLAPS: Bandwidth and delay-efficient distributed data searching in Fog-supported P2P content delivery networks. *J. Supercomput.* **2017**, *73*, 5239–5260. [\[CrossRef\]](#)
8. Majeed, M.F.; Ahmed, S.H.; Muhammad, S.; Song, H.; Rawat, D.B. Multimedia streaming in information-centric networking: A survey and future perspectives. *Comput. Netw.* **2017**, *125*, 103–121. [\[CrossRef\]](#)
9. Canali, C.; Chiaraviglio, L.; Lancellotti, R.; Shojafar, M. Joint Minimization of the Energy Costs from Computing, Data Transmission, and Migrations in Cloud Data Centers. *IEEE Trans. Green Commun. Netw.* **2018**, 1–16. [\[CrossRef\]](#)
10. Lan, K.; Fong, S.; Song, W.; Vasilakos, A.V.; Millham, R.C. Self-Adaptive Pre-Processing Methodology for Big Data Stream Mining in Internet of Things Environmental Sensor Monitoring. *Symmetry* **2017**, *9*, 244. [\[CrossRef\]](#)
11. Thalor, M.A.; Patil, S.T. Learning on High Frequency Stock Market Data Using Misclassified Instances in Ensemble. *Learning* **2016**, *7*, 283–288.
12. Pei, Y.; Li, X.; Yu, L.; Yu, L.; Li, G.; Ng, H.H.; Hoe, J.K.; Ang, C.W.; Ng, W.S.; Takao, K.; et al. A Cloud-Based Stream Processing Platform for Traffic Monitoring Using Large-Scale Probe Vehicle Data. In Proceedings of the 2017 IEEE Wireless Communications and Networking Conference (WCNC), San Francisco, CA, USA, 19–22 March 2017; pp. 1–6.
13. Lander, C.; Wiehr, F.; Herbig, N.; Krüger, A.; Löchtefeld, M. Inferring landmarks for pedestrian navigation from mobile eye-tracking data and Google Street View. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 2721–2729.
14. Simão, M.A.; Neto, P.; Gibaru, O. Unsupervised Gesture Segmentation by Motion Detection of a Real-Time Data Stream. *IEEE Trans. Ind. Inform.* **2017**, *13*, 473–481. [\[CrossRef\]](#)
15. Wang, H.B.; Hui, X.B.; Lin, J.F. The research of data stream mining and application in fault diagnosis of equipment. In Proceedings of the 2016 International Conference on Mechanical Engineering and Control System (MECS2016), Wuhan, China, 15–17 April 2016; pp. 101–107.
16. Costa, F.G.D.; Duarte, F.S.L.G.; Vallim, R.M.; de Mello, R.F. Multidimensional surrogate stability to detect data stream concept drift. *Expert Syst. Appl.* **2017**, *87*, 15–29. [\[CrossRef\]](#)
17. Ramírez-Gallego, S.; Krawczyk, B.; García, S.; Woźniak, M.; Herrera, F. A survey on data preprocessing for data stream mining: Current status and future directions. *Neurocomputing* **2017**, *239*, 39–57. [\[CrossRef\]](#)
18. Jankov, D.; Sikdar, S.; Mukherjee, R.; Teymourian, K.; Jermaine, C. Real-time High Performance Anomaly Detection over Data Streams: Grand Challenge. In Proceedings of the 11th ACM International Conference on Distributed and Event-Based Systems, Barcelona, Spain, 19–23 June 2017; pp. 292–297.
19. Zhang, L.; Lin, J.; Karim, R. Sliding window-based fault detection from high-dimensional data streams. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *47*, 289–303. [\[CrossRef\]](#)
20. Forestiero, A. Self-organizing anomaly detection in data streams. *Inf. Sci.* **2016**, *373*, 321–336. [\[CrossRef\]](#)
21. Xie, G.; Sun, Y.; Lin, M.; Tang, K. A Selective Transfer Learning Method for Concept Drift Adaptation. In Proceedings of the International Symposium on Neural Networks, Sapporo, Japan, 21–23 June 2017; Springer: Cham, Switzerland, 2017; pp. 353–361.
22. Hahsler, M.; Bolanos, M.; Forrest, J. Introduction to stream: An Extensible Framework for Data Stream Clustering Research with R. *J. Stat. Softw.* **2017**, *76*, 1–50. [\[CrossRef\]](#)
23. Guo, Y.; Xu, Q.; Li, P.; Sbert, M.; Yang, Y. Trajectory Shape Analysis and Anomaly Detection Utilizing Information Theory Tools. *Entropy* **2017**, *19*, 323. [\[CrossRef\]](#)
24. Gomes, H.M.; Barddal, J.P.; Enembreck, F.; Bifet, A. A Survey on Ensemble Learning for Data Stream Classification. *ACM Comput. Surv.* **2017**, *50*, 1–36. [\[CrossRef\]](#)
25. Tu, E.; Kasabov, N.; Yang, J. Mapping temporal variables into the neucube for improved pattern recognition, predictive modeling, and understanding of stream data. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 1305–1317. [\[CrossRef\]](#) [\[PubMed\]](#)

26. Ibidunmoye, O.; Rezaie, A.R.; Elmroth, E. Adaptive Anomaly Detection in Performance Metric Streams. *IEEE Trans. Netw. Serv. Manag.* **2017**, *15*, 217–231. [[CrossRef](#)]
27. Abbasi, B.; Calder, J.; Oberman, A.M. Anomaly detection and classification for streaming data using partial differential equations. *arXiv* **2016**, 1–23.
28. Roy, G.; Roy, G.; Roy, G.; Schrijvers, O. Robust random cut forest based anomaly detection on streams. In Proceedings of the International Conference on International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; pp. 2712–2721.
29. Lang, M. A Low-Complexity Model-Free Approach for Real-Time Cardiac Anomaly Detection Based on Singular Spectrum Analysis and Nonparametric Control Charts. *Technologies* **2018**, *6*, 26. [[CrossRef](#)]
30. Riaz, M.; Abbas, N.; Does Ronald, J.M.M. Improving the performance of CUSUM charts. *Qual. Reliab. Eng. Int.* **2011**, *27*, 415–424. [[CrossRef](#)]
31. Li, Y.; Yang, T.; Cheng, X.; Yang, R.; Xu, M. An Anomaly Detection Algorithm of Satellite Power System Based on CUSUM Control Chart. In Proceedings of the International Conference on Information Science and Control Engineering, Beijing, China, 8–10 July 2016; pp. 829–833.
32. Chen, S.H. The gamma CUSUM chart method for online customer churn prediction. *Electr. Commer. Res. Appl.* **2016**, *17*, 99–111. [[CrossRef](#)]
33. Chen, Q.; Chen, L.; Lian, X.; Liu, Y.; Yu, J.X. Indexable PLA for Efficient Similarity Search. In Proceedings of the 33rd International Conference on Very Large Data Bases, Vienna, Austria, 23–27 September 2007; pp. 435–446.
34. Gilbert, A.C.; Kotidis, Y.; Muthukrishnan, S.; Strauss, M.J. One-Pass Wavelet Decompositions of Data Streams. *IEEE Trans. Knowl. Data Eng.* **2003**, *15*, 541–554. [[CrossRef](#)]
35. Sadik, M.S.; Gruenwald, L. DBOD-DS: Distance based outlier detection for data streams. In Proceedings of the International Conference on Database and Expert Systems Applications, Bilbao, Spain, 30 August–3 September 2010; Springer: Berlin/Heidelberg, Germany, 2010; pp. 122–136.
36. Sadik, S.; Gruenwald, L. An adaptive outlier detection technique for data streams. In Proceedings of the Scientific and Statistical Database Management, Heidelberg, Germany, 30 June–2 July 2010; Springer: Berlin/Heidelberg, Germany, 2011; pp. 596–597.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).