International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

# COPY RIGHT

ELSEVIER SSRN



IJIEMR Transactions, online available on 12th Dec 2018. Link
:http://www.ijiemr.org/downloads.php?vol=Volume-07&issue=ISSUE-13

Title: DETECTING OUTLIER FROM SENSORS PRODUCING TIME-SERIES SIGNALS

Volume 07, Issue 13, Pages: 299–301.

Paper Authors

**RAJEESAB NADAF, M. VAZRALU**

Malla Reddy College of Engineering and Technology, Hyderabad, T.S, India


USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

# DETECTING OUTLIER FROM SENSORS PRODUCING TIME-SERIES SIGNALS

[1]RAJEESAB NADAF, [2]M. VAZRALU
[1]M.Tech Student, Dept of CSE, Malla Reddy College of Engineering and Technology, Hyderabad, T.S, India
[2]Associate Professor, Dept of CSE, Malla Reddy College of Engineering and Technology, Hyderabad, T.S, India

**Abstract:** In this paper, a novel method has been proposed to detect an outlier by analyzing the signal in the frequency domain. Before reaching to this stage, data preprocessing is very important. While sensing the signals, many times the device is unable to capture the signal. These data points appear as null or invalid. This creates worthless problems in the output while processing such signals. Hence, data preprocessing is important in these situations. In this paper, these null data points are removed by using the concept of moving window average. If any null data point is detected, then the value is replaced by the average of previous and the next value of the null data point. Then the signal is ready for further processing. Sensors detect the signal which is in time domain. FFT is performed on such signals to get the frequency domain of the signal. It is expected that the knowledge lies in the lower frequency. If the signal contains any outlier, then the knowledge moves in the higher frequency. Hence, the effort is to detect this knowledge in the higher frequency if any, to detect the outlier in the given signal.

**Keywords:** Outlier Detection; detecting outlier from sensors; Outlier Detection in Time-Series Data;

## 1. INTRODUCTION:

Outlier is unusual disturbance in the signal which is different from the usual conduct of the signal. Detecting an outlier is very important as this decides the performance of any industry. The quality of any product can be determined by detecting the outlier if any, in the performance of the machines that lead to the production of the materials. Blowing of the fuse is also an example of an outlier as the current level suddenly rises up. Short circuit happens because of sudden rise in the current level due to direct contact of positive and negative terminal. This again shows an outlier in the electric current flow. Sounding of security alarm is also an example of outlier detection as it detects some unusual behavior in the environment where the sensor has been installed. Whenever the threshold limit is crossed, this is a clear indication of a breach in the security. The popularity of any program in television is decided by an outlier of the strength of the frequency related to that program. Detecting an outlier becomes very sensitive in situations which can lead to heavy destructions if proper measures are not taken. One such example is nuclear reactors. Nuclear fission is kept under controlled limit. If any mistake happens here, then Chernobyl like incident can happen. MRI and CT scanners are helpful to detect the health of a patient. Any outlier can lead to death of the person.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

## 2. PREVIOUS APPROACH:

A lot of methods have been introduced to detect an outlier. Sabyasachi Basu et al. [1] proposed the method of one-sided and two-sided median methods. But, the author proposes that this method is inaccurate when there are consecutive outliers spanning longer than window width as it is difficult to differentiate between same value and actual signal. Kaouther Nouira et al. [2] gave the method of graphical approaches and Gibb's sampling approach. There is no experimentation performed using these two methods but, it has been promised that this information may be used to improve outlier detection. Chris E. Zwilling et al. [3] used the information from covariance of time series data to detect an outlier. This method becomes efficient because it provides the user to choose any number and type of features and the algorithm will correctly identify the outlier. The author also proposed Multivariate Voronoi Outlier Detection (MVOD) [4] method. It is also proposed that this method is accurate, sensitive and robust in multivariate time series data. Hermine N. Akouemo et al. [5] adduced an autoregressive integrated moving average with exogenous inputs (ARIMAX) model. But, the ARIMAX model needs to be trained on cleaner data at each step. Hui-xin Tian et al. [6] came up with a new method that combines density-based clustering algorithm with soft sensor modeling process. But, this method directs to detect outlier of soft sensor modeling in complex industrial processe

## 3. FINE-GRAINED METHOD:

A time series data is being analyzed in the frequency domain to detect an outlier. Step by step procedure is as follows:

1. A time series data can have null points at some time stamps. So, these null values are removed by replacing null values with the mean of adjacent points (mean of one value previous to the null value and other value next to the null value).
2. This data is divided into equal Window values.
3. For each Window, the magnitude of FFT and Window mean is calculated.
4. The threshold is set for both, the magnitude of FFT of Window and the mean of the Window.
5. For the magnitude of FFT, value at first index is not taken into consideration. This is done to neglect any information in the lower frequency range.
6. If any of the thresholds is violated, then the Window is considered to have an outlier.

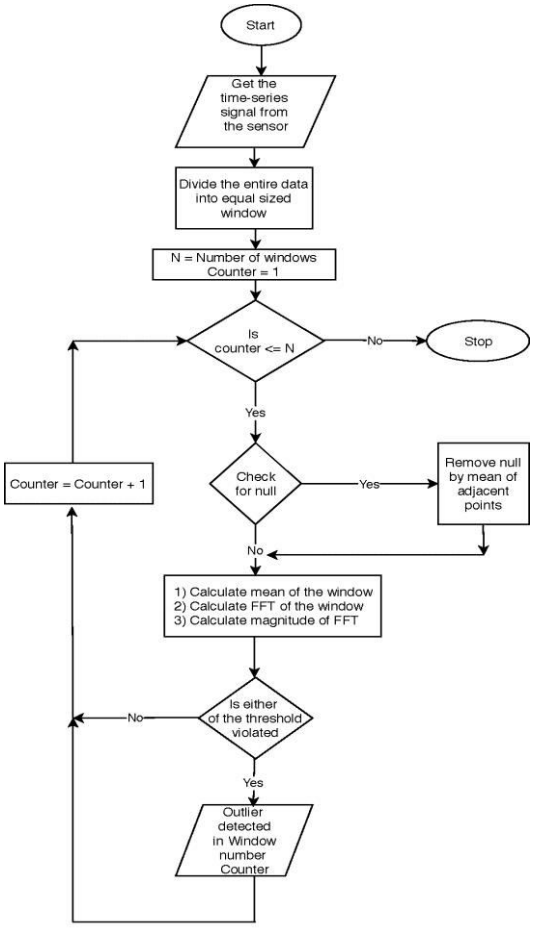


Fig. 1: Algorithm of detecting outlier from sensors producing time-Series Signals

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Fig.1: describes the whole procedure algorithmically.

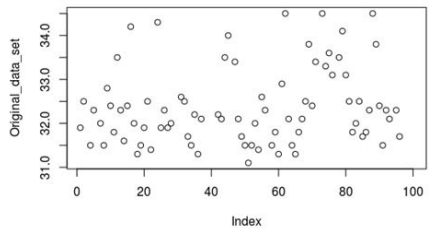


Fig. 2: Original Data values

In Fig. 2, NA data point is available at 3, 6, 23, 29, 30, 38, 39, 40, 41, 46, 57, 72, 77, 94, 97, 98, 99, 100. After data processing, the output is shown in Fig. 3.

This series of data points are divided into 10 equal windows with 10 data points each.

Each window is processed in the following ways:

1) Average mean is calculated for each window.

2) FFT is performed on each window.

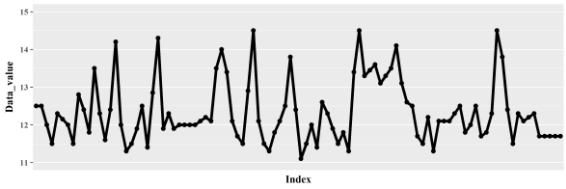


Fig. 3: Data set after removing NA

Now, the output is compared against the threshold in the time as well as frequency domain. If the threshold is violated in any of the domain, then the outlier is detected in that window. In frequency domain, first point is not considered for comparing against the threshold as it is assumed that the information will lie in the lower frequency domain. In this experiment, the threshold is 33 in the time domain and 2.5 in the frequency domain. In windows 2, 3, 4, 5 and 9, it can be observed that the threshold of 2.5 is being violated and hence, these windows have outlier.

## CONCLUSION:

There are a lot of papers describing about the outlier detection with different methods. But, generally these methods are specific to a particular sensor generating a particular form of signal. Hence, same method is generally not applicable to detect an outlier for the signals generated by other sensor devices. Analyzing the signal in frequency domain makes an impact to detect an outlier as all the signals are first in the time domain which is then converted to frequency domain. Hence, this detection method can be generalized to all the types of signals sensed by the detector.

## REFERENCES:

[Basu, Sabyasachi, and Martin Meckesheimer. "Automatic outlier detection for time series: an application to sensor data." Knowledge and Information Systems 11.2 (2007): 137-154.

Nouira, Kaouther, and Abdelwahed Trabelsi. "Time Series Analysis and Outlier Detection in Intensive Care Data." Signal Processing, 2006 8th International Conference on. Vol. 4. IEEE, 2006.

Zwilling, Chris E., and Michelle Y. Wang. "Covariance based outlier detection with feature selection." Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the. IEEE, 2016.

C.M. Bishop Pattern Recognition and Machine Learning (Corr.printing ed.)Springer, New York (US) (2007).

R. Bertani Installed capacity in 2015 worldwide Credit: Geothermal Power Generation in the World 2010–2014 Update Report (2015)

Tian, Hui-xin, Xing-jun Liu, and Mei Han. "An outliers detection method of time series data for soft sensor modeling." Control and Decision Conference (CCDC), 2016 Chinese. IEEE, 2016.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org