Explainable Anomaly Detection System for Categorical Sensor Data in Internet of Things

Peng Yuan[1], Lu-An Tang[1], Haifeng Chen[1, *], Moto Sato[1], and Kevin Woodward[2]

[1] NEC Labs America, Princeton, NJ, USA
[2] Lockheed Martin Space, Denver, CO, USA
{pyuan, ltang, Haifeng, moto}@nec-labs.com
kevin.woodward@lmco.com

Abstract. Internet of things (IoT) applications deploy massive number of sensors to monitor the system and environment. Anomaly detection on streaming sensor data is an important task for IoT maintenance and operation. However, there are two major challenges for anomaly detection in real IoT applications: (1) many sensors report categorical values rather than numerical readings; (2) the end users may not understand the detection results, they require additional knowledge and explanations to make decision and take action. Unfortunately, most existing solutions cannot satisfy such requirements. To bridge the gap, we design and develop an eXplainable Anomaly Detection System (XADS) for categorical sensor data. XADS trains models from historical normal data and conducts online monitoring. XADS detects the anomalies in an explainable way: the system not only reports anomalies' time periods, types, and detailed information, but also provides explanations on why they are abnormal, and what the normal data look like. Such information significantly helps the decision making for users. Moreover, XADS requires limited parameter setting in advance, yields high accuracy on detection results and comes with a user-friendly interface, making it an efficient and effective tool to monitor a wide variety of IoT applications.

Keywords: Explainable AI, Internet of things, Sensor data, Anomaly detection.

1 Introduction

Internet of things (IoT) integrates sensor devices with informational components to form a context sensitive system that responds intelligently to dynamic changes in real-world environments [9]. With rapid developments in recent years, IoT devices are widely used in different fields such as satellite, healthcare, transportation, and environment monitoring. A typical IoT application usually contains thousands of sensors to monitor its components and surrounding environment. Evaluating the streaming sensor data in real-time and detecting abnormal symptoms are critical for IoT maintenance and operation tasks.

In real applications, IoT sensors contain not only numerical readings but also categorical data representing the working status or operational mode. Unfortunately, most existing methods on anomaly detection are proposed to detect outliers and anomalies

* Corresponding author.

of numerical data [1-8]. They cannot be used on categorical data. In addition, many methods only provide a timestamp of the detected anomaly. Without enough context information, the users cannot understand such detection results. To bridge the gap, we design and develop an eXplainable Anomaly Detection System (XADS) to monitor the sensor data of IoT devices. The solution constructs a three dimensional histogram model of category, event duration, and frequency. XADS profiles the normal states by learning from historical data, and automatically determines the anomaly thresholds despite of noisy data. After training models from historical data, XADS monitors newly arrived data and detects the anomalies in real time. Once an anomaly is detected, XADS searches in the normal profiles and generates detailed reasons to explain the result. It also provides expected values as a normal baseline for comparison. With such detailed explanations, the users can understand the detected anomalies and take out correct actions.

Another major advantage of XADS is on the applicability and feasibility. The solution only requires limited parameter setting in advance and can be applied to a wide variety of IoT devices. In many real applications, it is difficult to obtain the abnormal or fault events as training data. XADS does not need such abnormal data for training. It trains the model only with normal data, which are much easier to collect. XADS can detect both seen and unseen anomalies (i.e., the types of anomalies that has not appear before) with high accuracy. The solution has been tested and applied in multiple real IoT applications including satellite and spacecraft [10][11]. A demo of XADS can be accessed from the project page at: https://github.com/pengyuan0106/eXplainable-Anomaly-Detection-System.

2 System Description

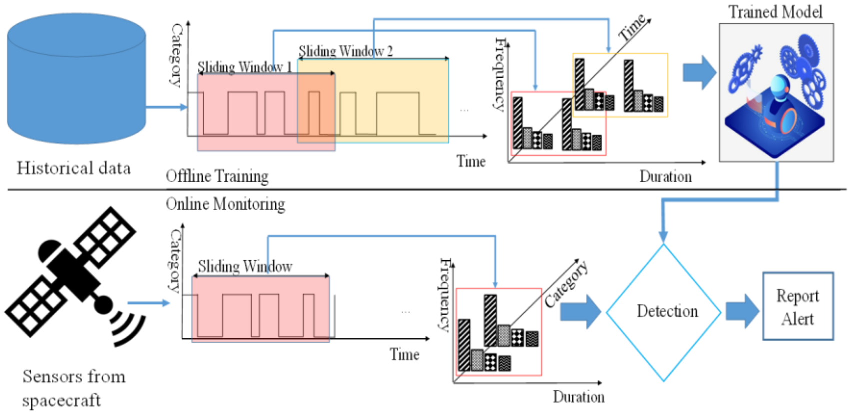

Fig. 1. System framework of XADS

As shown in Figure 1, the overall structure of XADS is consisted of two modules: (1) offline training from historical data and (2) online monitoring for streaming data.

In offline training module, XADS segments categorical sensor data into event sequences through sliding windows of adaptive length. The window length is learned from the distribution of events. After window separation, XADS learns the features of all historical segments and generates a 3D histogram model to profile the normal data from the dimensions of category, event duration, and frequency. At last, XADS computes the anomaly threshold by running trained model on historical data.

In online monitoring module, XADS first converts streaming data into a new sliding window and transforms the window into a histogram. The new histogram is then matched with trained model to calculate an anomaly score. The system raises an alert if the score is higher than anomaly threshold.

Figure 2 shows a snapshot of using XADS to monitor the telemetry sensor data from a soil moisture active passive satellite [3]. The dashboard of XADS includes two parts: a tool panel (left) and a set of view panels (right).

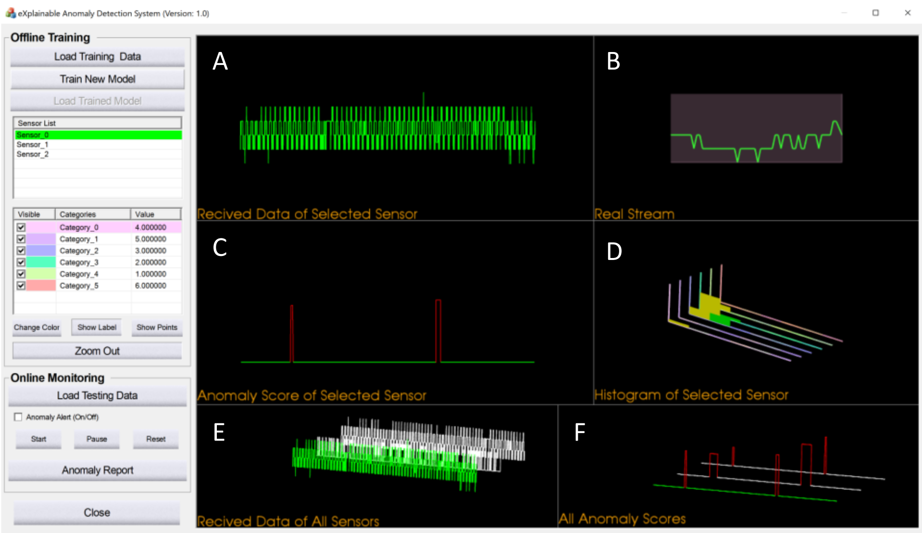

Fig. 2. Main interface of XADS

Tool Panel: It allows the user to upload historical data and train the model. Once the 3D histogram models are trained, XADS can either load testing data or receive streaming data from network and conduct online monitoring.

View Panels: XADS provides six different view panels for streaming data monitoring and anomaly detection. (1) As shown in Figure 2 (A), the streaming data panel plots the so-far arrived sensor data by time; (2) The sliding window panel (Figure 2 (B)) provides a zoom-in view of the current data; (3) The anomaly score panel (Figure 2 (C)) plots the computed anomaly scores in real time. It is aligned with the streaming data panel. The period labeled by orange color are with abnormal events. (4) The 3D histogram panel (Figure 2 (D)) shows the constructed histogram from streaming data in the new window; (5&6) To provide a global view to the users, XADS shows all the received data in Figure 2 (E) and anomaly scores of multiple sensors in Figure 2 (F).

Once an anomaly is detected, the users can check more details in anomaly report panel. XADS lists out the abnormal values and context information in an anomaly explanation panel, as shown in Figure 3 (B). Figure 3 (C) shows the histogram of abnormal data. The blue lines indicate the normal range of frequency in trained model. The orange rectangle represents the abnormal frequency of current window. Figure 3 (D) has two plots: the left one is a zoom in view of the observed anomaly, the orange color denotes the abnormal event. The right plot is a normal baseline, where the blue color denotes expected normal values during the abnormal period. In this way, XADS provides an explicit comparison to illustrate detected anomalies.

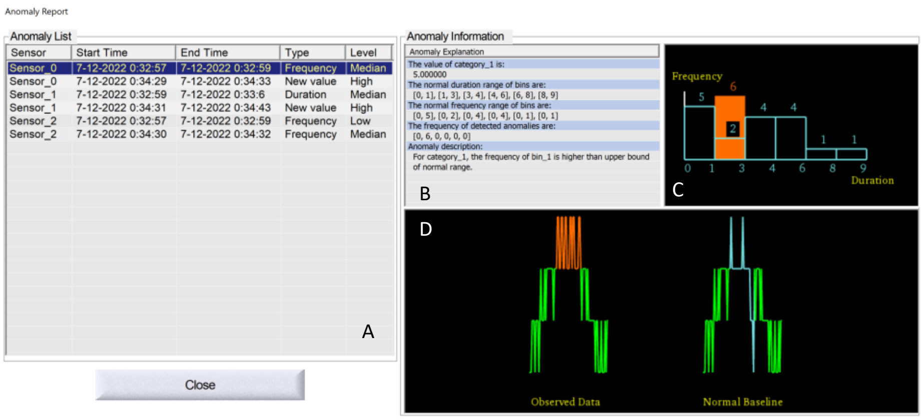

Fig. 3. Anomaly Report Panel of XADS

3 Conclusion and Future Work

In this paper, we present a novel eXplainable Anomaly Detection System (XADS) to monitor categorical sensors in IoT applications. XADS generates a histogram model on the dimensions of category, event duration, and frequency. It automatically determines the value's normal ranges and anomaly thresholds. The detected anomalies are reported in GUI interfaces with detailed explanations, as well as a normal baseline to help the user's understanding and decision making.

In the near future, we plan to extend XADS to complex IoT monitoring with both categorical and numerical sensors, and to test XADS on more applications such as weather forecasting and financial analysis.

Acknowledgements. The authors would like to thank Matthew Horak, Giovanni A Tobar, Sowmya S Chandrasekaran, and Sadananda Narayanappa from Lockheed Martin Space for the constructive comments and suggestions.

References

1. Li, X., Han, J.: Mining approximate top-k subspace anomalies in multi-dimensional time-series data. In: 33rd International Conference on Very Large Data Bases, pp. 447–458 (2007).
2. Gao, Y., Yang, T., Xu, M., & Xing, N.: An unsupervised anomaly detection approach for spacecraft based on normal behavior clustering. In: 2012 Fifth International Conference on Intelligent Computation Technology and Automation, pp. 478–481 (2012).
3. Hundman, K., Constantinou, V., Laporte, C., Colwell, I., Soderstrom, T.: Detecting Spacecraft Anomalies Using LSTMs and Nonparametric Dynamic Thresholding. In: Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining, pp. 387–395 (2018).
4. Tang, L. A., Cui, B., Li, H., Miao, G., Yang, D., Zhou, X.: Effective variation management for pseudo periodical streams. In: Proceedings of the 2007 ACM SIGMOD international conference on Management of data, pp. 257–268 (2007).
5. Wu, H., Salzberg, B., Zhang, D.: Online event-driven subsequence matching over financial data streams. In: Proceedings of the 2004 ACM SIGMOD international conference on Management of data, pp. 23–34 (2004).
6. Wu, H., Sharp, G. C., Salzberg, B., Kaeli, D., Shirato, H., Jiang, S. B.: A Finite State Model for Respiratory Motion Analysis in Image Guided Radiation Therapy. Physics in Medicine & Biology 49(23), 5357 (2004).
7. Wu, H., Salzberg, B., Sharp, G. C., Jiang, S. B., Shirato, H., Kaeli, D.: Subsequence matching on structured time series data. In: Proceedings of the 2005 ACM SIGMOD international Conference on Management of Data, pp. 682–693 (2005).
8. Schwabacher, M., Oza, N., Matthews, B.: Unsupervised anomaly detection for liquid-fueled rocket propulsion health monitoring. Journal of aerospace computing, information, and communication 6(7), 464–482 (2009).
9. Tang, S., Shelden, D. R., Eastman, C. M., Pishdad-Bozorgi, P., Gao, X.: A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends. Automation in Construction 101, 127–139 (2019).
10. Chen, H., Horak, M., Narayanappa, S., Woodward, K.: Integrating AI into Planning, Diagnostic, and Prescription Systems for Human & Robotic Deep Space Exploration Missions. In: AAAI 2018 Fall Symposium Series, (2018).
11. Yuan, P., Tang, L. A., Chen, H., Sato, M., Woodward, K.: 3D Histogram based Anomaly Detection for Categorical Sensor Data in Internet of Things. Submitted to VLDB Workshop on Very Large Internet of Things (VLIoT 2022).
12. Hodge, V., Austin, J.: A survey of outlier detection methodologies. Artificial intelligence review 22(2), 85–126 (2004).
13. Tamboli, J., Shukla, M.: A survey of outlier detection algorithms for data streams. In: 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 3535-3540 (2006).
14. Panjei, E., Gruenwald, L., Leal, E., Nguyen, C., Silvia, S.: A survey on outlier explanations. The VLDB Journal, 1–32 (2022).