The Thirty-Second AAAI Conference
on Artificial Intelligence (AAAI-18)

Towards Experienced Anomaly Detector through Reinforcement Learning

Chengqiang Huang,[1] Yulei Wu,[1] Yuan Zuo,[1] Ke Pei,[2] Geyong Min[1]
[1]University of Exeter, North Park Road, Exeter, Devon, UK, EX4 4QF
[2]Huawei Technologies Co. Ltd. Shenzhen, Guangdong, China, 518129
{ch544,Y.L.Wu,yz506,G.Min}@exeter.ac.uk, peike@huawei.com

Abstract

This abstract proposes a time series anomaly detector which 1) makes no assumption about the underlying mechanism of anomaly patterns, 2) refrains from the cumbersome work of threshold setting for good anomaly detection performance under specific scenarios, and 3) keeps evolving with the growth of anomaly detection experience. Essentially, the anomaly detector is powered by the Recurrent Neural Network (RNN) and adopts the Reinforcement Learning (RL) method to achieve the self-learning process. Our initial experiments demonstrate promising results of using the detector in network time series anomaly detection problems.

Introduction

Anomaly detection is a pervasive topic in various fields. In industry, it always serves as the first messenger to trigger more complicated procedures such as anomaly localization. As a result, anomaly detection is very significant and, ideally, it should be highly applicable to different scenarios and easily accessible by engineers. However, existing anomaly detection methods do not necessarily satisfy the requirements. **1)** A thorough survey of anomaly detection methods is nicely presented in (Chandola, Banerjee, and Kumar 2009). It clarifies the assumptions made by different types of anomaly detection methods, which reveals that methods with strong assumptions of the anomaly patterns, e.g., distribution-based methods, may not produce satisfactory results under scenarios where the assumptions do not hold. **2)** On the other hand, the anomaly detection methods are not always easily accessible. In 2015, Yahoo published their time series anomaly detection system EGADS (Laptev, Amizadeh, and Flint 2015). Within the system, a set of methods are implemented and integrated to generate anomaly detection results. Such a complex system requires the engineers to not only understand the components but also comprehend the set of methods so that being able to tune the parameters for each of them. **3)** Additionally, few methods used in the industry consider the evolvement of the anomaly patterns, which leads to static anomaly detection parameters that perform poorly under dynamic scenarios.

Objectives

In this paper, we consider the specific problem of time series anomaly detection and emphasize that an anomaly detector should have the following features:

- The anomaly detector makes **no assumption about the concept of the anomaly**, i.e., definition of the anomaly, but it learns the concept solely from the training datasets;
- The anomaly detector is **threshold-free**, which means the anomaly detector is a logical classifier with no tuning threshold. Preferably, except hyperparameters, e.g., the number of layers in the neural network, the detector does not have other tunable parameters.
- The anomaly detector is **dynamically improving** with the accumulation of the anomaly detection experience. In other words, the detector learns new anomalies and consistently enhances its knowledge for anomaly detection.

Problem Formulation and Solution

In this work, we propose a Recurrent Neural Network (RNN) based anomaly detector that is trained consistently through Reinforcement Learning (RL) to meet the objectives. Following the framework of RL, we cast the problem of time series anomaly detection as a Markov Decision Process (MDP) and define corresponding concepts.

Definition 1: Anomaly Detector π

An anomaly detector is defined as a conditional probability distribution $\pi := p(A|S)$, where S and A denote the sets of states in the target system and the set of actions respectively. Typically, $A = \{0, 1\}$ in which 1 means the given state is anomalous and 0 otherwise. And note that $\pi(s, a) = p(A = a|S = s)$ is the probability of action a given state s.

Definition 2: Anomaly Detector Performance V_π

The performance of an anomaly detector is measured through its capability of time series anomaly detection, which is formalized as:

$$V_\pi = \sum_{s \in S} d^\pi(s) \sum_{a \in A} Q(s, a) \cdot \pi(s, a),$$

where $d^\pi(s)$ is the probability of the target system being in the state s under the utilization of the anomaly detector π,

and $Q(s, a)$ represents the accumulated reward started from state s with action a. In other words, the performance is the average accumulated reward in anomaly detection following the anomaly detector π.

Definition 3: Optimal Anomaly Detector π^*

The optimal anomaly detector is the detector that satisfies:

$$\pi^* = \arg\max_{\pi} V_{\pi}.$$

Considering a deterministic optimal anomaly detector, it should maximize the performance, and, under cases where $d^{\pi}(s)$ is roughly the same for all $s \in S$, it has $\pi(s, a) = 1$ if $a = \arg\max_a Q(s, a)$. In other words, the optimal anomaly detector π^* is fully determined by the accumulated reward function $Q(s, a)$.

Definition 4: Experience E

The experience E is a set of tuples, each of which is defined as $\langle s, a, r, s' \rangle$. $s, s' \in S$ indicate the states of the target system before and after the action a, respectively. r is the instant reward obtained under the state s with the action a. In an anomaly detection system, the actions are picked by the anomaly detector π. Therefore, the experience records all the behaviors of the anomaly detector.

According to the definitions, an anomaly detector is to be improved consistently by learning from the experience, which in principle is to gain a better estimation of $Q(s, a)$. This process is actually a key target of RL systems and can be achieved by existing RL solutions. Specifically, we adopt Q-learning method to train an RNN for estimating $Q(s, a)$. It is worth noting that this formulation of the anomaly detector π makes no assumption for anomaly detection, refrains from the cumbersome work of threshold selection and is capable of consistently improving its capability.

Empirical Experiments and Results

Specifically, our initial experiments[1] adopt long short-term memory (LSTM) for RNN and standard Q-learning method with memory replay for RL.

The datasets used for training are Yahoo benchmark datasets (Laptev, Amizadeh, and Flint 2015) which includes 367 labeled time series. Each time series is transformed into a set of multi-dimensional data instances using the sliding window method. And the actions in Q-learning is $a \in \{0, 1\}$ where 0 represents no anomaly and 1 otherwise. On the other hand, a state in Q-learning is designed as the concatenation of the data and a fixed size record of actions performed earlier. To boost the process of model training, a binary tree strategy is used that the two states s'_0 and s'_1 generated by performing different actions 0 and 1 over the previous state s are both added to the set of experience for training. In other words, during our training, a state s will added two records, i.e., $\langle s, a = 0, r_0, s'_0 \rangle$ and $\langle s, a = 1, r_1, s'_1 \rangle$, to the experience. r_0 and r_1 are the rewards obtained by performing different actions under the state s. The reward function is designed according to the labels of the dataset.

[1] https://github.com/chengqianghuang/exp-anomaly-detector

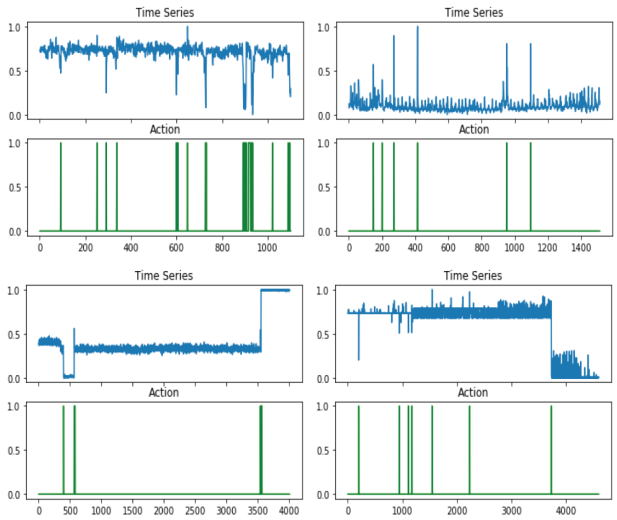

Figure 1: The sample performance of proposed method

Fig. 1 shows some sample performance of the anomaly detector in Numenta datasets (Ahmad et al. 2017) after training through memory replay of the anomaly detection experience of Yahoo datasets. The original time series data are marked as blue lines and the green lines indicate the actions performed by the anomaly detector. It is worth noticing that the anomaly detector is capable of identifying shift of means, point anomalies and anomalous patterns of the target time series, and achieves high-quality results in the testing datasets, e.g., the approximate 100% accuracy and 100% recall in the given samples.

Conclusion and Future Work

This abstract proposes a design of time series anomaly detector which is fully determined by the experience of anomaly detection without explicit definitions or assumptions of anomalies. No threshold is required for the detector. With growing experience, it is expected that the anomaly detector keeps evolving and is able to perform nicely in general and unseen anomaly detection problems. To extend the applicability of the method, the problem of generating accurately labeled time-series datasets of various types for anomaly detection training is considered as our next step.

References

Ahmad, S.; Lavin, A.; Purdy, S.; and Agha, Z. 2017. Unsupervised real-time anomaly detection for streaming data. *Neurocomputing* 262(Supplement C):134–147.

Chandola, V.; Banerjee, A.; and Kumar, V. 2009. Anomaly detection: A survey. *ACM Computing Survey* 41(3):15:1–15:58.

Laptev, N.; Amizadeh, S.; and Flint, I. 2015. Generic and scalable framework for automated time-series anomaly detection. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, 1939–1947. New York, NY, USA: ACM.