Applied Soft Computing 10 (2010) 859–867

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Sequential anomaly detection based on temporal-difference learning: Principles, models and case studies☆

Xin Xu

Institute of Automation, College of Mechatronics and Automation, National University of Defense Technology, Deya Road, Changsha 410073, PR China

ARTICLE INFO

Article history:
Received 3 January 2008
Received in revised form 11 May 2009
Accepted 3 October 2009
Available online 27 November 2009

Keywords:
Anomaly detection
Temporal-difference
Markov reward processes
Learning prediction
Computer security
Reinforcement learning

ABSTRACT

Anomaly detection is an important problem that has been popularly researched within diverse research areas and application domains. One of the open problems in anomaly detection is the modeling and prediction of complex sequential data, which consist of a series of temporally related behavior patterns. In this paper, a novel sequential anomaly detection method based on temporal-difference (TD) learning is proposed, where the anomaly detection problem of multi-stage cyber attacks is considered as an application case. A Markov reward process model is presented for the anomaly detection and alarming process of sequential data and it is verified that when the reward function is properly defined, the anomaly probabilities of sequential behaviors are equivalent to the value functions of the Markov reward process. Therefore, TD learning algorithms in the reinforcement learning literature can be used to efficiently construct anomaly detection models of complex sequential behaviors by estimating the value functions of the Markov reward process. Compared with other machine learning methods for anomaly detection, the proposed approach has the advantage of simplified labeling process using delayed evaluative signals and the prediction accuracy can be improved even if labeled training data are limited. Based on the experimental results on intrusion detection of host computers using system call data, it was shown that the proposed anomaly detection method can achieve higher or at least comparable detection accuracies than other approaches including SVMs, and HMMs.

1. Introduction

As a typical pattern recognition task, anomaly detection is to detect non-conforming or abnormal patterns from a given class of normal behaviors. These non-conforming patterns are often referred to as anomalies, outliers, exceptions, or surprises in different applications. Anomaly detection is widely used in a variety of domains, such as intrusion detection, fraud detection, fault detection, system health monitoring, and event detection in sensor networks. Although anomaly detection in data has been studied in the statistics community as early as the 19th century, there are still several open problems to be solved. As discussed in [1], one of the main challenges for anomaly detection techniques is that defining a normal region which encompasses every possible normal behavior is very difficult. The other challenge is that availability of labeled data for training/validation of models used by anomaly detection techniques is usually a major issue. In addition, the data usually contains noise which tends to be similar to the actual anomalies and hence is difficult to distinguish and remove. In recent years, aiming at the above challenges, a variety of anomaly detection techniques have been developed in the soft computing and machine learning communities. For a comprehensive survey on anomaly detection techniques based on machine learning, readers may refer to [1].

In general, existing soft computing approaches to anomaly detection can be grouped into three categories, i.e., supervised or classification-based, semi-supervised, and unsupervised anomaly detection methods. Supervised anomaly detection techniques learn a classifier using labeled instances belonging to normal or anomaly class, and then assign a normal or anomalous label to a test instance. Typical approach in such cases is to build a predictive model for normal vs. anomaly classes [2–4,28]. Semi-supervised anomaly detection techniques construct a model representing normal behavior from a given normal training data set, and then test the likelihood of a test instance to be generated by the learnt model. It is assumed that the training data have labeled instances for only the normal class. Since they do not require labels for the anomaly class, they are more widely applicable than supervised techniques [5,16]. Unsupervised anomaly detection techniques detect anomalies in an unlabeled test data set under the assumption that majority of the instances in the data set are

☆ This work is supported by the National Natural Science Foundation of China under Grant 60303012, 60774076, 90820302 the Fork Ying Tung Youth Teacher Foundation of China under Grant 114005, and the Natural Science Foundation of Hunan Province under Grant 07JJ3122.

E-mail addresses: xuxin_mail@263.net, xinxu@nudt.edu.cn.

1568-4946/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2009.10.003

normal [10,11]. The techniques in this category make the implicit assumption that normal instances are far more frequent than anomalies in the test data. If this assumption is not true then such techniques suffer from high false alarm rate.

Although anomaly detection techniques have been widely studied and applied in a variety of areas, there are still many challenges for detecting anomalies in sequential data which is very common in a wide range of domains where a natural ordering is imposed on data instances by either time or position. In anomaly detection literature, two types of data sequences have been popularly studied, i.e., symbolic and continuous sequences. In this paper, we will mainly focus on symbolic data sequences but the methodology developed may also be extended to continuous data sequences or time series. Due to the temporally related nature of sequential data, detecting anomalous subsequences is more challenging than anomaly detection in static patterns. In this paper, a novel sequential anomaly detection method based on temporal-difference (TD) learning [19], which can be called TD_SAD, is presented, where intrusion detection of multi-stage computer attacks is considered as a special application case. For anomaly detection of multi-stage cyber attacks, Markovian modeling of sequences has been a popular approach in this category. However, in our approach, a new Markov reward model is established for sequential data, which is different from previous works in that reward functions are defined as a feedback signal to indicate whether a long sequence of observation patterns is normal or abnormal. Furthermore, it is analyzed in theory that the sequential anomaly detection task can be implemented by predicting the value functions of the Markov reward process. In the proposed TD_SAD approach, by incorporating reward signals for every observation pattern in data sequences, the anomaly probabilities of sequential behaviors are equivalent to the value functions of the Markov reward process. Therefore, the TD learning and prediction algorithms developed from the reinforcement learning [15] literature can be employed to detect multi-stage cyber attacks. According to the authors' knowledge, this is the first attempt to apply TD learning and prediction in sequential anomaly detection, which is different from previous supervised learning or statistical methods.

As we will analyzed in the following sections, the anomaly detection method proposed in this paper provides a new framework for detecting anomalies in multi-stage cyber attacks and can also be applied to other anomaly detection tasks in sequential data. The main contributions of this paper include the following two aspects. The first aspect of innovation is that reward functions are designed in Markovian modeling of sequential data so that the anomaly detection problem can be formulated as an equivalent value function prediction task. In previous works, the Markov models only focused on statistical models of state transitions and no reward functions were considered, which range from Finite State Automatons (FSAs) to Hidden Markov Models (HMMs) without reward functions. FSAs have been used to detect anomalies in network protocol data and operating system call intrusion detection [5], where anomalies are detected when a given sequence of events does not result in reaching one of the final states. In [12], HMM-based techniques were proposed to detect anomalous program traces in operating system call data. In our approach, the reward functions can be viewed as indicative signals for learning and the teacher signals in supervised learning are special cases of reward functions. So, the TD_SAD approach proposed in this paper provides a more flexible and efficient framework than FSA and HMMs and more prior information can be used to improve the performance of sequential anomaly detection. The second aspect of contributions is that temporal-difference learning was applied in anomaly detection of multi-stage cyber attacks and very promising results have been obtained. Until now, there have been few research works on applying TD learning in anomaly detection of complex sequential data. Thus, the proposed anomaly detection method based on Markov reward process and TD learning not only provides a new direction for research on intrusion detection using reinforcement learning but also has lots of potential extensions to anomaly detection tasks in other areas. The performance of the proposed anomaly detection method using TD learning was evaluated on system call data of host-based intrusion detection, from the MIT Lincoln Lab. and the University of New Mexico (UNM) [17]. The experimental results illustrate that the proposed method can achieve higher or at least comparable detection accuracies than previous approaches.

The paper is organized as follows. In Section 2, the research background of anomaly detection in computer security and related works are introduced. In Section 3, the anomaly detection problem in sequential data is formulated and analyzed by using intrusion detection of multi-stage cyber attacks as an application example. In Section 4, the anomaly detection method based on Markov reward models and TD learning is presented. It is proved that by appropriately selecting the reward functions of the Markov reward model, there is equivalence between the estimation of anomaly probabilities and the learning prediction of value functions. In Section 5, experimental results on the system call data from the MIT Lincoln Lab. and the UNM are described to illustrate the effectiveness of the proposed method. And in Section 6, conclusions and discussions are provided.

2. Background and related works

Since the detection problem of multi-stage cyber attacks will be used as an application case for the proposed anomaly detection method, some research background and related works will be introduced in the following. The purpose of intrusion detection [6] is to find cyber attacks or non-permitted deviations of the characteristic properties in a computer system or monitored networks. Earlier intrusion detection techniques commonly made use of extracted signatures of known attacks and made decisions by comparing observation data with the signatures. This kind of detection strategy is usually called misuse detection. Nevertheless, it is almost impossible for misuse detection systems to find new attacks with unknown or deformed signatures. To overcome the shortcomings of misuse detection, anomaly detection techniques in computer security have attracted lots of research interests in the literature [3,4]. Anomaly detection is different from misuse detection techniques in that little prior knowledge on precise signatures of computer attacks is needed. So, one advantage of anomaly detection is the ability to detect novel attacks. However, since conventional anomaly detection techniques have to deal with a complete set of normal behaviors, which usually have large uncertainties and observation noises, it is very difficult to for anomaly detection systems to have high detection rates and low false alarm rates simultaneously. In addition, in order to use training data from recorded attack behaviors to improve performance, it is also desirable to develop systematic methods to incorporate attack behaviors in the framework of anomaly detection, i.e., to construct hybrid anomaly detection models using both normal behavior data and attack data.

Aiming at the above problems, soft computing methods have been widely studied for anomaly detection in computer security applications in the past decade [7–11]. In [7,8], several efforts have been devoted to designing anomaly detection algorithms using supervised learning algorithms, such as neural networks, etc. Some recent works have been focused on using supervised learning methods to construct hybrid anomaly detection models, i.e., models that trained both on normal data and attack data, such as the multi-class classifier approach based on support vector machines (SVMs) [10]. Another approach to anomaly detection is to use unsupervised