Pattern Recognition Letters 34 (2013) 1916–1927

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

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

On-line anomaly detection and resilience in classifier ensembles

Hesam Sagha *, Hamidreza Bayati, José del R. Millán, Ricardo Chavarriaga

Defitech Chair in Non-Invasive Brain–Machine Interface, Center for Neuroprosthetics and Institute of Bioengineering, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Available online 7 March 2013

Keywords:
Anomaly detection
Classifier ensemble
Decision fusion
Human activity recognition

ABSTRACT

Detection of anomalies is a broad field of study, which is applied in different areas such as data monitoring, navigation, and pattern recognition. In this paper we propose two measures to detect anomalous behaviors in an ensemble of classifiers by monitoring their decisions; one based on Mahalanobis distance and another based on information theory. These approaches are useful when an ensemble of classifiers is used and a decision is made by ordinary classifier fusion methods, while each classifier is devoted to monitor part of the environment. Upon detection of anomalous classifiers we propose a strategy that attempts to minimize adverse effects of faulty classifiers by excluding them from the ensemble. We applied this method to an artificial dataset and sensor-based human activity datasets, with different sensor configurations and two types of noise (additive and rotational on inertial sensors). We compared our method with two other well-known approaches, generalized likelihood ratio (GLR) and One-Class Support Vector Machine (OCSVM), which detect anomalies at data/feature level.

We found that our method is comparable with GLR and OCSVM. The advantages of our method compared to them is that it avoids monitoring raw data or features and only takes into account the decisions that are made by their classifiers, therefore it is independent of sensor modality and nature of anomaly. On the other hand, we found that OCSVM is very sensitive to the chosen parameters and furthermore in different types of anomalies it may react differently. In this paper we discuss the application domains which benefit from our method.

1. Introduction

The field of activity recognition has gained an increasing level of interest driven by applications in health monitoring and assistance, manufacturing and entertainment. In addition, given the advances in portable sensing technologies and wireless communication many of these systems rely on the fusion of the information from (heterogeneous) sensor networks. In this framework, the detection of anomalous (e.g. faulty or misbehaving) nodes constitute an important aspect for the design of robust, resilient systems. Besides activity recognition, anomaly detection (AD) is an important issue in the fields of control systems, navigation, and time series analysis. Generally, an anomalous pattern is one that is not desired or not expected. Therefore it often decreases the system performance or generates abnormal behavior in the data stream. It can be due to sensor failure, signal degradation, environmental fluctuations, etc. In most fields, such as health monitoring, a desirable characteristic is to process the data stream online and get a real-time anomaly detection to take an appropriate counteraction.

As mentioned above, there is a tendency toward the use of large number of sensors and with different sensor modalities to have more information about the observed environment. In a pattern recognition system, there are different levels to fuse sensor information; data, feature, or classifier. Generally, the goal of data fusion is to achieve more reliable data. Feature fusion concatenates different features from all the sensors before classification (Fu et al., 2008). Classifier fusion is applicable when an ensemble of classifiers is used and each classifier is assigned to different subset of sensors, and finally, the decision is made by a combination of the classifier decisions (Ruta and Gabrys, 2000). This architecture allows for a decentralized classification system where each classifier decides about a separate data stream. Therefore, if a particular stream is faulty or misbehaving we can remove the corresponding channel easily from the fusion in order to continue the classification, keeping the system performance as high as possible. A good example of such systems is human activity recognition (HAR) with on-body sensors mounted on specific limbs. A good practice is to devote a distinct classifier to each of the sensors (Roggen et al., 2011). In these systems, one inevitable anomaly is the rotation or sliding of the sensors. If these anomalies can be detected, it is easier to take an appropriate counter-action for the misbehaving sensor, without the need of

* Corresponding author. Tel.: +41 788757348.
E-mail address: hesam.sagha@epfl.ch (H. Sagha).

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.02.014

reconfiguring the whole system (Sagha et al., 2011a; Chavarriaga et al., 2011). This desirable characteristic is the core feature of *Opportunistic activity recognition systems* where elements in the network may appear, disappear or change during life time (Roggen et al., 2009).

Given a classifier ensemble, the anomaly detection process can be applied at different levels; raw signal or feature level, or at classifier/fusion level. Hereafter, we call both raw and feature level as low level and the detection process as *Low Level Anomaly Detection* (LoLAD), while for the later case we name the process as *Fusion Level Anomaly Detection* (FuLAD). The former case is the most commonly applied (Chandola et al., 2009), however, it is often not applicable as different sensor modalities may be available – requiring to design specific AD for each modality – as well as the energy and computational cost in the case of wireless sensor networks. Therefore in this paper we introduce an anomaly detection mechanism at the level of classifier fusion, based on the consistency of the classifier decisions. Applying the method on two human activity datasets, we show that, upon detection of anomalous classifiers, an adaptation strategy such as removing them from the fusion chain, leads to a graceful performance degradation.

The structure of the paper is as follows; the next section summarizes related work on anomaly detection and resilience, then we describe the proposed method in Section 3. The description of the experiments and the results are presented in Sections 4 and 5, respectively. Finally, we discuss about the use of the method and its pros and cons in Section 6, followed by a conclusion.

2. Related work

We can handle anomalies in two manners: through detection and isolation, or by anomaly resilience and adaptation. In the former case, the goal is to detect whether there is an anomaly in the data or not (detection) and which part of the system is affected (isolation). While for the latter one, the system is designed to be tolerant against anomalies, or to be able to take suitable counteraction whenever an anomaly is detected. For example, by changing the network structure in wireless sensor networks (WSN) or adapting parameters to the new data trend.

2.1. Anomaly detection and isolation

Numerous studies have been undertaken to detect anomalies at the data level. Chandola et al. (2009) survey methods to detect anomalous patterns in a pool of patterns. These methods (such as computing distance to the nearest neighbor or to a cluster center, estimating statistical models on the data, discriminating normal and anomalous patterns using artificial neural networks, or support vector machines) are used in fault detection in mechanical units (Jakubek and Strasser, 2002), structural damage detection (Brotherton and Johnson, 2001), sensor networks (Ide et al., 2007), etc.

Time series change detection (Basseville and Nikiforov, 1993) has been applied to fraud detection (Bolton et al., 2002), computer intrusion detection (Schonlau et al., 2001) and concept drift (Lane and Brodley, 1998). One of the well-known approaches is CUmulative SUM (CUSUM) (Page, 1954) which is particularly used when the parameters of the changed signal are known. CUSUM computes the cumulative sum of the log-likelihood ratio of the observations and once this value exceeds a threshold (which can be adaptive) it is considered a change. It is widely used in change and drift detection in time series (Wu and Chen, 2006), exerting neuro-fuzzy models (Xie et al., 2007), auto-regressive models (El Falou et al., 2000), and Kalman filters (Severo and Gama, 2010) to generate residuals whose changes should be detected. Li et al. proposed a subspace approach to identify optimal residual models in a multivariate continuous-time system (Li et al., 2003).

When the parameters of the changed signals are unknown, generalized likelihood ratio (GLR) (Lorden, 1971) and adaptive CUSUM are proposed. GLR maximizes the likelihood ratio over possible values of the parameter of the changed signal, assuming the normal data follows a Gaussian distribution. It is widely used for change detection in brain imaging (Bosc et al., 2003), diffusion tensor imaging for monitoring neuro-degenerative diseases (Boisgontier et al., 2009), for detecting land mines using multi-spectral images (Anderson, 2008), and target detection and parameter estimation of MIMO radar (Xu and Li, 2007). In the same class, Adaptive CUSUM is able to detect changes by suggesting a distribution for the unknown change model based on the distribution of the known model of unaltered data (Alippi and Roveri, 2006a,b, 2008). One-Class Support Vector Machine (OCSVM) is another common approach for anomaly detection (Das et al., 2011). Its rationale is to compute a hyperplane around train data and samples are considered anomalous if they fall outside that hyperplane.

Other approaches have also been proposed in control systems to detect abnormal sensors. One is to extract the model of the system and detect faults by monitoring residual error signal (Hwang et al., 2010). Another more complex way is to create the input–output model of the system by regression methods and detect potential faults when there is a change in the estimated parameters (Ding, 2008; Smyth, 1994).

There are many studies and methods toward detection of fault and changes in the sensor networks. In this area, anomaly detection can be done by having redundancies in the network. Either in the form of physical redundancy such as adding extra sensors, thus increasing the cost of the system deployment; or from analytical redundancies (Betta and Pietrosanto, 1998). For instance, Andrieu et al. (2004) discuss particle methods to model validation, change/fault detection, and isolation. In this scope, model validation is a process to ensure reliable operations. A survey on the approaches to fault detection and isolation in unknown environments has been done by Gage and Murphy (2010). Chen et al. proposed a probabilistic approach that recognizes faulty sensors based on the difference in the measurements of a sensor and its neighbors (Chen et al., 2006). Other approaches are based on the similarity between two time series (Yao et al., 2010), OCSVM (Rajasegarar et al., 2007) and clustering (Rajasegarar et al., 2006) to detect outliers in a sensor network, and kernel-based methods (Camps-Valls et al., 2008) to detect changes in remote imaging systems.

2.2. Anomaly resilience and adaptation

In sensor networks there are different strategies to introduce robustness against changes and drifts in the sensors. Luo et al. (2006) discuss how to optimize the parameters of the model of a sensor under both noisy measurement and sensor fault. In turn, Demirbas (2004) proposes scalable design of local self-healing for large-scale sensor network applications, while Koushanfar et al. (2002) propose to use an heterogeneous back-up scheme to substitute faulty sensors. Other different designs, principles and service managements have been proposed to provide self-healing services and diagnosing the true cause of performance degradation (Krishnamachari and Iyengar, 2003; Ruiz et al., 2004; Sheth et al., 2005).

Alternatively, the system can be adapted to dynamic changes. Snoussi and Richard (2007) model the system dynamics, including abrupt changes in behavior, and select only a few active nodes based on a trade-off between error propagation, communications constraints and complementarity of distributed data. Wei et al. (2009) used discrete-state variables to model sensor states and