Computer Networks 64 (2014) 195–207

Contents lists available at ScienceDirect

Computer Networks

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

An adaptive elliptical anomaly detection model for wireless sensor networks

Masud Moshtaghi [a,*], Christopher Leckie [a], Shanika Karunasekera [a], Sutharshan Rajasegarar [b]

[a] NICTA Victoria Research Laboratories, Department of Computing and Information Systems, University of Melbourne, Parkville, Melbourne, Australia
[b] Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Melbourne, Australia

ARTICLE INFO

Article history:
Received 6 June 2011
Received in revised form 30 November 2013
Accepted 3 February 2014
Available online 14 February 2014

Keywords:
Adaptive models
Anomaly detection
Clustering hyperellipsoidals
Wireless sensor networks
Elliptical anomaly detection

ABSTRACT

Wireless Sensor Networks (WSNs) provide a low cost option for monitoring different environments such as farms, forests and water and electricity networks. However, the restricted energy resources of the network impede the collection of raw monitoring data from all the nodes to a single location for analysis. This has stimulated research into efficient anomaly detection techniques to extract information about unusual events such as malicious attacks or faulty sensors at each node. Many previous anomaly detection methods have relied on centralized processing of measurement data, which is highly communication intensive. In this paper, we present an efficient algorithm to detect anomalies in a decentralized manner. In particular, we propose a novel adaptive model for anomaly detection, as well as a robust method for modeling normal behavior. Our evaluation results on both real-life and simulated data sets demonstrate the accuracy of our approach compared to existing methods.

1. Introduction

A Wireless Sensor Network (WSN) consists of a set of nodes, where each node is provided with a set of sensing devices, a processing unit and a wireless communication unit. WSNs provide a cost-effective platform for monitoring and data collection in environments where the deployment of wired sensing infrastructure is too expensive or impractical [1]. An important challenge in the management of such WSNs is the detection of anomalies, i.e., unusual measurements that are inconsistent with the distribution of the majority of observations. Anomaly detection has several important roles in the management of WSNs. For example, anomalous measurements can be caused by faults in sensors. In this context, it is important to detect and filter those erroneous measurements, to ensure the integrity of the collected data. Anomalies also correspond to events of interest. In this case, we may be interested in only collecting those unusual changes in order to avoid wasting resources on reporting normal measurements. The focus of this paper is on detecting changes from the normal behavior in the node.

In a centralized approach to monitoring, each node sends its data to a central location where the data is analyzed to detect anomalies. This is an inefficient approach since the communication cost of sending all the data over the network is high and this drains the batteries of sensor nodes. For example, transmitting one bit can consume as much energy as running several thousand instructions on a sensor's CPU [2]. While a centralized approach incurs a communication overhead, it potentially provides the best results in terms of accuracy [3], so it forms a baseline for comparing different algorithms. An alternative is a decentralized approach where each sensor performs local processing to detect and remove anomalies. However, due to the limited resources available in a wireless sensor node, we cannot

* Corresponding author. Present address: Faculty of IT, Monash University, Australia. Tel.: +61 450831677.
E-mail address: masud.moshtaghi@computer.org (M. Moshtaghi).

http://dx.doi.org/10.1016/j.comnet.2014.02.004
1389-1286/

use common anomaly detection techniques for this purpose. Hence our challenge is to detect anomalies in data that is distributed among a set of sensors, while minimizing communication between nodes. We also need to consider the limited computational capabilities of each node.

A naïve approach for detecting anomalies in WSNs that keeps the computational complexity low, is to define maximum and minimum bounds for the normal data [4,5]. However, this approach has low accuracy because minimum and maximum bounds are not flexible enough to create a tight boundary over the data. The limitations of maximum and minimum boundaries have motivated research into decentralized data modeling approaches to anomaly detection in WSNs, so that in-network processing can be used to reduce energy intensive communication within the network [6]. By effectively selecting critical measurements and forwarding them to the base station, we can increase the life span of the WSNs without affecting the decision making process. We aim to increase the lifetime of the network, which can make the application of WSNs more practical in large scale monitoring applications.

There are a variety of algorithms that have been proposed to detect anomalies by using different data models to accomplish this task [7–9]. However, they either incur high computational cost on each sensor [10,8,9] or they assume that the monitored environment is homogeneous [7], i.e., the underlying distribution of measurements seen by each sensor is approximately identical. In addition, in many applications the underlying system changes often, which limits the usage of the static methods. The authors of [11] demonstrated a need for adaptive modeling for anomaly detection in WSNs.

The main contributions of this paper are as follows: (1) we introduce two models for estimating the global decision boundary for robust anomaly detection, based on the minimum volume ellipsoid covering a set of ellipsoids from different sensors, and the Vysochanskij–Petunin inequality; (2) we evaluate the performance of these two methods on two real-life datasets; and (3) we propose an adaptive model for maintaining the decision boundary for anomaly detection without the need for re-training, and give an empirical evaluation of the performance of our approach on two real-life datasets. Our results demonstrate that our new approach can achieve higher accuracy compared to a state-of-the art approach [12], and it can adapt to changes in the environment, which makes it suitable for use in practical situations. In the next section we summarize related work in this field. In Section 3, we present the formal statement of the problem we address. We then present an overview of previous work in Section 4. Sections 5 and 6 are dedicated to our new approach to global boundary detection and our proposed adaptive model, respectively. A summary and conclusions are given in Section 7.

2. Background and related work

An important aspect of a monitoring system is to detect significant events or unusual behavior in that environment. Anomaly detection methods play an important role in modeling and detecting these anomalous events in the system.

Anomaly detection has been applied in a variety of applications [13], including intrusion detection [14,6], and event detection [15]. Numerous factors affect the use of anomaly detection in these applications, such as mobility in sensors, benign or adverse environments [3] and energy constraints.

WSNs are usually deployed to monitor environments such as an oil field, a forest or the perimeter of a military camp. In all these applications, detecting events or anomalies can help produce an early warning system. In [16], data from a large number of sensors deployed in modern oil fields are analyzed automatically for generating alerts. In [17], a forest fire detection scenario is explored, where sensor nodes build simple linear time series models, which are then used by the base station to estimate the fire danger classes and raise an alarm if necessary. In [18], the anomalies, which are referred to as ‘‘informative observations’’, are identified using statistical methods. The statistical model is created for the normal data distribution and is used to decide which observations or sensors are required to communicate the data to the base station for processing, i.e., which nodes or measurements are ‘‘censored’’ for communication.

A common approach in detecting anomalies in a monitored environment is to model the normal behavior of that environment over a period of time, and then label those measurements that do not conform to the normal model as anomalous. Most of the anomaly detection algorithms proposed in WSNs assume that all data is available at a single location for extensive processing. In this case, any outlier detection algorithm can be applied to the data (see [6,19] for surveys). In this paper we focus on approaches that do not require all the data to be stored centrally for processing. A survey of anomaly detection techniques proposed for WSNs in this context can be seen in [6]. Below, we review several prominent anomaly detection techniques proposed for use in WSNs.

The authors of [8,9] proposed one-class support vector machine models to separate normal data from anomalies. The main assumption in this approach is that all the training data is available at the sensors, and the training can be done in *batch mode*, i.e., all the measurements collected are processed as a single batch. Although these methods provide high accuracy, the main restriction of these algorithms are that they impose a computational overhead of $O(n^3)$ on each sensor, where n is the number of data samples.

In [20], clustering techniques are used to detect anomalies. The assumption in this approach is also that the training can be done in batch mode. The advantages of this algorithm are that it has lower computational complexity, and gives partitions of the data at the base station that help characterize the environment to the user. The restrictions of this algorithm are that its accuracy is sensitive to the fixed radius parameter of the hyperspheres, and its hyperspherical partitions do not always provide an accurate model of the distribution of normal data.

In order to create flexible boundaries in the model for the normal data, the authors of [7,12] used hyperellipsoids to model the normal behavior of the system. Both these models assume the measurements in a sensor node follow