Journal of Network and Computer Applications 66 (2016) 236–249

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

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

Big data analytic architecture for intruder detection in heterogeneous wireless sensor networks

Suvendu Kumar Mohapatra [a], Prasan Kumar Sahoo [b], Shih-Lin Wu [b,c,*]

[a] Department of Electrical Engineering, Division of Computer Science and Information Engineering, Chang Gung University, Kwei-Shan 333, Taiwan, Republic of China
[b] Department of Computer Science and Information Engineering, Chang Gung University, Kwei-Shan 333, Taiwan, Republic of China
[c] Center for Biomedical Engineering, Chang Gung University, Kwei-Shan 333, Taiwan, Repubic of China

ARTICLE INFO

Article history:
Received 28 August 2015
Received in revised form
14 December 2015
Accepted 7 March 2016
Available online 10 March 2016

Keywords:
WSN
Barrier coverage
Intruder detection
Big data
Spark

ABSTRACT

Barrier coverage in Wireless Sensor Networks (WSNs) is an important research issue as intruder detection is the main purpose of deploying wireless sensors over a specified monitoring region. In WSNs, excessive volume and variety of sensor data are generated, which need to be analyzed for accurate measurement of the image in terms of width and resolution. In this paper, a three layered big data analytic architecture is designed to analyze the data generated during the construction of the barrier and detection of the intruder using camera sensors. Besides, a cloud layer is designed for storing the analyzed data to study the behavior of the intruder. In order to minimize the number of camera sensors for constructing the barrier, algorithms are designed to construct the single barrier with limited node mobility and the barrier path Quality of Sensing (QoS) is maintained with a minimum number of camera sensors. Simulation results show that our algorithms can construct 100% of the barrier with fewer number of camera sensors and average data processing time can be reduced by using parallel servers even if for larger size of data.

1. Introduction

Advanced development and improvement in Micro-Electro-Mechanical Systems (MEMS) technology, mixed with low power, small in size, minimum cost sensors can be equipped in Wireless Sensor Network (WSN). Now-a-days various types of sensors such as microwave sensors, thermal sensors, laser sensors and camera sensors are available according to the applications and working environments. Those sensors are static or having limited mobility (Dantu et al., 2005; El-Moukaddem et al., 2013; Janansefat et al., 2013; Chellappan et al., 2007) to gather and process environmental information. Camera sensors are different from the general sensors in terms of image capture capability and are used for number of applications such as border surveillance (Tao et al., 2012; Cheng and Tsai, 2012), and intruder detection (Keung et al., 2012; Sahoo et al., 2013). Area, point and barrier coverage are critical coverage issues in WSN, and are the parameters to appraise the quality of surveillance. In this paper, we are interested to focus on the barrier coverage problem. Barrier coverage (Shih et al., 2010) is the line coverage to cover all possible crossing paths of the intruder within the monitoring region. In directional camera sensor (Tao et al., 2012; Cheng and Tsai, 2012; Wang and Cao, 2012), the sensing range is an arc having a field of view with a finite angle and the intruder is picked up within the arc range. However, previous barrier construction protocols do not consider the mobility of the sensors, camera rotation and Quality of Measurement altogether. Hence, we design here an efficient barrier construction algorithm by considering all the above parameters.

In our barrier construction algorithm, we use microwave sensors (Dual Technology Motion Sensor) and directional camera sensors. Microwave sensors detect the movement of the intruders whereas camera sensors are used to identify the image of an intruder. Intruder detection (Keung et al., 2012; Sahoo et al., 2013) is a part of border surveillance in which the intruder is captured by the sensors which are present along the barrier line. Once the barrier network is established, colossal amount of streaming data are generated by the camera sensors, which is difficult to handle and analyze using the traditional data processing platforms. Therefore, big data analytic platform is used here to process the gigantic image data.

In order to analyze the data generated by the camera sensors, big data is the best solution to manage those unstructured streaming data with a cost effective manner, which is very decisive in terms of volume, velocity, variety and value. Volume handles

* Corresponding author.
E-mail addresses: d0121007@stmail.cgu.edu.tw (S.K. Mohapatra), pksahoo@mail.cgu.edu.tw (P.K. Sahoo), slwu@mail.cgu.edu.tw (S.-L. Wu).

http://dx.doi.org/10.1016/j.jnca.2016.03.004
1084-8045/

enormous amount of data generated continuously by many different camera sensors. Velocity focuses on the tremendous speed at which the camera sensor data (bytes) are coming for processing. Variety defines that the diversified data format arrives from various sources. Value represents the meaningful information by converting the data insights. In the industry level data scale and cutting-edge network technology, new challenges force the researchers as well as developers to improve in solutions for data collection, transmission, processing and storage. Virtualization technology acts as a backbone of various big data analysis tools such as Hadoop, where chunks of data are processed in parallel. To support such batch processing parallel execution, Hadoop MapReduce framework (Dean and Ghemawat, 2008; Yang and Chen, 2015) is used. However, the Spark platform (Zaharia et al., 2012) is employed for the realtime coordinated streaming data processing for the intruder detection. A cluster of slave nodes are used in the data analysis with complex work flows, which are controlled by the Spark master nodes. The analyzed data are stored in Cassandra (Lakshman and Malik, 2010) distributed database in the data centers.

1.1. Motivations

Barrier construction and intruder detection using wireless camera sensor network is highly essential and is convenient in surveillance system. However, to the best of our knowledge no barrier construction protocol proposed so far considers all the three functionalities such as node mobility, rotation of the camera sensors and Quality of Measurement of WSN to detect the intruder efficiently. Moreover, camera sensors are normally expensive and efficient detection of an intruder with a minimum number of camera sensors is a challenging research issue. Hence, the main motivation of our proposed work is to reduce the number of camera sensors for constructing the barrier to detect the intruder by combining all those three parameters. Besides, high volume of streaming data are generated from the camera sensors once the barrier is constructed, which need to be analyzed for detecting the intruder properly. However, it is very difficult to analyze those low-latency and high volume of unstructured streaming data manually or in any batch processing big data framework such as Hadoop. In Hadoop, the complete batch data must be loaded before the processing is done, which encounters the startup delay with intermediate data shuffling overhead during computation. Therefore, we propose an in-memory data processing Spark platform to handle such real-time data sets.

1.2. Contributions

The main contributions of our work can summarized as follows.

- A barrier construction mechanism in heterogeneous WSN is designed with a minimum number of camera sensors to collect the huge amount of real time image data for the analysis.
- Quality of Sensing (QoS) is maintained throughout the barrier path with a minimum number of camera sensors.
- A big data analytic architecture is designed to analyze and store the low-latency big data generated from the wireless camera sensors.
- A Graphics Processing Unit (GPU) enabled Spark cluster is proposed for the in-memory data processing and frame-by-frame analysis of the realtime visual streaming data.
- Based on the analysis of the big data in our proposed Spark platform, intruder detection mechanism is also designed.
- Our proposed algorithms can provide 100% barrier coverage with a minimum number of camera sensors.

Remainder of this paper is organized as follows. Section 2 describes the related works on barrier coverage and intruder detection using heterogeneous sensors. In Section 3, the big data analytic architecture is described. The wireless sensor layer is described in Section 4, which includes the barrier construction algorithms. Big data analytic with cloud based storage layer is described in Section 5. Simulation results are given in Section 6, and Section 7 concludes the work.

2. Related work

Comprehensive studies have been carried out on barrier coverage issues in WSN. Directional sensor networks (Tao et al., 2012) use directional sensors to construct the strong barrier. The objective was to diminish the total number of sensors and save the energy by minimizing the maximum rotation angle. However, most of the existing solutions are centralized and take longer time to detect the intruder. In distributed barrier coverage with β-QoM (Cheng and Tsai, 2012), wireless visual sensors construct the barrier by maintaining β-breadth to increase the quality of monitoring (QoM). Authors have proposed two β-breadth belt-barrier construction algorithms without rotation of the sensors, in which barrier is constructed with β-breadth. Distributed β-breadth belt-barrier construction algorithm with rotation is proposed, in which barrier is constructed by camera sensors with rotation capability. Their main contribution is to minimize the number of visual sensors by maintaining the quality, the number of sensors can still be minimized by adding limited mobility to the sensors. Also resolution factor can be calculated by maintaining the distance from the location of the intruder.

It is to be noted that mobility in camera sensor networks (Dantu et al., 2005; El-Moukaddem et al., 2013) has heightened the monitoring quality. Sensors with controlled mobility (Vecchio and Lopez-Valcarce, 2015) can enhance the deployment strategy, adaptive sampling, hole detection and repair capability and event detection mechanism can even become better. MICAbot (Janansefat et al., 2013) is inexpensive, adaptable and modular mobile robots, which are used in large scale distributed sensor networks. By using those mobile sensors, we can build the barrier network in an efficient way. Now-a-days, many camera sensors (Mehta et al., 2009) are available for constructing the wireless sensor networks to detect the intruders. In Hoseini et al. (2012), the coverage problem of three dimensional objects by enabling the tilt, zoom and pan functionalities of the camera sensors is investigated. In their proposed solution, a circular target model is used to determine the full coverage. In Chen et al. (2010a), the object coverage problem with rotating capabilities of camera sensors is explored. In order to reduce the redundant image data, they map the proposed problem to the set coverage problem.

Authors in Chow et al. (2007) have analyzed the angle coverage problem in visual sensors and propose an algorithm to achieve full view of the target. By preserving 360° angle, they propose an energy efficient algorithm, which tries to minimize the transmission cost over the network. In Zanella et al. (2014), authors have focused on a smart city vision as an application of Internet of Things (IoT). The main goal is to collect environmental data and monitor the public street light. However, the massive IoT data storage is not considered. Authors in Jiang et al. (2014) have proposed a cloud based data storage framework for both structured and unstructured data. The data are collected by sensors and RFID readers and the main advantage of this framework is to combine and extend multiple databases with Hadoop to store. But no analysis is done on this huge stored data.

Recently, many computation intensive (Bhattacharya et al., 2014) and data intensive applications like border surveillance,