Poster Abstract

Warriach, Ehsan; Nguyen, Tuan Anh; Aiello, Marco; Tel, Kenji

Published in:
IPSN'12: PROCEEDINGS OF THE 11TH INTERNATIONAL CONFERENCE ON INFORMATION PROCESSING IN SENSOR NETWORKS

Publication date:
2012

Poster Abstract:

Fault Detection in Wireless Sensor Networks: A Hybrid Approach

Ehsan Warriach, Tuan Anh Nguyen, Marco Aiello
University of Groningen
Groningen, The Netherlands
{e.u.warriach, t.a.nguyen, m.aiello}@rug.nl

Kenji Tei
National Institute of Informatics
Tokyo, Japan
tei@nii.ac.jp

ABSTRACT

Wireless Sensor Network (WSN) deployment experiences show that data collected is prone to be imprecise and faulty due to internal and external influences, such as battery drain, environmental interference, sensor aging. An early detection of such faults is necessary for the effective operation of the sensor network. In this preliminary work, we propose a hybrid approach to the detection of faults and we illustrate its performance on data coming from a real sensor deployment. The proposal is a first step to have a hybrid method towards automated on-line fault detection and classification in context-aware WSNs middleware framework.

Categories and Subject Descriptors

B.8.1 [**Hardware**]: Metrics—*Performance and Reliability - Reliability, Testing and Fault-tolerance*

General Terms

Algorithm, Experimentation, Measurement, Reliability

Keywords

Data & System faults, Hybrid Fault detection

1 Introduction

Wireless Sensor Networks (WSN) has been widely employed for enabling various monitoring and control applications such as environment surveillance, industrial sensing, traffic monitoring, etc [4]. Many mobile and pervasive applications deliver us with information about phenomena or events at a much higher level of detail by continuously collecting and processing information from the physical world. The cornerstone for its success lies in the ability to draw meaningful and precise inferences from the collected data which in turn requires to have high data quality coming from the sensors. In general, wireless sensor nodes may experience two broad categories of faults that lead to the degradation of performance such as function and data faults. On the one hand, there is the data centric view comprising faults such as outlier, spike, stuck-at and noise. On the other hand, there is the system centric view with faults such as calibration, hardware, low battery and clipping [3].

To illustrate examples of fault, consider two specific sensors such as pressure sensor to detect chair occupancy status and designed by Advantic Systems uses the Tekscan® A201-100 FlexiForce® sensor. It provides force and load measurements. Another is Passive InfraRed (PIR) sensor detects the motion of users inside the office. The sensor uses the Perkin Elmer Optoelectronics® LHI878 sensor. We study the sensors in the context of an activity recognition effort part of the EU FP7 project GreenerBuildings [1]. For the project we set up a Smart Office Living Lab at the University of Groningen (RUG) where a number of simple sensors are used [2]. If we consider individual sensor readings over a period of time, we notice that faults can have a diversity of lengths, magnitudes, and patterns. For instance, some faults show a long-duration, relatively gradual change in sensor reading, whereas others exhibit a short duration, quite sudden change. Take for instance two PIR sensors. They can report value significantly deviating for short intervals from the expected values, resulting in a noticeable short in the graph (see Figure 1(a)). If instead we consider a pressure sensor, it reports value with much higher rate of change than expected over a period of time which may not return to the normal afterwords, resulting in a noticeable spike(see Figure 1(b)). So not a single mathematical fault model is perfect for detecting these kinds of faults.

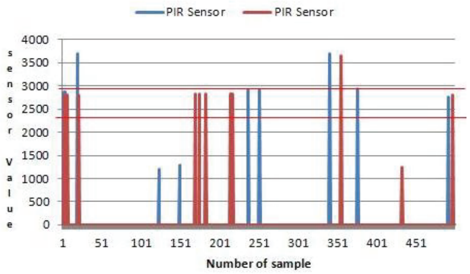
Given the general impossibility to install a perfectly calibrated and robust network of sensors, we consider the paramount problem of run-time detection and classification of faults. We couple it with root-cause analysis of sensor faults, as well as techniques that can automatically scrub collected sensor data to ensure high quality in context-aware WSNs middleware framework. A first step in this direction is to obtain an understanding of the occurrence of faulty sensor readings in existing real-world deployments by using more than one detection method. Here, we propose to firstly explore three qualitatively different techniques for automatically detecting faults from a trace of sensor readings: rule-based, estimation, and learning based methods. Our proposal is based on (1) by artificially injecting faults of varying intensity into sensor datasets, we are able to study the detection performance of these methods. We find that these methods sit at different points on the accuracy and robustness spectrum; (2) we evaluate the methods on real-world datasets, where we observe the actual frequency of fault occurrence; (3) we define heuristics rules to identify faults dynamically in WSNs. This study lays the foundations for the definition of a hybrid approach to fault detection in WSNs.

2 Faults Detection Approach

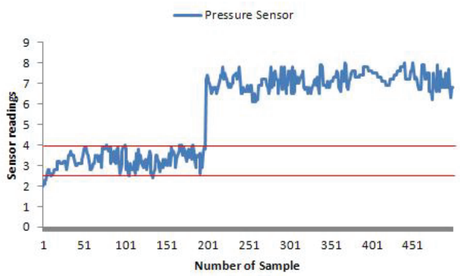
To define a hybrid approach, we begin by over-viewing three prominent methods.

Detection Methods: While there are numerous methods for detecting faults, we consider three qualitatively diverse methods Rule-

IPSN12, April 16–20, 2012, Beijing, China.
ACM 978-1-4503-1227-1/12/04.



(a) PIR sensors readings



(b) Pressure sensor readings

Figure 1: Visual faults in RUG dataset

based, Linear Least-Square Estimation (LLSE) and Hidden Markov Model (HMM - learning-based) to define heuristics rules for detecting and identifying faults dynamically. Rule-based methods are based on domain and expert knowledge to define heuristic rules for identifying and classifying faults. Estimation methods predict normal sensor behavior by leveraging sensor correlation. Finally, learning-based methods are able to statistically identify and classify classes of faults. The system heuristics rules built on expert and domain knowledge and data from activity recognition in smart office Lab. Next we describe such heuristic rules.

Rules: Upon getting a data point, the system applies the set of rules developed to identify inconsistent data. Consequently all rules are applied to all data points, a few of them are listed below:

• Noise: There is always some amount of noise in the measurements. We take the average of readings in the noise window N. If the standard deviation of points within the noise window is greater than some threshold specified by the noise parameter, the samples are corrupted by the noise fault. Clearly, the performance of this rule depends on the threshold value.

• Constant: Compute the standard deviation of sample readings within a window N. If it is zero, the samples are corrupted by constant faults. The window size N can be in terms of number of sample or time, so the performance of this rule depends on the window size N.

• Outlier: The most common features to model an outlier are distance from other readings and gradient but first needs to model the expected underlying behavior. For that purpose we use simple methods of determining the expected range based upon previous data sample points and also expert and domain knowledge. Contextual information about the sensor plays a larger role in this fault meanwhile we are modeling expected behavior in a struggle to identify outliers.

• Short: Compute the rate of change of the physical phenomenon being sensed between two successive sensors readings. If the rate of change is above a threshold, it is an instance of a short fault.

Once the rules have been applied to each data-point and recovery required for the various identified faults to resolve them dynamically. The recovery and reconfiguration techniques that would either isolate or resolve identified faults remain also to be investigate in future.

3 Preliminary Evaluation and Conclusion

Our preliminary evaluation of the approach on the occurrence of faults in real world datasets coming from the GreenerBuildings project shows promise. The major fault in the readings was of the type short. We applied the short rule, LLSE and HMM methods, and a Hybrid (combination of Rule, LLSE and HMM) method to identify short faults in pressure, acoustic and PIR sensors readings. Figure 2 shows the occurrence of short faults for different sensors in the GreenerBuildings dataset. The Hybrid method removes false positives reported by the Short, LLSE or HMM methods. From Figure 1(a) one notices that short faults are relatively common. They are most predominant in the PIR sensor data (almost 1 fault every 70 samples). In this data set, spike faults were infrequent. Only one node had spike fault but it last for long time. If we consider it as a system fault then it will be either low battery or node broken as shown in the Figure 1(b).

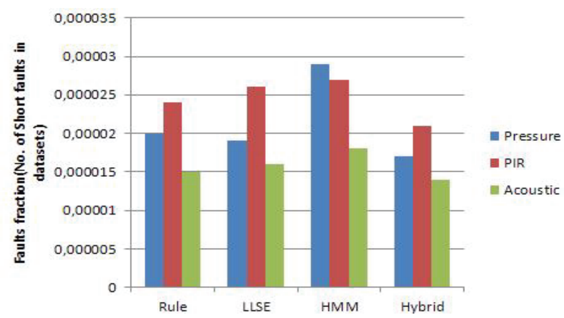


Figure 2: Short Faults in GreenerBuildings Dataset

In summary, we propose a hybrid approach for the detection of data faults in wireless sensor network which shows promise when applied to real sensor data. We plan to continue by integrating well-known mathematical fault models to make more robust and dynamic Hybrid method to identify data faults as well as system faults in context-aware WSNs and to thoroughly evaluate them.

4 References

[1] *GreenerBuildings*. http://www.greenerbuildings.eu.

[2] T. Nguyen and M. Aiello. Beyond indoor presence monitoring with simple sensors. In *Proc. 2nd Int. Conf. on Perv. & Embedded Comp. and Comm. Sys.*, 2012.

[3] K. Ni, N. Ramanathan, M. N. H. Chehade, and M. Srivastava. Sensor network data fault types. *ACM Trans. Sen. Netw.*, 5:25:1–25:29, June 2009.

[4] C. yee Chong, Ieee, S. P. Kumar, and S. Member. Sensor networks: evolution, opportunities, and challenges. *Proceedings of the IEEE*, 2003.