FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Smart Metering towards energy efficiency increasing in railways systems

Vítor A. Morais

DRAFT VERSION

Programa Doutoral em Engenharia Electrotécnica e de Computadores

Supervisor: António P. Martins

June 15, 2017

Smart Metering towards energy efficiency increasing in railways systems

Vítor A. Morais

Programa Doutoral em Engenharia Electrotécnica e de Computadores

June 15, 2017

Contents

Chapter 1

Introduction

This chapter presents the context, motivation and document structure of a study of outlier detection in a railways WSN-based smart grid.

1.1 Context and motivation of PhD

The railway system is responsible for 1.3% of entire European energy consumption, Birol and Loubinoux (2016). The discussion of the energy efficiency in railways is a grown topic due to its contribution to the global energy consumption.

The energy efficiency analysis and management requires a detailed mapping of the energy consumption/generation in the railway system.

This detailed mapping of the energy flows should include, not only the rolling stock level but also the traction substations and the auxiliary services.

The knowledge of all the load curves permits the load prevision, peak shaving and energy cost optimization for all global railway system.

1.2 Shift2Rail Framework

This work is supported by the iRail PhD programme – Innovation in Railway Systems and Technologies whose objectives are aligned with the Shift2Rail objectives, Shift2Rail Joint Undertaking (2015):

- 1. Cutting the life-cycle cost of railway transport by as much as 50%;
- 2. Doubling the railway capacity;
- 3. Increasing the reliability and punctuality by as much as 50%.

Framed on the Shift2Rail (S2R) Innovation Programme 3 (IP3) with the focus on "Cost efficient and reliable infrastructure", it is proposed to develop a Smart Metering Demonstrator (SMD) that reach a detailed monitoring and supervision of various energy flows on the premises of embrace the entire Railway System.

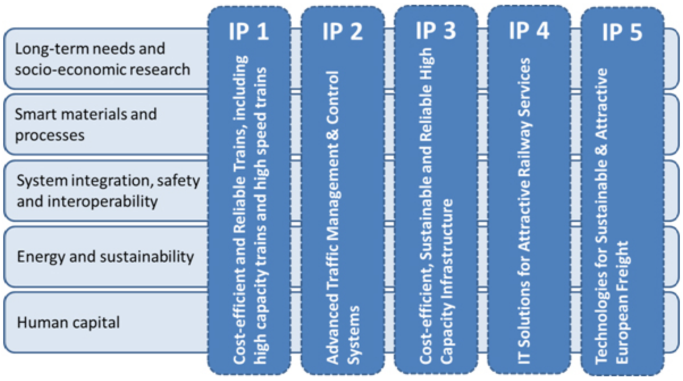

Figure 1.1: Shif2Rail Inovation Programmes.

The purpose of any energy management strategy is to build the dynamics of every loads and generators of the power system.

This should be performed based on an extensive knowledge of every energy flows.

This way, the SMD is required to propose and validate a standard metering architecture that involves the coordination of every measurements either in on-board and in ground. In advance, energy data analysis should be provided based on relevant stored data.

1.3 PhD state of the art

This section will cover a summary of the state of the art that supports this PhD.

Based on the state of the art, current metering systems focus on rolling stock on-board energy meters for energy billing purposes only, where the metering devices are located close to the pantograph, Shift2Rail Joint Undertaking (2015).

An advance beyond the state of the art is the expansion of the measurement system at railway system level, making it a distributed one, including both on-board and track-side measurements, thus achieving detailed mappings.

Other point in the state of the art is the intrusion level of currently used metering systems, that in one way, became a critical subsystem of the rolling stock and in other way, requires relatively long implementation, Shift2Rail Joint Undertaking (2015).

An advance beyond the state of the art is a solution based on non-intrusive technology. More detailed simulation models in conjunction with field measurements is the methodology to be investigated.

Specific challenges and requirements of this research are the development of non-intrusive Wireless Sensor Networks (WSN) in the railway environment. It is intended that this technology should be based on an open system and open interfaces for the data collection, aggregation and analysis. Issues like metering redundancy, outlier detection, fault tolerance and communication reliability, should be considered during the research. In addition, it is expected to design and specify a set of user applications. Those applications are focused in the energy analysis process with the aim of providing more information and detailed knowledge. It is expected that this detailed knowledge would be useful in a decision support system related with, in e.g., eco-driving strategies, timetable planning and preventive maintenance.

1.4 Influence of outliers in a railway remote monitoring system

Having in mind the state of the art that was previously presented in section 1.3, an important contribution of a wireless sensor network in the railway system is the availability of useful knowledge of the energy consumption to the decision support systems.

Therefore, such acquisition systems are required to provide accurate data regardless of the quality of the acquisition sensors, electromagnetic influences (EMI), sensor supply fluctuations, among others.

Through computational algorithms, the increasing of communication reliability and fault tolerance is possible. Those computational algorithms detect outliers or, in the scope of this PhD, detect erroneous data that will perturb the outcomes of decision support systems. Further on in chapter 2, this thematic is extensively explored.

1.5 Document structure

This document is divided in 5 chapters, each of them incorporate the relevant subsections to present the subjects mentioned.

Table 1.1: Document structure

Chapter	**Title**
1	Introduction
2	Outliers Detection
3	Future Research
4	Conclusions

Chapter 2

Railways Remote Monitoring Systems

In this chapter it is an overview of the railway system where the outliers detection is expected to be studied.

2.1 Smart Meters

2.2 Synthesis

Chapter 3

Railways Energy Model

In this chapter it is an overview of the railway system where the outliers detection is expected to be studied.

3.1 Smart Meters

3.2 Synthesis

Chapter 4

Outliers Detection

In this chapter it is made the study of the state of the art of outliers and it's relevance in railways.

In section 4.1 is defined what is an outlier either with base on the literature and with base on the scope of the PhD. In section 4.2 is covered the motivation, research opportunities and challenges in outlier detection for WSNs and for the scope of the PhD. In section 4.3 different aspects of outlier detection that has been used in the literature are presented. In section 4.4 the taxonomy to divide and classify the different techniques is presented.

The remaining sections will extensively cover the different techniques. Section **??** covers the classification techniques; Section **??** presents the statistical based techniques; In section **??** the nearest neighbor techniques are covered; Section **??** presents the cluster-based techniques and section **??** covers the Spectral-based techniques.

In section **??** is made a synthesis of the outlier detection techniques for WSNs.

4.1 Definition of outlier detection

Outlier detection is a computational task to detect and retrieve information from erroneous data values. The definition is usually close to anomaly detection or deviation detection.

Branch et al. (2006) identifies the outlier detection as an essential step to either suppress or amplify outliers and precedes most any data analysis routine. Abid et al. (2016) points the need of detecting aberrant data and sensors within an WSN. Zhuang and Chen (2006) extends the outlier definition to the case where the outliers introduce in sensing queries and in sensing data analysis.

In the scope of the PhD and as previously presented in chapter 1, an outlier is a data value or a data instance that do not represent the correct consumption status.

The threshold of what is an outlier or not (or a value that do represent the correct consumption status or not) is given by the output of the subsystem that is immediately after the acquisition of consumption status subsystem, the decision support subsystem, gave a correct output or not. Figure 4.1 illustrates the integration of the consumption acquisition subsystems with the decision support subsystem.

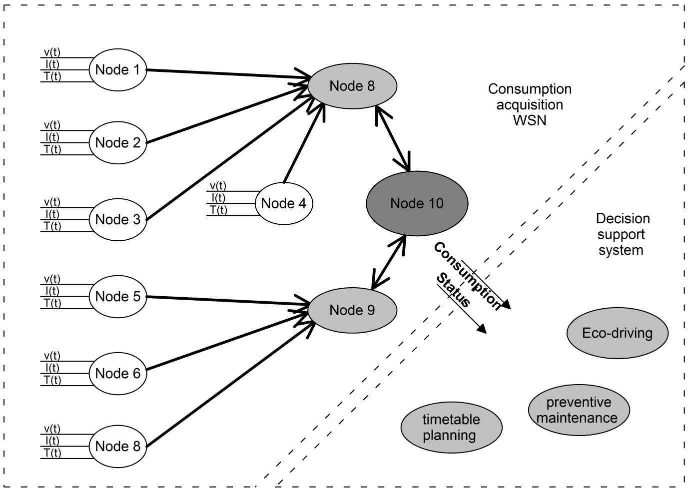

Figure 4.1: Integration of the WSN with a decision support system.

Without an outlier detection mechanism, the decision support subsystem may have the following outputs:

Input deviation from real value lower than a threshold The Decision Support Subsystem output is according to the real consumption conditions.

Input deviation from real value greater than a threshold The Decision Support Subsystem output is not according to the real consumption conditions.

The problem of taking decisions based on wrong considerations of the consumption status may lead to loss in desirable efficiency or increase of costs.

Let us consider a simple and hypothetical example where the DSS will provide an output towards suggesting an action in preventive maintenance based on the usage of a component. Considering that the usage of the component is depending on the counting of situations that the component is working above the nominal conditions. Without an outlier detection mechanism, the outliers will induce the DSS to count situations of overcharge of the component where the measurement is not related to the working above the nominal conditions but is related to external influences such as EMI or temperature. The output of DSS may suggest a preventive maintenance on a component that is working in proper conditions.

To conclude, with an outlier detection mechanism in the consumption acquisition subsystem the decision support subsystem may know if the value of consumption is an outlier or not and, with that information, the DSS output will be more accurate with the real conditions of operation.

4.2 Outlier detection in WSNs

Wireless sensor networks (WSNs) has been widely used in several applications in several domains such as industrial, scientific, medical and others. Those applications have been supported by the advances in wireless technologies as well as in the evolution of microcontroller technologies, with enhanced processing capabilities associated with reduced energy consumption.

4.2.1 Motivation

Rajasegarar et al. (2007) points an important motivation for the inclusion of computational algorithms, i.e. outlier detection algorithms, to reduce the transmission of erroneous data, since in WSNs, the majority of the energy consumption occurs in the radio communication. In particular, they present the case of Sensoria sensors and Berkeley motes where the energy consumption in communication exceeds in ranges from 1000 to 10000 the energy consumption of computation.

Thus, a research opportunity is raised to reduce the communication usage of μCs by adding processing features where the small increase in the computation will significantly reduce the energy consumption in the transmission. These processing features are, among others, the outlier detection algorithms.

On the field of the quality of the data acquired by WSNs, the motivation of detecting outliers in data acquired from WSNs has been extensively presented in the literature. The need for acquire data from harsh or "highly dynamic" environments as well as the need to validate and extract knowledge from the acquired data are one of the main points in the motivation to study the outlier detection in WSNs, Zhang et al. (2010); Chandola et al. (2009); Ghorbel et al. (2015); Martins et al. (2015).

4.2.2 Research areas

Zhang et al. Zhang et al. (2010) identifies the outlier detection research areas in three domains:

- Intrusion detection: Situation caused by malicious attacks, where the detection techniques are query-driven techniques;
- Fault detection: Situation where the data suffer from noise and errors and where the detection techniques are data-driven ones;
- Event detection: Situation caused by the occurrence of one atomic or multiple events and where the majority of the research has been developed due to its complexity.

Based on the division of this three domains, the upcoming research is intended to be focused on the event detection and fault detection techniques. Specifically, the main goal for this research will be the event detection algorithms.

4.2.3 Challenges

The challenges of outlier detection in WSNs are related to the quality of the acquisition of the sensors, the reliability of the modules in terms of energy or environmental susceptibility, and the communication requirements and restrictions.

Zhang et al. Zhang et al. (2010) lists the challenges as the following:

- Resource constraints;
- High communication costs;
- Distributed streaming data;
- Dynamic network topology,
  frequent communication failures,
  mobility and heterogeneity of nodes;
- Large-scale deployment;
- Identifier outlier sources;

Branch et al. (2006) identifies an important challenge, where the probability of occurrence of outlier events are extremely small. Abid et al. (2016) as well as Sheng et al. (2007) identifies the large amount of data as the main challenge for outlier detection in WSN. Zhuang and Chen (2006) identifies the inexpensive and low fidelity sensors as the main reason for the error generation and, the main challenge are identified on the distributed streaming data among a large amount of sensors. Ghorbel et al. (2015) points a main challenge as the processing of data from sensors that generates continuously data that is uncertain and unreliable.

To conclude, and in the scope of the PhD, the main challenges will be the usage of inexpensive and low fidelity sensors that will be affected by the rush railway environment. Complementary, the main challenge of using outlier detection mechanisms in the railway WSN is the balance between the detection accuracy and the influence that undetected data-instances will induce in other sub-systems (in particular in decision support systems dependent on data from the WSN). In addition, the detection accuracy is directly related with the memory usage, computational requirements, communication overhead, etc.

4.3 Classification of outlier

Zhang et al. (2010) presents aspects to be used as metrics to compare characteristics of different outlier detection techniques. In parallel, Chandola et al. (2009) presents a similar approach for the classification of outlier detection. In table **??** is present a comparison between two approaches to classify the nature of input sensor data.

Based on the work of Zhang et al. (2010) and Chandola et al. (2009), the table **??** identifies the different types of outliers. Those types differ on the objective of the outlier detection techniques: detect anomalies in individual data instances or in groups of data to detect irregularities, respectively, in local or in the global measuring system.

Table **??** continues the classification, focusing in three parts:

- The need of pre-classified data (to implement supervised, semi-supervised or unsupervised outlier detection techniques);
- The output of outlier detection techniques (binary labels for normal/abnormal data-set and a score for each data-set to evaluate the weight of being an anomaly)
- The identity of the outliers (detect errors, events or malicious attacks)

4.4 Taxonomy of Outlier Detection Techniques

The study of detection techniques requires a well-defined taxonomy framework that addresses the related work on the different areas. This taxonomy is well defined and solid in the literature, where the works of Zhang et al. (2010) and Chandola et al. (2009) reflect a similar approach on presenting a taxonomy for outlier detection techniques.

In the following sections a coverage in relevant techniques is presented:

- Classification based techniques.

  Bayesian Networks

  Rule-based techniques

  Support Vector Machines
- Statistical based techniques.

  Parametric — Gaussian based

  Non-parametric — Histogram based

  Non-parametric — Kernel function based
- Nearest Neighbor-based techniques.

  Using distance

  Using relative density
- Clustering based techniques.
- Spectral Decomposition based techniques.

Chapter 5

Future Research

In this chapter there are presented the future steps in research on outliers detection on railways WSN-based smart grid.

5.1 Outliers detection definition

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

5.2 Synthesis

Chapter 6

Conclusion

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Bibliography

Abid, A., A. Kachouri, and A. Mahfoudhi (2016, mar). Anomaly detection through outlier and neighborhood data in wireless sensor networks. In *2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*. Institute of Electrical and Electronics Engineers (IEEE).

Birol, F. and J.-P. Loubinoux (2016). 2016 edition of the uic-iea railway handbook on energy consumption and co2 emissions focuses on sustainability targets. Technical report, IEA - International Energy Agency; UIC - International Union of Railways.

Branch, J., B. Szymanski, C. Giannella, R. Wolff, and H. Kargupta (2006). In-network outlier detection in wireless sensor networks. In *26th IEEE International Conference on Distributed Computing Systems (ICDCS06)*. Institute of Electrical and Electronics Engineers (IEEE).

Chandola, V., A. Banerjee, and V. Kumar (2009, jul). Anomaly detection. *ACM Computing Surveys 41*(3), 1–58.

Ghorbel, O., W. Ayedi, H. Snoussi, and M. Abid (2015, jun). Fast and efficient outlier detection method in wireless sensor networks. *IEEE Sensors Journal 15*(6), 3403–3411.

Martins, H., L. Palma, A. Cardoso, and P. Gil (2015, may). A support vector machine based technique for online detection of outliers in transient time series. In *2015 10th Asian Control Conference (ASCC)*. Institute of Electrical and Electronics Engineers (IEEE).

Rajasegarar, S., C. Leckie, M. Palaniswami, and J. C. Bezdek (2007, jun). Quarter sphere based distributed anomaly detection in wireless sensor networks. In *2007 IEEE International Conference on Communications*. Institute of Electrical and Electronics Engineers (IEEE).

Sheng, B., Q. Li, W. Mao, and W. Jin (2007). Outlier detection in sensor networks. In *Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing - MobiHoc 07*. Association for Computing Machinery (ACM).

Shift2Rail Joint Undertaking (2015). Shift2rail joint undertaking multi-annual action plan. Technical report, Shift2Rail.

Zhang, Y., N. Meratnia, and P. Havinga (2010). Outlier detection techniques for wireless sensor networks: A survey. *IEEE Communications Surveys & Tutorials 12*(2), 159–170.

Zhuang, Y. and L. Chen (2006). In-network outlier cleaning for data collection in sensor networks. In *In CleanDB, Workshop in VLDB 2006*, pp. 41–48. APPENDIX.