Train Localisation using Wireless Sensor Networks

Adeel Javed

a thesis submitted for the degree of

Doctor of Philosophy

at the University of Otago, Dunedin,

New Zealand.

31 March 2016

Abstract

Safety and reliability have always been concerns for railway transportation. Knowing the exact location of a train enables the railway system to react to an unusual situation for the safety of human lives and properties. Generally, the accuracy of localisation systems is related with their deployment and maintenance costs, which can be on the order of millions of dollars a year.

Despite a lot of research efforts, existing localisation systems based on different technologies are still limited because most of them either require expensive infrastructure (ultrasound and laser), have high database maintenance, computational costs or accumulate errors (vision), offer limited coverage (GPS-dark regions, Wi-Fi, RFID) or provide low accuracy (audible sound). On the other hand, wireless sensor networks (WSNs) offer the potential for a cheap, reliable and accurate solutions for the train localisation system. This thesis proposes a WSN-based train localisation system, in which train location is estimated based on the information gathered through the communication between the anchor sensors deployed along the track and the gateway sensor installed on the train, such as anchor sensors' geographic coordinates and the Received Signal Strength Indicator (RSSI).

In the proposed system, timely anchor-gateway communication implies accurate localisation. How to guarantee effective communication between anchor sensors along the track and the gateway sensor on the train is a challenging problem for WSN-based train localisation. I propose a beacon-driven sensors wake-up scheme (BWS) to address this problem. BWS allows each anchor sensor to run an asynchronous duty-cycling protocol to conserve energy and establishes an upper bound on the sleep time in one duty cycle to guarantee their timely wake-up once a train approaches. Simulation results show that the BWS scheme can timely wake up the anchor sensors at a very low energy consumption cost.

To design an accurate scheme for train localisation, I conducted on-site experiments in an open field, a railway station and a tunnel, and the re-

sults show that RSSI can be used as an estimator for train localisation and its applicability increases with the incorporation of another type of data such as location information of anchor sensors. By combining the advantages of RSSI-based distance estimation and Particle Filtering techniques, I designed a Particle-Filter-based train localisation scheme and propose a novel Weighted RSSI Likelihood Function (WRLF) for particle update. The proposed localisation scheme is evaluated through extensive simulations using the data obtained from the on-site measurements. Simulation results demonstrate that the proposed scheme can achieve significant accuracy, where average localisation error stays under 30 cm at the train speed of 40 m/s, 40% anchor sensors failure rate and sparse deployment. In addition, the proposed train localisation scheme is robust to changes in train speed, the deployment density and reliability of anchor sensors.

Anchor sensors are prone to hardware and software deterioration such as battery outage and dislocation. Therefore, in order to reduce the negative impacts of these problems, I designed a novel Consensus-based Anchor sensor Management Scheme (CAMS), in which each anchor sensor performs a self-diagnostics and reports the detected faults in the neighbourhood. CAMS can assist the gateway sensor to exclude the input from the faulty anchor sensors. In CAMS, anchor sensors update each other about their opinions on other neighbours and develops consensus to mark faulty sensors. In addition, CAMS also reports the system information such as signal path loss ratio and allows anchor sensors to re-calibrate and verify their geographic coordinates. CAMS is evaluated through extensive simulations based on real data collected from field experiments. This evaluation also incorporated the simulated node failure model in simulations.

Though there are no existing WSN-based train localisation systems available to directly compare our results with, the proposed schemes are evaluated with real datasets, theoretical models and existing work wherever it was possible. Overall, the WSN-based train localisation system enables the use of RSSI, with combination of location coordinates of anchor sensors, as location estimator. Due to low cost of sensor devices, the cost of overall system remains low. Further, with duty-cycling operation, energy of the sensor nodes and system is conserved.

Acknowledgements

I am grateful to GOD for making this possible for me. Here, I want to thank many people who helped me during my journey of Ph.D.

First of all, I would like to express the deepest gratitude for all favours, support and guidance received from Associate Professor Zhiyi Huang. I am grateful to him for accepting me as a Ph.D. student. I would also like to thank Associate Professor Jeremiah D. Deng who jointly supervised my thesis and helped me in my tough times. I am especially thankful to Dr. Haibo Zhang who joined my supervisory team and I still remember I was excited a lot when he agreed to supervise me. He was always willing to guide me during my planned and sudden meetings with him. I have learned a great deal from my supervisors about research and I deeply cherish the time I have spent working with them. They always had time to answer my silly questions and always gave me a constructive feedback.

I would like to thank members of my "Systems Research Group" lab who were like a family to me and I had fun time with them. I am thankful to all my colleagues and friends especially Jason Mair and Leila Eskandari. They were always with me during my rough and tough days and I used to have very informative discussions with them. I am grateful to my friends Uzair Javed, Aftab Ahmed, Ammar Ahmed and Faran Shahzad for helping me in conducting field experiments at far places. I also have great regard for the staff of Computer Science department for their support throughout my stay. I want to express my deep gratitude for University of Otago for providing me scholarship and admission. In this university, I learned most important lessons and experiences of my life.

Finally, I would like to thank my parents and brother for their supportive role throughout my Ph.D. They remained source of inspiration for me in my life. I will also like to thank my wife, Saiqa who has been always very understanding and provided her faithful support in writing this work. A special thanks to my loving daughter, Eshaal Javed, for always keeping my spirits high.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Railway is used as a means of transport for passengers and goods. Its use has significantly increased in the last decades because of its low travelling cost and reliability (EuroStat, 2016; theGuardian, 2016; NetworkRail, 2016). In general, several attributes are linked with railway transport such as safety, reliability, comfort and security. All of these concerns are associated with the navigation system of trains, which further relates to the determination of accurate location of each train. The accurate estimation of a train's location is vital in the automation of railway transport systems, as it triggers the signals to close or open the railway crossing gates, inform the track-side workers, accurately updates the railway schedules and helps to raise the safety standards of the trains. Although existing systems, involving humans' inputs, are performing these tasks every day, incidents reports such as signal malfunctioning, coaches detachments, train parting, derailments, brakes failure and tracks' cracks, are quite high. The high rate of occurrence of such incidents is because of human errors and equipment's malfunctioning. In case of an incident, the location of train is required for rescue operations. Statistics show that the absence of accurate locations of trains minimise chances to mitigate emergency failure situations, which have claimed thousands of human lives and have damaged the infrastructure and properties (Amitabh, 2005). A lot of effort and funds have been invested in research to increase the safety of railway transport by improving the accuracy of train localisation systems.

Global Positioning System (GPS) is a satellite-based system designed to identify location of an object (Pace *et al.*, 1995). It can accurately identify the location within about 30 feet in any weather conditions (GPSHistory, 2016; Fales, 2003). Some other versions of GPS, such as differential GPS (DGPS) and RTK-GPS have improved the accuracy of several meters to sub-meter (Morales and Tsubouchi, 2007). GPS is designed as an outdoor positioning system and is being widely used for localisation in the transport systems, such as for railway, but it does not work well in several cases known as GPS dark regions, such as hilly terrains, tunnels, and forests. Moreover, as the GPS works on satellite communication, the accuracy of GPS also gets affected in extreme weather conditions and under ionospheric conditions (Grejner-Brzezinska *et al.*, 2007). High sensitivity GPS can penetrate into infrastructures but does not cover the GPS dark regions, making it unreliable. Localisation of trains or other means of transport in an outdoor environment without GPS is still an open and ongoing research problem (Chu and Jan, 2007; Chen *et al.*, 2013). Some technologies such as wireless local area network (WLAN), GSM, Inertial Sensors, or laser-based approaches provide very precise solutions for localisation in outdoors but the associated costs of such systems are high. Wireless Sensor Networks (WSNs) have proven to be a good low cost and reliable alternative for indoor as well as outdoor localisation in GPS-less scenarios (Stoleru *et al.*, 2005; Shen *et al.*, 2005; Zhang *et al.*, 2009; Constandache *et al.*, 2009), where sensor devices report different pieces of information such as signal strength, geographic coordinates, motion parameters or other phenomena of interest from the field of deployment. The analysis of the reported information is used to extract the information about the corresponding location. The basic idea is to extract meaningful information from the collected data from individual sensor devices and process it with some noise filtration technique to identify the current location of the vehicle such as a car or a train.

This thesis focuses on the problem of WSN-based train localisation systems. There is investigation of using received signal strength (RSS) as measurement model in proposed system. Investigation is an in-depth analysis of feasibility of using RSS measurements for localisation and understanding the degree of deviation of distance estimation from RSS values. RSS measurements are prone to environmental factors, therefore, investigation is followed by methods to improve RSS measurements by fusion of another data model such as location information of sensor motes. There are two main problems with most of the existing localisation solutions based on wireless sensor networks: (1) feasibility of using RSS in outdoor and harsh environments are not studied on a large scale, and (2) emphasis is placed on the combination of different techniques and technologies for localisation in outdoor environments, usually non-range-based methods–angle of arrival, time of arrival or time difference of arrival (discussed in Chapter 2), which increases the complexity of a system. Therefore, the existing approach cannot be adopted in the train localisation system. In the proposed WSN-based train localisation system, a range-based solution is proposed to address several research problems. The algorithms which together make the WSN-based localisation system include, (1) Beacon-driven sensor Wake-up Scheme (BWS), (2) Particle-Filtering-based train localisation scheme, and (3) Consensus-based sensor Management Scheme (CAMS). Initially, the feasibility of using RSS for train localisation is studied along with its usage with a distance estimation model and the log-normal path loss model. I have tested the system in a simulation environment by using real-world data collected from railway representative environments such as an open field environment, a railway station site and from within a tunnel. The obtained average localisation error is less than 30 cm (in one scheme) and manages to reduce it further to less than 13 cm (in an improved scheme) in all cases with the configuration of the sensor platform, train speed, sensor failure probability and deployment density. Experimental results show that RSS alone is not good, but the combination of RSS, location data, and Particle Filter is good for distance estimation in challenging outdoor railway environments. Moreover, the use of WSN in railway transport is a cost effective way to increase the safety of the system, navigation, location-based marketing and other services. A cost analysis of WSN with contemporary technologies is given in next chapter.

1.1 Motivation

The increase in the automation of railway transport systems has raised focus on the safety concerns (uic, 2016). With the traditional and semi-automated railway transport system, several unfortunate incidents in the history of railway transportation have claimed the loss of human lives, and public and private property. The cracks on the railway tracks, stress health of bridges or potential derailments can be detected through special laser sensors, mounted on the special purpose trains. In case of such incidents, the known location of a train can help to avoid or minimise damage (TheRegister, 2014). For example, if the position of a train is known, it can be signalled to stop before it reaches to broken track that may get damaged by a human threat such as a bomb blast (dailytimes, 2016). Studies have shown that a trained human driver reacts in a better way to unusual circumstances for the safety of passengers and assets. Therefore, there is an increasing drive to develop a safe automated railway transportation system. An automated transport system relies on the accuracy of its navigation system, which depends on the accuracy of the localisation system. The following are the motivation for a train localisation system:

- The unavailability or failure of an existing localisation system may lose control over the automated train system and increases the risk of the loss of human lives and assets worth million of dollars. An automated train can lose contact with control room (Railway-Technology, 2016). Here, trains can be programmed to stop if they lose contact with control room. However, the problem with such techniques is that, it may collide with another train following it or it may has to stop several times due to flaw of communication. Therefore, this sounds like an inefficient solution.

- There is an associated technology deployment cost and low-cost systems are always desirable that can also achieve high localisation accuracy. Sensor motes are low in price and become an attractive option to use in localisation schemes.

- The increase in the localisation accuracy is also associated with the power consumption and availability, which is not easy to provide on the long railway tracks in remote areas. Sensor motes can operate on low power listening mode and can perform duty cycling.

- Over the lifetime of a system deployment, the total cost of maintenance and upgrading exceeds the initial deployment cost.

The train localisation system based on any technology should at least meet some requirements: (1) cost should be low, which includes the costs of infrastructure components, locating devices, and installation; (2) location accuracy should be high, which means that the average error between true and estimated location should be minimum, the accuracy standard depends on the technology used. A list of accuracy standards is given in Table 2.1 (Song *et al.*, 2011; Gu *et al.*, 2009; Al Nuaimi and Kamel, 2011; Liu *et al.*, 2007; Khan, 2014); and (3) it should be easy to configure, to use, and provide full coverage on the target area. The target of this thesis is to develop a train localisation system and use it for safety, navigation and location-based services in GPS dark regions of harsh railway environments. Therefore, a cost-effective solution offering a good level of accuracy is required in order to be used in harsh railway environments.

Localisation systems based on infrared light (Want *et al.*, 1992), ultrasound (King *et al.*, 2006), WLAN (Chintalapudi *et al.*, 2010; Cavalieri, 2007) and Active-RFID (Huang *et al.*, 2006) provide good positioning accuracy but the cost is high. RFID devices have small communication range, therefore, may need large number of these devices to cover the deployment area. Further, their fragile nature will require replacements sooner and incur huge maintenance cost. Technologies other than wireless sensors, inertial sensors, vision and WLAN are quite expensive (more details follow in Chapter 2). WLAN solutions are cheaper for small scale localisation such as on a railway station. On the other hand, inertial solutions are cheap but are not reliable and provide low accuracy. However, some problems associated with WLAN and inertial sensors (discussed in Chapter 2) make WSN-based solutions preferable.

Unfortunately, GPS is the most commonly adopted technology for train localisation and in the absence of its signals in GPS dark regions, an automated railway system can lose its control. A similar incident was reported in London, where an automated train lost control for four miles due to the malfunctioning of a communication system. The operating company was unable to locate the train during that period and authorities were uncertain about the status of the train, that is, whether it is stopped or moving. London Underground was operating that train and company was lucky that no other railway traffic was scheduled on that route during that time (Railway-Technology, 2016). In such scenarios, if the location of a train is known, other trains on the same track can be signalled to stop. Such incidents are undesirable while the focus is being shifting towards automation of trains. Generally, to increase the accuracy of location estimation, hardware-based approaches are used, where current, inefficient hardware is replaced with newer, more accurate alternatives. Although almost every technology and system needs upgrades, upgrading a low-cost system implies less cost than upgrading a high-cost system. However, hardware upgrades may contribute a significant amount of investment and waste of previously developed solutions, making it an unsustainable solution. Though it depends on cost-benefit tradeoff, but generally, significant changes in a system are undesirable. Ideally, a train localisation system should incorporate several sub-systems based on different technologies. Each localisation sub-system operates independently and by their fusion, accuracy is improved. Further, in case of failure of one sub-system, another system is there to mitigate the challenge. Therefore, in the absence of GPS, there is a need for a train localisation system that has low installation, maintenance and upgrade costs along with a significant level of accuracy. WSN offers not only a cost-effective solution but also a reasonable positioning accuracy. Therefore, the main motivation of this research work is to:

> *produce a train localisation system based solely on WSN, which uses RSS measurements for location estimation in harsh railway environment particularly in the absence of other technologies such as GPS signals.*

Along with train localisation for railway safety, this thesis also addresses issues below:

- *Energy Saving* is a key concern for sensors that are deployed along the track and there is no infrastructure to power them up. In such a case, the common way for power supply is batteries. As the train's schedule is unknown, the idle listening for a long time to detect the incoming train will drain all battery power. Generally, duty-cycling is considered the solution to address the unnecessary idle listening problem. However, duty-cycling can be a compromise on the sensors being woken up while the train passes. Therefore, in order to minimise the energy consumption to prolong the battery and network life, there is a tradeoff for performance that needs to be dealt with carefully.

- *Sensor Management* is another motive to carry out this thesis. In particular, the sensors deployed in the remote areas with infrequent physical access are prone to be influenced by environmental effects. A lot of resources are required to sort out the faults in those sensor devices such as expensive human resource, rising fuel/travelling cost, and training cost. To mitigate management issues, if sensors can take care of their neighbour sensors and report it to the train, it can significantly reduce the cost and help to ignore the negative input of faulty sensors in the localisation process. Later on, trained staff members can sort out the reported faults with minimum effort and without diagnostic costs. The proposed management system works better in networks with dense sensor deployments.

- *Other benefits* of WSNs make it prominent choice from other technologies such as RFIDs. Sensors of different types are capable of detecting railway track faults and estimating the life of bridges. Such timely reports can help to save human lives by avoiding disastrous situation. For example, strain gauge sensors can be used to measure the stress on a bridge while a train passes (Bischoff *et al.*, 2009), or laser sensors can be used to detect cracks on tracks (Aboelela *et al.*, 2006; Ramesh and Gobinathan, 2012) and that can help to minimise the chances of accidents. Different type of sensor network is capable of providing such additional benefits of detecting these faults (Flammini *et al.*, 2010) along with identifying train's location.

1.2 Challenges

Most of the successful RSS-based localisation work has been done in indoor environments (Yang and Chen, 2009; Güvenc, 2003; Kaemarungsi and Krishnamurthy, 2004;

Feng *et al.*, 2012; Chen *et al.*, 2013; Feng *et al.*, 2010). Comparatively, less work has been done for train localisation and the existing pieces of work mostly rely on GPS or satellite systems for train localisation (Fararooy *et al.*, 1996; Fraile, 1999; Däubler *et al.*, 2003). It is quite challenging to develop a framework in which a train localisation system uses RSS measurements of the WSN. Amongst the challenges faced were these:

- To verify the feasibility of using RSS for localisation in harsh indoor environments, such as tunnels, several studies have been conducted (Xu *et al.*, 2013; Savic *et al.*, 2013; Chang *et al.*, 2011; Wang and Du, 2010), but not many pieces of work are available for railway environments. In the harsh outdoor environments there are several factors that affect the ability of RSS to estimate the distance of a transmitter. A railway environment is an example of a harsh environment that has interfering factors such metals, overlapping frequency signals, and weather impacts. To the best of my knowledge, there is no work done to verify the use of WSN-based RSS in dynamic railway environments. However, studies have been conducted to investigate the usage of RSS as estimator for location estimation in a few matching environments such as coal mines (Savic *et al.*, 2013), tunnels (Xu *et al.*, 2013) and for train integrity in railway environment (Scholten *et al.*, 2009). Therefore, it is quite challenging to analyse the feasibility of RSS in GPS dark regions such as open field, railway station and tunnel.

- The use of real datasets are more convincing, in the simulations and to evaluate the performance of proposed algorithms. To conduct the experiments on real railway systems is a challenging task and it involves the recording of datasets over short to long distances with several sensor deployment densities and types of sensor devices. In addition, these sorts of experiments require human resource to carry out several tasks with exact distance measurements between the sensor devices and angle of sensor antennas. Therefore, the experimentation people require training for the collection of credible datasets. Moreover, it is quite challenging to conduct experiments in such an environment that has minimum interference from other overlapping frequencies such as microwave.

- The collection of data itself is challenging in such harsh railway environments. The RSS from sender to receiver and vice versa may differ because of signal reflections from surrounding infrastructure. Moreover, a difference needs to be maintained in the transmitting sensors that reply to a transmission, and multiple transmissions of a single sensor in reply to a transmission from a sensor on the train. Furthermore, time of transmissions from the sender and receiver are also required to calculate the delay.

- The sensor devices are generally powered by batteries, which are hard to recharge in the remote rough environments. Therefore, to prolong the network lifetime, sensors operate on duty-cycling by turning their transceivers on and off frequently. On one hand, the duty-cycling operating patterns make it challenging to get sensors active at the time of the train passing, and there is need to guarantee the availability of sensors for communication with the train for train localisation. On other hand, the clock synchronisation is also an uphill task in large networks (Lin *et al.*, 2008). Therefore, it is a quite challenging task to design a suitable scheme that can guarantee the wake-up of sensors along with low energy consumption.

- The sensor devices deployed in the remote areas are prone to be affected such as weather extremities and theft. Manually, sorting out the faulty sensors along the track on the remote sites is a challenging task and incur a huge cost. Therefore, a scheme is required that can enable sensor devices to perform diagnostics in their neighbourhood and report the detected faulty sensors.

1.3 Contributions

Train localisation is often performed with GPS or other expensive infrastructure-based technologies. It becomes a challenge to perform train localisation with low-cost sensor devices to cover the GPS dark regions. Therefore, the performance of using RSS collected from WSNs for train localisation system needs to be carefully analysed in real-world railway environments. The literature shows that not much work is done on localisation in large-scale railway environments using RSS measurements from WSN. In this thesis, firstly, the feasibility of using RSS measurements for WSN-based train localisation is verified. Field experiments are conducted to collect RSSI measurements in railway environments followed by the detailed analysis on the collected datasets. After a careful analysis, it is observed that though RSS measurements are noisy, noise filters can be used to achieve a reasonable level of accurate distance estimation. Therefore, this thesis proposes a Particle Filtering based robust train localisation algorithm, which is a core component of a WSN-based train localisation system. In addition, a beacon-driven wake-up scheme is developed, which can guarantee to wake up sensor devices when the train is arriving without global knowledge of the train's schedule.

Finally, a sensor management scheme is proposed to report the faults in the sensors deployed along the track, and that helps to reduce the maintenance cost. The main contributions of this thesis are the following:

- A Beacon-driven sensor Wake-up Scheme (BWS) is proposed for train localisation. The BWS allows sensors to sleep for a maximum time within an upper bound and still guarantees the wake-up of sensors at the time of a train's arrival. BWS is analysed theoretically and through simulations and is found to be energy efficient (Javed *et al.*, 2014).

- The performance of sensor nodes is analysed based on BWS's ability to wake up, while operating on duty-cycling in the train localisation scenario (Javed *et al.*, 2013).

- An algorithm is developed that uses Particle Filtering techniques for train localisation. In the proposed algorithm, real-world RSS measurements are used to compute the location of the train.

- In the designed algorithm, a weighted RSSI-based likelihood function (WRLF) is developed which uses RSS measurements and geographic coordinates, transmitted by sensor nodes on trackside to the gateway sensor on the train. The WRLF estimates the likeliness of particles to represent the train's position.

- An algorithm to manage the sensors is developed, Consensus-based Anchor node Management Scheme (CAMS), to detect and report the faults and faulty sensor nodes in the network, which otherwise can be an expensive task to do manually. CAMS also assists the train localisation system by computing consensus-based path loss ratios to increase the accuracy of location estimation (Javed *et al.*, 2015).

1.4 Limit of Scope

As a whole, the localisation system presented in this thesis takes RSS measurements from the sensor nodes deployed along the track and the geographic coordinates received from those sensor nodes to identify the current location of the train. However, there are several factors that can affect the performance of the proposed system. It is not possible to address all challenges in a single PhD project. The limitations of this work are these:

1. Though the proposed train localisation system is based on the real-world data collected from the field experiments, I could not implement the solution prototype in the railway network. The prototype implementation requires permission from railway authorities and adjustment in the train schedules, which requires involvement of railway's top management and engineers; therefore, it needs a lot of time to go through this process. Given permission, the implementation of proposed scheme can help to test its performance and to highlight its shortcomings.

2. A set of homogeneous sensor devices were used for data collection in each experiment setup. However, a heterogeneous sensor network can be implemented easily to study the impact of different types of sensor devices, which is not the focus of this work.

3. The purpose of this work is to develop a WSN-based train localisation system that can be used in data fusion with train localisation systems based on other technologies such as RFID, WLAN and GPS (discussed in Chapter 2). However, this work's focus is on the design of a WSN-based train localisation system only and technology fusion is not discussed to combine several technology-based train localisation systems.

4. As the focus of this work is determining the current location of a train, I have not addressed the security issues in the communication between sensors. The security issues should be addressed at the infrastructure level.

5. There is no other available WSN-based train localisation system to compare with, though I compared the simulation-based performance of components of the system with the theoretical models, wherever possible.

1.5 Thesis Layout

This thesis describes the algorithms proposed for the WSN-based train localisation system. Experiments and simulations are carried out to analyse and evaluate the algorithms. The thesis consists of eight chapters and details are as follows:

- **Chapter 2** describes the metrics of a localisation system and then reviews the existing localisation technologies, their pros and cons, and localisation projects based on those techniques. This chapter presents an analysis of all these techniques in terms of cost, accuracy and adaptability. It also presents the wireless sensor-based localisation methods and highlights the research objective of train localisation by using received signal strength information from communication within entities of a wireless sensor network. At the end, existing projects about train localisation systems and their incorporated technologies are discussed in detail.

- **Chapter 3** shows the overview of the proposed components of WSN-based train localisation. In addition, the brief details about each component are discussed.

- **Chapter 4** verifies the feasibility of using RSS measurements for WSN-based train localisation. An analysis is performed on the real-world data collected from field experiments such as in an open field, railway station and tunnel. The analysis shows that though RSS measurements are noisy but still follow the model curve and with the use of some noise filtration algorithms, RSS can still be used for distance estimation.

- **Chapter 5** presents a beacon-driven sensor wake-up scheme (BWS) that enables sensors to wake up and schedule their communication with the gateway sensor at its arrival. BWS allows sensors to consume minimum energy by operating on asynchronous duty-cycling without global knowledge of train's arrival and still wake up in time for communication in an energy efficient way.

- **Chapter 6** proposes a Particle-Filter-based train localisation algorithm that uses noisy RSSI measurements and the geographic coordinates of sensors to develop the weighted RSSI likelihood function for the estimation of train's location in a recursive Bayesian way. The algorithm selects the particles and assigns weight to each particle based on its likelihood to represent the location of the train. Consequently, the location of a train with minimum error is estimated by averaging all the locations of particles with respect to their weights.

- **Chapter 7** presents the sensor management scheme, which enables sensors to develop consensus about the existing faults and faulty nodes in their neighbourhood and report them to the gateway sensor. Moreover, the proposed scheme also allows sensors to assist the gateway sensor in increasing the train localisation accuracy by estimating the path loss ratio.

- **Chapter 8** concludes with final remarks on the solutions provided by the proposed schemes for WSN-based train localisation system and includes suggestions for possible future research work.

1.6 List of Publications

- *A Javed, Z Huang, H Zhang, JD Deng, "CAMS: Consensus-based Anchor node Management Scheme for train localisation".* In the Proceedings of SPRINGER International Conference on Adhoc Networks and Wireless (ADHOC-NOW), 2015, Springer. This paper presents the sensor management scheme based on the consensus for train localisation using wireless sensor networks.

- *A Javed, H Zhang, Z Huang, JD Deng, "BWS: Beacon-driven wake-up scheme for train localisation using wireless sensor networks".* In the Proceedings of IEEE International Conference on Communications (ICC), 2014. This paper presents the sensor wake-up scheme using beacons for train localisation.

- *JVN Vijayakumar, H Zhang, Zhuang, A Javed, "A Particle Filter Based Train Localisation Scheme Using Wireless Sensor Networks".* The 11th IEEE International Conference on Embedded Computing (EmbeddedCom), 2013. This paper presents a Particle Filtering based train localisation algorithm using wireless sensor networks. (Best Paper Award)

- *A Javed, H Zhang, Z Huang, "Performance analysis of duty-cycling wireless sensor network for train localisation".* In the Proceedings of ACM Workshop on Machine Learning for Sensory Data Analysis (MLSDA), 2013. This paper evaluates the tradeoff between energy consumption and availability of sensors for communication with the train for train localisation while they operate on duty-cycles.

- **Note:** Part of text and results from these publications are included in thesis' chapters and copyrights' permissions allow authors to use this material in their thesis.

Chapter 2

Background

In this chapter, I begin by identifying the requirements of a train navigation system, followed by metrics to define a localisation system. I then review several localisation systems based on different technologies and analyse their usability in indoor and outdoor environments, and present their drawbacks. The details about WSN-based localisation methods are discussed in detail. After that I present my research goal and then discuss the existing train localisation systems.

2.1 Features of a Train Localisation System

Location of an object is a core component of any positioning system, navigation systems and localised services management systems. Based on the computed location of the train, a train localisation system should offer several features to its users, who are railway staff and passengers.

Railway staff can benefit from features of a train localisation system, such as obstacle detection, an alert at railway crossing, information about trackside workers, trains' schedules and emergency messages, to improve the operation of a train localisation system.

A train localisation system, which can offer a broad range of features as discussed above, eventually results in raising the standard of comfort and safety of passengers.

2.2 Metrics for Localisation Systems

Generally, a localisation system is evaluated on the basis of the level of accuracy it achieves. However, an extremely accurate localisation system might not be feasible to implement because of several reasons such as its cost, availability of trained staff and required error tolerance level of a system. Therefore, accuracy can be coupled with cost of a system to evaluate a localisation system (Song *et al.*, 2011; Gu *et al.*, 2009; Al Nuaimi and Kamel, 2011; Liu *et al.*, 2007). Khan (2014) has given a three-dimensional aspect of metrics such as accuracy, cost, and deployment complexity. The later two can be combined as a cost incurred because of investment on infrastructure and computation.

Cost of a localisation system is an essential metric to gauge the adoptability of a localisation system. It includes cost of infrastructure equipment, installation of infrastructure, and maintenance. The cost of a localisation system also depends on the types of equipment, such as satellite-based systems that can be used for specific services (paid/free) to compute the location of an object, whilst several systems require to develop the system and its services from the scratch. Cost heavily depends on the maintenance of a system. A new and complex technology-based localisation system requires highly skilled staff who may not be available or who are hard to train hence increasing the cost of a system as compared with a localisation system that is based on a commonly used technology and techniques. Maintenance cost also includes the scalability of the adopted technology. A system that can easily be upgraded with quick deployment is highly desirable as it saves time and funds. Otherwise, the opposite case can raise the cost of a system.

The metric of accuracy refers to the difference between the estimated location of an object and its actual location. If the average error between the estimated and actual locations of an object is close to zero, a system is considered to be more accurate. Generally, accuracy has a tradeoff with the cost of a system (Stoleru *et al.*, 2005). A highly accurate system is usually less cost efficient. The accuracy of a less accurate system can be increased by the addition of extra hardware or complex techniques but it increases the cost of the whole system.

2.3 Localisation Systems

In this section, several technologies are discussed along with the localisation system based on those technologies. Further, existing research is referred in each of the technologies. At the end, features of the existing techniques will be summarised in a table including the feasibility of usage in the closed space (indoor) environment, open space (outdoor) environment or both, cost of each technology, and their respective accuracy.

2.3.1 GPS-based

GPS is the most widely used technology in the development of a localisation system. A GPS-enabled device measures its location in accordance with satellites, which act as reference points. Initially, a GPS device calculates its distance from a satellite in the form of a sphere. The possible positions of the device are narrowed down by estimating device's distance from another satellite (Brain and Harris, 2011). Distance computation with at least three satellites can help to find out 2D position, known as triangulation. With the further increase in the number of reference points (satellites), such as four or more, a 3D position can be computed. GPS is helpful in estimating an object's position with certain accuracy. The average distance estimation error of standard GPS ranges from 3 m to 15 m. However, its accuracy is compromised in the indoor environments and in the GPS dark regions because of unavailability of satellite signals, thus making it unreliable in those cases.

Differential GPS (DGPS) is another variation of GPS that applies differential correction techniques to basic GPS. In DGPS, a stationary reference receiver with known position is added to the system to correct the timing of the mobile receiver (Stewart and Rizos, 2002). Such addition with known location helps to correct the timing information of receiver with unknown location. DGPS has two versions of implementation. The first DGPS implementation is easy to implement and in this DGPS, the location coordinates get corrected continuously. These corrected coordinates are then sent from reference station to the mobile receiver. On the other hand, the second implementation version corrects the ranges instead of coordinates. The corrected ranges are then used for computation of mobile receiver's positions. The second version is more suitable for real-time applications. DGPS is claimed to be more accurate than basic GPS and its measurement error stays within a few centimetres. DGPS commercial services are costly.

Pseudolites are devices which bridge gap in the scenarios where a few satellites are not available due to any reason. It transmits GPS-like signals (Drira, 2006). Pseudolites enable receivers to compute its location in the presence of minimum satellites. It has high accuracy with error as low as 1 cm (Bradford *et al.*, 1996).

Wide area DGPS (WADGPS) is another type of DGPS, in which an error vector is computed for each satellite. In WADGPS, there are several monitor stations and a master station (Drira, 2006). The GPS receivers in the monitor stations help to capture measurements, which are then transmitted to the master station. Master station, then computes GPS error which is used to correct the position of users. WADGPS errors range from 1 m to 8 m.

Another GPS is the wide area augmentation system (WAAS), developed by the aviation department of the USA. WAAS has raised its accuracy by deploying 25 reference stations and two master stations. WAAS is capable of dealing with ionospheric problems (Drira, 2006). WAAS has an average accuracy with error range from 1 m to 3 m.

Generally, GPS signals are not available in indoor railway environments, such as underground trains, tunnels and some railway stations in regions with rough terrain. To overcome this issue, GPS technology can be combined with other technologies such as a GSM network. Such an integration can compensate for the unavailability of GPS signals. Wireless Assisted GPS (AGPS) is the variation of GPS that works with a GSM network to perform the localisation tasks. AGPS computes the location of the device by receiving signals from GPS satellites and cellular network, and computation is performed at location servers (Giaglis *et al.*, 2003). AGPS can overcome the loss of GPS in indoor railway environments but at the cost of energy resources of mobile devices.

GPS is also being widely used in railway positioning systems (Burns *et al.*, 1992; Lemelson and Pedersen, 1999). Though it has reasonable accuracy and high cost for commercial services, it also has some shortcomings. The limitations of this technology are these:

- A-GPS can provide accurate localisation. It can reduce the GPS dark regions by penetrating into the walls, but in a tunnel railway environment, it still struggles to penetrate. A-GPS also needs clear sky for communication with the satellites which decreases the localisation accuracy in tunnels and hilly terrains.

- On one hand, commercial DGPS services are highly accurate but on the other hand, these services are quite expensive.

2.3.2 Infrared (IR)

An IR-based localisation system works on the principle of measuring the distance between the IR transmitter and receiver on the basis of delay in signal reception. A line of sight is required for such communication. AT&T developed one of first commercial IR-based localisation system, called Active Badge (Harter *et al.*, 2002; Want *et al.*, 1992). Active Badge was designed for indoor environments in which several sensors are fixed at known locations. An Active Badge, attached to the device with an unknown location, transmits frequent signals to sensors. Received IR signals' information is then forwarded to compute the location of an active badge/object. Extensive cabling to connect the sensor devices raises the cost of the system, which otherwise is a low-cost system. Because of IR's short range and wired connectivity, active badge is not suitable for outdoor environments and no longer available as a commercial product.

Cybernet System Corporation developed another IR-based positioning system, Firefly (Interactive, 2010). The Firefly system comprises a tag controller, several IR emitting tags and a camera array to track the 3D motion of a person. A person carries a tag controller and tags are mounted on several parts of his body. A camera array comprises three cameras are attached on a 1m bar that receives the IR signals. The Firefly system claims high accuracy and its position tracking is real-time performance. Though the Firefly system offers accuracy up to 3mm, it is quite expensive. Further, several small devices are not comfortable to wear on a human body and raises concerns about its adaptability.

States and Pappas (2006) proposed another IR-based localisation system, OPTOTRAK, for small businesses and workplaces. The proposed system uses an array of three cameras to compute the location of an object. The cameras receive IR light signals from the markers of the object and perform triangulation to determine the location of transmitter. This system is highly accurate and offers accuracy in millimetres. However, it requires line of sight between markers and cameras to perform effectively. The increase in the number of IR markers relaxes the strict requirement of line of sight communication but increases the cost of system. Such constraints make it less desirable in outdoor environments.

IR-based solutions are applicable in outdoor environments but with some limitations. In one piece of work, researchers have introduced a scheme that effectively detects cracks on the railway tracks (Kishor *et al.*, 2012). These systems are low-cost and use IR transmitters and receivers to detect cracks. The returned IR signal enables system to identify the difference between normal and abnormal track. Navaraja (2014) proposed another system that uses IR sensors along with ultrasound technology for a track's crack detection. GSM networks are then used to communicate the identification of faults to the authorities for rectification.

IR-based systems provide high localisation accuracy, often in a few millimeters. In conclusion, IR based systems have the following limitations:

- IR signals are affected by interference from sunlight and fluorescent light, hence such systems may get affected in the outdoors. This issue can be addressed with the use of optical filters and noise filters, but that will increase the cost of whole system.

- The IR-based localisation system is an attractive system due to its low-cost IR emitters but its cost increases with the use of an array of camera devices, sensors and wired connectivity. Further, to overcome the negative environmental effects, use of filters makes this system complex.

- The IR signals are prone to environmental degradation, thus making it unreliable in outdoor environments.

2.3.3 Ultrasonic

Ultrasound signal can also be used to determine the position of an object. The idea is taken from bats' communication mechanism. Bats transmit ultrasonic signals and determine their distance from an obstacle from reflection of transmitted signals. Initially, ultrasonic signals were used in medical applications to create images of internal organs and to locate a specific one.

In 1999, AT&T researchers designed one of the initial ultrasound-based localisation system, Active Bat system. Objects used to carry Active Bat tags and receivers are mounted on a ceiling in the form of a grid at known locations. A bat controller requests to locate its position by emitting an ultrasonic pulse to all receivers. A reset signal from connected wire is also sent by the controller to synchronise receivers. Each receiver sensor measures its distance from the bat by computing the time period between reset signal to ultrasonic pulse. Noise filters are used to remove the errors caused from signal reflections.

Cricket is another ultrasound-based localisation system (Priyantha *et al.*, 2000). It uses a few ultrasound emitters, mounted on the infrastructure such as ceiling or walls, and receivers are installed on objects that needs to be located. On a localisation request, emitters emit IR signals and receiver locally locates its position using the triangulation method. Like Active Bat system, Cricket also uses RF signals to synchronise between the components. Similarly, it uses reflective distance through time-of-flight data to compute its location. Unlike Active Bat system, it does not require a grid of sensors at known locations as fewer sensors serve the purpose and makes it a low-cost system. However, this is a compromise on the coverage area. Misra *et al.* (2011) attempted to overcome range limitation of Cricket by designing an omnidirectional receivers and increased the coverage range by 20%. Ultrasound-based systems have been put into selective use in outdoor environments, particularly in a railway network.

Ultrasound can be useful in detection of faults and testing in railway system (Oukhellou *et al.*, 2008). Though fault detection relates to the railway maintenance system, it lays foundation for successful implementation and operation of railway management and localisation system. Fault investigation has always been an important issue in the railway industry (Fan *et al.*, 2007). Successful and timely fault detection relates to the safety of railway assets and human lives. Ultrasound-based systems can be used to detect several faults such as cracks, broken track segments, metallic corrosion or corrugation. Generally, a combination of several methods are used to perform a complete rail track inspection (Oukhellou *et al.*, 2008). In the ultrasound-based technique, probes are used to slide and remain in contact with train's head (Lanza di Scalea *et al.*, 2005; IEM-RM, 2003). Fluid is used to keep the contact between them smoothly. In that particular case, the testing vehicle moves with limited speed. This method successfully detects major surface defects but is unable to detect minor cracks. In another ultrasound-based technique, electromagnetic acoustic transducers are used (Cerniglia *et al.*, 2006). In this technique, devices are linked without physical contact. Such improvement increases the speed of the fault detection process. However, this technique lacks low-level detection of faults because of low sensitivity.

The use of an ultrasound technology for train localisation may only be feasible because of its limited transmission range. However, it can effectively be used in railway system for other purposes such as track health maintenance and monitoring. Ultrasound technology offers an inexpensive solution compared with Infrared positioning systems. However, the associated problems are as follows:

- Ultrasound signals have limited range. Several individual efforts have been made to overcome such limitation. One such way is to combine ultrasound signals with radio frequency signals. Such fusion of technologies increases the coverage range along with cost of the system.

- The ultrasound signals are affected by negative environmental factors. Further, ultrasound signals' penetration ability is also lower compared with RF signals. Such limitation reduces system accuracy.

2.3.4 Radio Frequency (RF)

Radio Frequency (RF) signals is another technology that can be used to develop a localisation system. RF signals are more reliable compared with infrared and ultrasound because of its long range and ability to penetrate several types of materials. Such property makes it a desirable technology to design a localisation system. However, RF signals are also prone to reflections from several types of smooth surfaces. An advantage of an RF-based localisation system is that its infrastructure is low-cost and can be reused.

Landmarc (Ni *et al.*, 2004) uses RFID tags to localise objects. It matches profiles of objects with the reference tags at known locations. Landmarc uses several reference tags and nine RFID readers with several power levels to transmit signals. It compares the signal strength of reference tag with signal strengths of all readers to localise a tag. The location of an RFID tag is computed by the weighted average method of k-nearest neighbours. Landmarc acheives an average error of about 1m. However, Landmarc's accuracy, vulnerable to tags' orientations, increases with the number of mobile objects. VIRE (Zhao *et al.*, 2007) improves the accuracy of the Landmarc system by using a proximity map that limits the number of comparisons to the local neighbouring tags only. Zhang *et al.* (2009) improves the accuracy of the Landmarc system by introducing noise models that help to compare signal strengths of more reliable neighbouring tags. The Landmarc system does not deal with latency. Another drawback of this system is that the life of a tag is not long due to its fragile miniature structure.

WhereNet system offers real-time localisation. In this system, tags are attached to the object that needs to be tracked. A few location antennas, attached to a ceiling at a known location, forward location requests from tags to location servers. Location servers use information of location antennas, tag requests (signal strength information) and processing methods, such as triangulation, to compute the location of several tags simultaneously. WhereNet achieves accuracy in meters and can work in indoor and outdoor environments (WhereNet, 2008). However, installation of location antennas at multiple locations increases accuracy and system cost.

An RF-based system has been the choice of the railway industry for the outdoors in particular for railway industry. Railway operators place a strong focus on correct train location and management, which is directly related to safety of trains. It is undesirable to introduce major changes in the network. Therefore, with the evolution of telecommunication networks, it enables operators to use the combination of global systems for mobile communication (GSM) and GPS. In Portuguese railways, RF (from GSM) is used for track-train communication (Monte de Caparica, 2000). The communication is held between the regulator, installed on track segments, and drivers. In another piece of research (Santos *et al.*, 2005), authors proposed the combination of GSM and GPS system and claimed its better performance for secondary tracks where other networks are not available. Track-train communication is basic for train localisation and in the absence of either network, the other network tries to overcome the shortcomings. However, an RF from a GSM network has its own limitation.

Hofestadt (1995) proposed another approach and introduced the concept of a dedicated GSM network for railways. The proposed approach performs better for railway management because of minimised third party concerns about safety due to specific network according to railway system's requirement.

According to the best of our knowledge, RFID is not being used to localise trains. However, RF from RFID network is being used in the railway environment for selective purposes such as railway management and maintenance system (GAO-Inc., 2007; Char and Johns, 2006).

In RF-based localisation systems, RF readers can read many tags simultaneously with unique identity. Tags are small; therefore, the system is easy to implement. The associated problems with such systems are the following:

- RF-based localisation systems use proximity and absolute processing techniques and depend on many hardware parts in the deployment grid. It raises the cost for large-scale deployments.

- RF signal are subject to interference from other electromagnetic signals in the surroundings. Therefore, this makes it an unreliable technology.

- Size of tags makes it portable and handy but becomes vulnerable to physical damage.

- Use of GSM network may be alright for voice calls but might not support large data transfers, that may be required for communication in a localisation system, thus needing to acquire 4G or 5G services.

2.3.5 Wireless Local Area Network (WLAN)

WLAN is another technology that can be used to determine the position of an object. Generally, WLAN is available publicly at several places such as hospitals, train stations and airports. The reusable infrastructure of WLAN makes it a desirable low-cost localisation system. A WLAN-based localisation system is generally possible in indoor environments but in some cases it can be used in outdoor environments. A main component of this system is an access point (AP). Signal strength (RSS) is used to compute the location of a node. Another method to compute the location of a node is fingerprint. It stores the RSS footprint of several APs in a database and compares the received one to estimate the location.

RADAR is a WLAN-based localisation system developed by Microsoft (Pahl and Radar, 2000). RADAR uses existing infrastructure to reduce cost. It gathers signal strength measurements and performs triangulation to determine location of a computer node. The RADAR system achieves accuracy of about 2m.

Ekahau is another WLAN-based localisation system (Ekahau, 2008) that uses existing WLAN infrastructure. It observes the mobility of WLAN devices. It uses triangulation on the RSS measurements received at several APs. Ekahau offers 2D location estimation and can be useful for offering location-based services. Ekahau can achieve localisation accuracy up to 1m.

WLAN-based solutions are applicable in outdoor environments as well. Railway environment is a well-represented outdoor environment. Communication based train control (CBTC) is a system that provides solution for train and ground communication. It automates several processes that ensure the railway safety. Zhu *et al.* (2010) proposes a WLAN-based system that improves the availability of network to guarantee high level of availability for train-track communication in a CBTC system. An analysis is performed using Continuous Markov Chain Model on the availability of WLAN-based solution. Siemens (Lardennois, 2003) and Alcatel (Kuun and Richard, 2004) have proposed WLAN-based CBTC systems that are implemented on the New York City Canarsie Line and Las Vegas Monorail, respectively. Though WLAN-based solutions are used for specific purposes in the railway industry, such as train-ground communication, they can be used to locate the position of a train in the absence of GPS. The availability of a power source is still an issue for the devices on the track. Overall, these solutions will increase the cost of the whole system because of infrastructure installation due to the short transmission range of devices. Using devices with short transmission ranges implies more infrastructure installation and maintenance cost. On the other hand, WLAN-based localisation solutions are mostly feasible in indoor environments and provide low-cost solutions because of reusable infrastructure. The associated problems with this technology are the following:

- In outdoor environments, cost is the major concern because of unavailability of infrastructure in rough railway environments.

- In indoor environments, there are several sources of RF signals that can cause interference, hence it is a compromise on the accuracy of the WLAN-based localisation system.

- In the fingerprinting location estimation method, an offline location database is built, called calibration. The collected database is then used for online mapping and locating an object. The calibration step is expensive, requires extensive time, and needs frequent update in dynamic environment.

- WLAN signals share some common channels with microwaves and that can become a compromise in the performance of WLAN-based localisation systems.

2.3.6 Audible Sound

Audible sound is another technology that can be used to develop a localisation system. Generally, devices these days are capable of producing audible sound and can participate in the design of such a system. BEEP is an audible sound-based localisation system that offers a low-cost solution (Mandal *et al.*, 2005). In BEEP, microphone receivers are installed at several known locations and received sound is forwarded and processed at location servers. Location is computed by using the triangulation method on time of arrival data. BEEP claims an accuracy in sub-meters. To increase the performance of BEEP in the indoors, sound parameters such as rhythm, pitch and harmonics need to be configured carefully.

BEEP's performance is prone to the presence of audible noise in the environments. Therefore, such a limitation makes it a less desired candidate for localisation in outdoor environments. However, audible sound can be used as a part of a railway system that includes railway localisation, management and maintenance. It can be useful in selective parts such as to alert vehicles, pedestrians and trackside workers at rail-road crossings (Korve Engineering, 2007). An audible sound-based warning mechanism in a railway network can help to increase safety of humans and assets in light rail transit.

The limitations of such systems are these:

- These systems are prone to interference from other sources of sound, therefore, reduce the location accuracy. However, still audible sound systems can be used as warning or alert in localisation systems in outdoor environments.

- Sound waves have poor penetration capability through several materials and become an undesirable localisation system.

- These systems work on audible sound and can be noisy in any environment.

2.3.7 Inertial Sensors

An internal system works on the principle of inertia that deals with the displacement, velocity, momentum and impact of external forces on an object. There parameters can be observed if a device is equipped with an inertial measurement unit (IMU) that comprises compass, gyroscope, accelerometers or barometer. These days IMU-enabled mobile devices are available that can be mounted on wrist, waist, or arm, to record the number of steps a person has travelled at which speed (Foxlin, 2005). Such data is used by several applications such as health applications to compute the number of calories burnt or by navigation applications to show the path on a particular map. Inertial-based systems detect steps to compute displacement that is a basic part in a localisation and navigation system. In an IMU-based localisation system, there is a small error in computation, called drift. Such error adds up and become a large deviation in location estimation. There are several techniques involved in rectifying such erroneous and noisy data to raise the accuracy of a system (Godha *et al.*, 2006).

One of the inertial-sensor-based localisation system is FootSLAM (Robertson *et al.*, 2009). In this system, an IMU is attached to the foot of a person and a digital compass is attached to the person's pocket. Inertial sensors, accelerometer and compass are used to compute the stepping and location of a person. FootSLAM claims an accuracy of about 2m. However, FootSLAM also suffers from the general problem of an inertial system, that is, drift errors. Drift errors can be reduced by frequent synchronising components. Inertial-based localisation systems can be implemented in outdoor environments as well (Koch *et al.*, 2005).

Inertial sensors are a useful concept in outdoor environments. Generally, it is considered as an alternative of satellite-based systems for train localisation. Another approach suggests the use of onboard IMU sensors along with a combination of other sensor devices (Heirich *et al.*, 2013; Garcia Crespillo *et al.*, 2014). In the proposed scheme, train localisation and mapping was performed (RailSLAM) by using onboard sensors. A probabilistic filter takes input from several sensors to construct track map. This system was claimed to be low-cost and utilises GNSS along with IMU in its first implementation. In the implementation, sensors' data was recorded on a track with a train in Germany. The rail vehicle filter helps to limit the track deviations. The filter also estimates motion but with higher errors after IMU data updates. The errors were reduced after GNSS updates are received. The proposed concept successfully estimated the location of the train along with geometric track mapping. Authors further proposed a Bayesian train localisation approach using Particle Filter, loosely coupled GNSS, IMU and a track map (Heirich, 2016).

Inertial-based localisation systems offer low-cost solutions as huge infrastructure is not required. However, the accuracy of these systems reduces over time with the accumulation of drift errors. The limitations of this technology are these:

- The accuracy of these systems goes down because of drift with time. Old inertial-based localisation systems used to have localisation error of about 2 nautical miles (nm) but in the modern systems are improved up to 0.6nm (Savage, 2013; Skybrary, 2009).
- The drift errors are generally dealt with either by incorporation of another technology such as satellites (GNSS) or by using extra hardware, which increases the system's cost.

2.3.8 Vision-based

A vision-based localisation system take pictures and videos from cameras. It applies feature identification techniques on those images and match them with the contextual database to identify the location of an object. Vision-based localisation solutions are generally low in cost because of a low requirement of infrastructural equipment. Algorithms have been developed to extract useful features for successful image matching even from low quality cameras (Coetzee and Botha, 1993).

Rushant and Spacek (1997) have proposed a vision-based localisation system for vehicle navigation. In this system, features are extracted from images and context mapping is performed. A location is calculated through triangulation on the identified possible locations. This system face some processing delays in location estimation and it has a tradeoff between number of restraints and localisation accuracy.

Another vision-based localisation system was proposed by Se *et al.* (2001) for a mobile robot system. The proposed system simultaneously constructs a map of the surroundings and estimates its location. The location is estimated by identifying major landmarks from scale-invariant features of images. Once landmarks are identified through SIFT features, the robot's position is estimated by comparing features of current image with the landmark image's features. This is a low-cost system and its accuracy depends on the successful SIFT features matching. SIFT feature matching incurs large delays, making it infeasible to identify a fast moving object.

A high speed camera based localisation system was proposed to track the fast moving ping pong ball (Tian *et al.*, 2011). In the proposed scheme high speed cameras capable of capturing four colour images are used. Two offline calibrated stereo camera pairs are used on each opposite side of the table. After offline calibration, the system detects the 2D and 3D locations of fast moving ping pong ball. Colour, motion and trajectory features are used in identification of the ball's position. Once the 3D position is determined, the ball's trajectory is computed from data received from both stereo pairs.

Another high speed camera based tracking system is proposed by Zhang *et al.* (2009), in which a high speed smart camera is used to track the fast moving ping pong ball. In the proposed system, an efficient target tracking algorithm is designed which operates on grey images. Algorithm differentiates the ball from the background for tracking. The authors claim to have verified the robustness of the proposed algorithm by capturing ball quickly in experiments. The high speed cameras are an efficient method of vision that amalgamates the fast moving trains' scenarios. These cameras are capable of detecting motion, fast moving objects, trajectory prediction and so on.

In another approach, a platform monitoring system is developed. In the proposed system, video cameras are used to monitor the whole track in the platform to ensure human safety (Oh *et al.*, 2007). The system is supposed to detect obstacles and unexpected items that may cause accidents.

Song *et al.* (2012) proposed a vision-based train localisation scheme using fuzzy logic. In the proposed algorithm, frame difference and feature subtraction methods are used for train location estimation. The proposed algorithm claimed to have increased train safety. The following are the associated drawbacks with vision-based localisation systems:

- Accuracy of a vision-based localisation system depends on frequent update of database in dynamic environments: otherwise, accuracy of the location estimation can reduce.

- Cameras' images can be influenced by noise such as presence or absence of light.

- Localisation accuracy of these systems is low, and claimed to be in meters.

2.3.9 WSN-based

WSN-based solutions deploy small low-cost sensor devices at several places in a grid and record data transmission. Several data processing methods are used to get useful meanings out of received data. Wireless sensor devices are not only useful for communication but also offer benefits such as recording of several events of interest which broaden its application areas. One of potential application areas is the railway industry. In the railway industry train localisation is of prime importance for development of any application. Satellite-based solutions are commonly adopted by rail operators for such an application. To cover the loopholes of GPS-based solutions, such as GPS dark regions, technology fusion is the focus of researches. WSN-based solutions are adopted in several parts of the railway industry such as track health monitoring, train scheduling or fault detection in infrastructure.

Researchers have also proposed solutions for train localisation (Hu *et al.*, 2012). In the proposed system, wireless sensor nodes are deployed along the track at known positions. Gateway nodes are linked to central servers at railway stations. Sensor nodes detect the incoming train, its position and speed, and transfer the collected data for further processing to gateway and central servers. Later on, the position and arrival time of trains are updated at each station. A Bayesian filtering technique is used to filter RSS measurements and for further processing of localisation.

TrainSense is another solution proposed for train localisation and tested on a model train (Smeets *et al.*, 2013). In this system, wireless sensor motes are used. One mote is merged with model train and detector motes are linked with track at known locations. A controller was developed to transmit packets to trains. Received packets and data are then used to calculate the position of a train. Once the train passes the detector mote point on the track, a circuit is created and position of the train mote is identified with continuous packets transmitted from controller to detector mote. Dead reckoning is used to extend the method and to determine the position of train. The positioning system was claimed to have achieved centimetre-level accuracy. They used track energy to power motes.

The following are the associated drawbacks of WSN-based localisation systems:

- There are many data collection methods for WSN-based solutions but the RSS-based method suffers from signal deteriorations such as reflections. Such noisy data can be filtered using filtration algorithms, which may increase the complexity of a proposed algorithm.

- In the absence of AC power lines, limited power sources are available which may affect the lifetime of network. To overcome such an issue, an energy-efficient algorithm is required to allow necessary operations only.

Table 2.1: Comparison of Technologies for Positioning in Indoor and Outdoor Environments

(Song *et al.*, 2011; Gu *et al.*, 2009; Al Nuaimi and Kamel, 2011; Liu *et al.*, 2007; Khan, 2014)

Technology	Accuracy	Cost	Outdoor	Railway
Standard GPS	3 m to 15 m	High	Yes	Yes
Differential GPS	<10 cm	High	Yes	Yes
Infrared	0.1 mm - 10 m	High	No	Limited
Ultrasound	1 cm - 10 cm	High	No	Limited
Radio Frequency	5 cm - 5 m	High	Yes	Yes
WLAN	2 m - 100 m	Low - High	Yes	Limited
Audible Sound	1 m - 10 m	High	No	Limited
Inertial Sensors	1 m - 4 m	Low - High	Yes	Yes
Vision	1 m - 5 m	Low	Yes	Limited
WSN	0.5 m - 10 m	Low	Yes	Yes

2.4 Cost Analysis

The focus of this thesis is to propose a train localisation solution that can operate in GPS failure areas. Generally, GPS dark regions spread over many kilometres on tracks (Mazl and Přeučil, 2003). CPEC (China-Pakistan Economic Corridor) is one of huge investments made by Chinese government and one part of the project is to build railway track from Gawadar port in Arabian sea (Pakistan) to Kashgar city (China) across the Himalayas range. A map shows that western route of this railway track that will pass through Baluchistan province, KPK province and Gilgit Baltistan province comprise of hilly terrains, valleys and several tunnels (CPEC, 2016). Therefore, it is hard to mention the percentage of the GPS dark region on a track but on average it spreads over several kilometres (Mazl and Přeučil, 2003).

From Table 2.1, it can be seen that technologies such as radio frequency, inertial sensors and WSN are applicable in outdoor environments. These technologies are either being used for rail tracking applications or in railway supportive systems such as staff and inventory tracking in railway warehouses or fault detection in railway tracks.

In Table 2.2, a cost comparison of WSN with RFID and ultrasound-based solutions is given. Here, a sample track of 500 km is taken for cost analysis and GPS dark region comprises 5% of track length. In order to simplify the analysis, it is assumed that the GPS dark region is continuous. However, it can underestimate the cost analysis given in the Table 2.2, but this assumption will have proportional cost effects across all technologies. WSN motes are powered by AA batteries in the absence of AC power lines, a power source that is hard to guarantee in tough terrains. However, parts of such areas can still have AC power lines along railway tracks. The batteries of sensor motes on sleep mode can last up to a year (Datasheet, 2006). Although, Dron *et al.* (2014) suggested an emulation based model of battery life estimation of WSN and they claim to have raised the battery life up to 484 days. A reasonable and safe assumption under average traffic and MAC operations is that a sensor mote can survive up to several weeks.

There are some additional benefits that WSN-based solutions can offer and which are not explicitly available in the use of other technologies. These benefits include the capability of WSN to develop railway supportive systems such as to detect faults, cracks or obstacles on tracks. Generally, cracks on a track are detected by special laser sensors and UK railway is following the same system to detect faults on the tracks (TheRegister, 2014). However, Punetha *et al.* (2014) proposed a similar idea to detect cracks on the track by using WSN. The proposed architecture of the sensor mote includes an IR sensor, a photodiode, a GPS sensor and a GSM module. A robot carries the specified mote in such a way that the IR sensor and the photodiode are on the opposite sides of the railway track. The robot starts its motion and able to detect a crack when the IR light passes through the crack and reaches to the photodiode. The suggested system is capable of detecting major cracks. GSM module is then used to communicate the GPS based location of the detected crack to the control station. Other than the cost factor, WSN's ability to assist in monitoring the health of infrastructure and to diagnose faults makes it the preferred choice for a train localisation system.

The cost of a single AA battery is 0.10 c (Batteries, 2016a,b) and cost of a WSN mote is \$95 (Advanticsys, 2016). The costs of a RFID reader and an active tag are \$1500 and \$30, respectively (AtlasRFID, 2016b). The cost of a Ultrasound controller is \$500 (OmegaUltraSound, 2016) and receiver is almost \$20 (Receiver, 2016). The ranges of WSN mote, RFID reader and ultrasound controller are 800 m, 20 m and 10 m, respectively. The RFID active tag's battery lasts for 3 to 5 years but once battery depletes, tags are required to be replaced (AtlasRFID, 2016a). The ultrasound receiver's battery lasts for 8 months at specific configurations (Sonotronics, 2016).

It can be seen from Table 2.2 that the fixed cost of WSN-based solution is low, 32 devices required for 5% GPS-dark region of 500 km track, compared with the RFID and ultrasound systems. However, variable cost of WSN is higher than RFID, such as replacement of batteries, may be required to replace at a higher frequency in WSN-based solutions than in RFID and ultrasound systems. The fixed cost changes according to particular site and there may require dense deployment. In such case, the distance between anchor sensors will decrease and the number of devices required to cover the space will increase. The cost of skilled labour remains constant for all of these systems, so is ignored in comparison. The added benefits of WSN, as discussed earlier, makes it a better choice for a localisation system. Therefore, WSN-based solutions are applicable not only in the absence of GPS signals but also in fusion with GPS-based localisation system to increase the accuracy.

Table 2.2: Cost Analysis of WSN, RFID and Ultrasound

Features	**WSN**	**RFID**	**Ultrasound**
Hardware	WSN motes	Reader & Active tags	Controller & Receivers
Cost per Unit	$95	$1500 & $30	$500 & $20
Range	800 m	100 m	10 m
Track Length	500 km	500 km	500 km
GPS-Dark Region	5%	5%	5%
No. of Devices	32	250	2500
Cost of Hardware	$3040	$7500	$50000
Power Source	AA batteries	AA batteries	AA batteries
Battery life in Deep Sleep	1 Year	3-5 years	8 months

2.4.1 Maintenance Cost

An system, once operational, needs maintenance. The installation costs in GPS-dark regions of a localisation system based on several technologies are given in the Table 2.2. The associated maintenance cost of corresponding technologies in GPS-dark regions are given in the Table 2.3. Generally, a maintenance cost of a system is comprised of repair cost, replacement cost, labour cost, training cost and logistics cost.

It can be seen in the Table 2.3 that in order to do maintenance work in the GPS-dark regions, some of the cost is fixed such as, vehicles. Operating companies reuse these assets which further require maintenance. GPS-dark regions are assumed to be on tough terrains, with uneasy access, hence require more fuel to get labour and equipments there. This implies that the fuel cost will be high due to tough terrains and global increasing prices of fuels. The fuel cost will have the same impact on all localisation systems based on either WSN, RFID or Ultrasound. The requirement of skilled staff is low in a WSN-based localisation system and an RFID-based localisation system because of limited technical requirements to handle these devices. However, ultrasound is a sophisticated technology and equipment needs special attention, therefore, the skilled staff requirement is high.

Similarly, the training required for the maintenance staff is minimal in case of WSN and RFID. However, due to sensitive equipment and complex methods, the training cost is high for ultrasound-based localisations system. The batteries replacement cost is high in WSN-based system as compared to RFID and ultrasound because of the frequent requirement of replacement. The device replacement cost is high in the RFID-based localisation system because of the fragile nature of its devices. However, WSN and Ultrasound have a low requirement of device replacement.

Another important factor in the maintenance cost is the frequency of maintenance. It is an average frequency in all systems. The reason is that, in WSN it depends on battery replacement and in RFID, it may be device replacement. The Table 2.3 suggests that the WSN-based localisation system will not be the best in terms of maintenance cost. Therefore, it comes to the cost-benefit analysis of each technology. A technology can have high maintenance cost with low installation cost and huge benefits.

Table 2.3: Maintenance Cost of WSN, RFID and Ultrasound in a GPS-dark Region

Maintenance Types	**WSN**	**RFID**	**Ultrasound**
Fuel (GPS-dark Regions)	High	High	High
Vehicles	Fixed	Fixed	Fixed
Skilled Staff	Low	Low	High
Unskilled Staff	Low	Low	Low
Training Cost	Low	Low	High
Batteries Replacement Cost	High	Low	Average
Device Replacement	Low	High	High
Maintenance Frequency	Average	Average	Average

2.5 Data Collection Methods

In WSN-based localisation systems, wireless sensor devices are used to collect data. The collected data is then forwarded to the location servers to compute the position of an object. Data can be gathered by several methods such as Received Signal Strength (RSS), Angle-of-Arrival (AoA) and Time-of-Arrival (ToA).

2.5.1 Received Signal Strength (RSS)

In the RSS-based method, the decrease in the signal strength at receiver is used to compute its distance from the transmitter. The strength of the transmitted signal reduces over the travelled distance. RSS can be recorded from each communicated packet from transmitter to receiver and no special equipment is required for this purpose. Therefore, this is a low-cost method that needs no extra equipment, algorithm or technique to record data. RSS is prone to environment's negative impacts such as signal reflections, multi-path fading, antenna inadequacies or interference from other sources (Li, 2006; Liu *et al.*, 2006; Bahl and Padmanabhan, 2000). In RSS-based localisation systems, the distance between transmitter and receiver is calculated from the received signal strength. The attenuation of signal strength is attributed to the signal propagation characteristics, called path loss exponent (PLE) (Bahl and Padmanabhan, 2000) (discussed in detail in Chapter 4). Log normal path-loss model is used to model the attenuation in free space (Wang and Zhu, 2008; Tarrio *et al.*, 2008). Morávek *et al.* (2010) discussed the problem of uncertainty in RSS and its implications on the measurements. In addition, RSS uncertainty and its relation with log-normal distribution are analysed by several research articles such as Stoyanova *et al.* (2009) and Cho *et al.* (2007). Several statistical methods and noise filters are used to reduce the negative impact in RSS measurements. These methods take several measurements to process, such as particle filters and to identify the best measurements for the location computation. However, if the number measurements are higher, it increases the energy consumption and is a major drawback for energy-sensitive applications.

2.5.2 Angle of Arrival (AoA)

In the Angle-of-Arrival method, location of an object is computed by intersection of several sets of angles. An angle is considered between transmitted and reflected signals, given a fixed direction of propagation. An array of antennas can also be used in order to determine the location of an object through triangulation method. Use of an array also decreases error rate (BER) and multi-path effects along with other benefits (Giorgetti *et al.*, 2007). Erdogan *et al.* (2006) discussed the advantages of the AoA-based localisation method, such as energy conservation, that has an overall impact on network lifetime. Energy consumption is improved by a similar sort of technique that introduced a few sink nodes (Kalis and Dimitriou, 2005). The use of an array of antennas is considered as a limitation because it require huge resources. Further, the requirement of wide dimensions of antennas reduces with higher frequency communication. In another piece of work, electromechanical systems are used to increase the feasibility of array-based sensors (Giorgetti *et al.*, 2007; Kalis and Dimitriou, 2005). However, the use of antenna arrays increase the system cost as a whole (Kalis and Dimitriou, 2005). A distributed AoA-based localisation of acoustic sensors is presented by Arabaci and Strickland (2007).

2.5.3 Time of Arrival (ToA)

The ToA refers to the propagation delay of a transmission. In other words, it is the time taken by a transmitted signal from sender to receiver. As both sender and receiver are involved in determining the time consumed in such activity, they need to be synchronised. A slightly modified approach is to consider reflected signal back from receiver to initial sender, and it relaxes the strict requirement of clock synchronisation (Mao *et al.*, 2007). In the ToA approach, intersection points are obtained from ToA data measurements from at least three nodes to compute the location of an object (Peng and Sichitiu, 2006). The chances of getting unique intersection points are directly proportional to the number of transmitting nodes (Huang and Benesty, 2004; Mao *et al.*, 2007). The location of an object can be computed using the triangulation among several sensor nodes with known locations. ToA is prone to multi-path effects that increases with the number of obstacles. The ToA method is feasible for the indoor localisation scenarios with a few obstacles such as a hall or meeting room. ToA is also feasible in less dynamic outdoor environments.

2.6 Data Processing Methods

The measured data can be processed by several methods. Each method has its way to handle the collected data. The following are four different ways of a localisation system: Geometric, Fingerprinting, Dead Reckoning and Proximity.

2.6.1 Geometric

Location estimation can be performed on measured data using geometric properties (Lee *et al.*, 2009). Triangulation and trilateration are the methods that use geometric properties for localisation. In trilateration, the position of an object is computed by calculating its distances from several reference points. On the other hand, in the triangulation method, the integration of multiple sets of angles is used to compute the object's location. Drawbacks of geometric methods include mismatch of several radians because of blocking and multi-path errors.

2.6.2 Fingerprint

Fingerprinting is a method of matching the received data measurements, such as RSSI or ToA, with known measurements that are saved in the database, called data fingerprints (Gogolak *et al.*, 2011). The environment data recording process generates the fingerprint of data during the offline system calibration phase. On the other hand, in the online system localisation phase, the newly received data measurements are mapped to the saved fingerprints to estimate the position of an object (Saxena *et al.*, 2008). Drawbacks of the fingerprint method include deviation of the fingerprint because of modifications in an environment.

2.6.3 Dead Reckoning

The dead-reckoning method works with the last known location of an object and a set of inertial or non-inertial data values, such as time, speed, direction or acceleration, to estimate the new location of an object. (Jimenez *et al.*, 2009). Dead-reckoning is often used to predict the object's location in the next time stamp with respect to received inertial dataset.

2.6.4 Proximity

In the Proximity method, the measured data is processed with the reference location in close vicinity. The location is estimated with the best match to the closest known landmark (Patwari and Hero III, 2003).

2.6.5 Kalman Filtering for Localisation

One of the ways to implement a Bayes Filter is through Kalman Filters (KF). KF is an estimator to estimate an instantaneous state perturbed by noise that can be modelled in terms of Gaussian noise. KF uses Gaussian model in the prediction of posterior state (Grewal, 2011). In other words, KF is a tool to predict the likely future state of dynamic systems that are beyond control, commonly for the trajectories of celestial bodies. In KF-based Simultaneous Localisation And Mapping (SLAM) problems, Gaussian noise is added to the state transition and the measurement functions (Thrun *et al.*, 2004). In tracking problems, KF is used to estimate the location of a robot, where landmarks in a route are initialised. Landmarks are linked to the obstacles in the route, which restricts the movement of a robot. In addition, a data association problem associates the features in the route with the landmarks and identifies the change of the scenes. Thus, landmark initialisation and data association are important concerns in the tracking of a robot using KF.

One of the main drawbacks of the KF implementations is the fact that for long-duration tracking, the number of key tracking points increases and, at some stage, computational resources will not be sufficient to update the map in real-time.

The advantage of KF is that they provide optimal mean-square error (MMSE) estimates of the state, and its covariance converges convincingly.

2.6.6 Particle Filtering for Localisation

Particle Filter is a special type of recursive Bayesian Filter, also called sequential Monte Carlo (SMC) method. Particle Filtering is an estimator that starts with the random points, called particles. Initially, all particles hold the same weight and represent the exact location of an object with same likelihood. At each time instance, each particle moves to a new possible position and updates its weight, that show its chances of representing accurate location of an object. PF samples particles from a distribution to estimate the position of an object. This technique makes it reliable for nonlinear and non-Gaussian systems. PF is capable of handling computational complexity of the state that has grown with the addition of a landmark (Montemerlo *et al.*, 2002). Therefore, PF is highly suitable for localisation applications. However, it is still being studied for SLAM problems such as FastSLAM (Montemerlo *et al.*, 2002; Roller *et al.*, 2003).

The use of Particle Filtering in the train localisation system is discussed in detail in Chapter 6.

2.7 Existing Train Localisation Technologies

The railway environment can be divided into several parts such as open field, railway stations or tunnel. The technologies and methods used for train localisation can differ with the scenarios it is being developed for. The existing literature can be divided into technologies used in open field and tunnel (Fararooy *et al.*, 1996).

2.7.1 Traditional Technologies

Automatic train control is a system that places focus on train's safety by identifying its position. The identified position is then spread among trains within close proximity to avoid potential collisions. Since the beginning of railway, railway flare system was in practise in which flare was dropped from the backend of trains. The flare used to burn for sometime and it becomes indication for following train on that track to analyse the distance from next train and to adjust its speed in order to avoid potential collisions (Wiita, 1989).

Later on, the technology evolved in railway and block system was introduced, that used block of tracks to identify the location of train (Wiita, 1989). Signals were passed through attached wires to control room, upon entrance of train on specific track block. This system was inadequate during operations of high speed rails, multiple tracks and use of extensive wiring.

In the initial days of modern railway systems, on-board equipment were used for positioning. Signalling systems along with their coordination with track circuits were used to enable safety features in a train. Later on, systems were improved with the integration of electronic control boards, radio units and interlocking states. Track circuits were used in term of electric energy to detect the connectivity and location of train at any part of track.

Traditionally, another approach for train localisation was used, called axle counter system (Fararooy *et al.*, 1996). In this system, the train's wheels are detected. One set of equipment is placed on a section of track and another set of equipment, called evaluator, is installed in a control room. The trackside equipment detects the train's wheels once train enters or leave that section. Trackside equipment includes electronic junction tools and transducers to detect the wheels. Another approach uses axle rotation to detect train along with computer and radio-aided train control system (CARAT)

(Ikeda, 1993). In another approach, a speed sensor, called fail-safe, is introduced to detect the position of a train (Hill, 1981).

2.7.2 Technologies for Open-Field

An open field is one of environment in the railway. Generally, in railway industry, GPS or its variations are being used for localisation. DGPS is more accurate than basic GPS and railway companies often hire services based on DGPS. Currently, DGPS networks are being used in several parts of Europe (Fararooy *et al.*, 1996). The use of satellite-based systems are common in railway navigation projects. The limitation of such technology is disconnection from satellites in GPS dark regions such as skyscrapers, bridges, hills or tunnels. A solution to such a problem is integration of satellite-based solutions with other technologies such as sensors (Leahy *et al.*, 1993).

2.7.3 Technologies for Tunnels

Satellite-based systems do not perform in tunnels because of unavailability of satellite signals. Train-based or track-based solutions may not work as well solely. Mayhew *et al.* (1994), developed a hybrid solution for tunnels that focused on a train control system. Radio-based techniques, such as optoelectronic systems, use sensors and compute the location of train based on train-track communication. Military applications use frequency hopping spread spectrum techniques for such purposes (Fararooy *et al.*, 1996). These techniques involve radio beacons to transmit after regular intervals and a set of on-board transceivers are used to identify the position of a train.

In another approach, magnetic transponders are used. Magnetic transponders, placed on the track, transmit known information to the train and the train calculates its position. A speed sensor was developed based on Doppler's effect, as input to the navigation system (J and Faulkner, 1991). The Doppler sensor changes the frequency of signal to make it significantly noticeable for a receiver moving at relative speed. A slip-slide control system was proposed for modern trains that also benefits from Doppler sensors (Descamps *et al.*, 1991). Inertial train control system is often combined with traditional signalling system to identify the location of a train. In the past, laser diode and charge coupled devices are the focus of research because of their ability for restraining themselves to low error. In such techniques, wheel and rail inspections play their part in avoiding mechanical drifts (Seitz *et al.*, 1990). Further, errors are minimised by using Kalman Filtering techniques.

2.8 Existing Train Localisation Projects

In the railway world, GPS is a commonly used technology to track trains, its equipment, maintenance vehicles and trackside staff in real-time. GPS-based localisation and navigation systems improve their performance with the fusion of other technologies such as sensors or communication systems. Train localisation systems play a vital role in safe, timely and low-cost railway services. Other than GPS, the existing train localisation technologies are based on track circuits, on-board IMUs or trackside optical markers (Hill and Weedon, 1990). However, these technologies lack in precision and need huge infrastructure (Wikil, 2016; Johnson, 2016; Huffman *et al.*, 1977). In selected parts of the railway system other technologies are being used for localisation purposes such as RFID for railway inventory tracking and management, and WLAN for railway system management. A discussion about several train localisation projects and their adopted technologies is given in following sections.

2.8.1 RFID-based Railway System

An RFID-based localisation solution was proposed by the GAO group (GAO-Inc., 2007). The solution was proposed to specify the location of locomotives and railcars. In addition, the system is also useful for asset management and identification of equipment and staff. In the proposed system, an item can be located before and after it is loaded into a container for shipment purposes. Further, the location of a staff member can be tracked and well informed if the site needs to be cleared because of movement of trains. In a nutshell, the RFID-based solution proposed by GAO company increases safety features of railway staff and equipment, allows staff to do efficient stock maintenance and enables smooth operation of railway industry. Several other attempts have focused on transportation safety using RFID (Char and Johns, 2006), and RFID is considered to be an important application area of future (GRIFFIN *et al.*, 2006).

The RFID-based railway solution offers benefits for railway management in selected places such as railway stations or inventory stores of railways. However, the short-range RFID signals make it infeasible over long railway tracks where the speed of the train is high.

2.8.2 IMU-based Railway System

In modern railways, Positive Train Control (PTC) systems are being used to prevent railway accidents such as train derailments, mishaps with trackside workers, and wrong turns (Hansen, 2001). PTC works with a combination of technologies such as IMU–onboard inertial sensors and satellites. In this system, speed of train can be monitored, traffic routes can be adjusted and safety of the railway workers can be improved. This system is also useful for optimising the railway capacity by generating a global traffic view. In addition, this system incorporates and synchronises railroad communication to avoid any such mishaps.

2.8.3 Satellite-based Railway System

Generally, satellite-based solutions are adopted in the railway environment because of availability of services. On a commercial scale, services of several satellites are hired to develop a localisation and navigation system. In the past, European Train Control System (ETCS) was under focus in Europe (Rados *et al.*, 2007). ETCS consists of three levels and each level consists of standards, policies and techniques to develop a train control system. Majorly, ETCS takes care of international boundary policies as train networks are supposed to be deployed across Europe. European Rain Traffic Management System (ERTMS) is a commercial and industrial project of European rail (Midya and Thottappillil, 2008) and ETCS is a part of the ERTMS project. ERTMS is a satellite-based solution that works with GPS or GNSS. In one of its sub-project, EATS, the combination of information from several sources, such as Global Navigation Satellite System (GNSS), UMTS and GSM, are considered for localisation.

INTEGRAIL is another solution for the railway industry (Staton, 2005). It uses European Geostationary Navigation Overlay Service (EGNOS) signals to add safety features in existing railway navigation solutions (Umiliacchi *et al.*, 2006). INTEGRAIL offers several benefits such as reduced cost and reliability in the preceding systems which were based on onboard solutions such as odometers. The improvement is because of fusion of existing solutions with satellite-based solutions and EGNOS. INTEGRAIL provides an accurate localisation solution in several operational conditions as well.

ECORAIL (EGNOS COntrolled RAILway) project implements GNSS along with the ETCS and ERTMS system (Thevenot *et al.*, 2003). It offers advantages of GNSS and management standards of ERTMS system together in a reduced cost solution.

GRAIL is another project that introduces the use of GNSS for a railway system (Urech *et al.*, 2006). The use of GNSS and other technologies by different vendors create the problem of interoperability and compatibility. GRAIL takes into consideration such issues and provides smooth integration of a GNSS-based localisation solution with signalling and control systems proposed in main ERTMS and ETCS system in

Europe.

GADEROS is another project that is supposed to be integrated into ETCS and ERTMS. GADEROS offers safety features for life along with the integration of GNSS in the ETCS and ERTMS project (Urech *et al.*, 2002). GADEROS was under the directorate of European Union. The Railway User Navigation Equipment (RUNE) project offers the integration of GNSS and safety of life features offered by previous projects (Albanese *et al.*, 2005). The RUNE project involves extensive testing of its tasks in the laboratory and in the field. RUNE focuses on its object that is to enable a train to identify its position with limited or no support from trainside equipment. It also complies with the standards of the ERTMS project.

2.9 Train Localisation System

An ideal train localisation system should consist of several components, in which each component represents a train localisation subsystem based on different technologies such as Wireless Local Area Networks (WLANs), RFID, GPS and Wireless Sensor Networks (WSNs). WSN is a multi-dimensional technology that offers benefits in multiple domains along with localisation in many areas. Railway industry can also benefit from WSN, such as to identify track faults, management of inventory and staff, and strain measurement in bridges. In such a system, each localisation subsystem offers some benefits with some associated cost and reliability as discussed in the previous section. However, the integration of heterogeneous technologies increases the reliability of position information of trains through sensor fusion, cooperative localisation data and detection of poor localisation zones such as railway tunnels, underground trains, forests, and above-ground hilly terrains. Moreover, it will validate results and offer benefits of each incorporated technology-based localisation system in such complex environments with strong safety and security requirements.

The location information from each train localisation subsystem is collected by the communication server. The communication server provides the first platform to interact with estimated location information from each localisation system. Through the internal network, the estimated location information is sent to the Railway Localisation (RailLoc) fusion server. The fusion server takes the ingredients from each localisation system and uses fusion algorithms to make a more realistic estimation of train location at different time periods. The application servers use the train location information that is estimated by the fusion server in the related SLAM and tracking applications

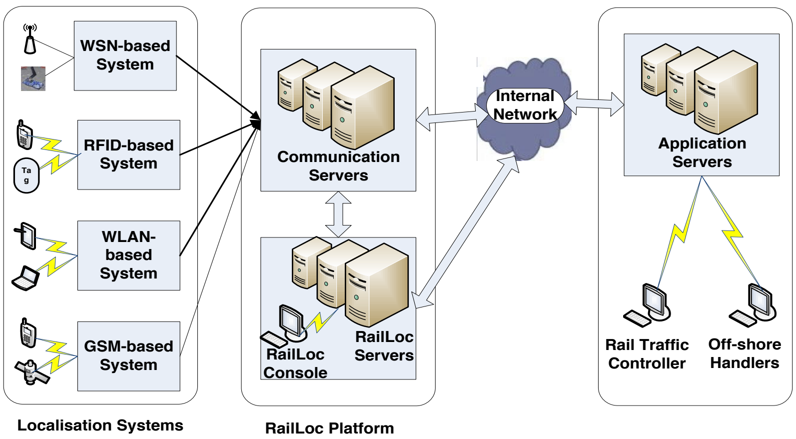


Figure 2.1: Overview of the Train Localisation System.

to fulfil the requirements of the train navigation system.

This thesis focuses on a WSN-based train localisation system, which can be incorporated in a global train localisation system, as shown in Figure 2.1, and can serve the purpose independently as well. Further, later chapters will discuss the rationale of WSN-based localisation system and its components. Though the proposed system can operate independently in GPS-dark regions, it becomes an adaptable sub-system solution in a global train localisation system as shown in the Figure 2.1.

2.10 Research Goals

The localisation is an essential component of the navigation system, which is directly related to the safety of the railway primarily, along with other benefits. I carefully analysed the strengths of different technologies for train localisation and their associated shortcomings. GPS is the most widely adopted technology in the train localisation projects. The requirement analysis from the train navigation system helped us to formulate the research goal of this thesis: "to develop a WSN-based train localisation system that can give effective performance where GPS is not available". The proposed train localisation system will provide a high accuracy at reasonable cost. This solution can be used in combination with solutions provided by other technologies by using data fusion techniques. I also identified a certain set of specifications for our WSN-based train localisation system that are summarised as follows:

- The developed system will use wireless sensor devices on the track, called anchor sensors, and another sensor device will be mounted on the train, called gateway sensor.

- The anchor sensors are powered by batteries that are hard and expensive to replace frequently in the remote areas. Therefore, anchor sensors sleep for sometime and wakeup to sense the incoming train for the communication. Batteries are common method to provide power in the absence of AC lines, as provided in track circuits (Nagel, 1979).

- The developed system will guarantee that anchor sensors will be awaken when a train passes by them. In addition, anchor sensors are unaware of their neighbour sensors' sleep schedule and train's arrival time. This approach has several benefits, such as, it saves memory of these miniature devices, in case of delay in train's schedule system does not fail, and clock synchronisation is not required.

- The developed system computes the maximum sleep time that an anchor sensor can follow and saves maximum energy along with the guarantee to wake up at the right time.

- The gateway sensor will use the geographic coordinates of anchor sensors and the RSS of the transmitted signals for the estimation of train's location.

- In the developed system, the noise of the RSS using Particle Filtering technique and position of the train is calculated along with the received location information from the anchor sensors. A weighted RSSI likelihood function (WRLF) is designed to identify the likelihood of particles representing the true location of the train.

- In our proposed system, real-world data is used in the simulations. The experiments were conducted to collect the real-world RSS data in the railway representative environments such as open field, railway station and tunnel. The use of real-world data increases the relevance of simulation with the real-world scenarios.

- In our developed system, anchor sensors cooperate with each other frequently to identify the faulty nodes among them, and to calibrate their location and pathloss ratio. A report is compiled by each anchor sensor based on the developed consensus and sent to the gateway sensor to eliminate the input from the faulty node.

2.11 Summary

In this chapter, I started with the identification of the metrics for localisation systems and then compared several localisation systems based on those metrics to identify the pros and cons of their underlying technologies as summarised in Table 2.1. The review of the literature shows that technologies such as Infrared, Ultrasonic and Radio Frequency offer reasonable positioning accuracy, but such systems require large infrastructure in outdoor environments that increases the overall cost of positioning system. In contrast, the technologies such as WLAN, Inertial Sensors, Machine Vision and WSN offer cheap solutions. In WLAN, the solution is cheap if infrastructure is reused; otherwise, it will raise the cost of deployment, and the fingerprinting method also increases the cost, time consumption and technical support. In Inertial-sensor-based systems, noisy measurements increase the localisation error, which can be reduced by adding sensors, at extra cost. In machine vision solutions, database maintenance and structural changes over time, increase its cost for large-scale network deployment. Mostly, the existing train localisation projects use GPS as primary technology to estimate the location of train in the railway system. However, GPS has several limitations such as GPS dark regions, signal penetration issues and large errors. Alternatively, WSN provides cheap solutions because of low-cost devices, but these devices are limited in resources because of miniature architecture. The associated low cost of WSN and its easy deployment and maintenance are features that make it a preferable choice for my work. The large transmission range of sensor nodes reduces (800 m) the number of devices required. However, the required number of devices depends on deployment density, which is a function of type of terrain. A broad concept of train localisation system is presented that has several technologies based on train localisation subsystems as shown in Figure 2.1.

The research objectives of this thesis are formulated based on the design of an accurate and low-cost train localisation system that can be opted as a solution in the absence of GPS-based localisation solutions. Such objectives lead to the design and development of a WSN-based train localisation system.

In the next chapter, I shall present a brief overview of WSN-based train localisation system in general and each of its modules in particular.

Chapter 3

Overview of WSN-based Train Localisation System

In this chapter, I begin by introducing the idea of train localisation using wireless sensor networks. I then describe each of the modules of the WSN-based train localisation system. The details about each module are briefly discussed in individual sections, which are later explained in detail in the next chapters.

3.1 System Models

The system models of WSN-based train localisation system consist of network, duty-cycling and train localisation models.

3.1.1 Network Model

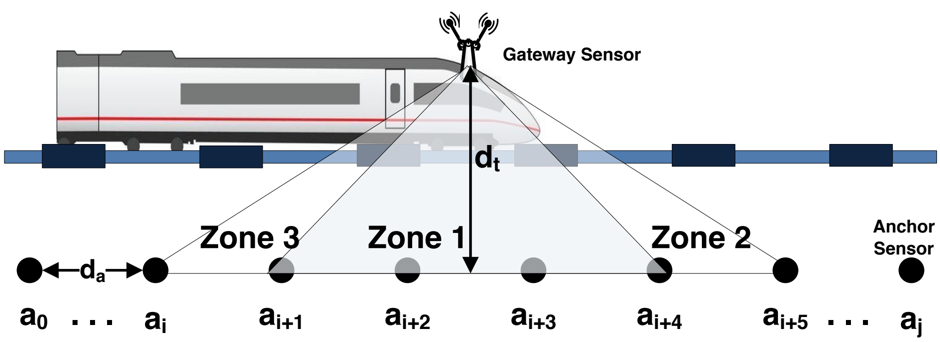


Figure 3.1: A WSN Architecture for Train Localisation

In the network model of the WSN-based train localisation system, there are two types of sensor nodes: anchor sensors and gateway sensors, as shown in Figure 3.1. A set of anchor sensors is uniformly deployed along a straight track with equal distance between any two consecutive anchor sensors. The distance varies and depend on the deployment requirement such as in sparse network (straight and smooth terrain), distance can range from 100 m to 800 m, and in dense deployment (terrains with hills and turns), there can be multiple sensors, deployed within 100 m. The anchor sensors are powered by the batteries that can deplete rapidly, a few days (Royo *et al.*, 2009), if anchor sensors stay in idle listening mode all the time. Therefore, in a WSN-based train localisation system, anchor sensors operate on duty-cycles. It is assumed that each anchor sensor has hard-coded its geographic coordinates before deployment. This assumption is reasonable as there can be a few sensors with known locations and rest of sensors can compute their location with trilateration method after communication with each other. A single gateway sensor is installed on the train. Multiple gateway sensors have their own pros and cons such as there will be mechanism required to avoid data collision during communication. Therefore, in a simplistic model, single gateway sensor serves the purpose. The gateway sensor is equipped with two radio transceivers: one transceiver has long transmission range and is used to continually broadcast beacon packets to wake up anchor sensors, called beacon-transceiver; other transceiver has short transmission range and is used to communicate with the anchor sensors that are woken up by long-range transceivers, called communication-transceiver. To avoid interference, beacon-transceiver and communication-transceiver operate on non-overlapping frequency channels. Each anchor sensor operates on both channels and once a beacon packet is received, it switches its channel to communicate with the communication-transceiver of the gateway sensor. As shown in Figure 3.1, zone 1 is the region covered by communication-transceiver and zone 1, zone 2 and zone 3 are the regions covered by beacon-transceiver.

3.1.2 Duty-Cycling Model

All anchor sensors operate in an asynchronous duty-cycling mode in which each anchor sensor switches between sleep and wake-up states independently without global synchronisation. Figure 3.2 shows one duty-cycle, in which an anchor sensor first sleeps for t_{sleep} duration with its radio turned off, and then wakes up and turns its radio on to perform clear channel assessment (CCA) to detect incoming signals. If an incoming signal is detected, the anchor sensor will keep in the active state until the scheduled communication between the anchor sensor and the gateway sensor is completed; otherwise it switches back to the sleep state and repeats another duty-cycle.

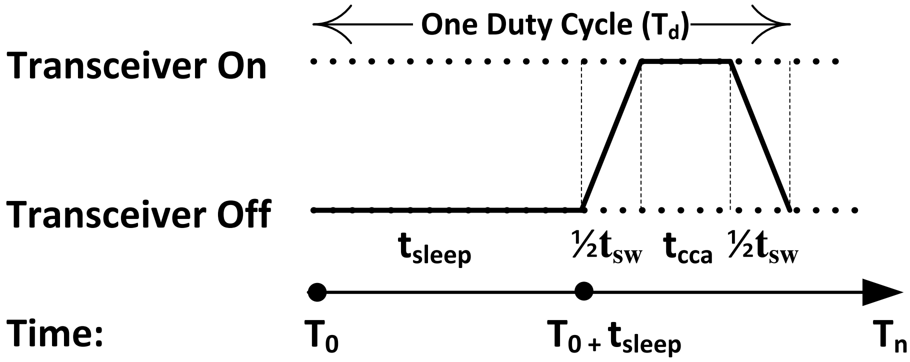


Figure 3.2: Illustration of One Duty-Cycle

3.1.3 Train Localisation Model

As the train moves, the gateway sensor continually broadcasts beacon packets. Each beacon packet contains information about the current train location (represented by the location of the gateway) and speed. Once an anchor sensor receives a beacon packet, it stops duty-cycling and prepares for communication with the gateway sensor. When an anchor sensor goes into the transmission range of the gateway sensor, it sends its geographic coordinates to the gateway sensor. After an anchor sensor finishes the communication with the gateway sensor, it resumes duty-cycling. Based on the geographic coordinates received from anchor sensors as well as the RSSI information of the transmissions from anchor sensors, the train location will be computed at the gateway in a real-time manner.

3.2 System Design of WSN-based Train Localisation

The WSN-based train localisation system consists of several important modules such as Sensors Wake-up Scheme, Train Localisation Scheme, and Sensors Management Scheme as shown in Figure 3.3. Each module serves a specific purpose in the WSN-based localisation system. However, the integration of these modules consolidates the benefits that each module offers and makes it a WSN-based train localisation system.

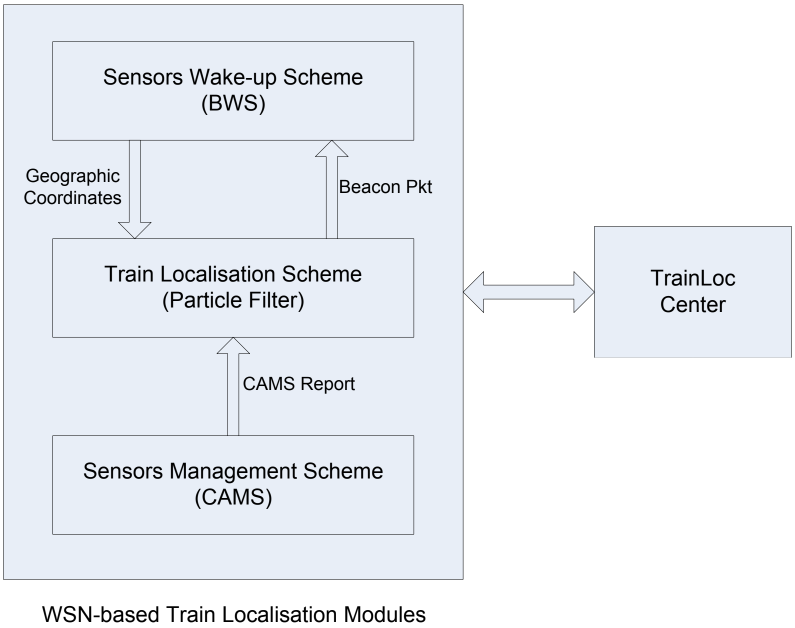


Figure 3.3: Overview of the WSN-based Train Localisation System

In the first module, a beacon-driven sensor wake-up scheme (BWS) is proposed to address the problem of waking up anchor sensors in an energy efficient way when the train is approaching. To prolong the lifetime, anchor sensors operate on asynchronous duty-cycles without the global knowledge of sleep schedules of their neighbour anchor sensors. Moreover, it becomes challenging to wake up anchor sensors at right time in the absence of train's arrival schedule. Therefore, a wake-up scheme is designed in which a train assists anchor sensors to wake up upon its arrival by broadcasting beacon packets and allowing anchor sensors to sleep most of the time otherwise.

In the second module, the Particle-Filtering-based train localisation scheme is presented. This scheme uses the RSSI measurements and anchor sensors' geographic coordinates to estimate the train's location. A weighted RSSI-based likelihood function is developed to compute the train's location. The likelihood function updates the weight of the particles based on RSSI measurements and geographic coordinates received from the anchor sensors. The particle weight represents a particle's probability of representing correct train's location. Consequently, average probability of all particles is computed that represents the estimated train's location.

Finally, in the last module, a consensus-based anchor sensor management scheme is presented to identify the possible anomalies in the system. Anchor sensors are the vital ingredient of the WSN-based train localisation system, and they must be maintained in order to maintain a high performance of the system. Anchor sensors can suffer from faults due to their software or hardware problems such as low battery, thermal effect, and dislocation. These anomalies can result in incorrect data input to the train localisation system that can compromise the system's performance. Therefore, the consensus-based anchor sensor management scheme enables anchor sensors to mutually identify such faults.

The detailed information about each module is presented in the next sections.

3.2.1 Beacon-driven Wake-up Scheme (BWS)

In the WSN-based train localisation system, anchor sensors are deployed along the track. They communicate with the gateway sensor once they go into the transmission range of the gateway sensor. The train's schedule is unknown to the anchor sensors, it is hard to guarantee the availability of anchor sensors for communication with the gateway sensor at the time of train's arrival. One simple solution to guarantee the availability of anchor sensors is to keep them in an idle listening state forever. Though, such scheme can serve the purpose but it raises another problem of rapid battery drainage. Therefore, keeping anchor sensors in idle listening state has the worst impact on the system lifetime. The ultimate solution to prolong the network lifetime is to enable anchor sensors to follow duty-cycles. In duty-cycling, anchor sensors switch between wake-up and sleep states by periodically turning their radios on and off. Though the duty-cycling solution can help to minimise the energy consumption the anchor sensors, it still cannot guarantee the availability of anchor sensors for communication with the gateway sensor at the arrival of train.

To guarantee the timely wake-up of the anchor sensors, the Beacon-driven Wake-up Scheme (BWS) offers a cost and energy efficient solution. In BWS, the two gateway transceivers TS_b and TS_c assist anchor sensors to wake-up and communicate, as shown in Figure 3.1. Each anchor sensor, once it has received the beacon packet from TS_b, stays active and prepares to communicate with the transceiver TS_c of the gateway sensor. In BWS, the duty-cycling parameter t_{sleep} plays an important role in the timely waking up of anchor sensors. If t_{sleep} is small, each anchor sensor needs to frequently turn on and turn off its radio, thereby wasting too much energy. From an energy saving perspective, the larger the t_{sleep}, the more energy each anchor sensor can conserve. However, if t_{sleep} is too large, an anchor sensor may miss the chance to detect the beacon packet broadcast by TS_b and fail to wake up in time.

The BWS module of this thesis derives the upper bound on t_{sleep}, which enables each anchor sensor to stay in sleep state as long as possible while still guaranteeing that each anchor sensor can wake up in time once the train approaches. Secondly, it designs an energy-efficient wake-up scheme, which guarantees that each anchor sensor can wake up in time once it goes into the transmission range of TS_c, and resumes low power duty-cycling once it finishes communication with the gateway. The designed scheme is evaluated through both theoretical analysis and simulations.

3.2.2 Particle-Filtering-based Train Localisation Scheme

The gateway sensor collects the geographic coordinates of the anchor sensors in its communication range and corresponding RSSI measurement of each transmission. Though the RSSI measurements are used to estimate the distance of the sender and can be useful to estimate the location of the train, they can fluctuate because of multi-path fading and signal reflections from the surrounding infrastructure. Therefore, it may not stay reliable to estimate the location of the train from noisy RSSI measurements, alone. This problem can be dealt in two phases, in one phase, RSSI can be used to estimate distance with large errors and in second phase, estimated distance can be fine tuned by using another type of data such as geographic coordinates of anchor sensors. The problem of noisy RSSI measurements need to apply noise filtering to minimise the location estimation error.

Particle Filter, which implements recursive Bayes Filter, is an efficient solution for nonlinear/non-Gaussian tracking problems. The Particle Filter can be used to filter out the noise from RSSI measurements because of its noise tolerant property. In WSN-based train localisation, Particle Filtering is used to compute the location of the train. In this scheme, two models are required in this filter: the movement model that describes the evolution of the state with time (i.e., the train movement model in our case), and the observation (measurement) model that relates the noisy measurements to the state (i.e., the RSSI measurement model in our case). Particle Filter relies on the construction of the posterior probability density function of the state based on the set of received measurements, and recursive filtering is performed by taking into account new measurements once they are available. The Particle Filter consists of the *prediction* and *update* stages. In the prediction stage, the location of the train is estimated using the movement model of particles. The update stage uses the measurement model to modify the predicted particles' locations and weights. The particles are then filtered with the likelihood of being the exact representation of the location of the train. The particles with the highest weight are more likely to represent the current train location. Once the RSSI measurements are available, the observation model is used to update the weights of the particles.

In the Particle-Filtering-based train localisation module, an algorithm is developed to estimate the location of the train by using noisy RSSI measurements and the location information received from the anchor sensors. A weighted RSSI likelihood function is developed that computes the likelihood of the particles to represent the train location. The designed scheme is evaluated through extensive simulation using real-world datasets collected from the field experiments.

3.2.3 Consensus-based Anchor sensor Management Scheme (CAMS)

The anchor sensors along the railway track may suffer from the location errors caused by software or hardware bugs. Therefore, they need to be re-calibrated for their geographic coordinates and the path loss of the signals sent by the anchor sensors. Moreover, the presence of faulty sensors in the system can also reduce the accuracy of the location estimation. All these issues should be addressed in the WSN-based train localisation system. Manually sorting out such problems by human beings incurs significantly high cost. Here, a cost can be categorised as the number of times a technical team may need to visit to check faults. However, CAMS generates a report by sensors, then targeted effort is required to rectify the faults, thereby, reduces the cost. The management and maintenance of the anchor sensors with the help of each other play an important role in the stability of the whole localisation system.

Therefore, the need for a management scheme comes into play that can enable anchor sensors to detect the faulty sensors among themselves. The faults should be reported to the gateway sensor for further analysis. Furthermore, the management scheme should assist anchor sensors to estimate the path loss ratio of their signals, which depends very much on the surrounding environment and affects directly the distance estimation based on RSSI. Such a management scheme can significantly improve the accuracy of train localisation by excluding the faulty sensors and re-calibrating the parameters of the anchor sensors like path loss ratio.

In this module, the CAMS scheme is proposed for the WSN-based train localisation system. CAMS allows anchor sensors to share their opinions about the trustworthiness of their neighbour sensors and develop consensus to detect the faulty sensors. The anchor sensors can be automatically re-calibrated in terms of path loss ratio and geographical coordinates. CAMS is implemented in a simulated environment using

MATLAB. The simulation is based on the real data collected from field experiments in various environments such as open field, train station and a tunnel.

3.3 Summary

In this chapter, I presented an overview of the WSN-based train localisation system and discussed its components that are the focus of this thesis. A system design of WSN-based localisation system is presented, which elaborates the network model. Moreover, an introduction is presented to explain the overall working of WSN-based train localisation system.

The remainder of the chapter presented the overview of each of the three modules of the WSN-based train localisation system. As the anchor sensors are powered by batteries, they operate on asynchronous duty-cycling to save energy to prolong their operational life. However, duty-cycling raises concerns about guaranteeing the availability of anchor sensors for the communication with the gateway sensor. Therefore, an overview of the first module, beacon-driven wake-up scheme, is presented, which enables the anchor sensors to sleep for a maximum time and still guarantee their wake-up at the time of train passing by them. The wake-up of anchor sensors is vital as their input (geographic coordinates and RSSI) are required by the gateway sensor to estimate the train's location. Such measurement data from the anchor sensors are utilised by the second module of WSN-based train localisation system, that is, the Particle-Filtering-based train localisation scheme. An overview of Particle Filter and its usage to compute the train location is presented briefly. The anchor sensors are required to be maintained as they are prone to environmental, device-ageing (need to re-calibrate), safety and security effects. In the last module, I have given a consensus-based anchor sensor management scheme, which reports the existing faults and faulty sensors in the network to the gateway sensor. Such a scheme helps to maintain the network, enhance the overall lifetime of the localisation system, and reduce the cost of manual diagnostics and maintenance.

In the next chapter, I shall validate that RSSI follows the log-normal path loss model, a known signal propagation model, in harsh railway environments, that are different from general open field environments. Moreover, I shall present an analysis to determine the feasibility of RSSI measurements usage for WSN-based train localisation.

Chapter 4

Experiments to Validate the Feasibility of Using RSSI for WSN-based Train Localisation

In this chapter, I begin by introducing the idea of train localisation using wireless sensor networks and use of signal strength measurements to estimate the location of the train. I then describe the log-normal path loss model in detail, which is a known model to map the signal strength over the distance of the transmitter. Railway is a harsh and a different environment from other open field environments as it is influenced by the involvement of infrastructure, metals and radio frequencies. Therefore, it is required to verify that signal strength follows the log-normal path loss model in railway environments. In the remainder of the chapter, I present the details of the experiments to collect the RSS measurements in an open field, railway station, and tunnel. Tunnels and open fields in GPS dark regions in hilly terrains. However, railway station is considered as another example railway environment in which WSN-based localisation system can provide redundancy to other localisation systems, thereby, increases localisation accuracy. The existing system of track circuits is not very precise and WSN-based solution can be an option in remote railway stations. Further, WSN can be deployed for multiple purpose: train localisation, track monitoring, track side worker alarming, etc. Therefore, can be a good choice for railway stations. Later on, I analyse the feasibility of using RSS measurements for train localisation.

4.1 Train Localisation using Wireless Sensor Networks

The key idea of using wireless sensor networks for train localisation is to estimate the location of the train using received signal strength measurements and the geographic coordinates of the anchor sensors. In the ideal medium, RSS measurements can be trusted, but in this physical world, RSS gets influenced by some other factors along with attenuation. Radio signals get affected by the surrounding obstacles such as reflections from walls or metals, diffraction from sharp-edged surfaces, interference from overlapping frequencies and shadowing effects due to antenna properties. Such phenomena induces noise in RSS measurements, which can deviate them from their actual values. Therefore, there is a need of another data model, such as location information of anchor sensors, to overcome the deficiencies of RSS measurements by data fusion technique. Such incorporation of location data increases the credibility of RSS measurements and minimises estimation error that would have emerged otherwise.

The log-normal path loss model is used to analyse the power loss of the signal during transmission in several railway environments. Such analysis is useful to analyse the distribution of the noise in the collected datasets, which is necessary to select the noise filter that will be used for train localisation.

4.2 Log-Normal Path Loss Model

The free-space model (Abhayawardhana *et al.*, 2005) and the two-ray model (Sommer *et al.*, 2012) estimate the received power of a transmission as a deterministic function of distance. Both models represent the communication range as a perfect circle. In reality, the received power at a certain distance is a random variable due to multi-path propagation effects, which is also called the fading effects. In fact, the above two models predict the mean received power at distance d. A widely used and more generic model is called the log-normal path loss.

The path loss model (Xu *et al.*, 2010) is a well-known radio propagation model that predicts the path loss a signal encounters over distance, and it has been widely used for distance estimation. The log-normal path loss model can be expressed as

$$RSSI(d) = P_{Tx} - PL(d_0) - 10\eta \log_{10} \frac{d}{d_0} - X \tag{4.1}$$

where $RSSI(d)$ is the received signal strength in dBm at a given distance d from the transmitter, P_{Tx} is the power in dBm of the transmitted signal, $PL(d_0)$ is the path loss at a reference distance d_0, and η is the path loss ratio. In Eq. (4.1), X is a random variable that reflects the noise in the signal strength due to different environmental factors such as reflection and fading.

4.3 Experimental Constraints

There were several constraints which were faced during field experiments. Though, omnidirectional antennas were used with sensor motes, the antenna and sensor motes deployment dimensions can affect the outcome. The intensity of such impact varies, increases or decreases the sensitivity of devices. Following are the constraints which were encountered with possible ways during experiments.

- Despite RF design is not the focus of the thesis, I have explored a number of different settings. Further, in my exploration, I tried several deployment dimensions of sensor motes and their antennas such as by mounting sensor motes on a stand and on the ground, in an open field and a railway station. Similarly, in a tunnel, sensor motes were placed in the middle of the tunnel ground, along tunnel wall and on stands. The best design according to high packet delivery rate was then incorporated for data collection experiments.

- For experiments in which large distance is considered between anchor sensors, the availability of volunteers, at the same time, was an uphill task. Moreover, I asked different group of friends to help in different times. Each time, a small training was given to each person about handling devices, sensitivity of directions and its impacts on results. Due to such issues, several experiments were redone with improved approach.

- Permission was required to conduct experiments on railway station, open field and tunnel. After permission, experiments were conducted on railway station, a verbal permission was taken, from personals of a sensitive instalment in the Ravensnbourne suburb of the Dunedin city, for experiments in an open field.

- The tunnel experiments were conducted in Lauder, central Otago. It was a remote tunnel (15 km walking track off the main road), which was once used by Kiwi rail and now this rail trail is used by cyclists. It was hard to get a team of volunteers ready on a common day for experiments.

- The experimental RSSI measurements at 40 m in tunnel with short range sensor devices at high transmission power level (PL31) are missing, shown in Figure 4.14(a). This was observed after coming back from experimental site while analysing data. The reason was due to malfunctioning of a particular sensor device. It was hard to conduct experiment again due to lack of logistics. Moreover, our findings suggest that data received at PL31 is not better than low power transmissions at PL7. So, the plan to re-conduct experiments for missing RSSI measurements at single distance point was dropped.

- Though RSSI can be recorded with small size preambles without requirement of a whole packet that contains information such as train location, but for consistency sake, whole packet is used as that is required in train-anchor communication.

- We assume antennas on the train will be mounted on the engine head so as to reduce possible impairments caused by the train itself. Due to safety and regulation constraints, on-track live experiments were limited. However, I managed to conduct experiments on railway station in the presence of trains. During experiments on the railway station, a train was at the platform and another one arrived at the station. I have compared the recorded RSSI measurements with and without the arriving train and found that the RSSI measurements were affected by reflections from large metallic bodies. However, the average RSSI was not found to have significant differences in either case.

4.4 Experiments in Railway Representative Environments

4.4.1 Wireless Sensor Platforms and Motes

To validate the feasibility of using RSSI for train localisation, experiments are carried out in three representative railway environments: an open field, a railway station, and a tunnel. In experiments, a series of Maxfor's MTM sensor platforms (MTM, 2012) are used that are equipped with the CC2420 radio chipset and different types of antennas. The antenna types of the motes used in the experiments include external dipole antennas (short range), external dipole antennas with amplifier (long range), and internal PCB antennas.

The long-range sensor device with amplifier is used in the open field experiments.

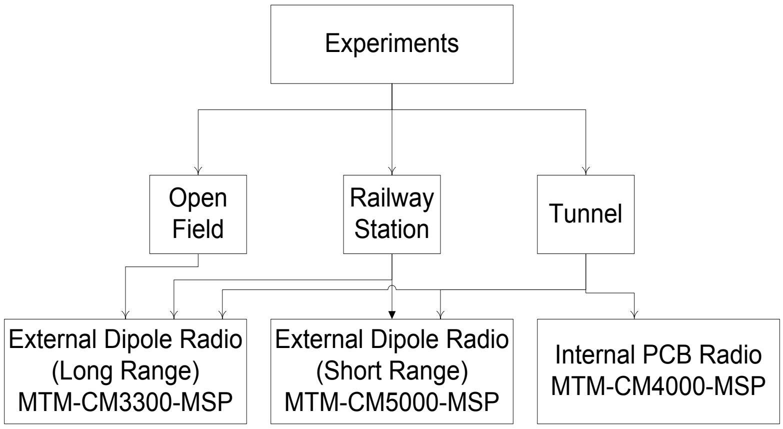


Figure 4.1: Trichotomy of Sensor Motes Used in Experiments

The purpose to use only a large-range sensor device is to use a minimum number of devices over large distances that are natural for the plain open fields. However, in the hilly terrains, the number of devices can increase to cover the dark spots in the coverage areas. The use of a fewer long-range sensor devices (transmits up to 800 m) is more cost effective than using a large number of short-range sensor devices (with short-range external dipole antennas or internal PCB antennas, transmits up to 150 m). In the railway station experiments, the datasets are collected using both short-range sensor devices with external dipole antennas and long-range sensor devices with external dipole antennas and amplifier. These devices are carefully selected for the railway station environment in accordance with the external dipole antenna dynamics, existing infrastructure and other operating radio frequencies. In the experimental setup of the tunnel environment, all three types of sensor devices are used as shown in Figure 4.1. The rock structure of tunnel causes multi-path fading, shadowing and intense signal reflections.

In the experimental setups, the anchor sensors are placed along a line with equal distance. Another sensor device called the gateway sensor is placed on a stand and connected to a laptop. Though a gateway sensor that is mounted on a metallic train is not equivalent to mounted on a stand, but experiments do incorporate existence of metallic trains during experiments on railway station. Therefore, it is the best closest experiments from real-world that was possible. The gateway can be moved to different locations. At each location, the gateway broadcasts a packet, and the anchors will send packets back to the gateway in sequence using different transmission powers after receiving the gateway's packet. The gateway then measures the RSSI of each packet transmitted by each anchor sensor, and the laptop will record these RSSIs as well as the corresponding transmission powers.

4.4.2 Analysis of Datasets

I use the collected RSSI dataset to validate if the received signal strength follows the log-distance path loss model in the three representative environments. In addition, log-distance path loss model is used to find out the path loss ratio. The path loss ratio (exponent) η is a key parameter in the log-normal path loss model, which varies in different environments. A method of least square fitting is used to compute the best η that minimises the sum of squares of the difference between the experimental RSSIs and their corresponding values in the log-distance path loss model, by solving the following optimisation problem:

$$\begin{array}{ll} \text{minimize} & \sum_{d\in\mathcal{D}} \sum_{i=1}^{n_d} (RSSI_e(d,i) - RSSI_m(d))^2 \\ \text{subject to} & 1 \leq \eta \leq 4, \end{array} \tag{4.2}$$

where $\mathcal{D}$ is the set of transmitter-receiver distance, and n_d is the number of RSSIs with transmitter-receiver distance of d. $RSSI_e(d,i)$ represents the i^{th} RSSI measurement with transmitter-receiver distance of d, and $RSSI_m(d)$ denotes the RSSI value computed based on log-distance path loss model with transmission distance of d. The value of η is usually in the range of 1 to 4. Here, RSSI noise is defined as the difference between each experimental RSSI and its corresponding value from the log-distance path loss model. Once getting the best η, Anderson-Darling test (Anderson and Darling, 1954) is applied on the dataset obtained at each location to characterise the distribution of RSSI noise, which is essential to select a feasible noise filter in designing localisation scheme.

4.5 Experiments in an Open Field Environment

In this section, the experimental setup of the open field is presented with the details of the devices used and their transmission ranges. Moreover, an analysis is presented in detail to discuss the nature of received dataset, log-normal model fitting and distribution of noise, to conclude the feasibility of using RSSI in the open field for the train localisation.

Figure 4.2: Experimental Tests in the Open Field Environment

4.5.1 Experimental Setup

In this set of experiments, the MTM-CM3300-MSP sensor motes with external Dipole antennas are used, which have a transmission range around 800m. As shown in Figure 4.2, the anchor sensors are deployed along a track of 700m, and the distance between two adjacent anchor sensors is 25m. A gateway sensor is attached to a laptop to record the RSSI measurements. The gateway sensor transmits a beacon packet and anchor sensors reply with a number of packets. The gateway sensor records the RSSI of each received packet along with the other data it contains.

4.5.2 Analysis of Open Field Datasets

Figure 4.3 shows the collected RSSI measurements as well as the best-fitted curve. The best-fitted log-normal path loss curve is obtained as discussed, in Equation 5.22. The black vertical stripes in the figure show the experimental RSSI collected from packets received from the anchor sensor at each location. The best log-normal model curve is obtained with $\eta = 3.8$, $d_0 = 25$, and $P_T - PL(d_0) = 1.2\ dBm$. It can be seen that, there are some fluctuations and the RSSI measurements follow the curve of log-normal path loss model with fluctuations. The variations in RSSI measurements are attributed to both the multi-path fading and the $\pm 6\ dBm$ error margin in RSSI measurement for CC2420 radio transceiver (Texas Instruments, 2003). These and other sources of errors induce noise in RSS measurements, which may compromise the use of RSS as

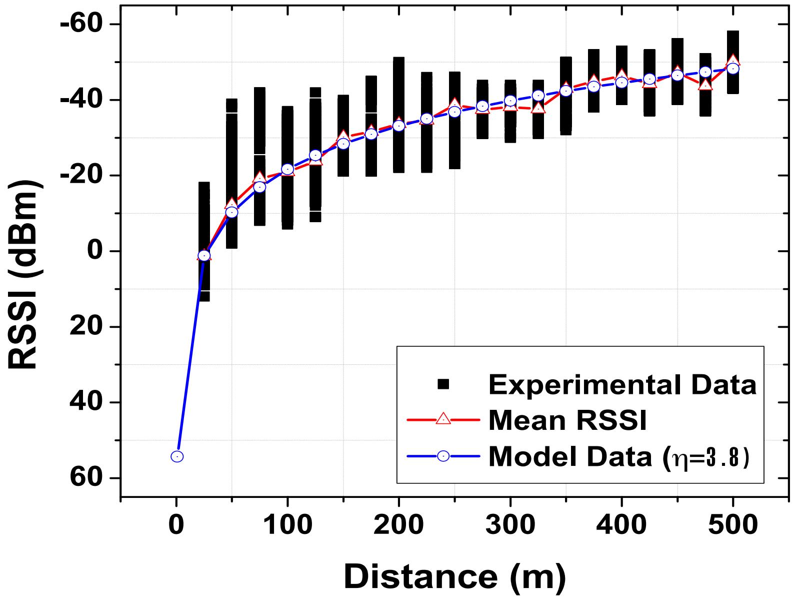
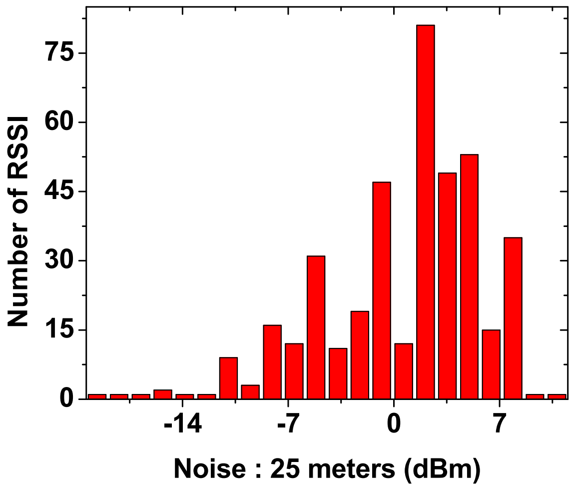


Figure 4.3: Model Fitting for Open Field Experiments

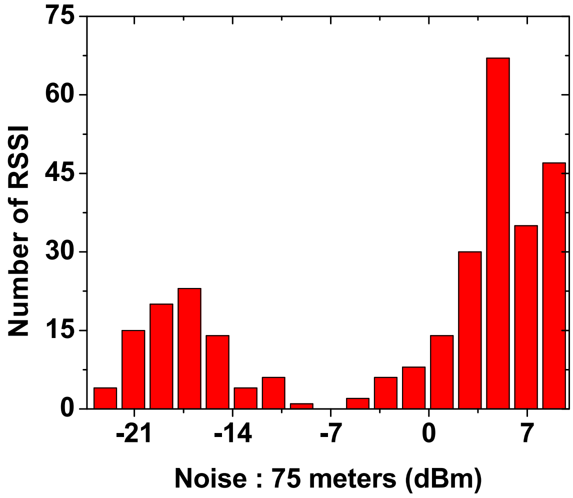
an estimator. Therefore, measurement model and a noise filtering algorithm, such as Particle Filter is required for location estimation.

Figures 4.4 and 4.5 show the noise distribution of the RSSI measurements collected from anchor sensors at selected locations of the total deployment area. Figure 4.4(a) explains the noise distribution for the RSSI measurements collected with a transmission distance of *25 m*. The maximum absolute deviation goes up to *18 dBm*, but *95%* of the noise is in the range between -8 *dBm* and *8 dBm*. The noise distribution for transmission distance of *75 m* is shown in Figure 4.4(b). It can be seen that there are two peaks in the figure, which represent the signal reflections and multi-path fading due to the surrounding infrastructure. The impact of such signal deteriorating factors can be seen in the noise distribution, where the absolute deviation reaches to *21 dBm*. However, still the maximum number of RSSI measurements lie between -8dBm to 8dBm. The noise distribution of the RSSI measurements collected from anchor sensors that are deployed at the distance of *125 m* and *175 m* are shown in Figures 4.4(c) and 4.4(d), respectively. Similarly, the noise distribution of RSSI measurements from anchor sensors deployed at distance of *225 m*, *275 m*, *325 m*, *375 m*, *425 m*, and *475 m* are shown in Figures 4.4(e), 4.4(f), 4.5(a), 4.5(b), 4.5(c), and 4.5(d), respectively.

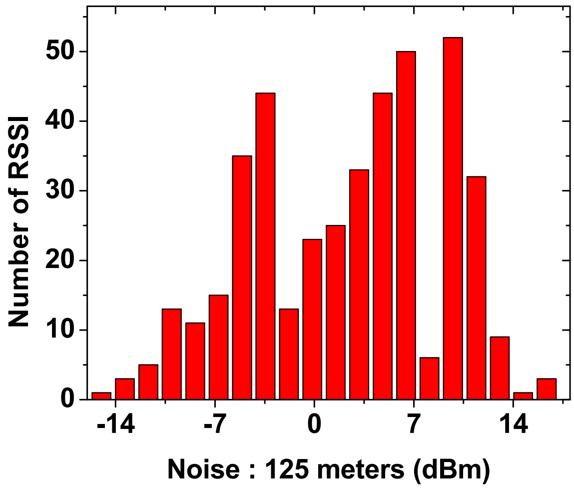
The observed behaviour shows that the noise distribution for larger transmission distance shows much smaller deviations. However, the short-range transmissions are more prone to multi-path fading caused by signal reflections from the surrounding en-



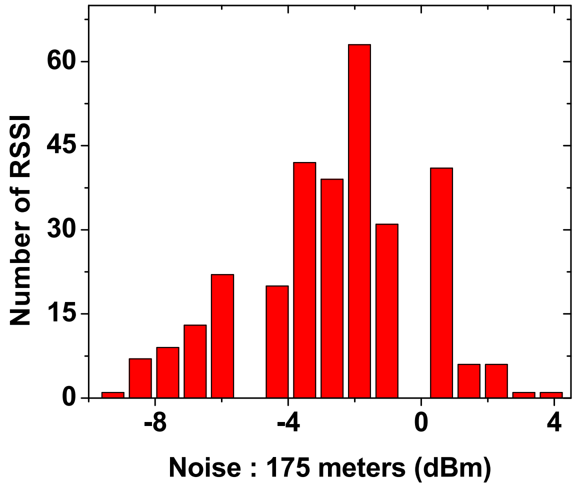
(a) Noise (25m)



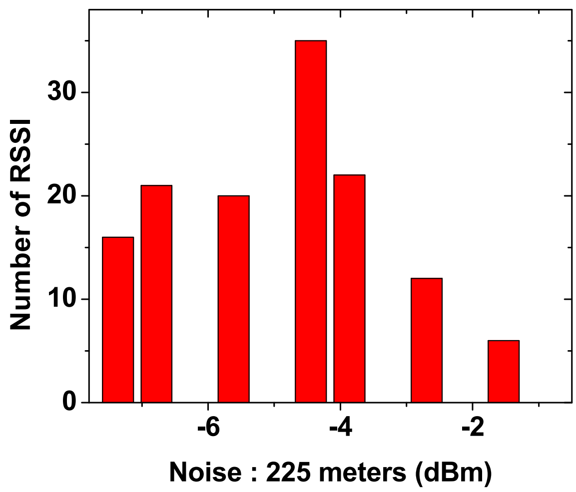
(b) Noise (75m)



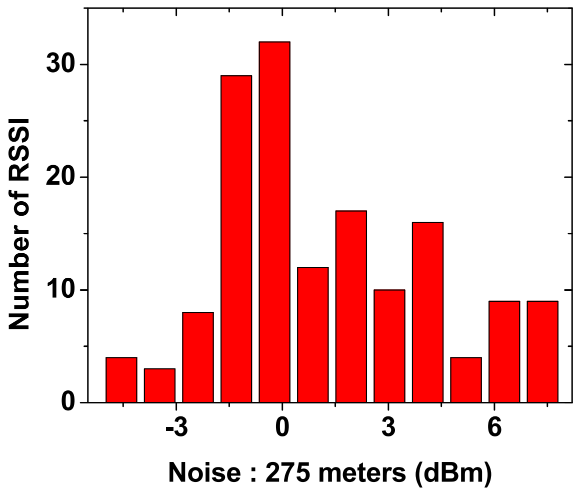
(c) Noise (125m)



(d) Noise (175m)

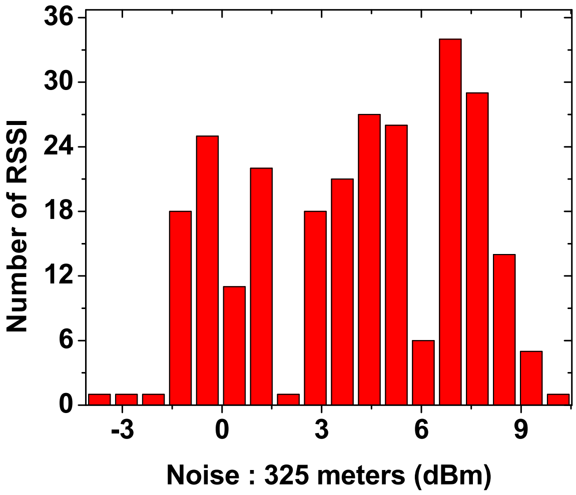


(e) Noise (225m)

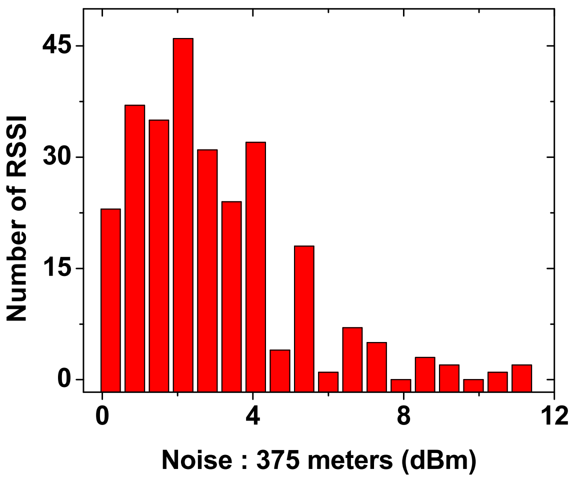


(f) Noise (275m)

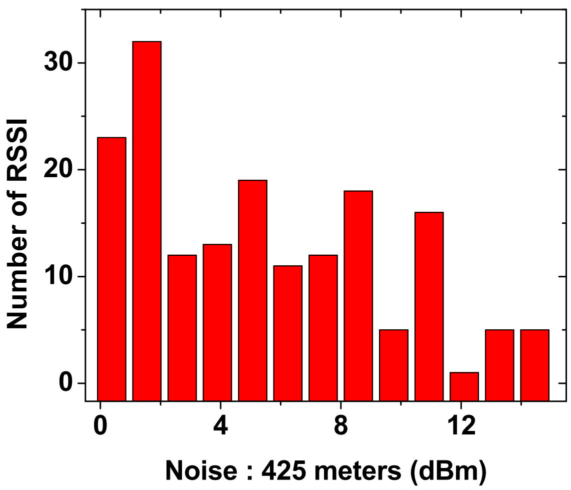
Figure 4.4: Open Field Noise Distribution (Figure 1)



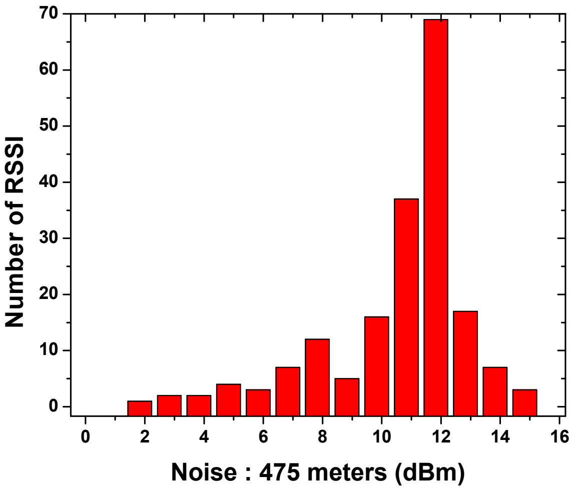
(a) Noise (325m)



(b) Noise (375m)



(c) Noise (425m)



(d) Noise (475m)

Figure 4.5: Open Field Noise Distribution (Figure 2)

vironmental objects such as the neighbouring guard rail, tracks and buildings. After further analysing the noise distribution for measurements collected with other transmission distances in open field, it is not found from normal distribution. The Figures 4.4 and 4.5 reveal that none of the RSSI noises follows the normal distribution. The noise distribution describes the deviation of RSSI measurements from the log-normal model RSSI. This deviation pattern is essential to be known to select a noise filter method in the design of train localisation algorithm. The non-Gaussian distribution of the noise is attributed several factors such as outlier RSSI measurements, constructive or destructive interference of signal strength by reflected signals, data discrimination by hardware error limits, RSSI measurements lost due to the environment that might have contributed as inliers, and redundant RSSI measurements. Therefore, non-Gaussian noise is more realistic in a railway environment, which means Particle Filter can be used for distance estimation.

The RSSI measurements received from closest anchor sensor is stronger than anchor sensors at large distance from gateway sensor. The rate of change of received signal strength decreases in the measurements received from farther locations. This implies that RSSI may not be trustworthy for distance estimations over large distances. However, as RSSI follows log-normal path loss model, it is possible to increase accuracy on train localisation by combining it with location information of anchor sensors. Such data fusion minimises the error range and increases the weightage of more correct RSSI values, resulting in a robust localisation scheme.

4.6 Experiments in Railway Station Environment

In the second set of experiments, anchor sensors are deployed in the railway station environment. This section presents the details of this experiment setup such as devices used and their transmission ranges. Moreover, an analysis is presented in detail to discuss the nature of received dataset, log-normal model fitting and distribution of noise, to conclude the feasibility of using RSSI in the railway station for the train localisation.

4.6.1 Experimental Setup

Railway Station is a representative environment that needs to be considered for train localisation due to the presence of trains, railway tracks, platform offices and other infrastructures that may affect the propagation of the wireless signals. The existing

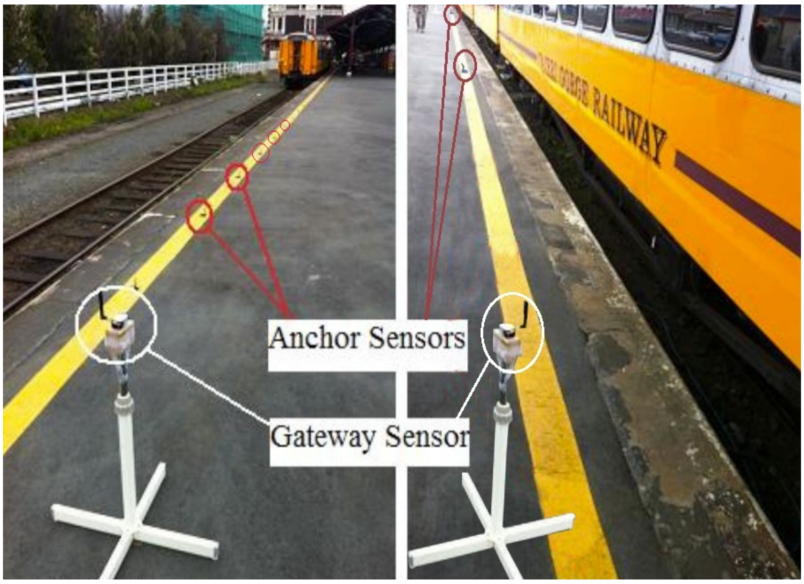


Figure 4.6: Experimental Tests in the Railway Station Environment

solutions are generally based on GPS and track circuits. GPS may not be available on remote railway stations and track circuits are not precise and its failure has caused several accidents in the past (Johnson, 2016; Wikil, 2016). Therefore, railway station is an important environment for localisation. As shown in Figure 4.6, two sets of experiments are carried out to validate the feasibility of using RSSI for distance estimation in a railway station environment: (a) *sparse deployment*, in which the MTM-CM3300 motes with long-range external Dipole antennas (with amplifier) are used, and the distance between two adjacent anchor sensors is *25 m*; (b) *dense deployment*, in which the MTM-CM5000 motes with external Dipole antenna (without amplifier) are used. The maximum transmission range of MTM-CM5000 mote is around *150 m*, and the distance between two adjacent anchor sensors is *2 m*. The 2 m distance is expensive deployment, but here it is considered due to several reasons, such as, to counter the large number of obstacles in congested areas within railway stations and to minimise the negative effects on radio signals. As the area of railway stations is limited as compared to open field, such dense deployment can be used. Further, deployment density within railway station can be decreased depending on specific dynamics of particular railway station.

4.6.2 Analysis of Railway Station Datasets

Figure 4.7 shows the relationship between the raw RSSI measurements, the mean RSSI, and the best log-normal path loss model curve for sparse deployment. The black

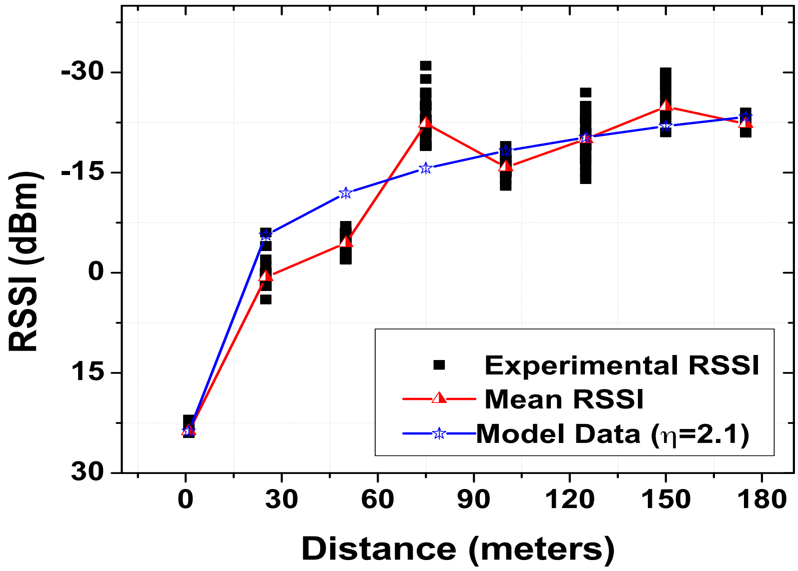
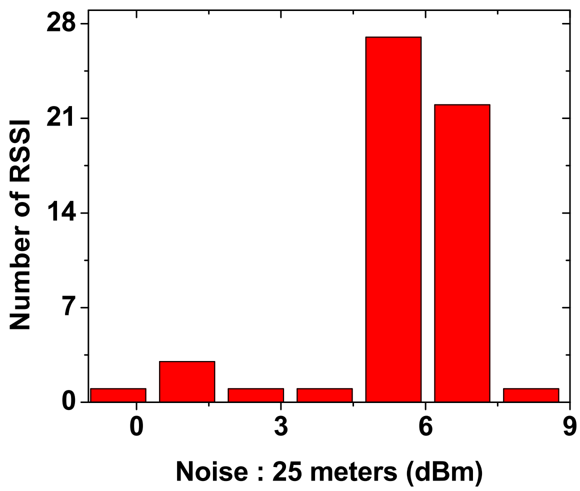


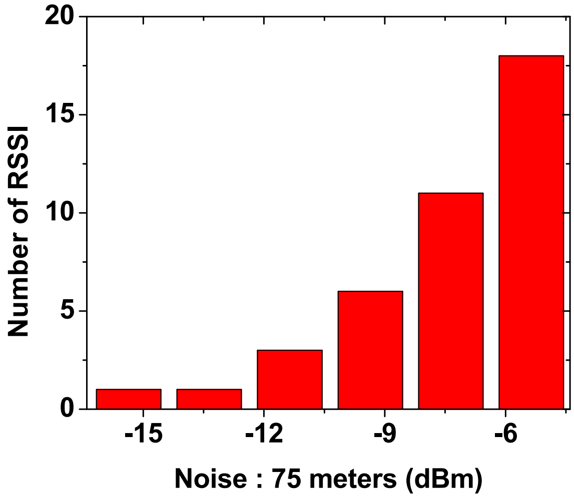
Figure 4.7: Model Fitting with Sparse Deployment for Railway Station Experiments

vertical stripes in the figure show the experimental RSSI collected from packets received from the anchor sensor at each location. The best log-normal model curve is obtained with $\eta = 2.1$, $d_0 = 25$ and $P_T - PL(d_0) = 0.66\ dBm$. It can be seen that the RSSI measurements roughly follow the log-normal model due to larger variations, especially for the measurements collected for short transmission distances (large difference at two points between model and mean RSSI). In this case, RSSI to distance estimation and vice versa is not equivalent because of noise. Therefore, the use of RSSI measurements only for location estimation will be compromised. Another set of measurements, such as location data can be helpful in location estimation.

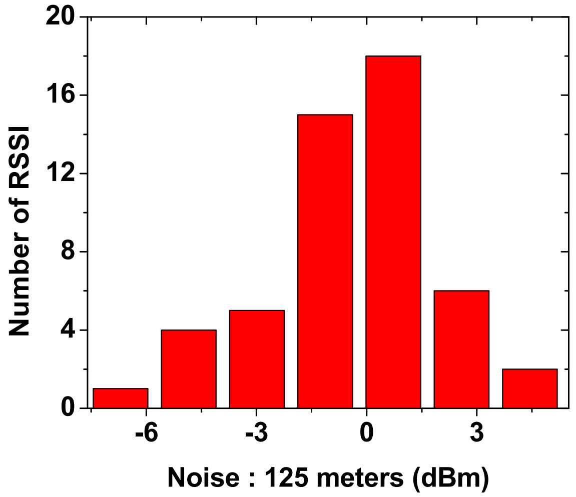
Figures 4.8(a), 4.8(b), 4.8(c), and 4.8(d) show the noise distribution for RSSI data collected from anchor sensors at transmission distances of $25\ m$, 75, $125\ m$ and $175\ m$, respectively. The observation about the noise distribution is the same as the findings in the open field experiments, which is, the RSSI data collected from long distance transmissions has smaller average deviations as shown on the x-axis. The absolute noise variation can be seen on x-axis, which is low, around 8 dBm at $25\ m$ and 2 dBm at $175\ m$, but the number of RSSI measurements (histogram peaks) at different noise levels do not follow any pattern. Such irregular behaviour suggests the distribution of noise as non-Gaussian. This inferred result is then further verified by the application of Anderson-Darling normality test, which suggests that none of the noise distributions lie in the definition of the Gaussian distribution. The noise distribution is important to select noise filter such as Particle Filter.



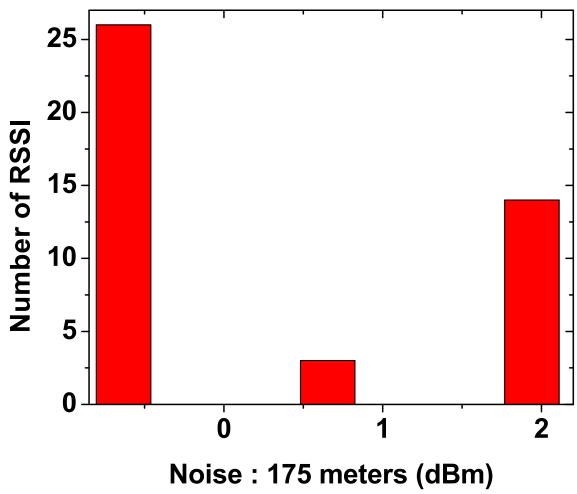
(a) Noise (25m - Sparse)



(b) Noise (75m - Sparse)



(c) Noise (125m - Sparse)



(d) Noise (175m - Sparse)

Figure 4.8: Railway Station Noise Distribution with Sparse Deployment

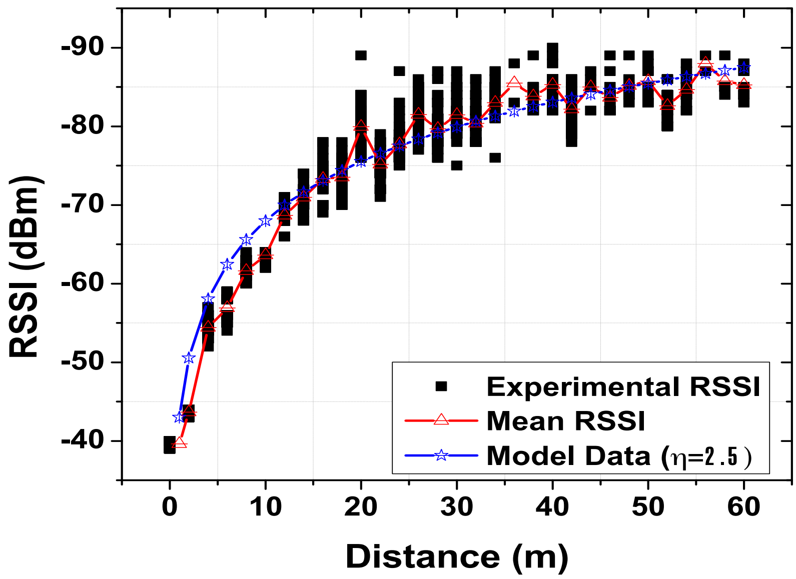
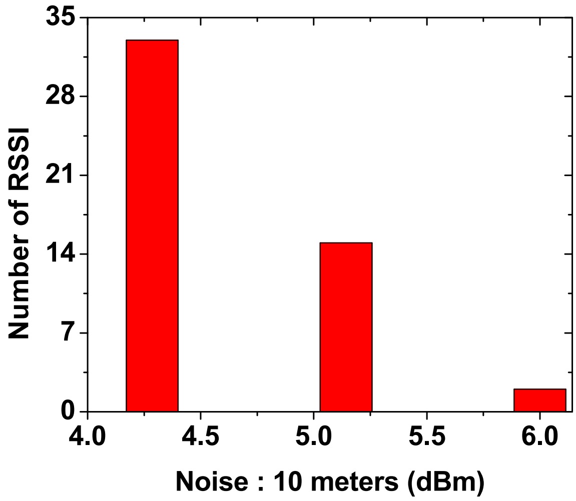


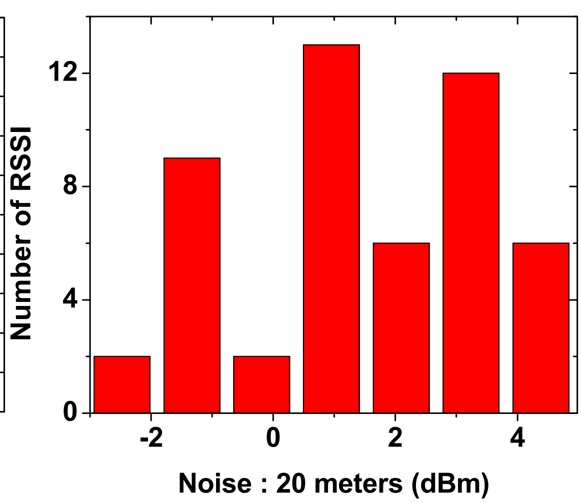
Figure 4.9: Model Fitting with Dense Deployment for Railway Station Experiments

Figure 4.9 shows the relationship between raw RSSI, mean RSSI, and the best log-normal model curve for dense deployment. Like model fitting graphs of other datasets, the black vertical stripes in the figure show the experimental RSSI collected from packets received from the anchor sensor at each location. The best log-normal curve is obtained with $\eta = 2.5$, $d_0 = 1$ and $P_T - PL(d_0) = -39.62\ dBm$.

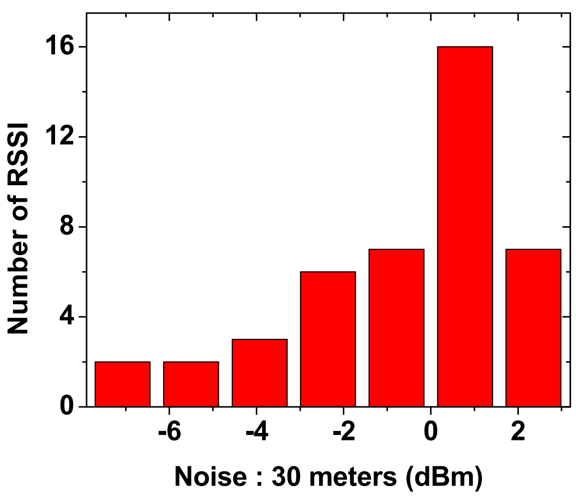
Figures 4.10(a), 4.10(b), 4.10(c), 4.10(d), and 4.10(e) show the noise distribution for RSSI data collected from anchor sensors at transmission distances of *10 m*, *20 m*, *30 m*, *44 m*, and *54 m*, respectively. It can be observed that the noise variation is much smaller in comparison with sparse deployment because amplified power transmission is more robust to signal reflections. It can be seen that with the increase of the distance, the maximum number of RSSI measurements, which is shown in the y-axis scale, decreases. This trend means that several low power multi-path transmissions either do not reach the receiver or stays under acceptable power level to be considered as a transmission. We used Anderson-Darling Normality test to examine the noise distribution, which suggests that none of them follows the normal distribution. It can be seen that even in railway station environment, in almost all cases, the anchor sensor that is closest to the gateway sensor gets the strongest RSSI and signal strength reduces from the anchor sensors deployed at farther places. Log-normal path loss model also suggests that received signal strength decreases with the square of the distance between transmitter and receiver, which validates that RSSI follow log-normal model, though there are large differences at a few points between model and mean RSSI. This



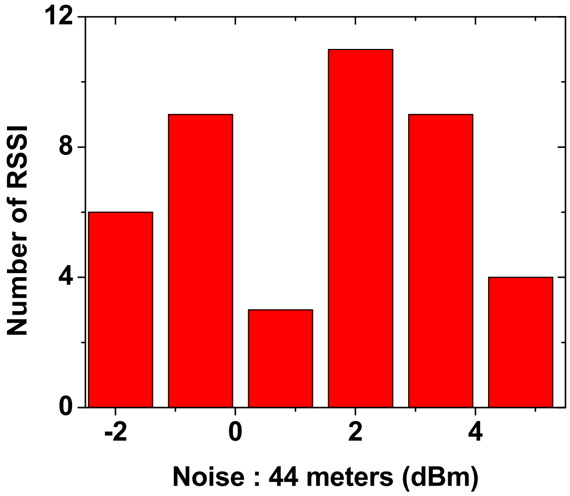
(a) Noise (10m - Dense)



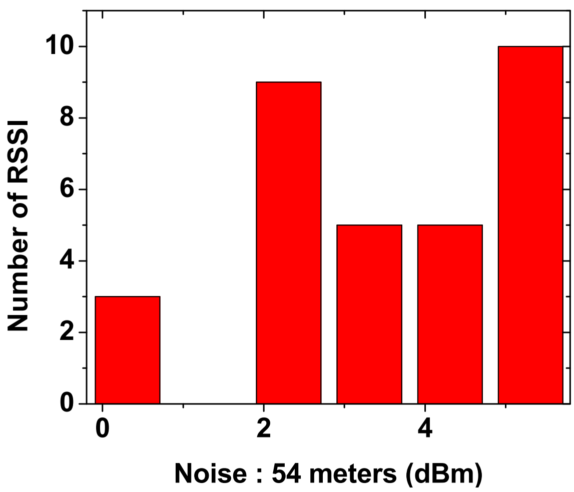
(b) Noise (20m - Dense)



(c) Noise (30m - Dense)



(d) Noise (44m - Dense)



(e) Noise (54m - Dense)

Figure 4.10: Railway Station Noise Distribution with Dense Deployment

observation suggests that RSSI may not be an effective metric for distance estimation and therefore, noise filtration algorithms such as Particle Filter will be required for distance estimation.

4.7 Experiments in Tunnel Environment

In the third set of experiments, anchor sensors are deployed in the tunnel environment. This section presents the details of this experimental setup such as devices used and their transmission ranges. Moreover, an analysis is presented to discuss the nature of received dataset, log-normal model fitting and distribution of noise, to conclude the feasibility of using RSSI in the tunnel for the train localisation.

Figure 4.11: Experimental Tests in the Tunnel Environment

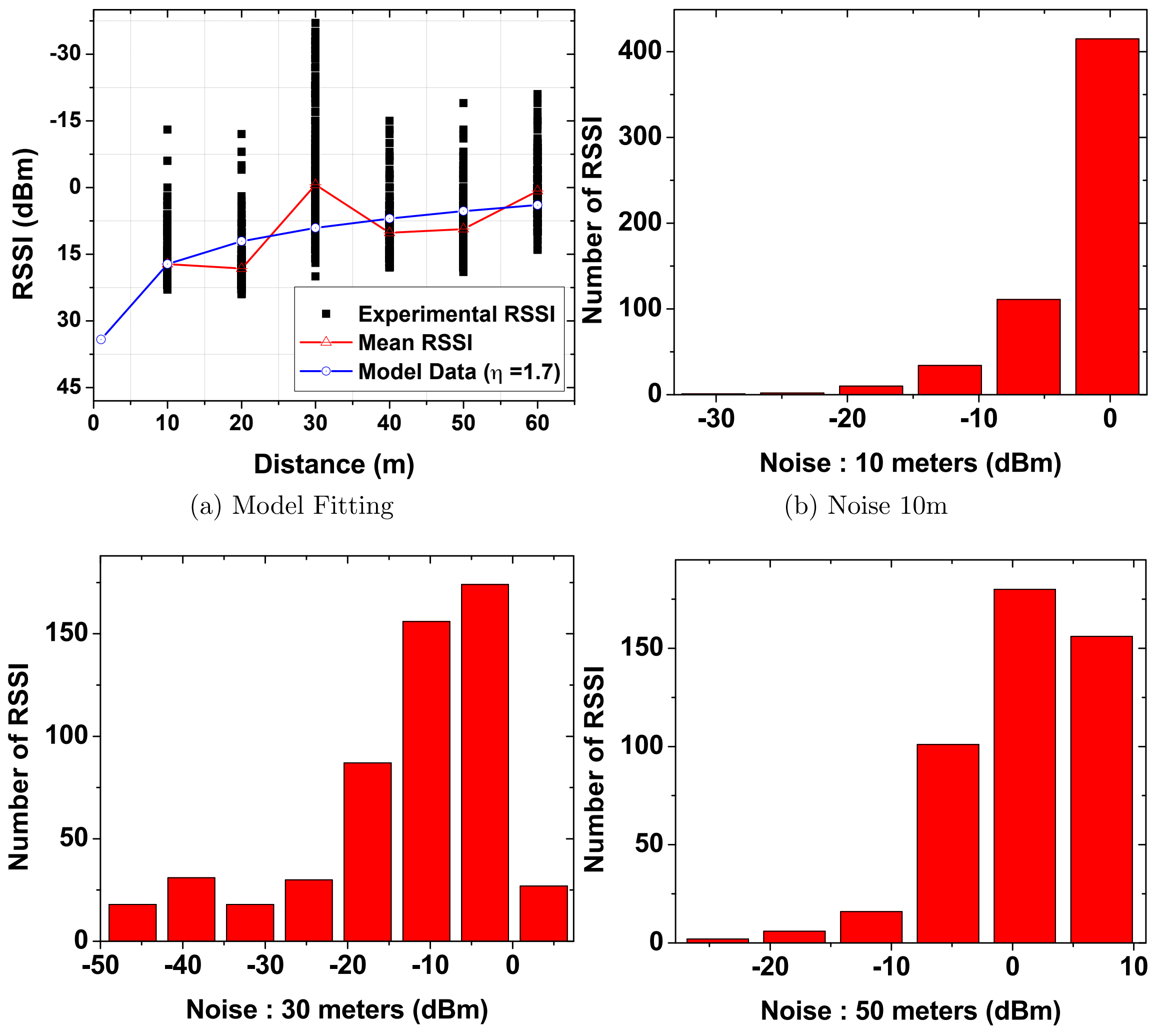
4.7.1 Experimental Setup

A tunnel is another representative environment that needs to be considered for train localisation due to its distinctive characteristics such as the absence of GPS signals and serious multi-path fading. The rough rock structure can cause serious signal reflections, and each copy of the signal may experience different attenuation, delay and phase shift, thereby resulting in large variations on RSSI measurements. In such environments, a long-range transmission with a large transmission power can lead to much serious signal reflections. However, to study the impacts of several sensor motes, the following three types of sensor platforms are used: (a) MTM-CM3300 platforms with external Dipole antennas (with amplifier) that have maximum transmission range of around *800 m*; (b) MTM-CM5000 platforms with external Dipole antennas (without amplifier) that have maximum transmission range of around *150 m* (c) MTM-CM4000 platforms with internal antennas that have maximum transmission range of around *150 m*. As shown in Figure 4.11, anchor sensors along the central line of the tunnel are deployed, and the distance between two adjacent anchors is *10 m*. In these experiments, deployment was tested with several patterns such as sensors on the side of tunnel, on stand in the middle and the one which is discussed here, that is, on the central line. The experiments received large number of RSSI measurements in central line deployment. These issues are discussed in experimental concerns in section 4.3. In the experiments, the transmission power levels are adjusted to investigate the impact of transmission powers on RSSI measurements. The highest transmission power is at PL31, which operates at 0 *dBm*, and lowest transmission powers is at PL7, which operates at -15 *dBm*.

4.7.2 Analysis of Tunnel Datasets

Figure 4.12(a) shows the RSSI measurements and the best log-normal path loss fitting on the data collected with external Dipole antennas with amplifier using high transmission power at power level PL31. The black vertical stripes in the figure show the experimental RSSI collected from packets received from the anchor sensor at each location. The best log-normal model curve is obtained with $\eta = 1.7$, $d_0 = 10$ and $P_T - PL(d_0) = 12.5\ dBm$. It can be seen that the RSSI measurements roughly follow the log-normal model with larger variations. An important common observation between experimental RSSI measurements and log-normal path loss model based RSSI values is that, received signal strength reduces with the increase of the distance between receiver and transmitter. The large fluctuations are attributed to tunnel properties. However, there is need to improve RSSI as an estimator by combining it with other data such as anchor sensor locations to filter out noisy measurements. The noise distribution is calculated in the RSSI measurement collected from the anchor sensors at each location. Figures 4.12(b), 4.12(c), and 4.12(b) show the noise distribution for RSSI data at transmission distance of *10 m*, *30 m*, and *50 m* at *PL31*, respectively.

Similarly, the experiment is repeated with the same set of sensor motes to analyse the impact of low power transmissions at power level PL7. The Figure 4.13(a) shows the RSSI measurements and the best log-normal path loss fitting for the data collected with external Dipole antennas with amplifier using high transmission power at PL7. The best log-normal model curve is obtained with $\eta = 1.8$, $d_0 = 10$ and

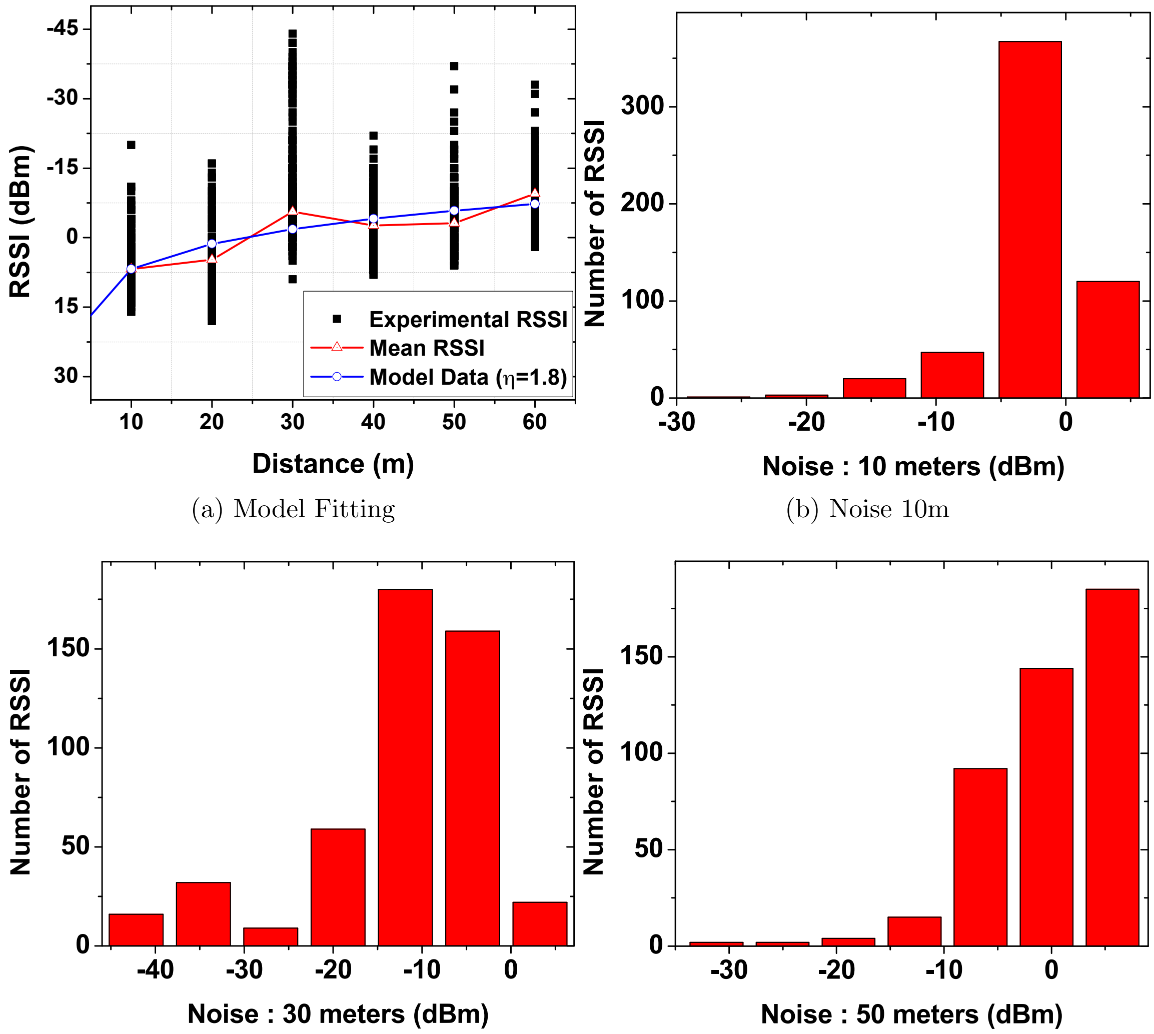


(a) Model Fitting
(b) Noise 10m
(c) Noise 30m
(d) Noise 50m

Figure 4.12: Tunnel Experiments with Long Range Sensor Devices at PL31

$P_T - PL(d_0) = 6.8\ dBm$. It can be seen that the RSSI measurements follow the log-normal model more closely compared to Figure 4.12(a), even though there are still prohibitively large variations. However, the large density of RSSI measurements are close to the mean RSSI, and large number of deviated RSSI measurements are not received due to low power transmissions at PL7, which were received in PL31 transmissions. This observation suggests that largely deviated RSSI measurements are not recorded due to signal strength beyond acceptable level. The tunnel structure results as a wave guide and leads to a large number of signal reflections. Therefore, the use of low power transmissions significantly reduces the number of signal reflections.

The RSSI measurements and the analysis of the noise distribution with the short-range external dipole antenna MTM-CM5000 motes are shown in Figures 4.14 and



(a) Model Fitting

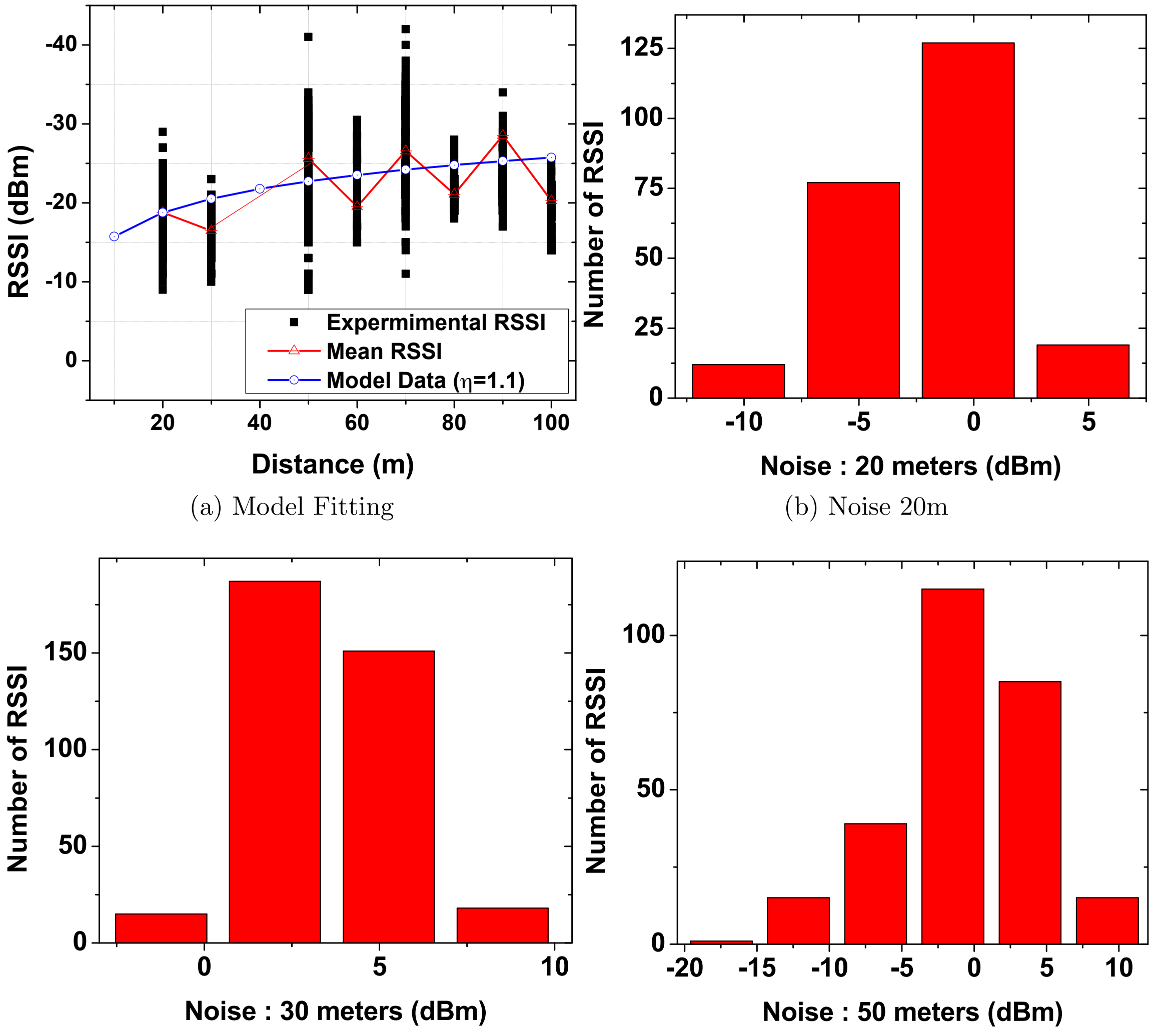
(b) Noise 10m

(c) Noise 30m

(d) Noise 50m

Figure 4.13: Tunnel Experiments with Long Range Sensor Devices at PL7

4.15 at PL31 and PL7, respectively. The missing RSSI measurements at distance 40m in figure of model fitting at PL31 is discussed in section 4.3. Figure 4.14(a) shows the RSSI measurements and the best log-normal path loss fitting for the data collected with external Dipole antennas using high transmission power at power level PL31. The best log-normal model curve is obtained with $\eta = 1.1$, $d_0 = 20$ and $P_T - PL(d_0) = -19.1\ dBm$. It can be seen that the RSSI measurements follow the log-normal path loss model with significant variations and can not be used as an estimator directly. However, to improve localisation accuracy, location coordinates of anchor sensors can be combined with RSSI measurements to lower weights of noisy RSSI measurements. Figures 4.14(b), 4.14(c), and 4.14(d) show the noise distribution with transmission distance of $20\ m$, $30\ m$ and 50 m at $PL31$, respectively. The variations in the noise increase in the RSSI data collected from anchor sensors at the farther locations, which means multi-path fading increases due to a large number of signal reflections. The RSSI measurements are severely affected by large number of signal reflections by dense deployment of anchor sensors. The reflected signals are directly proportional to the number of transmitters. In this set of experiments, *10* sensors transmit their signals to the gateway sensor and result in the increase of multi-path fading because the number of paths for each transmitter accumulate to increase the multi-path fading.



(a) Model Fitting (b) Noise 20m

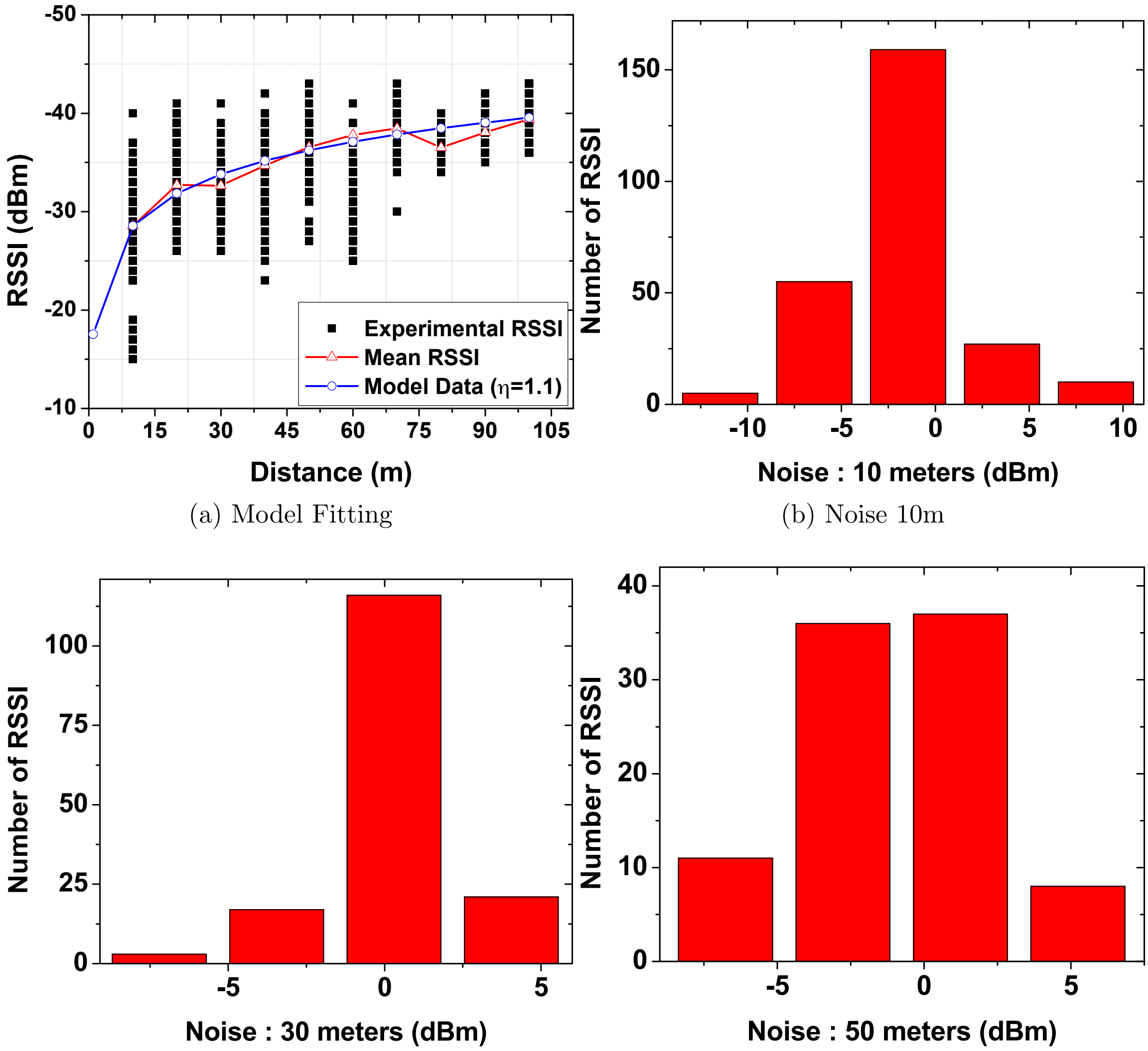
(c) Noise 30m (d) Noise 50m

Figure 4.14: Tunnel Experiments with Short Range Sensor Devices at PL31

The RSSI dataset is also collected with the same motes but through low power transmissions to study the impact of reducing the transmission power on the noise distribution in the RSSI data. Figure 4.15(a) shows the RSSI measurements, mean

RSSI and the log-normal path loss model curve for low power transmissions at power level *PL7*. The best log-normal model curve is obtained with $\eta = 1.1$, $d_0 = 10$ and $P_T - PL(d_0) = -28.5\ dBm$. It can be seen that the RSSI measurements follow the log-normal path loss model more closely compared to transmissions at high power level PL31 as given in Figure 4.14(a). The variations in the RSSI, length of black vertical stripes, decrease with the increase in the distance between anchor and the gateway. However, the deviation of mean RSSI from the log-normal curve decreases with the decrease in the transmission power, thus implying the reduction of signal reflection and scattering into the multiple paths. Figures 4.15(b), 4.15(c), and 4.15(d) show the noise variation in RSSI measurements received at *PL7* from anchor at *10 m*, *30 m*, and *50 m* distance from the gateway. The noise variation, x-axis scale, in RSSI from anchor sensor at *50 m* is reduced to *15 dBm* (at PL7) from *31 dBm* (at PL31). Therefore, noise variation reduces significantly with the transmission power and low transmission power results in low interference. According to the Anderson-Darling Normality test, none of the noise distributions at any power level follows a normal distribution. RSSI from low power transmissions follows the log-normal path loss model more closely as compared with high power transmissions. It can be seen that, despite signal reflections, anchor sensors closest to the gateway get the strongest RSSI and signal strength reduces from the anchor sensors deployed at farther places, which validates that RSSI follows log-normal model and can be used for distance estimation even in tunnel environments. However, as RSSI is prone to signal reflections and other signal deterioration factors, it may not be useful to use RSSI alone as distance estimator. There is a need to eliminate the noisy measurements by using another type of data, that is, location coordinates. Therefore, RSSI, as an estimator, needs to be improved by using Particle Filter.

Another type of motes, the MTM-CM4000, are used to collect the RSSI measurements and to perform the analysis of the noise distribution. This device mentioned above operates with PCB internal antennas. The results are shown in Figures 4.16 and 4.17 for PL31 and PL7, respectively. Figure 4.16(a) shows the RSSI measurements and the best log-normal fitting for internal antennas where the transmission power is set to the maximum power level (PL31). The best log-normal path loss model curve is obtained with $\eta = 1.2$, $d_0 = 10$, and $P_T - PL(d_0) = -9.5\ dBm$. Figures 4.16(b), 4.16(c), and 4.16(d) show the noise distribution with transmission distance of *10 m*, *30 m* and *50 m* at *PL31*, respectively. It can be seen that RSSI measurements collected from farther anchors tend to have larger variations. The signals from long-range transmissions can experience more reflections, and each reflected copy with different



(a) Model Fitting

(b) Noise 10m

(c) Noise 30m

(d) Noise 50m

Figure 4.15: Tunnel Experiments with Short Range Sensor Devices at PL7

signal strength influences the original signal over the time it takes to reach the gateway.

Similarly, for low power transmission, Figure 4.17(a) shows the results obtained with a lower transmission power at PL7. The best log-normal path loss model curve is obtained with $\eta = 1.8$, $d_0 = 10$, and $P_T - PL(d_0) = -23.61\ dBm$. It can be seen that the RSSI measurements follow the same trend, but the noise is much smaller in comparison with large transmission powers. Figures 4.17(b), 4.17(c), and 4.17(d) show the noise distribution with transmission distance of $10\ m$, $30\ m$ and $50\ m$ at $PL7$, respectively. In the Figure, 4.17(c), the noise distribution with transmission distance of $30\ m$ at low power transmission ($PL7$) is shown. Moreover, in this figure, the noise variation is decreased to $19\ dBm$ ($-10\ dBm$ to $9\ dBm$) in RSSI measurements received by low power transmission compared with noise variation in high power transmission

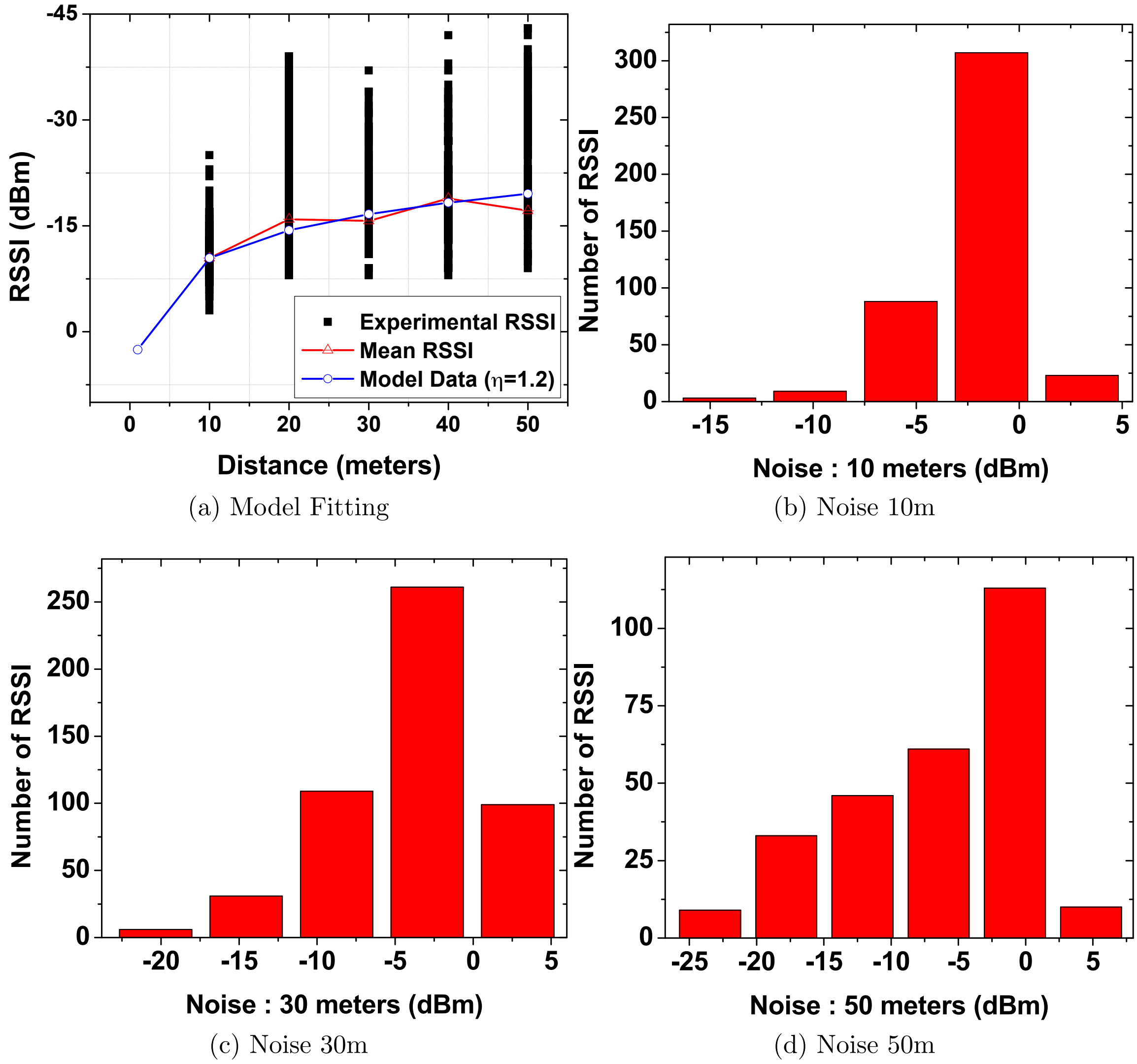
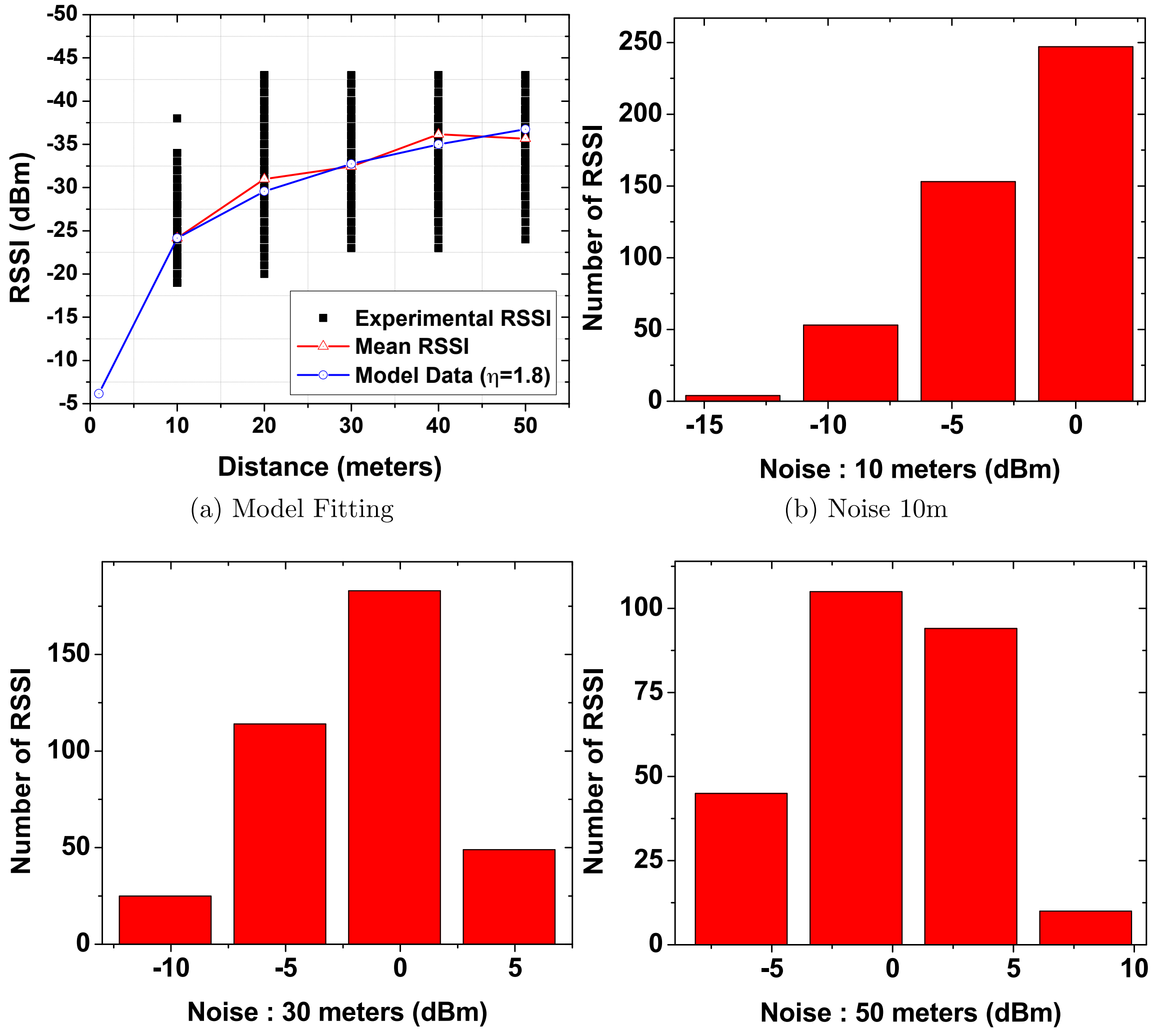


Figure 4.16: Tunnel Experiments with Internal Radio Sensor Devices at PL31

(PL31) which is *27 dBm* (*−18 dBm* to *9 dBm*) in Figure 4.16(c). According to the Anderson-Darling Normality test, none of the noise distributions at any power level follows a normal distribution. Though RSSI follows the log-normal path loss model while using high power transmissions, it has more fluctuations compared with low power transmissions. Therefore, the low power transmission can yield more accurate distance estimation.

4.8 Concluding Remarks

The feasibility analysis of the datasets collected in the experiments conducted in all three railway environments yields the following observations: (a) In an open field en-



(a) Model Fitting (b) Noise 10m

(c) Noise 30m (d) Noise 50m

Figure 4.17: Tunnel Experiments with Internal Radio Sensor Devices at PL7

vironment, the use of long-range sensor motes, MTM-CM3300, are feasible to cover large transmission areas by a few devices. The RSSI data obtained from closer sensors from the gateway are more trustworthy, but it may not serve the purpose of sole estimator and there will still be need of another type of data to become a useful distance estimator; (b) In the railway station environment, the use of short-range sensor motes, MTM-CM5000, is more feasible to be used with dense deployment settings. The collected datasets have better log-normal model data fit with fewer fluctuations compared with the datasets obtained through long-range sensor motes, but it is still not precise and there is need to incorporate secondary data, anchor sensor locations, to develop better estimator; (c) In the tunnel environments, use of internal antenna sensors has a better fitting of RSSI measurements with the log-normal path loss model, because the internal antenna has a gain of *4.5 dBi* (Andersen, 2008) whereas the external Dipole antenna has a gain of *1.9 dBi* (Jonsrud, 2008). Therefore, internal antennas have better signal reception and low power transmission energy loss compared with external antenna sensors. Though outcome of sensors with internal radios are better options among existing devices, it is still a rough outcome and needs filtering algorithm for train localisation; (d) The noise variation increases with the distance from the gateway because long distance transmissions are more prone to reflections from tunnel walls; and (e) The noise variation reduces in the RSSI measurements received using low power transmissions from the anchor sensor at the same distance. Therefore, the use of low power signals are useful for more accurate distance estimation.

The experimental results in the above three environments demonstrate that, by choosing proper devices and appropriate configurations, the difference between experimental RSSI and log-normal path loss model improves. The experiments also show that RSSI measurements are noisy, which implies that the RSS is not a good choice for distance estimation due to its fragile nature. Therefore, there is a need to filter the noisy measurements to improve RSSI as an estimator. The location coordinates of anchor sensors can help to determine the noisy measurements while using a Particle Filter. Such data fusion improves feasibility of using RSSI as distance estimator and becomes a good metric for distance estimation for train localisation.

4.9 Summary

The railway environment is significantly different from other open field environments, because it is harsh and involves metals, rough terrains, a large number of uneven surfaces for multi-path fading and interference from other frequency channels (WLAN or microwave on railway station or tunnel). Therefore, in this chapter, I conducted experiments to collect the RSSI datasets to fulfil the need to validate the feasibility of using RSSI for distance estimation in train localisation system. The extensive experiments are conducted with the variation of several sensor motes and power levels in open field along the railway track, railway station and tunnel environments.

In the remainder of the chapter, I performed an analysis on the recorded datasets from each of the railway environment setup to get log-normal path loss model fitting, which is a well-known signal propagation model used for distance estimation, by using different path loss ratios and minimum mean square error (MMSE) method. In future, with improved experimental designs, it needs to be examined against multiple configurations per environment. Further, an analysis including a normality test is performed to evaluate the noise distribution of RSSI data received from anchor sensors at each location. It is observed that though the RSSI datasets are noisy and non-Gaussian, it still follows the log-normal path loss model with fluctuations, which validates the idea that RSSI can be used effectively in fusion with other measurements (geographic coordinates of anchor sensors) for distance estimation by using noise filtration technique.

Finally, I presented some experimental observations to select the most feasible dataset from each railway environment experiments. These datasets are then used in later chapters for WSN-based train localisation simulations. In the next chapter, I shall present the proposed beacon-based anchor sensors' wake-up scheme.

Chapter 5

Beacon-driven Wake-up Scheme for Train Localisation using Wireless Sensor Networks

In this chapter, I shall present the beacon-based wake-up scheme for anchor sensors in the absence of a train' schedule. I first give the upper bound of sleep time which can guarantee the wake-up of anchor sensors when the train arrives. Then, I present the energy analysis of the proposed scheme. In the remainder of the chapter, I describe the simulation setup and results to verify the feasibility of the scheme.

5.1 System Model

The wireless sensor network consists of two types of sensor nodes: anchor sensors and gateway sensors, as shown in Figure 5.1. A set of anchor sensors $\{a_0, a_1, ..., a_n\}$ are uniformly deployed along a straight track with equal distance d_a between any two consecutive anchor sensors. The uniform deployment of anchor sensors offer several benefits, such as, it results in uniform battery drainage, network life increases, a few nodes can provide coverage of target area and it is considered to be a non-complex deployment strategy (Bendigeri and Mallapur, 2015). Each anchor sensor is equipped with a single radio transceiver with transmission range of R_c. It is assumed that each anchor sensor is hard-coded with its geographic coordinates before deployment. A single gateway sensor is installed on the train. The gateway sensor is equipped with two radio transceivers: TS_c and TS_b. TS_c is used to communicate with the anchor sensors that fall into its transmission range, and TS_b is used to continually broadcast

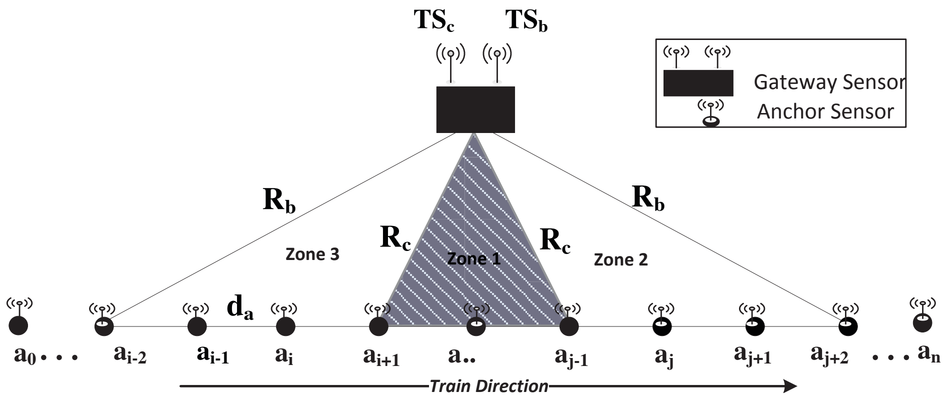


Figure 5.1: A WSN Architecture for Train Localisation

beacon packets to activate the anchor sensors before they go into the transmission range of TS_{c}. The transmission range for TS_{c} and TS_{b} is R_c and R_b, respectively. It is assumed that R_b is larger than R_c. To avoid an interference it is assumed that TS_c and TS_b operate on two non-overlapping channels ch_c and ch_b respectively. Each anchor sensor operates on both channels, that is, uses ch_b during duty-cycling and switches to ch_c to communicate with TS_c. As shown in Figure 5.1, zone 1 is the regions covered by TS_c, and zone 1, zone 2 and zone 3 are the region covered by TS_b.

The train localisation scheme works as follows: as the train moves, TS_{b} continually broadcasts beacon packets. Each beacon packet contains information of the current train location (represented by the location of the gateway) and speed. Once an anchor sensor receives a beacon packet, it stops duty-cycling and switches to channel ch_c to prepare for communication with TS_c. When an anchor sensor goes into the transmission range of TS_{c}, it sends its geographic coordinates to the gateway sensor. After an anchor sensor finishes the communication with the gateway sensor, it switches back to channel ch_b and resumes duty-cycling. Based on the geographic coordinates received from anchor sensors as well as the RSS information of the transmissions from anchor sensors, the train location will be computed at the gateway in a real-time manner.

5.2 Duty Cycling Model

All anchor sensors operate in an asynchronous duty-cycling mode in which each anchor sensor switches between sleep and wake-up states independently without global synchronisation. Figure 3.2 shows one duty-cycle, in which an anchor sensor first sleeps for t_{sleep} second with its radio turned off, and then wakes up and turns its radio on to perform clear channel assessment (CCA) to detect incoming signals. If an incoming signal is detected, the anchor sensor will stay in the active state until the scheduled communication between the anchor sensor and the gateway sensor is completed; otherwise it switches back to the sleep state and repeats another duty-cycle. The length of one duty-cycle is represented by T_d, and the time for turning on/off radio and performing CCA is denoted by t_{sw} and t_{cca}, respectively.

5.3 Problem Statement

In our train localisation scheme, each anchor sensor must be in wake-up state and reports to the gateway once it goes into the transmission range of TS_c. However, each anchor sensor runs an asynchronous duty-cycling protocol and can be woken up only if it detects the transmission signal from TS_b by performing CCA. The duty-cycling parameter t_{sleep} plays a significant role in the timely waking up of anchor sensors. If t_{sleep} is small, each anchor sensor needs to frequently turn on and turn off its radio, thereby wasting too much energy. From an energy saving perspective, the larger the t_{sleep}, the more energy each anchor sensor can conserve. However, if t_{sleep} is too large, an anchor sensor may miss the chance to detect the beacon packet broadcast by TS_b and fail to wake up in time. The first issue that will be addressed in this chapter is to derive the upper bound on t_{sleep}, which ensures that each anchor sensor can stay in a sleep state as long as possible while still guaranteeing that it can wake up in time once a train approaches.

The second issue that will be addressed is to design an energy-efficient wake-up scheme, which guarantees that each anchor sensor can wake up in time once it goes into the transmission range of TS_c, and resume low power duty-cycling once it finishes communication with the gateway. The designed scheme will be evaluated through both theoretical analysis and simulations.

5.4 BWS: Beacon-driven Wake-up Scheme

The BWS scheme computes an upper bound on t_{sleep}, which is then used to guarantee the availability of anchor sensors that communicate with the gateway sensor. The rest of the section explain BWS in detail.

5.4.1 Upper Bound of t_{sleep}

As shown in Figure 5.2, suppose that anchor sensor a_i enters into the transmission range of TS_b at time t_b and enters into the transmission range of TS_c at time t_c. Since to guarantee that anchor sensor a_i will be active in the communication range of gateway, that is communication range of TS_c, anchor sensor a_i must be active at time t_c to communicate with TS_c, it must wake up during the period $t_c - t_b$. To wake up, anchor a_i should receive at least one beacon from TS_b. Therefore, the following constraint on t_{sleep} has to be satisfied:

$$t_{sleep} \leq t_c - t_b, \tag{5.1}$$

otherwise, a_i may just start sleeping at t_b, and will remain in the sleep state at time t_c, thus will fail to wake up. Though, an anchor sensor may wake up in communication range and can communicate with the gateway sensor, but it is not guaranteed. Therefore, the above constraint on t_{sleep} guarantees the train-anchor communication at any time in Zone 1.

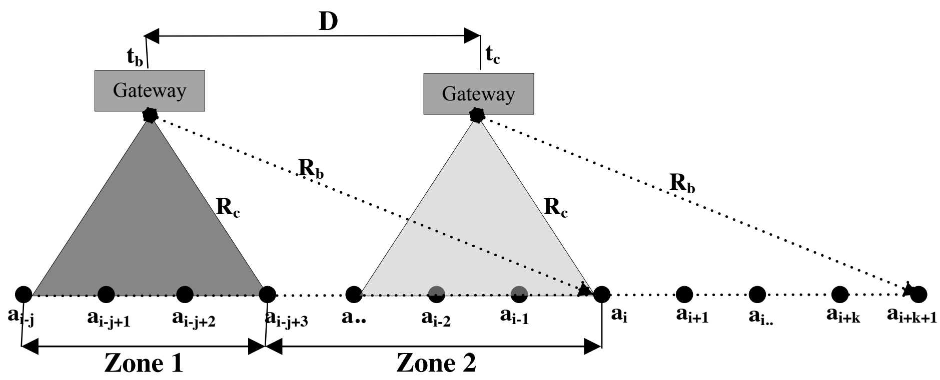


Figure 5.2: BWS: Illustration of Sensor-Train Communication

Let D denote the distance travelled by train during the period of $t_c - t_b$, and d_T represent the direct distance (on 2D plane) from the gateway (train) to the line along which the anchor sensors are deployed. To find the upper bound for t_{sleep}, the size of Zone 2 is determined first, as follows:

$$D_{z2} = \sqrt{R_{\text{b}}^2 - d_{\text{T}}^2} - \sqrt{R_{\text{c}}^2 - d_{\text{T}}^2}.$$

As d_T is very small, such as 2m, as compared to R_b (almost 800m) and R_c (500m on average), this is negligible and it will have no effect on the upper bound of t_{sleep}. The size of Zone 2 can be calculated as follows:

$$D_{z2} = R_{\mathrm{b}} - R_{\mathrm{c}}. \tag{5.2}$$

Let $S_{\max}$ represent the maximum train speed, at which the distance travelled by the train in the period $t_c - t_b$ is $D = S_{\max}(t_c - t_b)$. To guarantee that anchor a_i must perform CCA at least once in the period $t_c - t_b$ regardless of the actual train speed, the following condition must be satisfied:

$$D_{z2} \geq S_{\max}(t_c - t_b) \tag{5.3}$$

Based on Equations (5.1), (5.2) and (5.3), the relationship of t_{sleep} can be formulated with size of Z2 and maximum train speed as follows:

$$t_{sleep} \leq t_c - t_b \leq \frac{D_{z2}}{S_{\max}} \leq \frac{R_b - R_c}{S_{\max}}. \tag{5.4}$$

5.4.2 Design of Communication Protocol in BWS

The key idea behind the BWS protocol is to let the gateway broadcast beacon packets to wake up the anchor sensors. Specifically, the TS_b radio continuously broadcasts beacon messages that contain the following information: (a) the gateway ID (GW_ID), (b) the current train speed (S_T), and (c) the current train location (Loc_T). Once an anchor sensor receives a beacon packet from the gateway, it performs the following three tasks: *Duty-cycling Suspension*, *communication with* TS_c and *Duty-cycling resumption*, which are described in detail below. The pseudocode for BWS is given in Algorithm 1.

Duty-cycling Suspension

Upon receipt of a packet, the anchor sensor first checks if the packet is a beacon packet (line 2 in Algorithm 1). As the received beacon packet contains the current train location, the anchor sensor can check in which zone it is located by comparing its location with the train location. If the anchor sensor is located in Zone 2, it should first suspend the duty-cycling protocol and stay active (lines 3 and 4).

Communication with TS_c

Once an anchor sensor a_i finds itself in Zone 2, it starts preparing to communicate with TS_c. Anchor sensor a_i first estimates the amount of time it takes to enter Zone

1, denoted by β_i, as follows:

$$\beta = \frac{\sqrt{(x_{a_i} - x_T)^2 + (y_{a_i} - y_T)^2} - R_c}{S_T}, \tag{5.5}$$

where (x_{a_i}, y_{a_i}) and (x_T, y_T) are the coordinates of a_i and the train, respectively. Then a_i starts a timer with the timeout value set to β_i (line 5 in Algorithm 1). Once the timer expires, anchor sensor a_i will start communicating with TS_c.

Anchor sensors in Zone 1 can report multiple times, which is controlled by TS_c in the following way: TS_c periodically broadcasts *data_request*. The interval between two adjacent *data_request* packets is called one report round in which all anchor sensors in Zone 1 can communication with TS_c. Due to the presence of multiple anchor sensors in Zone 1, the communication between anchor sensors in Zone 1 and the gateway needs to be scheduled to avoid collisions. The key idea for scheduling anchor sensors in Zone 1 is to let the anchor sensor that is going to leave Zone 1 soonest transmit first. It is more important to receive reports from maximum sources because it increases the accuracy of a localisation system. A sensor mote which is about to leave communication Zone 1, transmits with highest priority. Each anchor sensor a_i in Zone 1 is associated with a priority p_{a_i}, which is computed as follows:

$$p_{a_i} = \frac{x_{a_i} - (x_T - R_c)}{d_a}, \tag{5.6}$$

where d_a is the distance between two adjacent anchor sensors. The x-coordinates of location of an anchor sensor a_i (x_{a_i}) and the train (x_T), and communication range R_c ensures that the anchor sensor which is about to leave the Zone 1 will have highest priority to communicate with the gateway sensor as expressed in Eq. 5.6. As illustrated in Fig. 5.3, at time t_i anchor sensor a_i in the Zone 1 receives the *data_request* packet from the TS_c and calculates its priority. This is followed by initiation of two timers, denoted by *start_data* and *stop_data* to trigger the start and stop of data transmission to TS_c. The *start_data* and *stop_data* timers are initialised by value of $p_{a_i} t_{report}$ and $p_{a_i} t_{report} + t_{report}$, respectively. Anchor sensors report back multiple packets, and the RSSI value in each packet can vary. The time taken by an anchor sensor to report multiple packets is t_{report}. TS_c calculates the average of received RSSI data to be used in localisation algorithm. Similarly, TS_c broadcasts *data_request* packet to initiate another data collection round and anchor sensors keep on reporting back to TS_c until they leave Zone 1.

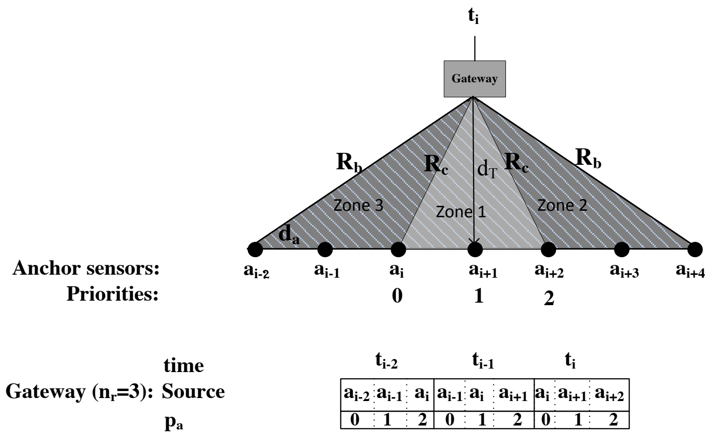


Figure 5.3: BWS Communication with the Train's Transceiver TS_c

Duty-cycling resumption

Once an anchor sensor a_i goes out of the transmission range of TS_c (i.e., Zone 1 in Figure 1), it should resume the duty-cycling protocol. To achieve this, each anchor sensor maintains another timer called *stop_Z1* (line 6). Once *stop_Z1* is expired, the node should resume duty-cycling. The *stop_Z1* is initialised with value γ as given in Eq. 5.7, which is the amount of time elapsed since the node wakes up till the time it goes out of Zone 1.

$$\gamma = \frac{\sqrt{(x_{a_i} - x_T)^2 + (y_{a_i} - y_T)^2} + R_c}{S_T} \tag{5.7}$$

BWS enables an anchor sensor to accomplish these three tasks and guarantees the wake up of an anchor sensor for communication with TS_c. However, it is possible for an anchor sensor to get beacon packets when it is located in Zone 3 because of large omnidirectional transmission range of TS_b. BWS adaptively avoids unnecessary wake-ups, but it uses the gateway's location information received in the beacon packet to calculate its zone. If an anchor sensor lies in Zone 3, BWS ignores such beacon packets and allows an anchor sensor to continue following duty-cycles. Although the train location Loc_T may not be always accurate at a particular point of time, the associated localisation error is acceptable by an anchor sensor to calculate its zone. However, for correct decision making for wake-up, it is assumed that localisation error, due to time drift between anchor sensors and the train, will never be larger than distance d_a due to the frequent location broadcast by the gateway sensor.

Algorithm 1: Beacon-driven Wake-up at Anchor Sensor $\mathrm{a_i}$

1 **On receiving beacon packet:**
2 **if** *SourceID* $= GW\_ID$ **then**
3   **if** a_i *locates in Zone 2* **then**
4   *duty_cycling = False* /* Pause duty-cycling */
5   **start_Z1** $\left(\frac{\sqrt{(x_{a_i}-x_T)^2+(y_{a_i}-y_T)^2}-R_c}{S_T}\right)$ /* Zone-1 start timer */
6   **stop_Z1** $\left(\frac{\sqrt{(x_{a_i}-x_T)^2+(y_{a_i}-y_T)^2}+R_c}{S_T}\right)$ /* Zone-1 end timer */
7   **else**
8   **Ignore Beacon**
9   *duty_cycling = True*
10 **else**
11   **Ignore Beacon**
12   *duty_cycling = True* /* Resume duty-cycling */
13 **On start_Z1 timer expiry (Zone-1 starts):**
14 **if** *start_Z1 is expired* **then**
15   *Channel* $= ch_c$ /* channel switch */
16   **start_data** $(p_{a_i} t_{report})$ /* transmission start timer */
17   **stop_data** $((p_{a_i} t_{report}) + t_{report})$ /* transmission stop timer */
18   **set priority** $(p_{a_i}) = \frac{x_{a_i}-(x_T-R_c)}{d_a}$ /* priority calculation */
19 **On stop_Z1 timer expiry (Zone-1 ends):**
20 **if** *stop_Z1 is expired* **then**
21   **Stop Sending Packets to Gateway**
22   *Channel* $= ch_b$ /* channel switch */
23   *duty_cycling = True* /* Resume duty-cycling */
24 **On start_data timer expiry:**
25 **if** *start_data is expired* **then**
26   **Send_Packet**(x_{a_i}, y_{a_i})
27   **Repeat process at line 25 & 26, unless** *stop_data* **is expired.**
28 **On stop_data timer expiry:**
29 **if** *stop_data is expired* **then**
30   **if** *Zone 1* **then**
31   *Wait* **for beacon from** TS_c **for next round.**
32   **else**
33   **Wait for** *stop_Z1* **timer to expire.**

5.5 Energy Analysis of BWS Scheme

In our system, the energy consumed at each anchor sensor can be divided into two parts: energy consumed in duty-cycling and energy consumed in wake-up state. Table 5.1 gives the list of states in which an anchor sensor can operate and the corresponding power level for each state.

Table 5.1: Anchor Sensor's States and Corresponding Power Level

States	**Power Level**	**Energy Consumed**
Transmission	P_{tx}	$E_{tx} = t_{tx}P_{tx}$
Idle Listening	P_l	$E_l = t_lP_l$
Packet Reception	P_{rx}	$E_{rx} = t_{rx}P_{rx}$
Radio Switch	P_{sw}	$E_{sw} = 2t_{tsw}P_{sw}$
CCA	P_{cca}	$E_{cca} = t_{cca}P_{cca}$
Sleeping	P_{sleep}	$E_{sleep} = t_{sleep}P_{sleep}$

5.5.1 Energy Consumed during Wake-up

If an anchor sensor goes to sleep at the point when it just goes into Zone 2, the amount of time that the anchor sensor will sleep throughout Zone 2 is t_{sleep}. However, if the anchor sensor wakes up at the point when it just goes into Zone 2, it will receive a beacon packet and stay active. In this case the amount of sleeping time throughout Zone 2 is 0. Since duty-cycling is not synchronised among all anchor sensors, an anchor sensor may wake up at any time between the above two extremes when it is in Zone 2. The amount of time that an anchor sensor sleeps in Zone 2 follows a uniform random distribution between 0 and t_{sleep}. Hence, the average amount of time that an anchor sensor stays in sleep state throughout Zone 2 is $t_{sleep}/2$.

The average amount of time that an anchor sensor stays in Zone 2 can be calculated by

$$\frac{R_b - R_c}{S_{avg}}, \tag{5.8}$$

where S_{avg} is the average train speed.

Let T_{z_2} denote the average amount of time that an anchor sensor stays active when it is in Zone 2 for one train pass. Then

$$T_{z_2} = \frac{R_b - R_c}{S_{avg}} - \frac{t_{sleep}}{2} \tag{5.9}$$

Let T_{z_1} denote the average time that an anchor sensor stays in Zone 1 for one train pass. Then

$$T_{z_1} = \frac{2R_c}{S_{avg}} \tag{5.10}$$

Let T_{wk} denote the average time that an anchor sensor stays in active state for one train pass. Since each anchor sensor will resume duty-cycling at a point when it enters into Zone 3, it will be

$$\begin{aligned} T_{wk} &= T_{z_1} + T_{z_2} \\ &= \frac{R_b + R_c}{S_{avg}} - \frac{t_{sleep}}{2} \end{aligned} \tag{5.11}$$

To simplify our analysis a reliable communication between anchor sensors and the gateway sensor is assumed. So each anchor sensor will receive one beacon packet and send one report packet in one data collection round. Through, this assumption does not lead to fair comparison with other approaches, but it is a reasonable approach to compare the baseline energy consumption and to compare the simulation-based results with theoretical-based energy consumption. Let t_{tx} and t_{rx} denote the time for transmitting and receiving a packet respectively. Therefore, the amount of time that an anchor sensor stays in idle listening state for one train pass, which is represented by t_l, can be computed as follows:

$$t_l = T_{wk} - t_{tx} - t_{rx}, \tag{5.12}$$

Let E_{wk} denote the amount of energy consumed at an anchor sensor during wake-up state for one train pass. According to Table 5.1,

$$\begin{aligned} E_{wk} &= t_{tx}P_{tx} + t_{rx}P_{rx} + t_l P_l \\ &= t_{tx}P_{tx} + t_{rx}P_{rx} + (T_{wk} - t_{tx} - t_{rx})P_l \end{aligned} \tag{5.13}$$

5.5.2 Energy Consumed during Duty-Cycling

As shown in Figure 3.2, one duty-cycle includes three parts: sleep (t_{sleep}), CCA(t_{cca}) and state switch ($2t_{sw}$). Let E_{dc} denote the energy consumption for one duty-cycle. Then from Table 5.1 It will be,

$$E_{dc} = 2t_{sw}P_{sw} + t_{cca}P_{cca} + t_{sleep}P_{sleep} \tag{5.14}$$

The time required for switching radio between on and off states and the time for CCA check are constants, therefore the amount of energy consumed by state switching and CCA check is fixed for one duty-cycle. For simplicity, e_x is used to denote this amount of energy, that is,

$$e_x = 2t_{sw}P_{sw} + t_{cca}P_{cca} \tag{5.15}$$

Then

$$E_{dc} = e_x + t_{sleep}P_{sleep} \tag{5.16}$$

5.5.3 Total Energy Consumption for a Period

Let L be the total length of the time that a anchor sensor operates and T_d be the length of one duty-cycle. λ is used to denote the total number of times that a train passes by an anchor sensor. Let E_{dc}^{total} be the total energy consumed during duty-cycling for the whole period L. Then

$$E_{dc}^{total} = \frac{L - \lambda T_{wk}}{T_d} E_{dc}, \tag{5.17}$$

where $\frac{L-\lambda T_{wk}}{T_d}$ is the total number of duty-cycles in time period L. Let E_{wk}^{total} be the total energy consumed during wake up for whole period L. Then

$$E_{wk}^{total} = \lambda E_{wk} \tag{5.18}$$

Let E_L^{total} represent the total energy consumed by an anchor sensor in time period L. Then

$$E_L^{total} = E_{dc}^{total} + E_{wk}^{total} \tag{5.19}$$

Based on Equations (5.17) and (5.18), Equation (5.19) can be expressed as,

$$E_{total}^{L} = \frac{L - \lambda T_{wk}}{T_d} E_{dc} + \lambda E_{wk} \tag{5.20}$$

By substituting Equations (5.11), (5.13) and (5.16) in Equation (5.20), it will be,

$$\begin{aligned} E_L^{total} = & \frac{1}{2t_{sw} + t_{cca} + t_{sleep}} \left(L - \lambda \Big(\frac{R_b + R_c}{S_{avg}} - \frac{t_{sleep}}{2} \Big) \right) (e_x + t_{sleep}P_{sleep}) \\ & + \lambda \Big(t_{tx}P_{tx} + t_{rx}P_{rx} + \Big(\Big(\frac{R_b + R_c}{S_{avg}} - \frac{t_{sleep}}{2} \Big) - t_{tx} - t_{rx} \Big) P_l \Big) \end{aligned} \tag{5.21}$$

5.5.4 Optimal t_{sleep} for minimising energy consumption

The minimisation of energy consumed at each anchor sensor can be formulated as the following optimisation problem:

$$\begin{aligned} \text{minimize} \quad & E^L_{total} \\ \text{subject to} \quad & 0 < t_{sleep} \leqslant t^{ub}_{sleep} \end{aligned} \tag{5.22}$$

where t^{ub}_{sleep} is the upper bound for t_{sleep} which is given in Section 5.4.1. As can be seen from Equation (5.21), the only variable is t_{sleep}, and it can be proved that E^L_{total} is strictly decreasing with the increase of t_{sleep}. The optimal t_{sleep} in terms of minimising the total energy consumption at each anchor sensor is $t^{ub}_{sleep} = \frac{R_b - R_c}{S_{\max}}$.

5.6 Simulation Setup

To evaluate our proposed beacon wake-up scheme (BWS) and average energy consumed by an anchor sensor, extensive simulations are carried out to evaluate the performance of the BWS scheme. In our simulations, *145* to *4000* anchor sensors are deployed with various distances between the adjacent anchor sensors; called deployment density, d_a. Moreover, the maximum train speed, S_{max}, ranges from *10 m/s* to *40 m/s*. The wireless channel model has *10%* packet loss rate with no requirement of packet re-transmission because there are no contenders for channel Ch_b as only gateway uses this channel to transmit beacon packets. Similarly, anchor sensors communicate with the gateway sensor by sending multiple packets without requirement of acknowledgement packets.

5.6.1 Parameter Configurations

The detailed parameter configuration used in our simulation setup are given in Table 5.2 along with their values.

In BWS, the successful wake-up of anchor sensors in Zone 2 is guaranteed for any t_{sleep} less than given t^{ub}_{sleep}. We conduct four set of simulations with different S_T and t_{sleep} settings. The size of Zone 2 is *40 m* and the size of Zone 1 is *500 m* which means a maximum of 6 anchor sensors with $d_a = 100\ m$ can stay in Zone 1. The set of $\{S_T, t^{ub}_{sleep}\}$ can be calculated by Eq. 5.4 such as $\{10\ m/s, 4\ s\}$, $\{20\ m/s, 2\ s\}$, $\{30\ m/s, 1.33\ s\}$ and $\{40\ m/s, 1\ s\}$. In each set of the simulation, the configuration shows the average percentage of anchor sensors that stay awake for specific percentage

Table 5.2: Simulation Parameters (BWS)

Parameters	Values
Simulation time period L (s)	10000
Simulation iterations	50
Train trip frequency λ	1
S_{max} (m/s)	10, 20, 30, 40
No. of Anchor Sensors	145-4000
t^{ub}_{sleep} (s)	1-4
Size of Zone-1(m)	500
Dist b/w Sensors (m)	100-700

of time in Zone 1 ($S_T\{t^{ub}_{sleep}, Avg\%\}$). For example, if on average 5 out of 6 anchor sensors stay active throughout Zone 1 then it can be stated as the percentage of time at least *83%* anchor sensors stay active in Zone 1. Similarly, on average 4 out 6 active anchor sensors throughout Zone 1 makes the percentage of *66%* of anchor sensors that stayed awake in Zone 1. It can be seen that in all cases when the $t_{sleep} \leq t^{ub}_{sleep}$, BWS nearly achieves the theoretical performance thresholds and wakes up *99.5%* to *100%* of anchor sensors. Figure 5.4 also shows that for a given increase in the t_{sleep}, the average percentage of active anchor sensors decreases, while the chances of observing at least *83%* active sensors stays high. The requirement for gateway sensor to communicate with multiple anchor sensors is intended to increase the localisation accuracy. However, if at least one anchor sensor can communicate its location information and RSSI value, the gateway can still calculate a reasonably accurate location. The rationale for this compromise is the minimisation of energy consumption.

5.6.2 Number of Active Anchor Sensors in Zone 1

The results in Figure 5.4 can be seen in another way in Figure 5.5, which shows the switching pattern of anchor sensors between sleep and wake-up states in Zone 1 with several settings of t_{sleep} for initial $900\ s$ of simulation time. According to Equation 5.4, the maximum t_{sleep} that can guarantee timely wake up of anchor sensors is 4 s at $10\ m/s$ train speed. It can be seen that, for all cases where t_{sleep} is not larger than $4\ s$, the number of active anchor sensors that are active in Zone 1 fluctuates between 3 and 4. For the case where $t_{sleep} = 8\ s$, the number of active anchor sensors in Zone 1 varies from 0 to 4, and most of the time there are only 1 or 2 active anchor sensors. This is because the value of t_{sleep} (i.e., 8 s) exceeds the upper bound t^{ub}_{sleep}. This figure validates the finding given in the Figure 5.4 that though the increase in the sleep time

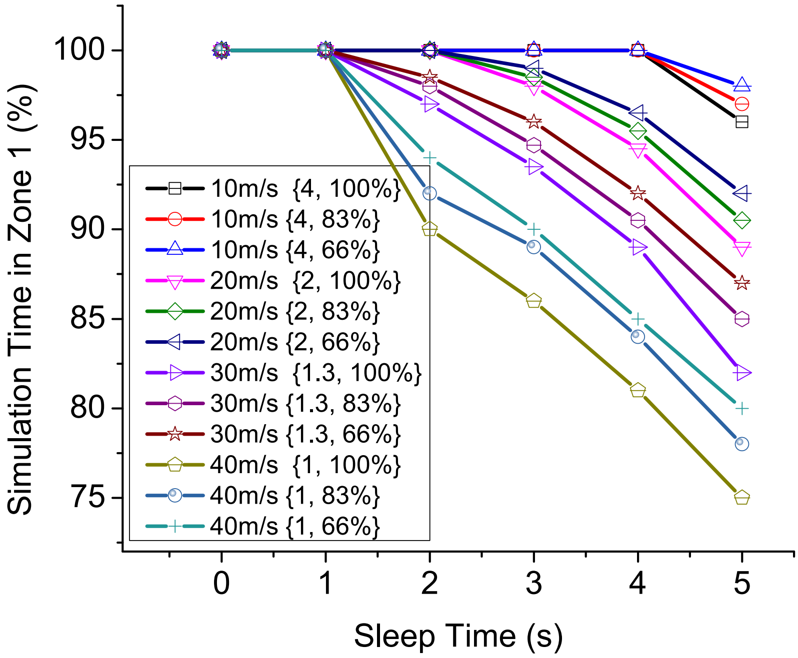


Figure 5.4: Percentage of Wake-up Anchor Sensors in Zone 1 using the BWS Protocol

results in the compromise in the number of active anchor sensors, it can save energy by letting anchor sensors sleep for a long time.

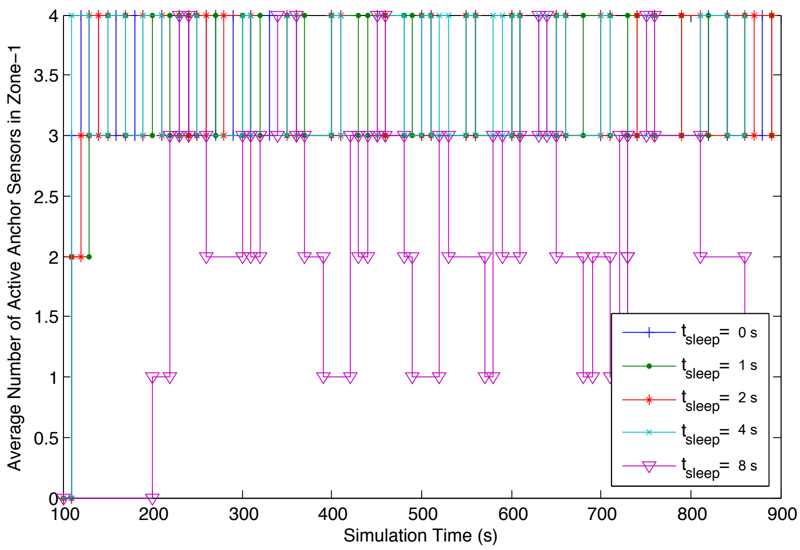


Figure 5.5: Number of Active Anchor Sensors in Zone 1 at Different t_{sleep}

The impact of multiple gateway sensors or trains on the average number of active anchor sensors is also studied and shown in the Figure 5.6. In the simulation, trains arrive with uniform distribution and an anchor sensor may wake up again to serve a gateway sensor on the train after doing so for another train. The inclusion of multiple

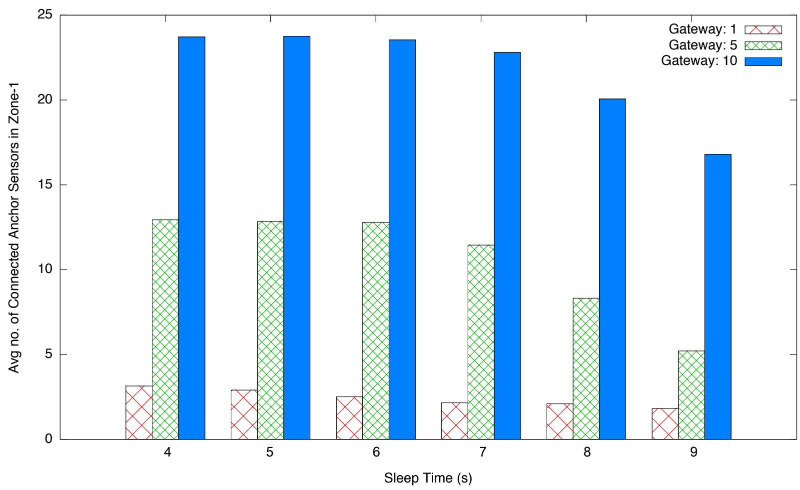


Figure 5.6: Number of Active Anchor Sensors in Zone 1 under Multiple t_{sleep} and Gateway Settings

trains is useful to analyse the BWS ability to wake-up anchor sensors for more than one time. The communication of an anchor sensor with several gateway sensors still follows the pattern of communication with a single gateway sensor at a time. In the Figure 5.6, the number of active anchor sensors in Zone 1 is shown for several sleep times and the gateway sensors with $d_a = 100\ m$. It can be seen that as the t_{sleep} being followed by the anchor sensors exceeds the $t^{ub}_{sleep} = 4\ s$, the average number of active anchor sensors in Zone 1 drops. However, the total average number of active anchor sensors in Zone 1 increases with the rise in the number of gateway sensors (trains) due to multiple Zone 1.

The rationale of gateway sensor to communicate with maximum number of anchor sensors is to increase the localisation accuracy level, which is directly proportional to the number of inputs from anchor sensors. However, if at least one anchor sensor can communicate its location information and RSS value, gateway can still calculate significantly accurate location. However, there is a tradeoff between the number of anchor sensors involved in the communication with the gateway sensor and the total energy consumed in the network of anchor sensors.

5.6.3 Energy Consumption

Figure 5.7 shows the average energy consumption at each anchor sensor in simulation compared with their theoretical counterparts during simulation duration of 10,000s. All calculations are based on the current and voltage specifications of CC2420 radio chipset data sheet. Each anchor sensor stays in Zone 1 for a long time when the train

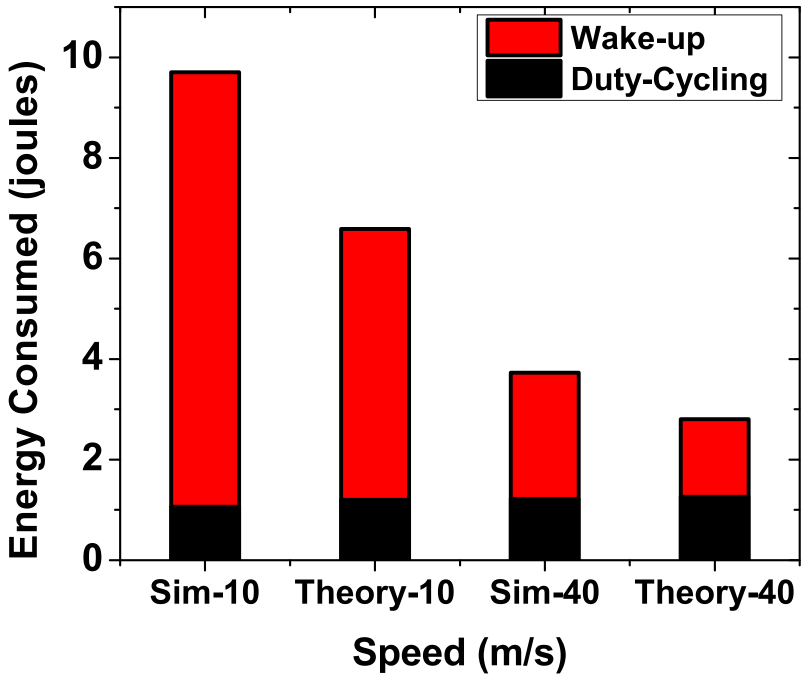


Figure 5.7: Comparison of Theoretical and Simulation based Energy Consumed by an Anchor Sensor at Different Train Speed Values for simulation duration of 10000s

speed is slow, such as *10 m/s*. However, when the train speed is fast, such as *40 m/s*, it passes through an anchor sensor rapidly and the anchor sensor stays in Zone 1 for a shorter time. Therefore, the energy consumed by an anchor sensor in wake-up state drops with the increase in the train speed because the active time duration reduces. So there is a tradeoff between the sleep time and the energy consumption. It can be seen from the Figure 5.7 that energy consumed by an anchor sensor is high when the train speed is *10 m/s*. Moreover, with the increase in the train speed up to *40 m/s*, the energy consumed drops in both theoretical-based and simulation-based calculations. Here it is worth mentioning that the simulation-based results are elevated because theoretical results are based on average calculations and considering the single packet transmission in the reliable transmission mode for the sake of simplicity. However, the theoretical-based and simulation-based results verify that the time an anchor sensor stays in wake-up state to communication with the gateway sensor decreases with the increase in the train speed.

The impact of the presence of multiple gateway sensors or trains on the energy consumption by an anchor sensor is shown in the Figures 5.8 and 5.9 at train speed of 10 m/s. In these figures, 5.8 and 5.9, the average energy consumption is shown over multiple t_{sleep} settings and gateway sensors 5 and 10 respectively. The multiple gateway sensors represents the passing of multiple trains on track at different times. Vertical

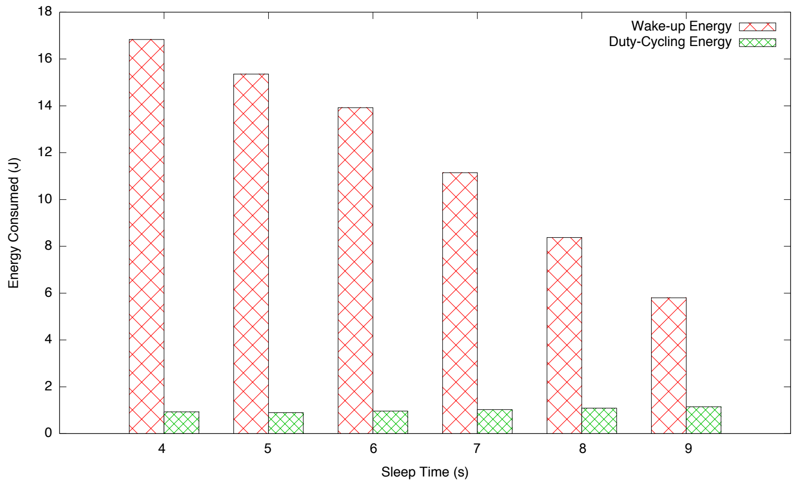


Figure 5.8: Energy Consumed by an Anchor Sensor at Different t_{sleep} Values with 5 Gateway Sensors

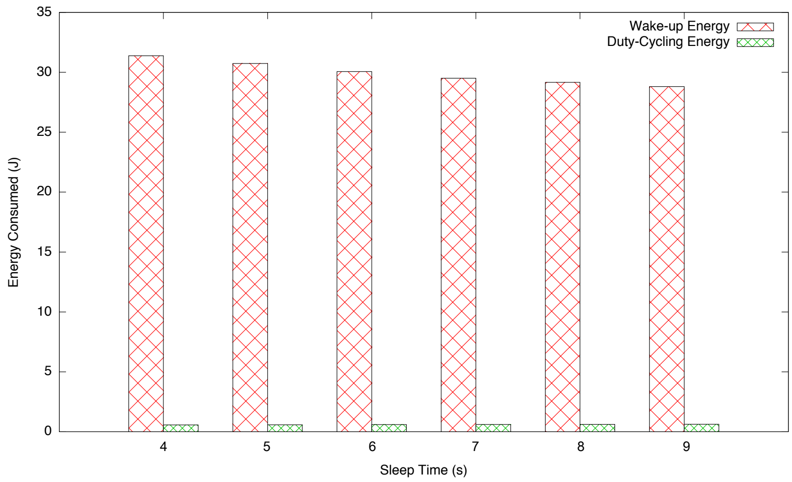


Figure 5.9: Energy Consumed by an Anchor Sensor at Different t_{sleep} Values with 10 Gateway Sensors

axis represents the energy consumed (joules). It can be seen that, in all cases, as t_{sleep} increases, energy consumed during wake up decreases and energy consumed during duty cycling increases. However, the rate of reduction in the wake-up energy consumption depends on the number of gateway sensors and their configuration. In our simulation, gateway sensors' inter-arrival time follows a uniform distribution. Therefore, there is possibility that when an anchor sensor finishes communication with one gateway sensor, it falls into the Zone 1 of another gateway sensor. This wake-up pattern of an anchor sensor increases the time spent in the wake-up states.

One aspect of train-anchor communication is that, if there is at least one active anchor sensor available for communication with train during Zone 1, gateway sensor can estimate train's location. In such case, the estimation errors will be large, therefore, there is an associated compromise in the system performance against a few active anchor sensors in Zone 1. However, such compromise results in a decrease of average energy consumption at each anchor sensor, which in return increases the network lifetime. In the simulations, energy consumption is recorded for 10000 s. On the other hand, if due to performance compromise, some of the anchor sensors are allowed to sleep for a longer time, the amount of energy consumed during duty-cycling slightly increases. This increase is because of the increase in the amount of time that each anchor stays in sleep mode, whereas, the amount of energy consumed in the active state is subject to the number of trains passing through anchor sensors.

5.7 Related Work

In the train localisation, the sensors deployed along the track report to the gateway sensor on the train, the gateway sensor then uses these noisy measurements to compute its location through localisation schemes. Typically, sensors follow duty-cycling to enhance their battery life, which makes them unreliable to be available for communication. Therefore, a wake-up scheme can guarantee the wake-up of the sensors when train is passing by them. The existing wake-up schemes can be divided into two classes: synchronous wake-up and asynchronous wake-up.

In synchronous wake-up schemes, sensor nodes synchronise their duty-cycles in such a way that they wake-up and sleep at the same time. The benefit of such protocols is that they enable sensor nodes to be available for any communication. However, there is a an associated shortcoming with such schemes, that is, the synchronisation cost, overhead, may exceed the available resources and thus make such schemes unreliable for large networks. Synchronous protocols cut down the idle-listening period of sensors nodes, which is one of the major causes of energy consumption. SMAC is one of the synchronous protocols (Wong *et al.*, 2007). SMAC allows sensor nodes to exchange SYNC packets to synchronise the duty-cycling sleep and wake-up intervals of neighbour sensor nodes (Ye *et al.*, 2004). Thus, reduces the power consumption. TMAC is another synchronous wake-up protocol that enhances the functionality of SMAC protocol by allowing sensor node to immediately returning to duty-cycling if the even of interest is not detected while performing CCA checks (Van Dam and Langendoen, 2003). Likewise TMAC, ADMAC (Kim *et al.*, 2008) allows unintentional receivers of packets to get back to their duty-cycling without receiving complete packets. DWMAC synchronises transmitter and receiver by exchanging scheduling (SCH) and its confirmation packets

(Sun *et al.*, 2008). Time duration is reserved between both sensor nodes and for successful reservation from sleep period, the chances of collision at destination node is almost zero. In TSMP (Pister and Doherty, 2008) time is divided into several small slots. Transceiver of nodes sense the potential transmission at the start of each slot. For a transmission for itself, it stays active; otherwise resumes its duty-cycling.

In asynchronous wake-up schemes, sensor nodes follow independent sleep schedules without having a global view of duty-cycles of their neighbour nodes. In such protocols, control over head is minimised and therefore energy consumption is reduced. However, performance efficiency also suffers due to unavailability of sensor nodes at the same time for communication. Several protocols are proposed by researcher to overcome this issue. One of initial efforts towards this issue was BMAC protocol (Polastre *et al.*, 2004). In BMAC protocol, a potential transmitter tries to wake-up its intended receiver by transmitting preambles. A receiver detects energy level in the medium and cooperates for a successful transmission. Though BMAC saves huge control overhead traffic for synchronisation, but it generates large overhead by continuously transmitting preambles and becomes infeasible in large networks. Buettner *et al.* (2006) proposed improvements in BMAC and developed XMAC. XMAC uses small frequent preamble packets that reduces the overhead traffic. Though XMAC improves overhead cost but still it can deplete bandwidth resources in high traffic network. WiseMAC protocol deals with this shortcoming by enabling sensor nodes to remember the sleep schedules of neighbouring nodes to avoid preambles in frequent destination nodes (El-Hoiydi and Decotignie, 2004). WiseMAC significantly improves network performance. ContikiMAC (Dunkels, 2011) considers the energy efficiency of MAC protocol and allows sensor nodes to optimise their duty-cycle intervals. RI-MAC (Sun *et al.*, 2008) proposed a significantly improved system by eradicating the need of preambles. Potential receivers advertise their availability once they are in wake-up state and senders start transmission. A hybrid approach was introduced by Chen *et al.* (2001), in which sensor nodes with high resources stay active and coordinate between other nodes. Coordinator nodes buffer the packets and allow receivers to get their packets from coordinator once they are active.

5.8 Summary

In this Chapter, I have presented a new scheme, a beacon-based wake-up scheme, to wake-up the anchor sensors that operates on the asynchronous duty cycles. The WSN-based train localisation system heavily depends on the communication between anchor sensors and the gateway sensor, where the gateway sensor computes the location of the train by using transmitted geographic coordinates of the anchor sensors and the RSSI measurements of corresponding transmissions. The arrival time of the train and sleep times of anchor sensors are unknown, and none of the entities have this global knowledge of the system. In such a situation, the availability of anchor sensors for communication plays a vital role in the accuracy and reliability of WSN-based train localisation system. To achieve this task, keeping anchor sensors always active and in idle listening state is expensive in terms of energy consumption, whereas, the duty-cycling does not guarantee the timely activation of the anchor sensors to communicate with the gateway sensor. Another approach, in which, anchor sensors wake up each other by communication before arrival of train, is not a feasible approach. We studied this approach and found infeasible as the train's speed, train location and duty-cycling pattern of neighbouring sensors are unknown. Therefore, a relation can not be derived between the speed of train and speed of packets transmitted to neighbouring sensors to wake them.

BWS provides a new solution to ensure the wake-up of the anchor sensors before they reach in the communication range of the gateway sensor. This scheme enables the gateway sensor to broadcast the beacon packets that contain the recent location of the train. Moreover, BWS also computes the upper bound on the sleep time that an anchor sensor can sleep in a duty-cycle. During the regular channel scanning period (CCA), anchor sensors wake up to detect the incoming packets and if they receive a beacon packet, they stay active and prepare to communicate with the gateway sensor. Based on the received geographic information and signal strength measurements, the gateway sensor computes its new location and broadcasts it again. Finally, the proposed scheme is analysed theoretically and with the simulations for its ability to wake up the anchor sensors and the energy consumption.

In the next Chapter, I will present the Particle-Filter-based train localisation scheme that uses the RSSI measurements and geographic coordinates of anchor sensors, and develops a weighted likelihood function to compute the location of the train.

Chapter 6

Particle Filter based Train Localisation

In this chapter, I begin by introducing the Particle Filtering technique, its components, and propose a novel Particle-Filtering-based train localisation scheme. Moreover, a weighted RSSI-based likelihood function is introduced to estimate the likelihood of the particles for best representation of the train's location. In the remainder of the chapter, extensive simulations, that take real-data, are used to evaluate the developed scheme. The real-data is collected from field experiments.

6.1 Introduction

The associated benefits of using WSN technology for train localisation include cost effectiveness and feasible alternative in the absence of GPS technology. However, RSSI measurements are prone to noise, caused by the infrastructure in the surrounding environment and other overlapping frequencies such as microwave. To counter the fragile nature of RSSI, a measurement model is introduced that comprises of data such as RSSI readings and the geographic coordinates transmitted from anchor sensors.

The Particle Filtering technique is used to smoothen the noise elements in the RSSI measurements and to increase the accuracy of train location estimation. In the Particle Filtering technique, a large number of particles are spread in the target area and weight is assigned to each particle. The weights are assigned based on the likelihood function and represents the likeliness of particles to represent the location of the train. The details of the Particle Filtering technique and train localisation algorithm are given in the following sections of this chapter.

6.2 Bayesian Filtering

In train localisation or any other tracking problem, system has few variables such as target, state and data measurements. An object that needs to be identified is called target. The parameters of an object define its state such as its location, speed and movement angle. Data measurements are readings received from sensor nodes and demonstrate the evolution of state of an object at new time interval.

Generally, to represent the evolution of target state at any time t, hidden markov model (Ghahramani, 2001) is used as shown in Figure 6.1. Z_t refers to data measurements received at time t.

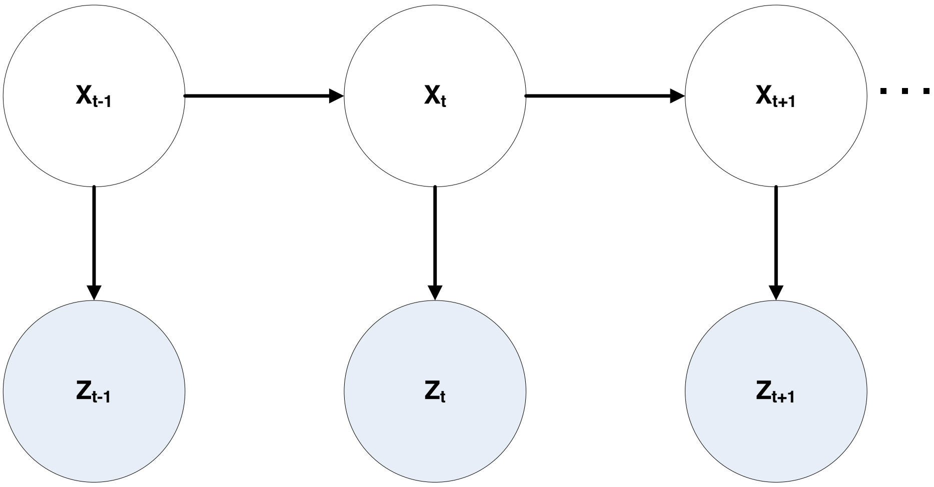


Figure 6.1: Hidden Markov Model

In the Bayesian Filtering, the estimation of new state of a target is related to recursive approach. A probability density function (PDF) $P(X_t|Z_t)$ is required, from which new state will emerge and represented by attributes with highest probability. PDF comprises of two stages such as prediction and update stages. Chapman-Kolmogrov developed a model of equations for these stages. The proposed model was discussed in detail by Klepal *et al.* (2007a,b) and its equations are expressed in Equations 6.1 (predict stage), 6.2 (update stage), 6.3 (normalise) and 6.4 (new state). Later on, in development of Particle Filter based localisation scheme, models are developed on these basic equations.

$$P(X_t|Z_{t-1}) = \int P(X_t|X_{t-1})P(X_{t-1}|Z_{t-1})dX_{t-1}. \tag{6.1}$$

When data measurements Z_t are received at time t, the update stage can be expressed as follows:

$$
\begin{aligned}
P(X_t|Z_t) &= P(X_t|Z_t, Z_{t-1}) \\
&= \frac{P(Z_t|X_t, Z_{t-1})P(X_t|Z_{t-1})}{P(Z_t|Z_{t-1})} \\
&= \frac{P(Z_t|X_t)P(X_t|Z_{t-1})}{P(Z_t|Z_{t-1})}
\end{aligned} \tag{6.2}
$$

$$
P(Z_t|Z_{t-1}) = \int P(Z_t|X_t)P(X_t|Z_{t-1})dX_t \tag{6.3}
$$

and here, the normalising factor is given in equation 6.3.

As discussed before, a PDF is developed in the update stage once data measurements are received. New state X_t and its error is then computed from the developed PDF.

$$
X_t = \int X_t P(X_t|Z_t)dX_t \tag{6.4}
$$

The equations as discussed by Klepal *et al.* (2007a,b), analyse the model basics and state evolution with the help of data measurements from sensor nodes. However, for non-linear and Gaussian noise, analytical solution is hard. Therefore, an approximate solution is desired and for which noise filters play a vital role in estimation of target's location.

6.3 Basics of Particle Filter

Particle Filter offers a solution, for non-linear problems ideally for non-Gaussian noise, by implementing Bayes Filter recursively. It works with the set of particles, or weighted samples, to represent probability density. The set of particles and their associated weights are used to compute the posterior probability. The posterior probability estimation of the state is expressed as

$$
P(X_t|Z_t) \approx \sum_{i=1}^{N} w_t \sigma(X_t - X_t^i) \tag{6.5}
$$

where X_t^i is the i-th particle ($1 < i < N$) and w_t is the associated weight of the particle. Every single particle represents the state X_t at particular time t, with the probability of its correctness as its weight.

In the proposed scheme, it takes two models as input to estimate the position of a train. First one is the motion (movement) model that describes the new state emergence from the old one with time (i.e., the train movement model). The second model is observation (movement) model that relates to the data received at particular time and state (i.e., the RSSI measurement model). The principle of Particle Filter is to develop a PDF on existing set of measurements for a particular state. Afterwards, it filters the calculated position recursively, once new set of measurements are received. In the proposed scheme, the estimated location of a train is denoted by $X_t = (x_t, y_t)$, a set of measurements (RSSI and anchor sensor location coordinates) are denoted by Z_t at particular time t. Further, Particle Filter constitutes of two stages such as prediction and update stages.

6.3.1 Prediction Stage

In the prediction stage, the PDF of new state, at next time interval, is predicted based on the movement model. The movement model facilitates determination of the position of particles at every time instant in the prediction stage. It includes the noise factor to estimate the realistic particle positions.

6.3.2 Update Stage

The predicted PDF gets corrected in update stage through measurement model. This happens after new set of data measurements are received. In the train localisation system, measurements model comprises of RSSI measurements and location information received from anchor sensors in anchor-gateway communication in Zone 1. The rationale to use multiple pieces of information is to minimise the interim impact of environmental factors on RSSI measurements.

6.4 Particle Filtering based Train Localisation Algorithm

A posterior PDF function $P(X_t|Z_t)$ is computed at time t by using particles, $\{X_t^{[i]}\}$. Each particle is linked with a weight $w_t^{[i]}$. Each particle $X_t^{[i]}$ has an associated weight $w_t^{[i]}$. The particles and their associated weights are updated using latest measurements and predictions, respectively.

A Particle-Filter-based train localisation algorithm is developed that consists of five steps, given in Algorithm 2: initialisation, prediction, update, resampling, and location estimation. However, with the addition of another step, an improved Particle Filtering algorithm is developed for train localisation, that is, Particle-Filter-based Signal Strength Rectification (PF-SSR), which contains pre-processing of raw measurement model before using it for the particles' weight update and a novel weighted RSSI likelihood function that considers the probability function to update the particles' weights. The PF-SSR-based train localisation algorithm is given in Algorithm 3.

Algorithm 2: PF based Train Localisation Algorithm

Step 1: Initialisation
t = 0
for $i \leftarrow 1$ **to** N **do**
    Initialise $x_0^{[i]}$ and $y_0^{[i]}$
    $w_0^{[i]} = \frac{1}{N}$
Step 2: Prediction
t = t + 1
for $i \leftarrow 1$ **to** N **do**
    $x_t^{[i]} = x_{t-1}^{[i]} + S_t^x T + n_{t-1}^x$
    $y_t^{[i]} = y_{t-1}^{[i]} + S_t^y T + n_{t-1}^y$
Step 3: Update
for $i \leftarrow 1$ **to** N **do**
    update $w_t^{[i]}$ based on Equation (6.16)
Normalise Weights $\overline{w}_t^{[i]} = \frac{w_t^{[i]}}{\sum_{j=1}^{N} w_t^{[j]}}$
Step 4: Resampling
Generate a set of $\overline{N}$ new particles from $\{X_t^{[i]}, w_t^{[i]}\}$.
Initialise Random variable r such that $r \in (0, \sum w_t)$.
for $i \leftarrow 1$ **to** $\overline{N}$ **do**
    $\overline{N}$ new samples are based on low variance sampling $[(r + \frac{i \sum w_t}{\overline{N}})\ mod\ \sum w_t]$.
Step 5: Estimate the location
Return $X_t = (x_t, y_t)$, where $x_t = \sum_{i=1}^{N} w_t^{[i]} x_t^i$, $y_t = \sum_{i=1}^{N} w_t^{[i]} y_t^i$;
$t = t + 1$ and goto step 2;

6.5 Initialisation

Initially, N number of particles are evenly distributed over L meters length of track. Therefore, L can be maximum possible error in initially estimated location at time $t = 0$. Whereas, the actual location of train can be anywhere within L meters.

Algorithm 3: Improved PF-SSR-based Train Localisation Algorithm

Step 1: Initialisation
t = 0
for $i \leftarrow 1$ **to** N **do**
    Initialise $x_0^{[i]}$ and $y_0^{[i]}$
    $w_0^{[i]} = \frac{1}{N}$
Step 2: Prediction
t = t + 1
for $i \leftarrow 1$ **to** N **do**
    $x_t^{[i]} = x_{t-1}^{[i]} + S_t^x T + n_{t-1}^x$
    $y_t^{[i]} = y_{t-1}^{[i]} + S_t^y T + n_{t-1}^y$
Step 3: PF-SSR- RSSI Rectification
if $AnchorCommunicated$ **then**
    for $l \leftarrow 1$ **to** k **do**
        $RSSI_{avg} = \frac{1}{k} \sum_l KalmanFilter(RSSI_l)$
else
    $RSSI = RSSI_{LN} + RSSI_{err}$
Step 4: Update
for $i \leftarrow 1$ **to** N **do**
    update $w_t^{[i]}$ based on Equation (6.16)
Normalise Weights $\overline{w}_t^{[i]} = \frac{w_t^{[i]}}{\sum_{j=1}^{N} w_t^{[j]}}$
Step 5: Resampling
Generate a set of $\overline{N}$ new particles from $\{X_t^{[i]}, w_t^{[i]}\}$.
Initialise Random variable r such that $r \in (0, \sum w_t)$.
for $i \leftarrow 1$ **to** $\overline{N}$ **do**
    $\overline{N}$ new samples are based on low variance sampling $[(r + \frac{i \sum w_t}{\overline{N}}) \ mod \ \sum w_t]$.
Step 6: Estimate the location
Return $X_t = (x_t, y_t)$, where $x_t = \sum_{i=1}^{N} w_t^{[i]} x_t^i$, $y_t = \sum_{i=1}^{N} w_t^{[i]} y_t^i$;
$t = t + 1$ and goto step 2;

6.6 Prediction

In train localisation system, the motion model is used to determine the new position of particles. The estimated positions of new particles are then used to estimate the location of a train. The adopted motion model in the proposed train localisation system takes into account the train speed and noises.

$$X_t^{[i]} = \begin{bmatrix} x_{t-1}^{[i]} + S_t^x T + n_{t-1}^x \\ y_{t-1}^{[i]} + S_t^y T + n_{t-1}^y \end{bmatrix} \tag{6.6}$$

where $x_{t-1}^{[i]}$,and $y_{t-1}^{[i]}$ represents the location coordinates of particle $X_{t-1}^{[i]}$. Speed parameters with respect of x-axis and y-axis are represented by S_t^x and S_t^y, respectively. The noise in the train speed at both axis are represented by n_{t-1}^x and n_{t-1}^y with Gaussian distribution. T is the particle update interval.

6.7 Update

$P(Z_t|X_t)$ is the likelihood function that shows the chances of particles representing the actual train's position, provided Z_t measurements are received. In the proposed train localisation system, the likelihood function is developed based on weights, called novel Weighted RSSI Likelihood Fuction (WRLF) Before going into the details of WRLF, some notations are required to represent entities. Let $RSSI(j,t)$ be the RSSI reading received from an anchor sensor a_j at time instant t, and Z_t are RSSI readings received from M anchor sensor nodes at time t.

6.7.1 WRLF for Particle Update

In order to update the weights of the particles, WRLF is used. The Equation 6.7 defines the weighted RSSI likelihood function that is used for particles' weight update.

$$P(Z_t|X_t) = \epsilon + \exp\left(\frac{(X_t - L_t)^2}{2\sigma^2}\right), \epsilon << 1 \tag{6.7}$$

where ϵ is the model parameter and a constant value and σ is the deviation of particles. L_t is the location that can be computed as in Equation (6.8). The value of ϵ ranges from 0 to 1. I have simulated this system variable with several sets of values and used

0 as most suitable value.

$$L_t = \begin{bmatrix} \sum_{j=1}^{M} h_t^j L_{t,j}^x \\ \sum_{j=1}^{M} h_t^j L_{t,j}^y \end{bmatrix} \tag{6.8}$$

where h_t^j is the associated weight of RSSI reading $RSSI(j,t)$, and $L_{t,j}$ is the location computed based on the log-normal path-loss model and the anchor sensor locations.

6.7.2 Weighted RSSI Function

The following four methods are used to calculate h_t^j for $RSSI(j,t)$ and are important to analyse the effects of weight assignment methodology on the localisation accuracy.

Equally Weighted RSSI Function

In this scheme, equal weights are assigned to all RSSI measurements from anchor sensors. The total weight that is divided among all RSSI readings is 1. Hence,

$$h_t = \frac{1}{M} \tag{6.9}$$

and the train location can be computed as,

$$L_t = \begin{bmatrix} \frac{\sum_{j=1}^{M} L_{t,j}^x}{M} \\ \frac{\sum_{j=1}^{M} L_{t,j}^y}{M} \end{bmatrix} \tag{6.10}$$

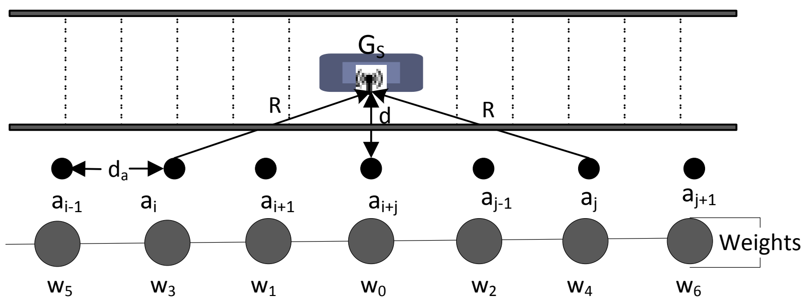


Figure 6.2: Equally Weighted RSSI Method

Strength Weighted RSSI Function

In this scheme, higher weights are assigned to stronger RSSI, as illustrated by Figure 6.3. I have,

$$h_t^j = 1 - \frac{RSSI(j,t)}{\sum_{i=1}^{M} RSSI(i,t)} \tag{6.11}$$

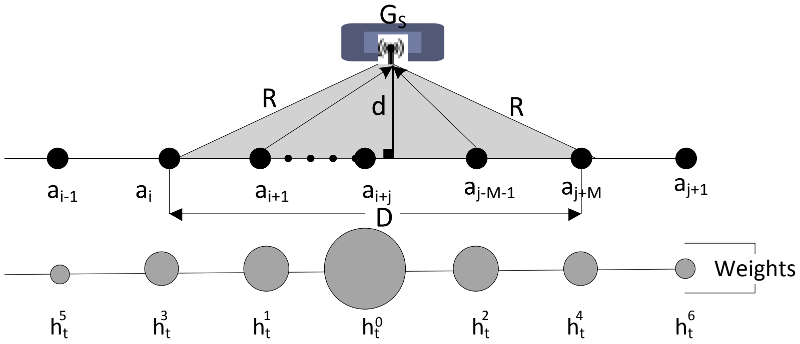


Figure 6.3: Strength Weighted RSSI Method

Single Strongest RSSI Function

In this scheme, only the strongest RSSI reading is used to compute L_t. Suppose $RSSI(k,t) = \max_j^M RSSI(j,t)$. Then, 1 is assigned to h_t^k, and 0 to all the other weights.

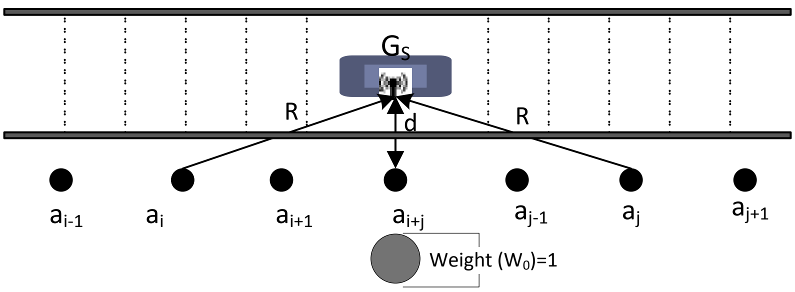


Figure 6.4: Single Strongest Weighted RSSI Method

Gaussian Weighted RSSI Function

In this scheme, weights are assigned based on Gaussian distribution to the RSSI received from the anchor sensors. Therefore, the stronger RSSI measurements contribute

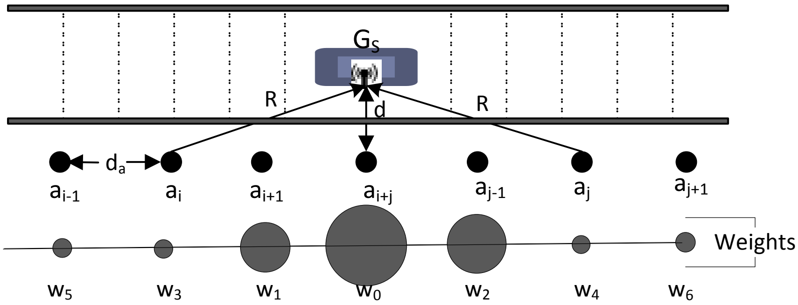


Figure 6.5: Gaussian Weighted RSSI Method

more to the WRLF and weaker RSSI measurements contribute less to the WRLF.

$$\begin{aligned}
\mu_{RSSI_j} &= \frac{\Sigma_{j=1}^{M} RSSI_j}{M}, \\
\sigma^2_{RSSI_j} &= \frac{1}{M}\Sigma_{j=1}^{M}(RSSI_j - \mu_{RSSI_j})^2, \\
h_t^j &= \frac{1}{\sigma_{RSSI_j}\sqrt{2\pi}} exp(-\frac{(RSSI_j - \mu_{RSSI_j})^2}{2\sigma^2_{RSSI_j}})
\end{aligned} \tag{6.12}$$

6.7.3 Weighted Particles Update Function

The WRLF, as expressed in Eq. 6.7, is used to update the particle weight of each particle and the weight represents the likelihood of true representation of train's location by that particle. The initial weights of the particles' are equally distributed at time $t = 0$. Following are the methods that can be used to update the weight of the particles.

Squared Weighted Particle Update

In this scheme, weight w_t^i is calculated by dividing the square of the difference between the particles $X_t^{[i]}$ and computed location L_t to the square of total transmission distance D. The rationale of squared weighted particle update scheme is that it increases the strong weights and decreases the weak weights. Therefore, particles with low probability gets eliminated sooner.

Gaussian Weighted Particle Update

In this scheme, weight w_t^i follows the Gaussian distribution for the given particle's distance $X_t^{[i]}$. The function mean $\mu(L_t)$, and standard deviation $\sigma(L_t)$, are obtained by using following equations. The rationale of Gaussian weighted particle is that it increases the weights of particles that are close to train's actual location and decreases the weights of particles that have large differences with the train's actual location. The difference between Gaussian and Squared method is that Gaussian method converges particles quickly than squared method.

$$\mu(L_t) = \frac{\Sigma_{j=1}^{M} L_t}{M} \tag{6.13}$$

$$\sigma(L_t)^2 = \frac{1}{M}\Sigma_{j=1}^{M}(L_t - \mu(L_t))^2 \tag{6.14}$$

$$w_t^i = \frac{1}{\sigma(L_t)\sqrt{2\pi}} exp(-\frac{X_t^{[i]} - \mu(L_t))^2}{2\sigma(L_t)^2}) \tag{6.15}$$

By following the Gaussian distribution, there is a clear split between the particles associated with the higher weights and RSSI values associated with lower weight. In this procedure, low weighted RSSI readings get eliminated and only strong RSSI readings becomes candidate in the calculation of X_t.

$$w_t^{[i]} = P(Z_t|X_t^{[i]}) \times w_{t-1}^{[i]} \tag{6.16}$$

Weights are normalised after update stage in such a way that $\sum_{i=1}^{N} w_t^{[i]} = 1$. The weighted average of particles' location gives the estimated train location.

6.8 Particle Filter based Signal Strength Rectification (PF-SSR)

The gateway sensor receives signal strength measurements when it initiates the data collection rounds. It is possible that the RSSI measurements vary due to noise, also, it is possible that the gateway sensor may not receive measurements from anchor sensors due to some failure, such as battery outage, physical damage or changed antenna dynamics. In such a scenario, the lack of available measurements or noisy measurements can affect the location estimation. There are many techniques to deal with the noisy measurements and outliers such as the Dixon method (Feng *et al.*, 2012), Grubs method (Grubbs, 1969), Tukey's rule (Anscombe, 1960), but some of these methods work on Gaussian data assumption and some have sample size limitations. In addition, it is hard to segregate the noisy measurements from the outliers. Therefore, in PF-SSR, a Kalman Filter is used to rectify the noisy measurements. Moreover, in PF-SSR, measurements are generated based on log-normal path loss model for those anchor sensors that failed to wake-up and communicate with the gateway sensor. The generated measurements are the boundary measurements (weakest measurements) within the acceptable range to minimise the fabricated effects. The Kalman Filter works in two stages: (a) prediction stage, and (b) correction stage. In the prediction stage, the system state is predicted based on error covariance, and the measurement stage corrects the predicted system state by calculating the trust factor of received noisy measurements. The trust factor is known as Kalman gain, which computes the gain of received measurements. It is worth noting that here the system state (RSSI measurements) is not modified with any control input. Also, it is assumed that processed noise in the RSSI generation is neglectable as compared to measurement noise, from sender to the receiver. These assumptions results in the simplified Kalman Filter equations for prediction stage.

$$RSSI'_k = RSSI^{est}_{k-1} \tag{6.17}$$

$$P'_k = P_{k-1} \tag{6.18}$$

and the correction stage equations are,

$$K_k = \frac{P'_k}{P'_k + R} \tag{6.19}$$

$$RSSI^{est}_k = RSSI'_k + K_k(RSSI^k_{a_j} - RSSI'_k) \tag{6.20}$$

$$P_k = (1 - K_k)P'_k \tag{6.21}$$

Let k be the number of RSSI measurements received at time t_i from anchor sensor a_j. The $RSSI'_k$, $RSSI^{est}_k$ and $RSSI^k_{a_j}$ are the predicted RSSI, estimated RSSI at correction stage and original RSSI measurement received at gateway sensor, respectively. K_k is the Kalman gain computed for k^{th} noisy measurement, R is the measurement noise covariance and 0.1 is its value. A Kalman Filter does not expect R to be accurate as it converges with time and number of samples. P'_k and P_k are the priori and posteriori error variance estimates, respectively. Eq. 6.19 & 6.20 suggests that with the high gain, the Kalman Filter trusts more on the received measurements, and with a low Kalman gain, the Kalman Filter trust the prediction more than measurements, thus making it tolerant and less responsive to the noisy measurements. Eq. 6.21, P_k re-estimates the error variance which refers to the environmental noise factors related to the current time and location of train. It changes with the train location and improves significantly with the decrease in the added factors of measurements noise, such as reflection due to terrains and infrastructure. Similarly, the gateway sensor knows the deployment density and generates the weakest RSSI measurement in the acceptable range for the anchor sensor failed to communicate. This incorporation of RSSI measurements against failed sensors improves the accuracy of estimation because its unavailability may result in increase of error. Information received from several anchor sensors helps to consolidate the assumed signal measurements based on log-normal path loss model RSSI and minimises the location estimation error. PF-SSR improves the particle filter ability to estimate location of the train as given in Algorithm 3. PF-SSR relies on the input from sensor nodes and its accuracy improves with number of replies from multiple sensor nodes.

6.9 Resampling

In the process of particles' update, particles with lower weights get eliminated and number of particles decreases. In such case, the accuracy of location estimation suffers in such a way that the likelihood of a particle to represent accurate train's location may fluctuate large. Resampling enables the procedure to increase the number of particles by duplicating high weight particles. In Particle-Filtering-based schemes, a low variance sampler scheme (Baker, 1987) is used, that focus on to duplicate particles of higher weights. The samples are drawn in such a way that particles are kept in a list and occupy the length of list according to their weights. After that, It draws $\overline{N}$ samples using a single randomly generated number. Let the random number generated be r, such that $r \in (0, \sum W_t)$. It is assumed that r is drawn from a uniform distribution. Then, the first sample is drawn from location r in the list, the second from location $[(r + \frac{2\sum w_t}{\overline{N}}) \mod \sum w_t]$, the i'th sample from location $[(r + \frac{i\sum w_t}{\overline{N}})\ mod\ \sum w_t]$ and so on, until $\overline{N}$ new samples have been generated.

6.10 Train Location Estimation

Estimation of the location $X_t = (x_t, y_t)$ is calculated based on the updated weights of the particle at time t by following equations.

$$x_t = \frac{\Sigma_{i=1}^{N} w_t^{[i]} x_t^i}{\Sigma_{j=1}^{N} w_t^{[j]}} \tag{6.22}$$

$$y_t = \frac{\Sigma_{i=1}^{N} w_t^{[i]} y_t^i}{\Sigma_{j=1}^{N} w_t^{[j]}} \tag{6.23}$$

All the weights are normalised so that $\sum_{i=1}^{N} w_t^{[i]} = 1$. The weighted average of the particles location gives the estimated train location. By simplifying the Equations 6.22 and 6.23 can be expressed as,

$$x_t = \Sigma_{i=1}^{N} w_t^{[i]} x_t^i \tag{6.24}$$

$$y_t = \Sigma_{i=1}^{N} w_t^{[i]} y_t^i \tag{6.25}$$

Therefore, (x_t, y_t) gives the estimated train location at time t.

6.11 Performance Analysis of Particle-Filtering-based Train Localisation Schemes

In this section, the performance of Particle-Filtering-based train localisation schemes, PF and PF-SSR, are analysed through extensive simulations in the OMNET++ simulator (Köpke *et al.*, 2008), using real-world RSSI measurements.

6.12 Performance Metrics

The Kalman Filter is explicitly implemented to process the received measurements for the PF-SSR-based train localisation scheme. In addition, the PF-based localisation scheme is implemented and results are compared with the PF-SSR-based train localisation scheme. I use maximum localisation error and average localisation error as metrics.

Definitions

- **Maximum Localisation Error:** The range of maximum localisation error is the maximum absolute difference between the actual and estimated train location in every 100s. An average of maximum error range is computed over range of maximum error.

- **Average Localisation Error:** The range of average localisation error is the average absolute difference between the actual and estimated train location in every 100s. An average of average error range is computed over range of average localisation error.

6.12.1 Simulation Setup

In the simulations, 145 to 4000 anchor sensors are deployed with various distances between the adjacent anchor sensors, called the deployment density, d_a. Unless specified otherwise, the default values are used for several parameters, such as d_a, S_T and anchor sensor failure percentage, to evaluate the reliability of the proposed schemes (Table 6.1). The reliability here is defined as the successful wake-up of anchor sensors in Zone 2 or chances of hardware or software based failure in an anchor sensor that fails an anchor sensor to wake-up and communicate with the gateway sensor. I use $d_a = 100\ m$, $4\ m$, and $10\ m$ for open field, railway station and tunnel, respectively. Moreover, the train speed, S_T is $40\ m/s$, $10\ m/s$, and $40\ m/s$ for open field, railway station and tunnel, respectively, reflect the realistic train speeds in corresponding environments. The impacts of train speed on localisation error is studied with multiple train speed values, such as $10\ m/s$, $20\ m/s$, $30\ m/s$ and $40\ m/s$. However, when looking into impacts of other parameters, such as reliability and deployment density, on localisation errors, default train speed is used as mentioned earlier. Though, simulations were conducted over other speed values, but these speed values are presented here because these speed values represent these environments. Furthermore, the default anchor sensor failure percentage is 0%. The impact of failed sensor nodes is studied in which several setting of anchor sensor failure are simulated independently. Based on the analysis in Chapter 4, real-world datasets are used in the simulations, which are long-range sensors, short-range sensors, and internal radio sensors datasets for open field, railway station, and tunnel, respectively. There is an unreliable wireless channel model with 10% packet loss rate (Liu *et al.*, 2014). Similarly, anchor sensors communicate with the gateway sensor by sending multiple packets without a requirement for acknowledgement packets.

In simulation, RSSI measurements are selected from a pool of real-world datasets of particular distance between anchor sensor and gateway sensor. For example, in the case of $500\ m$ as distance between adjacent anchor sensors, RSSI measurements are recorded by gateway sensor at each $10\ m$ and RSSI measurements are used from that pool randomly.

6.12.2 Parameter Configurations

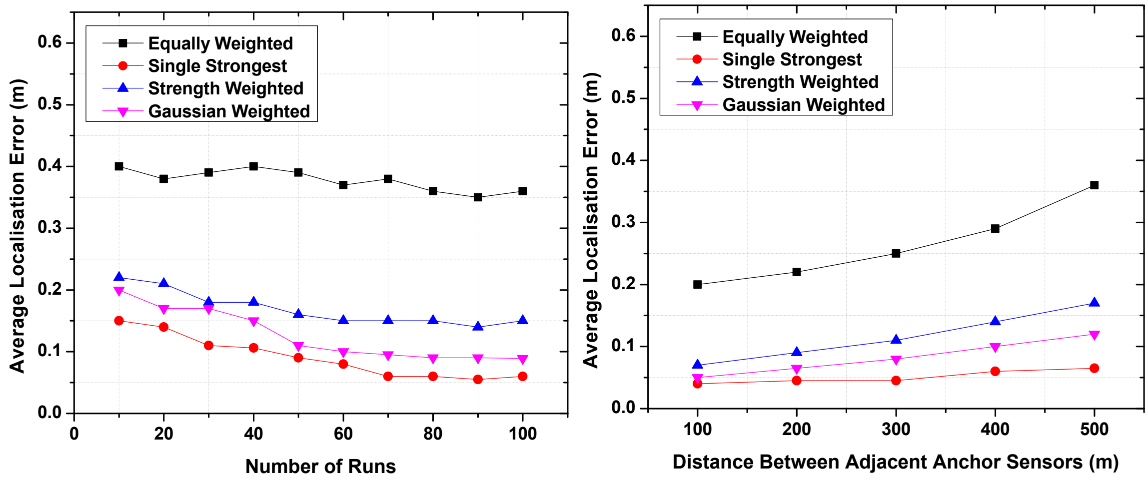
The detailed parameter configurations used in simulation setup are given in Table 6.1 along with their values.

Table 6.1: Simulation Parameters (PF-based Localisation)

Parameters	**Train Speed**			**Reliability**			**Deployment Density**		
	OF	**RS**	**TL**	**OF**	**RS**	**TL**	**OF**	**RS**	**TL**
L (s)	10000	10000	10000	10000	10000	10000	10000	10000	10000
λ	1	1	1	1	1	1	1	1	1
Number of Anchors	145 to 4000	145 to 4000	145 to 4000	145 to 4000	145 to 4000	145 to 4000	145 to 4000	145 to 4000	145 to 4000
Zone 1 (m)	500	500	500	500	500	500	500	500	500
d_T (m)	2	2	2	2	2	2	2	2	2
Simulation iterations	50	50	50	50	50	50	50	50	50
S_{max} (m/s)	10, 20 30, 40	10, 20 30, 40	10, 20 30, 40	40	10	40	40	10	40
Anchor Failure (%)	0	0	0	0,10,20, 30,40	0,10,20, 30,40	0,10,20 30,40	0	0	0
d_a (m)	100	4	10	100	4	10	100-700	2-14	10-50
t^{ub}_{sleep} (s)	1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4

6.12.3 Performance of Weighted RSSI Functions

The weighted RSSI functions play an important role in the particles update stage. Assigning weight to a RSSI measurement develops the credibility of the geographic coordinates of an anchor sensor. In other words, an anchor sensor closer to the gateway sensor will have more impact on the particle's weight update. The impact of weighted RSSI functions on location estimation (PF-based scheme) are shown in Figure 6.6. In this simulation, 500 m are used as distance between any two adjacent anchor sensor nodes, and train speed is 10 m/s. Figure 6.6(a) shows the average localisation error for different simulation runs. Among all weighing schemes, the Single Strongest RSSI function outperforms other schemes with average error of 0.1 m. It can also be seen that the Gaussian Weighted RSSI and Single Strongest RSSI are very close. However, Equally Weighted RSSI function performs worst because it assigns equal weights to unreliable RSSI measurements from anchor sensors at far locations, compared with reliable RSSI measurements from anchor sensors at closer locations. Figure 6.6(b)



(a) Average Localisation Error for different Simulation Runs

(b) Average Localisation Error for different Deployment Densitites

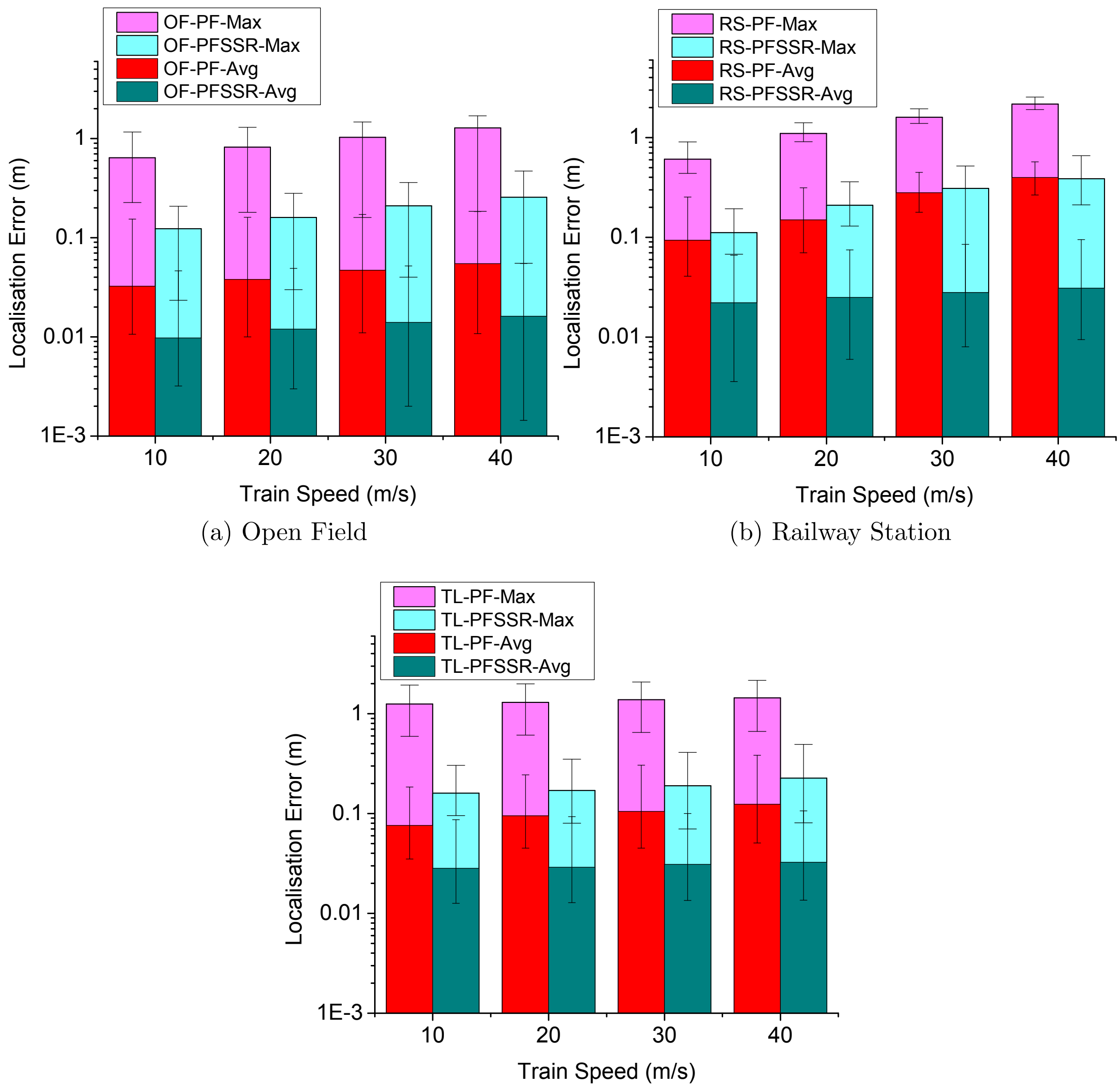
Figure 6.6: Average Localisation Error with Different Weighted RSSI Functions.

plots the average localisation error for same weighted RSSI functions in context of deployment densities. An important observation is that, from the collected datasets, no large difference can be seen in performance of weighted RSSI functions. The Single Strongest RSSI schemes performs better than other schemes. However, it will perform worst in the presence of unreliable anchor sensors (fail to wake up) and sparse networks. However, Gaussian Weighted RSSI function keeps average localisation error low and will not suffer much in the presence of unreliable anchor sensors as it does not rely on single RSSI measurement.

6.12.4 Performance of Particle-Filtering-based Train Localisation Schemes

The performance of PF-based and PF-SSR-based localisation schemes are evaluated in three train representative environments such as an open field (OF), a railway station (RS), and a tunnel (TL). three sets of simulations are conducted to evaluate the impact of train speed S_T, anchor sensor failure (reliability), and anchor sensors' deployment density, on the localisation error. Along with the presentation of maximum and average localisation errors, the range of the maximum localisation error, represented by error bars with wide caps, and the range of the average localisation error, represented by error bars with short caps are given in the resulting figures.

Impact of Train Speed (S_T) on Localisation Errors



(a) Open Field

(b) Railway Station

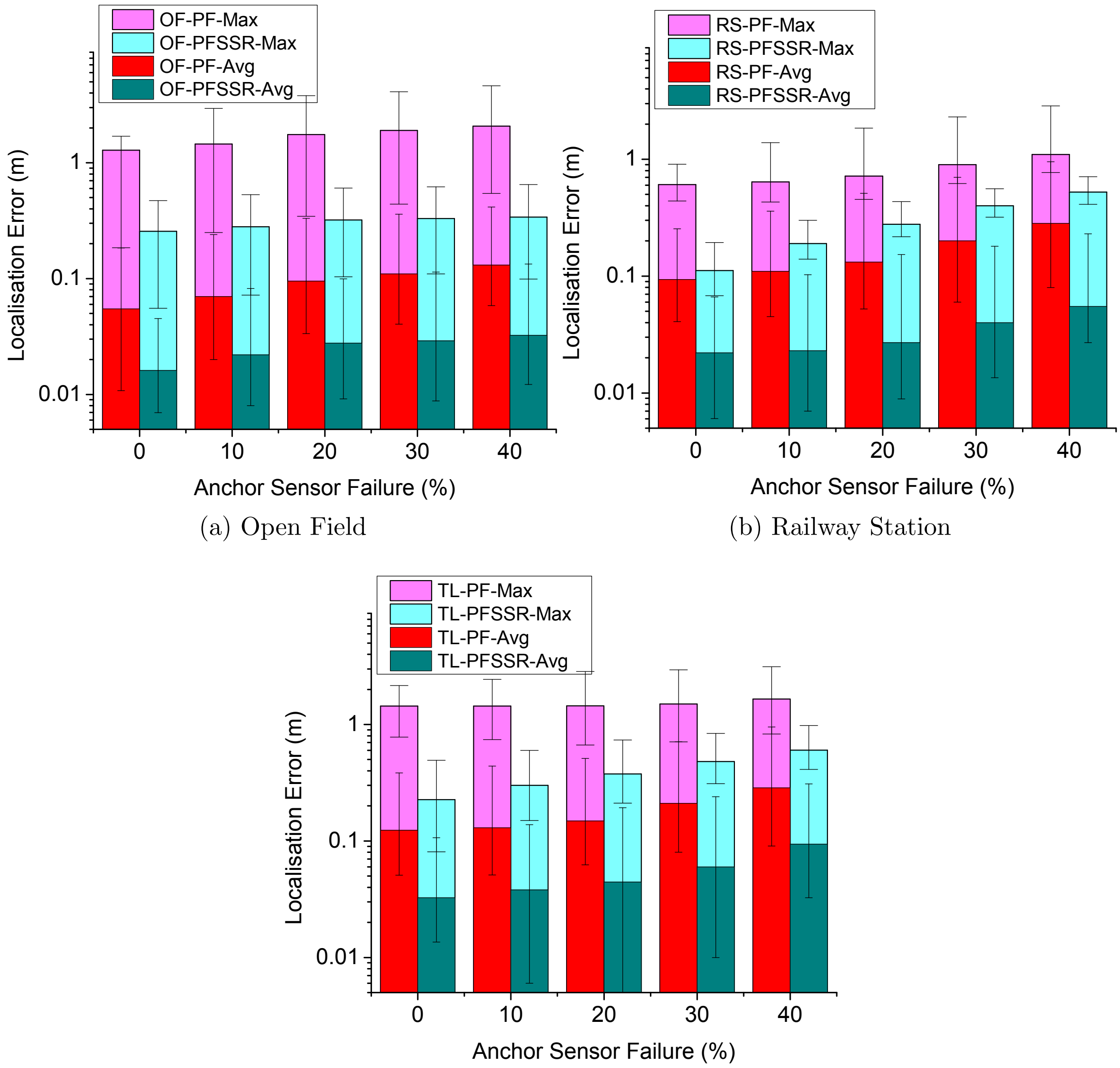
(c) Tunnel

Figure 6.7: Localisation Error at Several Train Speed Settings

Figures 6.7(a), 6.7(b), and 6.7(c) show the localisation error by PF and PF-SSR at various train speed settings. It can be seen that the PF-SSR outperforms PF by minimising both maximum and average localisation error in OF, RS and TL environments. The maximum localisation error by PF-SSR protocol increases with the train speed, but remains under 1 m at all times, that is, 60 cm in OF, 70 cm in RS, and 75 cm in TL, at train speed of 40 m/s. Moreover, the localisation error is higher in tunnel environment compared with open field and railway station, which is because of high noise in RSSI measurements due to signals fading and reflections from the tunnel walls.

The localisation error bar shows the range of error, which suggests that the PF-SSR estimation is more accurate in open field environment. However, in all environments, the average localisation error stays under 0.15 m, that is, 8 cm in OF, 9 cm in RS, and 14 cm in TL, at train speed of 40 m/s.

Reliability of Particle-Filtering-based Train Localisation Schemes



(a) Open Field

(b) Railway Station

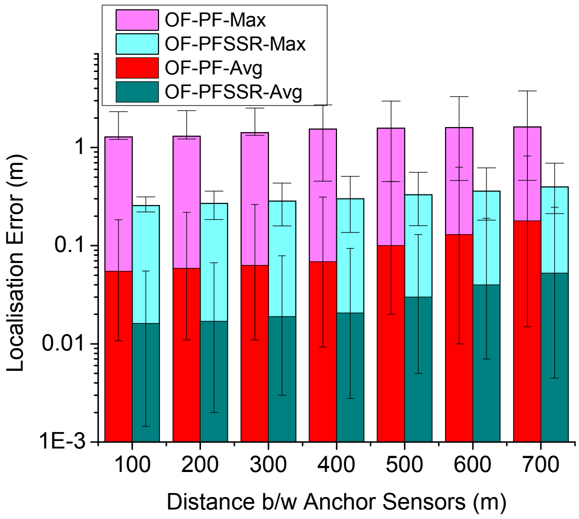
(c) Tunnel

Figure 6.8: Localisation Error at Several Anchor Sensor Failure Settings

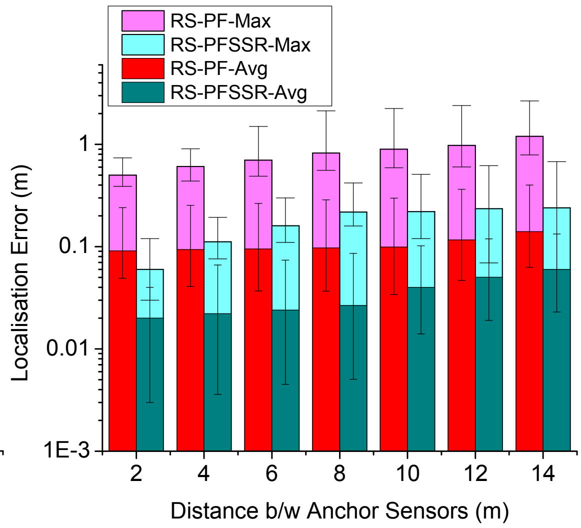
The gateway sensor estimates location of the train after communication with anchor sensors. Therefore, the timely wake-up of anchor sensors and successful communication is vital in the location estimation. In real-time scenarios, anchor sensors may fail to wake up even after the guarantee of BWS protocol because of some hardware or software faults that may affect the accuracy of location estimation. The impact of such reliability issues on the location estimation errors are shown in Figures 6.8(a), 6.8(b), and 6.8(c), for OF, RS and TL environments, respectively. It can be seen that the average and maximum localisation errors increase with the number of failed anchor sensors. The range of maximum and average localisation errors in PF and PF-SSR schemes also increase with the number of failed anchor sensors. However, because of the signal reconstruction feature of PF-SSR, the performance deterioration is minimal and PF-SSR performs better than PF-based localisation scheme. From the figure, it can also be seen that despite high reflections from tunnel walls, PF-SSR is more reliable in tunnel environment, which is due to the low power measurements and dense deployment settings. Consequently, with the dense deployment, if some percentage of anchor sensors fails to wake up, there are still some anchor sensors that can communicate in Zone 1 at that time. Therefore, reliability is related to the cost of deployment. Similarly, among all environments, the PF-SSR scheme estimation error is high in open field, which is due to amplified high power transmission by anchor sensors that are prone to the noise, and the distance estimation from long-range signal power is not very accurate. Moreover, the failure of anchor sensors from close proximity in an open field ebbs the accurate location estimation due to unavailability of active anchor sensors at a reliable distance to communicate with gateway sensor. In a nutshell, PF-SSR outperforms PF and keeps the maximum localisation error at 80 *cm* in OF, 75 *cm* in RS and 100 *cm* in TL, at 40% anchor sensor failure rate. Also, the average localisation errors of PF-SSR stays at 15 *cm* in OF, 25 *cm* in RS, and 29 *cm* in TL, even when 40% of anchor sensors fail to wake up, thus do not communicate with the gateway sensor.

Impact of Deployment Density (d_a) on Localisation Errors

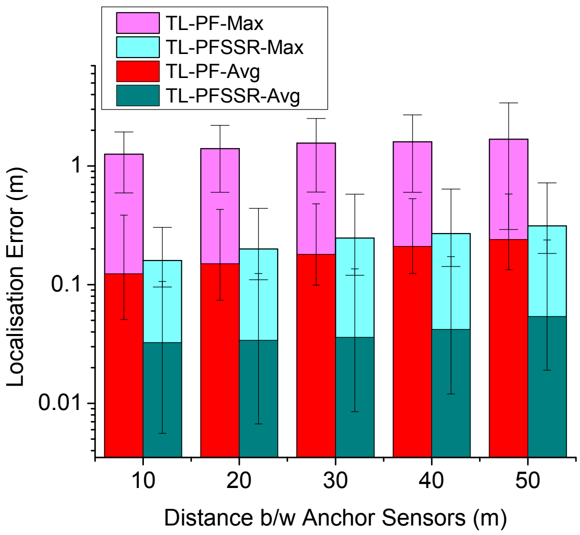
The accuracy of distance estimation from signal strength relates to the distance between sender and receiver, which means that the closer the sensors, the better the accuracy. The impact of noisy measurements from any train localisation environment can be reduced by having several measurements from anchor sensors at reliable, close distance from the gateway. In the Free-space path-loss model, the signal attenuation rate is linked with the square of the distance between the transmitter and receiver. The significant drop in the signal power results in accurate distance estimation and that phenomenon is observed from anchor sensors within close proximity. The impact of



(a) Open Field



(b) Railway Station



(c) Tunnel

Figure 6.9: Localisation Error at Several Deployment Density Settings

deployment density on the train location estimation are shown in Figures 6.9(a), 6.9(b) and 6.9(c) for open field, railway station and tunnel, respectively. In all figures, there is distance between the adjacent anchor sensors (d_a) on x-axis. It can be seen that the range of maximum and average localisation error increases with the distance between the anchor sensors in all cases. However, the increase in the error is high in the tunnel environment because of special noisy features of tunnels. In all cases, the range of upper maximum localisation error with PF-based scheme exceeds 1 m, however, the range of maximum error in PF-SSR-based scheme stays under 1 m even in sparse deployment of anchor sensors, that is, 82 cm in OF with 700 m between sensors, 75 cm in RS with

14 m between sensors, and 88 cm in TL with 50 m between sensors. From all figures, it can be seen that dense deployment reduces the average error and its range but it has an inverse relationship with the cost of anchor sensors deployment. PF-SSR manages to keep the average location estimation error under 0.3 m with sparse deployment in all cases, that is, 26 cm in OF, 17 cm in RS, and 28 cm in TL.

6.13 Related Work

In localisation or tracking problems, an initial example is to track a mobile target with known initial location. A much harder problem is when a mobile target doesn't know its initial location. In such problems, there is need to estimate the location of target by minimising location errors. Particle Filtering is a technique that is commonly preferred in such scenarios (Montemerlo *et al.*, 2003). It allows target to extract part of probable space from developed PDF. Particle Filtering is efficient once space is minimised. The benefit of Particle Filtering technique is that it allows to develop PDF from any distribution, unlike Kalman Filter which develops PDF from normal distribution (Fox *et al.*, 2001). Gustafsson *et al.* (2002) presents an overview of Particle-Filter-bsed localisation approaches. A distributed particle filter approach focuses on improving robustness by introducing constant set function (Wu and Pei, 2013). In their approach, data fusion techniques are used to increase focus on more relevant particles. Such addition increases the accuracy of location estimation.

In another piece of research a hybrid approach is proposed for partial linear and non-linear problem space. The proposed solution discussed the use of Kalman Filter for linear part of a problem and incorporates it with Particle Filter. This approach reduces the cost and complexity of the system (Schön *et al.*, 2005).

In Simultaneous Localisation And Mapping (SLAM) problems, map of target area needs to be learned and updated. Howard (2006) proposed a multiple robot-based scheme to learn the map. This approach is quick and more accurate map can be generated. However, use of multiple robots incur high resource consumption. Ren and Meng (2009) further investigated this issue and proposed a solution based on multiple power based transmissions. The proposed scheme increased efficiency of map learning.

The SLAM problem is investigated by few researchers and proposed a FastSLAM algorithm (Montemerlo *et al.*, 2002). In FastSLAM, the benefits of Kalman Filters and Particle Filters are combined in such a way that Particle Filters represent posteriors for several paths of mobile target and multiple Kalman Filters are linked to a each Particle

Filter's attribute, that is, feature of a map. FastSLAM algorithm also considers less important features of a map and lowers its impact on estimation by reducing its weight (Durrant-Whyte *et al.*, 2003; Majumder *et al.*, 2000).

Particle Filtering also offer solutions to develop service robots, where identification of their position is prime requirement for rest of navigation. Schulz *et al.* (2001); Montemerlo *et al.* (2002) focused in service robot solutions. In each of the developed approach, Particle Filters were used to track objects. In the former approach, factored Particle Filter was introduced that compute likelihood of particles based on received measurements. In the later approach, features are extracted by comparison of maps, that is, new map acquired by range measurements and previously constructed maps. Another effort in localisation using Particle Filters in maps was proposed by Avots *et al.* (2002).

Conditional Particle Filters are proposed by Montemerlo *et al.* (2002), in which pose of a mobile target is located against several number of people sitting in the surroundings. In this approach, large distribution function is constructed for people space and mobile target's position is located by comparing its small distribution function. This approach is efficient as it is tolerant to sensor noise and uncertainty.

6.14 Summary

In this chapter, I introduced the Particle Filtering technique, its derivation, stages and models. Based on the Particle Filtering technique, I developed novel PF-based and PF-SSR-based localisation algorithms and presented their several steps in detail. Particle Filtering is a common solution for tracking object problems that predicts the state of a target through a motion model and updates the state estimation through a measurement model. In train localisation, the location of the train is predicted through the motion model. In the measurement model, two types of measurements are used such as the geographic coordinates of anchor sensors and the signal strength of their corresponding transmissions. RSSI reduces with the increase of distance between the sender and the receiver as suggested by log-normal path loss model. However, there are several environmental factors that can also influence the RSSI measurements such as interference of other frequency channels, reflection of signal or multi-path fading. The proposed Particle-Filtering-based train localisation algorithms relies on RSSI measurements received from anchor sensors. Due to the fragile nature of RSSI, the localisation accuracy gets affected and Particle Filter is used to increase the accuracy. However, Particle Filter uses location information of anchor sensors along with RSSI in measurement model. Therefore, this novel measurement model enables the gateway sensor on the train to minimise the negative influences on RSSI measurements by using the geographic coordinates. The particles represent the location of the train with some probabilities as their associated weights. A weighted RSSI-based likelihood function is developed to assign and update the weights of the particles. The PF-based localisation train algorithm is then improved by adding another step of signal strength rectification in PF-SSR-based train localisation algorithm.

The proposed schemes are then evaluated through extensive simulations using real-world RSSI measurements collected from experiments conducted in open field, railway station and tunnel. The proposed train localisation algorithms are then compared in context of their ability to minimise the localisation error under several deployment density variations, train speed and reliability parameters. The results suggest that by using Particle Filtering algorithm and suitable weighted RSSI likelihood function, the location estimation error can be significantly reduced in all railway representative environments. In the next chapter, I will present the consensus-based anchor sensor management scheme to manage the anchor sensors deployed along the track.

Chapter 7

Consensus-based Anchor-sensor Management Scheme for Train Localisation

In this chapter, I present the Consensus-based Anchor-sensor Management Scheme (CAMS) for a WSN-based train localisation system. The proposed scheme enables anchor sensors to dynamically figure out the faulty sensors among them by developing consensus, and to report to the gateway sensor. Moreover, this scheme also helps anchor sensors to develop consensus about the estimated path loss ratio to increase the accuracy of WSN-based train localisation. In the remainder of the chapter, I describe the simulation setup and results of CAMS based on the real-world collected data from the field experiments.

7.1 Introduction

Railway systems have provided an important means of transport over the past hundred years, with significant investments having been made in safety infrastructure. In recent years, real-time train localisation is becoming more essential in meeting the need for safety.

The anchor sensors along the railway track may suffer from the location errors caused by software or hardware bugs. Therefore, they need to be re-calibrate their geographic coordinates and send calculate the path loss of the signals. Moreover, the presence of faulty sensors (Kaligineedi *et al.*, 2010) in the system can also deteriorate the accuracy of the location estimation. All these issues should be addressed in the

WSN-based train localisation system. Manually sorting out such problems by human beings incurs significantly higher costs. The management and maintenance of the anchor sensors with the help of each other play an important role in the stability of the whole localisation system.

Therefore, the need for a management scheme comes into play, which can enable anchor sensors to detect the faulty sensors among themselves. The faults should be reported to the gateway sensor for further analysis. Furthermore, the management scheme should assist anchor sensors to estimate the path loss ratio of their signals, which depends very much on the surrounding environment and affects directly the distance estimation based on RSS. Such a management scheme can clearly improve the accuracy of train localisation by excluding the faulty sensors and re-calibrating the parameters of the anchor sensors like path loss ratio.

In this chapter, I propose a management scheme called CAMS (Consensus-based Anchor-sensor Management Scheme) for our WSN-based train localisation system. CAMS allows anchor sensors to share their opinions about trustworthiness of their neighbour sensors and develop consensus to detect the faulty sensors.

The anchor sensors can automatically re-calibrate path loss ratio and geographical coordinates. The main contributions of this work are summarised as follows.

- I propose a new CAMS for management and maintenance of anchor nodes in WSN-based train localisation systems. CAMS uses a consensus-based approach to manage anchor nodes in train localisation. Additionally our consensus algorithm uses history data as well as the current data to reduce false detection ratio of faulty nodes and increases the accuracy of the re-calibrated path loss ratio.

- CAMS is implemented in a simulated environment using MATLAB. The simulation is based on the real data collected from field experiments in various environments such as open field, train station and a tunnel.

- From the results collected from the simulations, I find that CAMS can effectively detect the presence of faulty nodes in the system. The results show that, with the re-calibration of the path loss ratio of the anchor nodes, the accuracy of train localisation can be improved up to 15%.

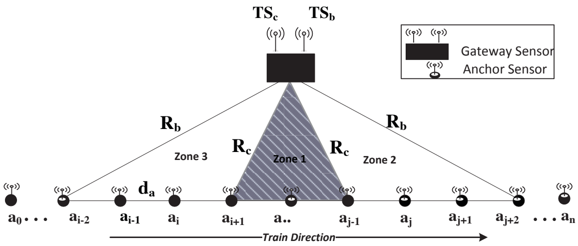
7.2 Challenges in the Sensors' Management

The small size of sensor devices encouraged scientists and researchers to deploy large networks in harsh environments. In such environments, aerial drop of sensor devices is used as a deployment option. Along with associated benefits of small size devices, there are some limitations on the capability of hardware resources, such as, limited memory and power supply, and processing power. In rough environments, the limited power supply is generally considered an important issue. Sensor devices self-configures the network and lasts for limited network lifetime, days to months. Due to these constraints, faults are expected to occur frequently, compared with wired networks. The network availability and energy consumption has a trade-off and that is maintained through the required objective of a network. Further, sensor devices are prone to environmental conditions, such as fire, snow or rain, which may turn sensor devices faulty. These faults can be malicious sensing data, inactive or delayed responses, etc. For such reasons, sensor management in WSNs is an important aspect of research.

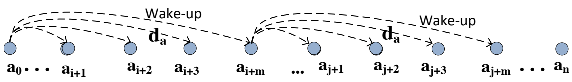
In the WSN-based train localisation system, sensors devices are exposed to the harsh environment and are prone to several faults. Sensor devices can do self assessment to report faults such as low battery power. Contrarily, if the hardware or software is not capable enough to detect the existing faults, neighbouring sensor devices can detect faults and faulty sensors among themselves. Sensor devices are also useful for estimating the path loss ratio to assist the gateway sensor to increase the accuracy of train localisation. However, a faulty sensor node can elevate the results, consequently, compromising the accuracy of train localisation. In such scenarios, a sensor management scheme can better serve the purpose in a cost effective way, compared with manually sorting out faults in sensor network.

7.3 Problem Statement

In this chapter I consider a set of m single hop anchor sensors denoted as $\{a_1, a_2, \ldots, a_m\}$, as shown in Figure 7.1. I model the network as an undirected graph $G = (V, E)$, where V is the set of anchor sensors and E the set of communication links between the sensors. An edge exists between any two sensors that are in each other's communication range. As the anchor sensors in the network follow asynchronous duty-cycling without any knowledge of sleep schedules of neighbouring anchor sensors, they must wake up to perform the faulty sensor detection and calibration. Each anchor sensor



(a) Gateway and Anchor sensors



(b) CAMS Architecture

Figure 7.1: WSN-based Train Localisation System

sleeps for t^{ub}_{sleep} time, which is computed as in Chapter 5. In CAMS, periodically, any anchor sensor that wakes up continually broadcasts beacons for t^{ub}_{sleep} to wake up its neighbour sensors to detect faulty sensors and to calibrate. Once awoken, the anchor sensors develop opinions about their neighbour sensors. The opinion of a neighbour sensor includes the evaluation of the claimed location of the sensor. If the claimed location of the neighbour sensor is found to be at the estimated distance according to RSS with acceptable threshold, the opinion score of the neighbour increases; otherwise opinion score decreases. The acceptable threshold depends upon the radio chipset's noise range, which is $\pm 6dB$ for CC2420 radio chipset (Texas Instruments, 2003). Each anchor sensor stores opinions about its neighbours in Neighbour Opinion Table (NOT).

The first problem I address is "how to make anchor sensors to communicate and develop opinions about their neighbour sensors in the presence of asynchronous duty-cycling". The opinions are important for analysing the trustworthiness of neighbour sensors. The untrusted anchor sensors, for example, faulty sensors must be reported to the gateway sensor to eliminate their false input, which may affect the accuracy of the WSN-based localisation system. To calculate opinion scores about the neighbour sensors, they must be in wake-up state to communicate. In CAMS, BWS (as discussed in Chapter 5) is used to enable anchor sensors to wake up.

The second problem is "how to evaluate their NOT in order to detect the faulty sensors based on consensus of the received individual opinions". Anchor sensors are required to report NOT to each other. Each anchor sensor develops consensus based on received NOTs and marks and eventually eliminates the faulty anchor sensors from the system. The detection of such anchor sensors allows the gateway sensor to ignore the inputs of the faulty anchor sensors and reduce the effects of their biased opinions.

The third problem I address is "how to assist the gateway sensor to estimate the consensus-based re-calibration of path loss ratio estimated by the individual anchor sensors". The consensus-based re-calibration of path loss ratio helps the gateway sensor to improve the accuracy of train localisation.

7.3.1 System Assumptions

CAMS is proposed under the following assumptions, though these assumptions can be relaxed with slight modification of our existing system.

- Each anchor sensor is anchored at a fixed location along the railway track. They cannot move without physical intervention. If a sensor is maliciously removed to a different location, it will be treated as a faulty sensor in CAMS.

- The geographic coordinates of each sensor are hard-coded before deployment. However, the location information could be different later from the real location due to sensor malfunction or malicious dislocation.

- The ID of an anchor sensor is unique and encrypted with the message using a shared key. In the case of intrusion where the shared key is cracked and the ID is forged, the sensor with the forged ID will be detected by CAMS as a faulty sensor due to its peculiar behaviour such as incorrect claimed location. Though this is an unsophisticated attack, but CAMS is capable of looking in to only these aspects due to its focus on detection of faulty nodes. Other than this, CAMS is not capable of detecting malicious nodes. All reports on faulty sensors will be sent urgently to the management centre for immediate human response.

7.4 CAMS: Consensus-based Anchor sensor Management Scheme

CAMS enables anchor sensors to work in their consensus-based management. The following sections discuss CAMS in detail.

7.4.1 Impacts of Faults in Anchor sensors

Anchor sensors provide gateway sensor with information such as geographic coordinates. Therefore, to infer meaningful conclusions with the received information, its quality must be ensured. The potential occurrence of faults in the sensor network can affect the data integrity, which can lead to wrong estimation of the train's location. There are several features that lead to faults in the sensors such as transducer input-output detection range, sensor age, battery state, noise, sensor response hysteresis, and dislocation. Input from anchor sensors with aforementioned faults can affect the accuracy of WSN-based train localisation systems.

Over time, anchor sensors may not remain consistent with their hard-coded geographic coordinates due to change in their position by physical, unauthorised intervention or because of sensor malfunction. This inconsistency leads to dissemination of misinformation. Moreover, low battery states can also make anchor sensor's parameters such as path loss ratio deviate from the expected range. Similarly, environmental effects such as weather can change those parameters to an unacceptable range.

CAMS allows anchor sensors to communicate with each other to detect such faulty sensors based on consensus and assist the gateway sensor to neglect their inputs. In addition, it allows anchor sensors to re-calibrate their path loss ratio to improve the accuracy of WSN-based train localisation. It is worth noting that the assumption of encrypted ID keeps the system safe from the intrusion of malicious sensors. However, if a malicious sensor successfully breaks into the system with a forged duplicated ID, it can be detected as a faulty sensor as soon as it starts to transmit wrong location information. However, with any sophisticated attack, CAMS will not be able to detect malicious behaviours. Here, it is worth mentioning that CAMS focus is to detect faulty nodes and the security section of CAMS is part of future work.

7.4.2 Computation of Opinion Score

Periodically, anchor sensors perform calibration and exchange their opinion about each other. When an anchor sensor wakes up to perform CAMS tasks, it continually broadcasts beacons for t_{sleep}^{ub} time to wake up its neighbour anchor sensors as given in the BWS protocol. Once its single-hop neighbours are awoken and broadcast their location information, it computes and stores the opinion scores of its neighbour sensors in its NOT table. Initially, the opinion score computed by an anchor sensor about any of its neighbours is zero as it knows nothing about the neighbour.

Suppose a neighbouring sensor a_i broadcasts a packet to a sensor a_j. The receiver sensor a_j develops its opinion about a_i after examining the location information Loc_{a_i} sent by a_i. Let d be the distance between a_j's location Loc_{a_j} and a_i's claimed location Loc_{a_i}, and d' be the distance estimated according to the log-distance path loss model in Eq. 7.1.

$$RSS(d') = P_{tx} - PL(d_0) - 10\eta \log_{10} \frac{d'}{d_0} - X, \tag{7.1}$$

The path loss model in Eq. 7.1 is a well-known radio propagation model (Xu *et al.*, 2010) that predicts the path loss a signal encounters over distance, and it has been widely used for distance estimation.

Sensor a_j computes its opinion about a_i as follows. If the difference between the estimated distance and the claimed distance of a_i is within acceptable threshold (θ_1), a_j gives a positive vote ($v = 1$) to a_i; otherwise, a_j casts a negative vote ($v = 0$) to a_i. The acceptable threshold depends upon the radio chipset's noise range, which is $\pm 6dB$ for CC2420 radio chipset (Texas Instruments, 2003). The opinion is computed using Eq. 7.2. It depends on the historic opinion as well.

$$O'_{j \to i} = pO_{j \to i} + (1 - p)(1 - \frac{|d - d'|}{\theta_1})v, \tag{7.2}$$

where, $O'_{j \to i}$ is the update of the opinion about the trustworthiness that a_j develops about a_i, p is the weight assigned to the historic opinion. The term $(1 - \frac{d-d'}{\theta_1})$ decreases with the increase of deviation between d and d' if the deviation is within the acceptable threshold. If the deviation is larger than the threshold θ_1, v is zero and the opinion is gradually decreased to zero.

The above process is carried out in each sensor periodically. However, there is a tradeoff between energy saving and the frequency of the process. Assuming the sensors are not faulty very often, the process could be less frequent, say once a day. However, the process could also be triggered by the gateway on the train if the gateway finds the deviation of a sensor's claimed location and its estimated location based on Eq.7.1 is larger than the threshold.

7.4.3 Consensus-based Faulty Sensor Detection

After the information exchange between the anchor sensors, each anchor sensor updates its NOT. Then each anchor sensor broadcasts its location information, opinion about its neighbours (NOT), transmission power (P_{tx}), date of deployment, and residual battery level (which can be estimated as given in (Zhao *et al.*, 2002)). Each sensor then develops consensus according to the received opinions from other anchor sensors using Eq. 7.3, where CS'_{a_i} is the consensus score about the anchor sensor a_i. It includes the averaged received opinions from other sensors and also incorporates historic consensus score CS_{a_i} with a weight q.

$$CS'_{a_i} = qCS_{a_i} + (1-q)\frac{\sum_{j=1}^{m} O_{j\to i}}{n} \tag{7.3}$$

After the anchor sensors communicate and share their opinions with each other, if the consensus score of a sensor is below a threshold (θ_c), that sensor is identified as a faulty sensor and will be excluded from the trusted sensor list. This information is communicated in a report to the gateway sensor during anchor-gateway communication. Similarly, each anchor sensor also marks the anchor sensors with residual battery power under the acceptable threshold (θ_2) as faulty anchor sensors. The ages of anchor sensors are computed from the date of deployment and their recommended operational period. Each anchor sensor receives the date of deployment of other sensors and computes their age. A list of anchor sensors with their expiry date under a threshold (θ_a) is reported back to the gateway sensor for possible replacement.

Algorithm 4: Detection of faulty anchor sensor by a_j

for *Each anchor sensor* a_i **do**
    if $|d - d'| \leq \theta_1$ **then**
        $O'_{j\to i} = pO_{j\to i} + (1-p)(1 - \frac{d-d'}{\theta_1})$
    else
        $O'_{j\to i} = pO_{j\to i}$
for *Each anchor sensor* a_i **do**
    $CS'_{a_i} = pCS_{a_i} + (1-p)\frac{\sum_{j=1}^{m} O_{j\to i}}{n}$
    if $CS_{a_i} < \theta_c$ **then**
        Enlist a_i as faulty anchor sensor
    else
        a_i is a trustworthy anchor sensor

7.4.4 Consensus-based Calibrated Path Loss Ratio Estimation

The anchor sensors can also calibrate the path loss ratio (η) in Eq. 7.1 to assist the gateway sensor in the WSN-based train localisation. Each anchor sensor broadcasts its transmitting power (P_{tx}) along with geographic coordinates, which help the receiving anchor sensors to estimate the attenuation rate of signal strength, called path loss ratio.

Eq. 7.4 defines the simplified equation for an anchor sensor a_j to calculate η_{a_i}:

$$\eta_{a_i} = \frac{P_{tx}^{d_0} - P_{rx}^{d'}}{10logd'}, \tag{7.4}$$

where $P_{rx}^{d'}$ is the receiving power. Each anchor sensor receives the calculated η_{a_i} from other anchor sensors and calculates the consensus η'_{cs} as shown in Eq. 7.5. The equation considers the weighted averages of the received η according to the consensus scores of the corresponding sensors. This gives more weight to the trusted anchor sensors' estimation. In addition, an anchor sensor also considers the historic path loss ratio η_{cs} when updating η'_{cs} and r is the weight assigned to the historic path loss ratio.

$$\eta'_{cs} = r\eta_{cs} + (1-r)\frac{\sum_{i=1}^{n}\eta_{a_i}CS_{a_i}}{\sum_{i=1}^{n}CS_{a_i}} \tag{7.5}$$

7.4.5 Reporting to Gateway

Each anchor sensor updates the gateway sensor about the new path loss ratio η, and updates a list of faulty sensors when train passes. The gateway sensor later on updates the human resource involved in the management of the anchor sensors to undertake the necessary steps in rectification of faults or removal of faulty sensors. In the long run such management improves the accuracy and lifetime of the system, as shown in the next section.

7.4.6 Analysis of CAMS in the Localisation System

In the CAMS, anchor sensors are capable of detecting the faulty nodes among each other. The Particle Filtering based train localisation system depends on the input of anchor sensors. The proposed localisation scheme operates on several functions such as, weighted RSSI function. One of the weighted RSSI function, that is, single strongest RSSI function, allows the Particle Filtering based train localisation system to estimate the location of a train even if the RSSI is received from one sensor within communication Zone 1. This implies that if a single anchor sensor is working fine and others become faulty, localisation system can work fine. However, the accuracy of location estimation will be reduced. On the other side, railway operators will be required to put a threshold on the number of faulty nodes before a maintenance operation can start.

7.5 Performance Evaluation

In this section, I will discuss the implementation of our CAMS scheme to assist the overall maintenance of the train localisation system. In our simulation, CAMS scheme is implemented, which enables the anchor sensors to coordinate and mark the faulty sensors among them, notify the faults, and calculate the path loss ratio to assist the gateway sensor to increase the accuracy of the train localisation.

The radio characteristics of the anchor and gateway sensors used in our analysis are taken from CC2420 chipset data sheet (Texas Instruments, 2003). All simulations are run independently and their results are averaged for all iterations. The performance metrics I have evaluated are consensus score, opinion score, path loss ratio estimation, and distance estimation error. The other simulation parameters are deployment density, average number of received packets, weights for historic opinion and consensus score, and weights for historic path loss ratio. In order to perform these computations, sensor devices are capable of doing such computations.

7.5.1 Simulation Setup

Simulations have been done on a MATLAB simulator. In simulations, I have used real data collected from field experiments in railway representative environments such as open field, railway station, and tunnel. However, the faulty sensors are incorporated in simulations through a random sensor failure model on top of RSSI measurements. Our experiments are based on several Maxfor's MTM sensor platforms (MTM, 2012) such as MTM-CM3300, MTM-CM5000, and MTM-CM4000 with transmission ranges of 800 m, 150 m, 150 m, respectively. Moreover, in our field experiments, I have deployed anchor sensors in several deployment density settings to collect RSS measurements. The simulation results show how different variables shape the opinion and consensus score of anchor sensors that assist each anchor sensor to mark the faulty anchor sensors. The consensus threshold $\theta_c = 0.4$ is calculated based on average difference between consensus scores of trusted and faulty sensor over 500 iterations. In the simulation results, I have simulated opinion score with different values of p, q and r to show the pros and cons of large and small historic weights. However, a systematic calculation of these values is yet not incorporated. However, the systematic calculation of these variables depends on the system parameters, such as, number of neighbours, or number of received packets, which is related to time allowed for communication between anchor sensors. The longer time period is better but it has a tradeoff with energy consumption.

The detailed configurations for the simulation parameters are given in Table 7.1.

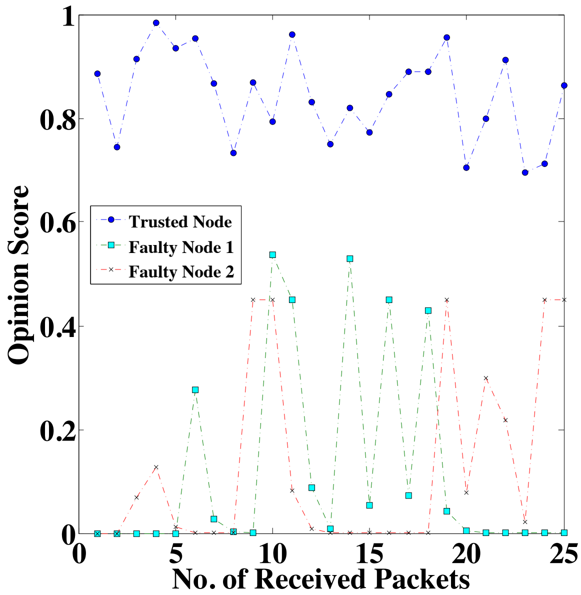
Table 7.1: Simulation Parameters (CAMS)

Parameters	Values
No. of 1-hop anchor sensors	25
Average Received Packets	25
p	0.5
r	0.5
θ_c	0.4
Voltage	3.0v
θ_1	5 m
q	0.5
θ_2	5%
0.4	θ_a

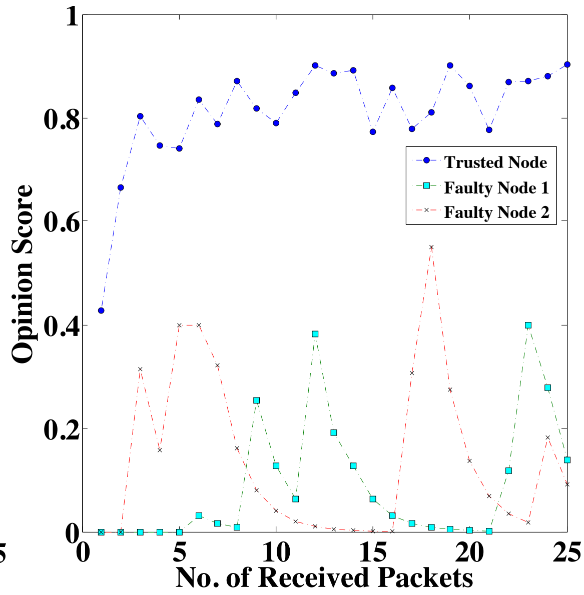
7.5.2 Detection of Faulty Anchor Sensors

Each anchor sensor develops an opinion about its neighbour anchor sensors by receiving packets. The opinion score categorises anchor sensors as trusted or faulty based on difference between the claimed and estimated distance. The number of received packets also affects the computation of opinion score about sending anchor sensor. In the first set of results, I have 2 faulty anchor sensors in the system. Fig. 7.2 shows the opinion scores of both faulty anchor sensors and a trusted anchor sensor calculated by a trusted anchor sensor. It can be seen that the opinion score fluctuates a lot when the weight of the historic opinion is small such as $p = 0.1$ as shown in Fig. 7.2a. However, opinion about other anchor sensors becomes more stable with the increase in the weight assigned to historic opinion score to 0.5 and 0.9, respectively, as shown in Figs. 7.2b and 7.2c. This takes more number of packets by sender to develop an opinion score about it and thus takes longer to pass the threshold of trustworthiness (θ_c). Here, it can also be seen that small number of packets transmitted may not yield consolidated results. Therefore, large number of packets need to be transmitted between anchor sensors. Further, the cause of fluctuations in opinion scores is because of lack of system maturity. Here, system maturity means the availability of data. System maturity increases (fluctuations will be reduced) with increasing history score or by increased number of received packets.

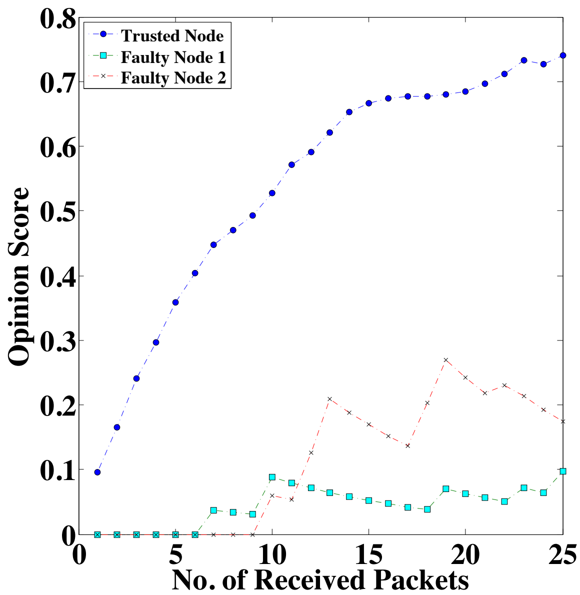
Each anchor sensor compiles the consensus score from opinions received from other anchor sensors. In our simulation, the number of neighbouring anchor sensors varies



(a) $p = 0.1$



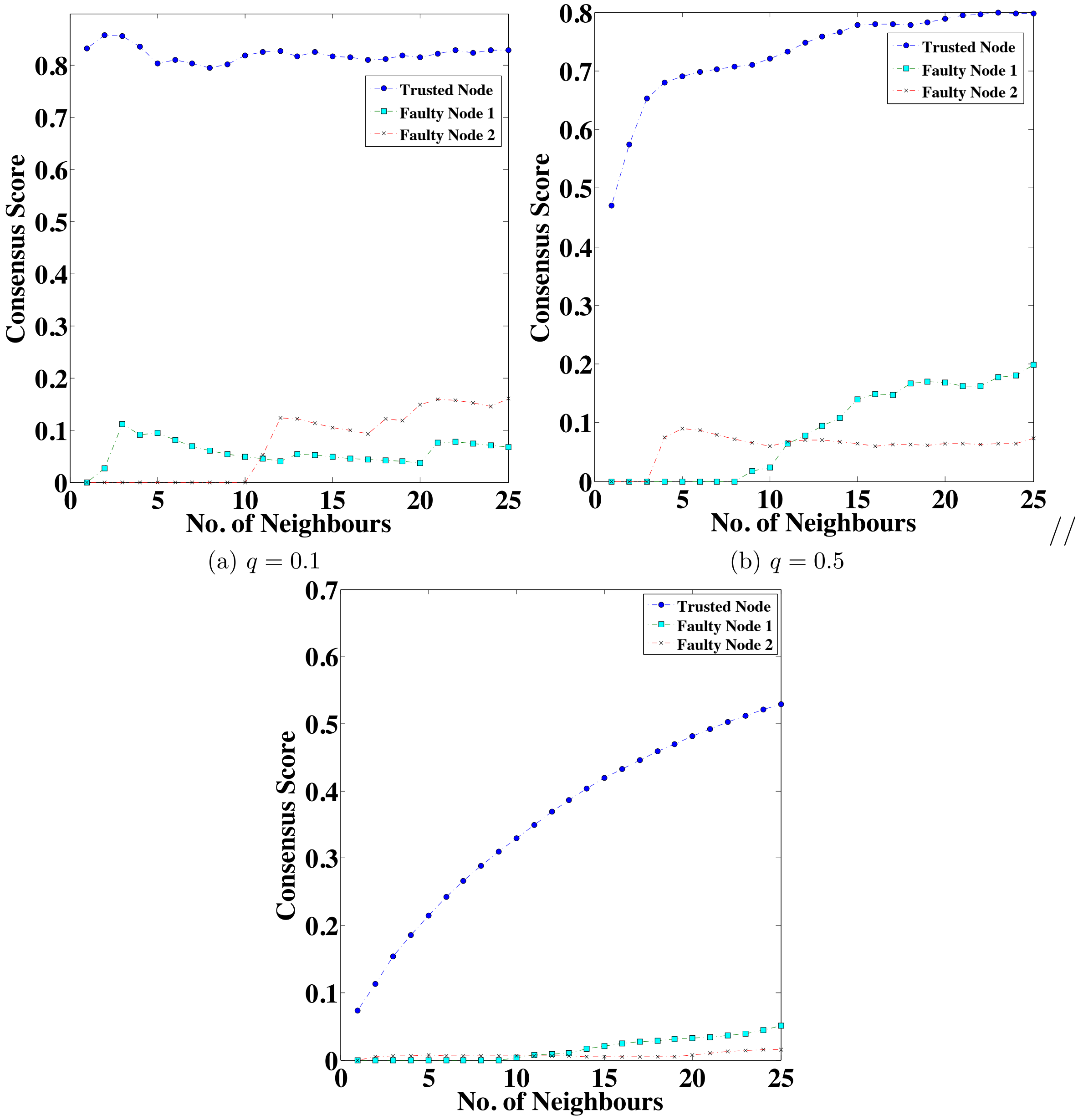
(b) $p = 0.5$



(c) $p = 0.9$

Figure 7.2: Opinion Score Computed by an Anchor Sensor

from 5 to 25, which is called deployment density. As sensor devices have transmission ranges of 150 m and 800 m, these number of neighbouring sensor devices can develop sparse to lightly dense network deployment scenarios. In Fig. 7.3, I have shown consensus score computed by a trusted sensor about 3 anchor sensors: 1 trusted anchor sensor, and 2 faulty anchor sensors. In Figs. 7.3a, 7.3b, and 7.3c, it shows the impact of weight factor given to the past consensus score, which is 0.1, 0.5, and 0.9, respectively.



(a) $q = 0.1$

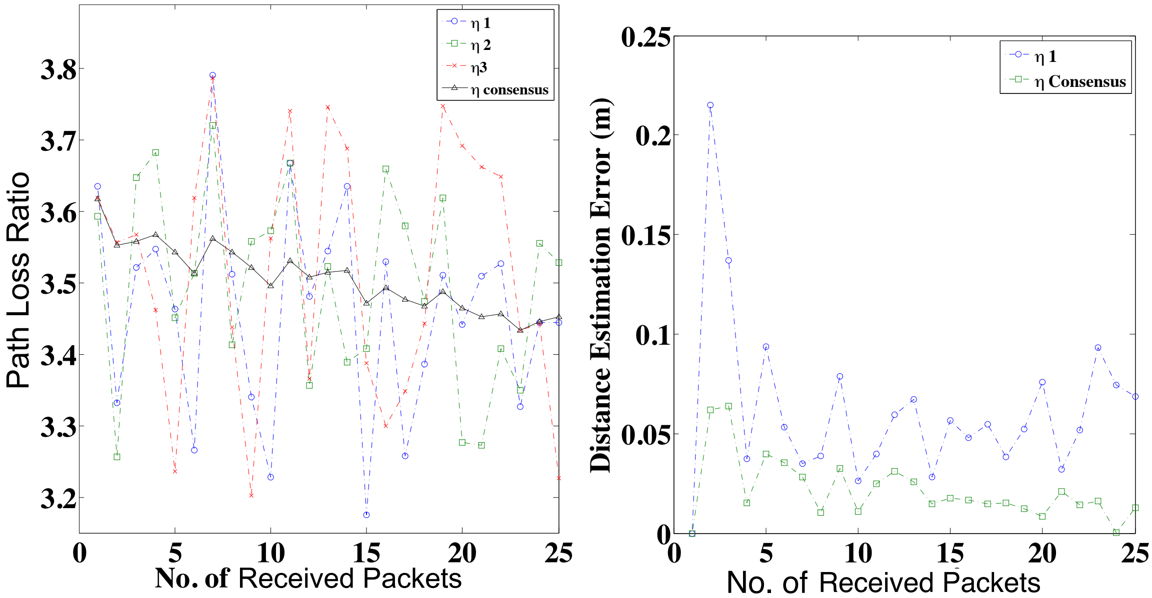
(b) $q = 0.5$

(c) $q = 0.9$

Figure 7.3: Consensus Score to Mark the Faulty Anchor Sensor

It can be seen that if the consensus score, CS_{a_i}, relies least on the historic data the score is significantly high for trusted anchor sensor even if the claimed location deviates to the far extent within the acceptable margin. However, if I increase the weight percentage to 0.5 and 0.9 it takes more opinions from multiple anchor sensors to develop the consensus score and it takes more time to see the decline in the consensus score of a trusted sensor, which turns out to be faulty later on.

7.5.3 Estimation of Path Loss Ratio



(a) η'_{cs} (b) Distance Estimation

Figure 7.4: Impact of Consensus-based Path Loss Ratio Estimation

The estimation of consensus-based path loss ratio represents the real-world signal attenuation rate in a particular environment. Such a parameter assists the gateway sensor to improve the accuracy of WSN-based train localisation. The consensus-based path loss ratio based on calibrated path loss ratio estimated by each anchor sensor and its impact on the accuracy of the localisation system are shown in Figs. 7.4a and 7.4b, respectively. Fig.7.4a presents the path loss ratio calibrated by three trusted anchor sensors. It can be seen that because of signal reflections, each anchor sensor's path loss ratio fluctuates significantly, which is based on RSS. However, I can see that the consensus based path loss ratio is relatively more stable and it improves with the increase in the number of packets received. The impact of consensus-based path loss ratio for localisation error is shown in Fig 7.4b. The localisation error ranges from 0.22 m to 0.04 m and then drops down to 0.06 m to 0.001 m while using consensus-based path loss ratio. This improves the localisation accuracy from almost 5% to 15%, which means the error range decreases with use of more appropriate path loss exponent.

7.6 Related Work

Management of sensors in the WSNs is an important research issue for the stability of the system (Yu *et al.*, 2007). In particular the safety related application, such as train localisation, creates significant increase in the importance of sensors post deployment maintenance and management. It includes the calibration of anchor sensors, detection of possible faults and malicious entity in the system. Along with these typical research benefits, anchor sensor management schemes can also assist the gateway sensor to estimate the attenuation rate of signal strength to increase the localisation accuracy.

Marti *et al.* (2000) proposed watchdog and pathrater techniques to monitor the compromised sensors in the ad-hoc networks scenarios. In the proposed scheme, watchdog identifies the existence of compromised nodes and pathrater technique tries to establish route for communication by ignoring those malicious nodes. Together watchdog and pathrater improve the system performance overall.

Though the cooperative sensing is an important technique to observe the phenomenon of interest in the WSNs, it raises new concerns about the reliability and the security, as expressed by Mishra *et al.* (2006). The malicious users may get access to the network and can affect the aggregated data because by default every sensor trusts the other neighbouring sensors. Moreover, the work in (Song and Zhang, 2008; Zhou *et al.*, 2010) discuss the drawback of cooperative sensing for large scale networks in which synchronisation is another uphill task to achieve, and it gets worse with the duty-cycling sensors.

Kaligineedi *et al.* (2008) discuss the detection of outlier values to filter out prior to manipulation. It computes the trust factor to rate the reliability of user that is used as a weighting factor in the calculation of mean values of received data. However, authors extended their idea of cooperative sensing in (Kaligineedi *et al.*, 2010) and used it to detect the malicious users by the received outlier values.

Srinivasan *et al.* (2006) uses an interesting idea of a majority voting scheme in which each beacon sensor with a known location computes the repute of other beacon sensors and caste votes upon request by other sensors to judge the trust factor of that beacon. However, if a beacon sensor pretends to be a legitimate sensor for some time until it gains positive repute, it may not be detected as malicious if later it starts to spread misinformation.

The sensors can also inform the gateway sensor about their residual energy level, which is one of indication of sensor failures. Zhao *et al.* (2002) proposed a technique to monitor every level using local energy level aggregations. Each individual sensor scans its energy and reports the range of residual energy to the gateway which aggregates the results and send it to the servers. However, in (Mini *et al.*, 2004) the authors presented a technique to generate the energy map in which a sink uses local information received from sensors to update energy level based on the activity performed at each sensor.

Sensors can also assist the gateway sensor to estimate the path loss ratio. Srinivasa and Haenggi (2009) presented a technique that takes empirical and analytical distribution of received mean power to estimate the path loss ratio by comparing the empirical and theoretical distribution. However, in (Mao *et al.*, 2007) Cayley-Menger determinant was proposed to determine the geometric constraints that are used to estimate the path loss ratio.

7.7 Summary

In this chapter I have presented a novel consensus-based anchor sensor management scheme to assist the WSN-based train localisation system for management of anchor sensors deployed along the track. CAMS works on the mutual cooperation and consensus based theory to detect the faulty anchor sensors, report the faults, and assist the gateway sensor in the estimation of path loss ratio. CAMS is implemented in a simulated environment using MATLAB. The simulation is based on the real data, RSS measurements, collected from field experiments in various environments such as open field, train station and a tunnel. Sensor node failure mode is executed in simulation on top of real-world RSS measurements. From the results collected from the simulation, it is observed that CAMS can effectively detect the presence of faulty sensors in the system. Our results show that, with the re-calibration of the path loss ratio of the anchor sensors, the accuracy of train localisation can be improved. Moreover, it is shown that the proposed scheme is robust in dense networks and can detect the presence of faulty sensors.

Chapter 8

Conclusions and Future Work

In this dissertation, I have presented a WSN-based train localisation system using Particle Filtering technique and incorporating RSSI measurements. Simulation results have shown that the proposed system is able to estimate the train's location with high accuracy. Moreover, WSNs provide additional benefits such as assisting train with path loss ration estimation and fault diagnostics among sensor nodes along the track. In this chapter, I summarise my contributions and discuss potential future research.

8.1 Conclusions

A train localisation system is a core component to ensure the safety and reliability in railway transportation. The commonly used technology is GPS, which heavily depends on line-of-sight with the satellites. There are several scenarios when GPS devices are unable to get clear sky for connectivity and location computation, known as GPS dark regions. In this research work, a train localisation system has been envisioned that could provide the position of a train when localisation systems based on GPS or other technologies are not feasible. This PhD project has progressed to the point of proposing and testing through simulations a WSN-based train localisation that is low in cost and conserves energy by opting the duty-cycling. The system architecture of the WSN-based train localisation system includes two type of sensor nodes: gateway sensor node and anchor sensor nodes, where a gateway sensor node has rich resources and is installed on the train, and anchor sensor nodes with their known locations are deployed along the railway track and have low-capacity in terms of energy and computation.

This thesis has described three components of a WSN-based train localisation system: Beacon-driven Wake-up Scheme (BWS), Particle-Filter-based train localisation scheme and Consensus-based Anchor sensor Management Scheme (CAMS).

The first component of the WSN-based train localisation system, BWS scheme is presented to wake up the anchor sensors when train is approaching them. As the anchor sensors are powered through batteries and operate on duty-cycles to conserve the energy that could have been wasted in unnecessary waiting in the idle listening state. The beacon-driven sensor wake-up scheme (BWS) is developed to guarantee timely communication between anchor sensors and the gateway sensor with minimum energy consumption, which is a challenging problem for WSN-based train localisation. BWS plays an essential role in minimising the energy consumption by each of the anchor sensors and, consequently, prolonging the network lifetime. BWS establishes the upper bound on the anchor sensor sleep time within one duty-cycle in order to guarantee timely wake-up. Furthermore, a theoretical analysis of the energy efficiency of BWS is presented and performance of the scheme is evaluated through extensive simulations.

In the second component of the WSN-based train localisation system, Particle-Filtering-based train localisation algorithms (PF and PF-SSR) are developed, which use the geographic coordinates from anchor sensors and the received signal strength information of the corresponding transmissions to compute the location of the train. The developed localisation schemes use the combination of RSSI-based distance estimation and particle filtering techniques. In addition, a novel Weighted RSSI Likelihood Function (WRLF) is developed for particle update, based on the special characteristics the train movement. To evaluate the performance of the presented schemes, extensive simulations are performed on the data obtained from the on-site experiments. Simulation results show that the proposed schemes can achieve significantly high localisation accuracy, such as under 10 cm. Moreover, proposed scheme is robust to the changes of train speed and the deployment density of the anchor sensors. The proposed schemes are evaluated in several conditions such as sparse deployment, unreliable anchor sensors' wake-up (40% anchor sensors fail to wake up) and high train speed up to 40 m/s. PF-SSR scheme manages to keep the average localisation error under 10 cm while evaluating each one of these factors' worst test cases independently. The PF-based train localisation scheme keeps the average location error under 30 cm in all cases.

In the third component of this thesis, the CAMS scheme is developed to assist the train localisation system for management of anchor sensors deployed along the track. CAMS works on the mutual cooperation and consensus-based theory to detect the faulty anchor nodes, report the faults, and assist the gateway node in the estimation of path loss ratio. Such additional help improves the performance of the developed system. It is shown through simulation that the proposed scheme is robust in dense networks, can detect the presence of faulty sensor nodes and calibrate the path loss ratio.

In conclusion, this thesis has shown that WSN can be used in an alternative localisation system that can operate in GPS-dark regions. This system can achieve high accuracy and duty-cycling helps to reduce energy consumption at each node. Such energy saving impact at individual nodes has cumulative effect on the whole system. Particle Filter enables accurate location estimates for trains in a wide range of difficult railway environments, through the use of a measurement model such as RSSI and geographic coordinates of anchor sensors. This has been shown to increase the ability of distance estimation through RSSI compared with typical approaches that do not perform well with RSSI in certain environments. Furthermore, the WSN-based train localisation system was able to be evaluated on several datasets, collected from field experiments. Finally, the WSN-based train localisation system is able to provide additional benefits, such as assisting gateway with path loss ratio estimation and automatic diagnostics to report faults in the system.

8.2 Future Work

The contributions of this thesis raise the following issues for future research:

- In future, I plan to implement my proposed system in the railway system to validate my methods on a real system.

- A study on the impact of using a heterogeneous sensor network can be conducted. Such study will help to explore the insights of cross platform sensor nodes' performance in WSN-based train localisation systems.

- In future, the development of train localisation systems based on technologies other than WSN such as RFID, WLAN and GPS is an important milestone. The estimated location from each of the implemented technologies can then be incorporated through the data fusion technique to raise accuracy.

- I have not addressed the security issues in the management and communication between sensors. The security issues can be addressed at the infrastructure level in future to make a WSN-based localisation system more secure.

- I plan to develop an application for a train localisation system that will display the current location of the train on a map (Heirich *et al.*, 2013).

References

Abhayawardhana, V., Wassell, I., Crosby, D., Sellars, M., and Brown, M. (2005). Comparison of empirical propagation path loss models for fixed wireless access systems. In *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, Volume 1, 73–77. IEEE.

Aboelela, E., Edberg, W., Papakonstantinou, C., and Vokkarane, V. (2006). Wireless sensor network based model for secure railway operations. In *Performance, Computing, and Communications Conference, 2006. IPCCC 2006. 25th IEEE International*, 6–pp. IEEE.

Advanticsys (2016). 802.15.4 Mote prices. http://www.advanticsys.com/shop/mtmcm4000msp-p-8.html. Accessed = 09-03-2016.

Al Nuaimi, K. and Kamel, H. (2011). A survey of indoor positioning systems and algorithms. In *Innovations in Information Technology (IIT), 2011 International Conference on*, 185–190. IEEE.

Albanese, A., Marradi, L., Labbiento, G., and Venturi, G. (2005). The RUNE project: The integrity performances of GNSS-based railway user navigation equipment. In *ASME/IEEE 2005 Joint Rail Conference*, 211–218. American Society of Mechanical Engineers.

Amitabh, A. (2005). Rail Accidents due to Human Errors-Indian Railways Experience. In *International Railway Safety Conference, Cape Town, South Africa.*

Andersen, A. (2008). Small size 2.4 GHz PCB antenna. *Texas Instruments, Application Note AN043*.

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, *49*(268), 765–769.

Anscombe, F. J. (1960). Rejection of outliers. *Technometrics*, *2*(2), 123–146.

Arabaci, M. and Strickland, R. N. (2007). Direction of arrival estimation in reverberant rooms using a resource-constrained wireless sensor network. In *Pervasive services, IEEE international conference on*, 29–38. IEEE.

AtlasRFID (2016a). RFID active tags life. http://blog.atlasrfidstore.com/active-rfid-vs-passive-rfid. Accessed = 09-03-2016.

AtlasRFID (2016b). RFID Reader and tags prices. http://www.atlasrfidstore.com/active-rfid/. Accessed = 09-03-2016.

Avots, D., Lim, E., Thibaux, R., and Thrun, S. (2002). A probabilistic technique for simultaneous localization and door state estimation with mobile robots in dynamic environments. In *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, Volume 1, 521–526. IEEE.

Bahl, P. and Padmanabhan, V. N. (2000). RADAR: An in-building RF-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, Volume 2, 775–784. Ieee.

Baker, J. E. (1987). Reducing bias and inefficiency in the selection algorithm. In *Proceedings of the second international conference on genetic algorithms*, 14–21.

Batteries (2016a). AA Batteries. http://www.bestbuy.ca/en-CA/product/dynex-aa-batteries-48-pack/10058015.aspx?path=fc65413a7bbb4520da22841990af80d8en02. Accessed = 09-03-2016.

Batteries, A. (2016b). AA Batteries. http://www.cheap-batteries.com/AA.html . Accessed = 09-03-2016.

Bendigeri, K. Y. and Mallapur, J. D. (2015). Multiple Node Placement Strategy for Efficient Routing in Wireless Sensor Networks. *Wireless Sensor Network*, *7*(8), 101.

Bischoff, R., Meyer, J., Enochsson, O., Feltrin, G., and Elfgren, L. (2009). Event-based strain monitoring on a railway bridge with a wireless sensor network. In *Proceedings of the 4th International Conference on Structural Health Monitoring of Intelligent Infrastructure, Zurich, Switzerland*, Volume 2224, 7482.

Bradford, P. W., Spilker, J., and Enge, P. (1996). Global positioning system: theory and applications. *AIAA Washington DC*, *109*.

Brain, M. and Harris, T. (2011). How GPS receivers work.

Buettner, M., Yee, G. V., Anderson, E., and Han, R. (2006). X-MAC: a short preamble MAC protocol for duty-cycled wireless sensor networks. In *Proceedings of the 4th international conference on Embedded networked sensor systems*, 307–320. ACM.

Burns, R. D., Elliott-Bryan, S., and Turner, D. B. (1992). Rail vehicle positioning system. US Patent 5,129,605.

Cavalieri, S. (2007). WLAN-based outdoor localisation using pattern matching algorithm. *International Journal of Wireless Information Networks*, *14*(4), 265–279.

Cerniglia, D., Garcia, G., Kalay, S., and Prior, F. (2006). Application of laser induced ultrasound for rail inspection. In *Proceedings of the World Congress in Railway Research. Montreal, Canada*.

Chang, D. T. T., Tsai, Y. S., Lee, H. C., Guo, L. L., and Yang, K. C. (2011). Wireless sensor network (WSN) using in tunnel environmental monitoring for disaster prevention. In *Advanced Materials Research*, Volume 291, 3401–3404. Trans Tech Publ.

Char, R. and Johns, H. (2006). Improving Transportation Safety & Security Using RFID. In *Research Opportunities in Radio Frequency Identification (RFID) Transportation Applications Conference*.

Chen, B., Jamieson, K., Balakrishnan, H., and Morris, R. (2001). Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. In *Proc. of ACM/IEEE 7th International Conf. on Mobile Computing and Networking (MobiCom)*. ACM.

Chen, Y., Crespi, N., Lv, L., Li, M., Ortiz, A. M., and Shu, L. (2013). Locating using prior information: wireless indoor localization algorithm. In *ACM SIGCOMM Computer Communication Review*, Volume 43, 463–464. ACM.

Chen, Z., Xia, F., Huang, T., Bu, F., and Wang, H. (2013). A localization method for the internet of things. *The Journal of Supercomputing*, *63*(3), 657–674.

Chintalapudi, K., Padmanabha Iyer, A., and Padmanabhan, V. N. (2010). Indoor localization without the pain. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, 173–184. ACM.

Cho, H., Kang, M., Park, J., Park, B., and Kim, H. (2007). Performance analysis of location estimation algorithm in ZigBee networks using received signal strength. In *Advanced Information Networking and Applications Workshops, 2007, AINAW'07. 21st International Conference on*, Volume 2, 302–306. IEEE.

Chu, H.-C. and Jan, R.-H. (2007). A GPS-less, outdoor, self-positioning method for wireless sensor networks. *Ad Hoc Networks*, *5*(5), 547–557.

Coetzee, L. and Botha, E. C. (1993). Fingerprint recognition in low quality images. *Pattern Recognition*, *26*(10), 1441–1460.

Constandache, I., Gaonkar, S., Sayler, M., Choudhury, R. R., and Cox, L. (2009). Enloc: Energy-efficient localization for mobile phones. In *INFOCOM 2009, IEEE*, 2716–2720. IEEE.

CPEC (2016). CPEC railway project. https://cpecnews.wordpress.com/2016/01/15/3/. Accessed = 20-01-2016.

dailytimes (2016). Damaged railway tracks. http://www.dailytimes.com.pk/national/28-May-2015/bomb-blasts-damage-railway-tracks. Accessed = 16-03-2016.

Datasheet, M. (2006). Crossbow technology inc. https://www.eol.ucar.edu/isf/facilities/isa/internal/CrossBow/DataSheets/mica2.pdf. *San Jose, California*.

Däubler, L., Becker, U., and Schnieder, E. (2003). Availability Analysis of Satellite-Based Train Localisation Systems. *WCRR Proceedings, S*, *224*.

Descamps, P., El Bekkali, M., and Vindevoghel, J. (1991). An integrated microwave Doppler cinemometer for transport applications. *Review of scientific instruments*, *62*(8), 2016–2018.

Drira, A. (2006). GPS navigation for outdoor and indoor environments. *University of Tennessee, Knoxville*.

Dron, W., Duquennoy, S., Voigt, T., Hachicha, K., and Garda, P. (2014). An emulation-based method for lifetime estimation of wireless sensor networks. In *2014 IEEE International Conference on Distributed Computing in Sensor Systems*, 241–248. IEEE.

Dunkels, A. (2011). The contiki mac radio duty cycling protocol. *Technical Report, T2011:13, Swedish Institute of Computer Science*.

Durrant-Whyte, H., Majumder, S., Thrun, S., De Battista, M., and Scheding, S. (2003). A bayesian algorithm for simultaneous localisation and map building. In *Robotics Research*, 49–60. Springer.

Ekahau (2008). Ekahau Systems. http://www.ekahau.com/. Accessed = 11-02-2016.

El-Hoiydi, A. and Decotignie, J.-D. (2004). WiseMAC: An ultra low power MAC protocol for multi-hop wireless sensor networks. *Algorithmic Aspects of Wireless Sensor Networks*, 18–31.

Erdogan, A., Coskun, V., and Kavak, A. (2006). The sectoral sweeper scheme for wireless sensor networks: Adaptive antenna array based sensor node management and location estimation. *Wireless Personal Communications*, *39*(4), 415–433.

EuroStat (2016). Railway passenger transport statistics. http://ec.europa.eu/eurostat/statistics-explained/index.php/Railway_passenger_transport_statisti _quarterly_and_annual_data. Accessed = 16-03-2016.

Fales, D. (2003). ALONE AGAIN. *Motor Boating*, *191*(1), 30.

Fan, Y., Dixon, S., Edwards, R. S., and Jian, X. (2007). Ultrasonic surface wave propagation and interaction with surface defects on rail track head. *Ndt & E International*, *40*(6), 471–477.

Fararooy, S., Allan, J., Jin, J., Maleki, K., and Pang, F. (1996). Accurate train localisation in open space and tunnels.

Feng, C., Au, W. S. A., Valaee, S., and Tan, Z. (2010). Compressive sensing based positioning using RSS of WLAN access points. In *INFOCOM, 2010 Proceedings IEEE*, 1–9. IEEE.

Feng, C., Au, W. S. A., Valaee, S., and Tan, Z. (2012). Received-signal-strength-based indoor positioning using compressive sensing. *Mobile Computing, IEEE Transactions on*, *11*(12), 1983–1993.

Feng, W.-J., Bi, X.-W., and Jiang, R. (2012). A novel adaptive cooperative location algorithm for wireless sensor networks. *International Journal of Automation and Computing*, *9*(5), 539–544.

Flammini, F., Gaglione, A., Ottello, F., Pappalardo, A., Pragliola, C., and Tedesco, A. (2010). Towards wireless sensor networks for railway infrastructure monitoring. In *Electrical Systems for Aircraft, Railway and Ship Propulsion (ESARS), 2010*, 1–6. IEEE.

Fox, D., Thrun, S., Burgard, W., and Dellaert, F. (2001). Particle filters for mobile robot localization. In *Sequential Monte Carlo methods in practice*, 401–428. Springer.

Foxlin, E. (2005). Pedestrian tracking with shoe-mounted inertial sensors. *Computer Graphics and Applications, IEEE*, *25*(6), 38–46.

Fraile, J. M. (1999). Development of a GPS-based train localisation system for low traffic rail lines. In *Proceedings of the European Symposium on Transport Telematics, Potsdam, Germany.*

GAO-Inc. (2007). GAO RFID Solutions for Railway Industry. railway.gaorfid.com. Accessed = 01-02-2015.

Garcia Crespillo, O., Heirich, O., and Lehner, A. (2014). Bayesian GNSS/IMU tight integration for precise railway navigation on track map. In *Position, Location and Navigation Symposium-PLANS 2014, 2014 IEEE/ION*, 999–1007. IEEE.

Ghahramani, Z. (2001). An introduction to hidden Markov models and Bayesian networks. *International Journal of Pattern Recognition and Artificial Intelligence*, *15*(01), 9–42.

Giaglis, G. M., Kourouthanassis, P., and Tsamakos, A. (2003). Towards a classification framework for mobile location services. *Mobile commerce: technology, theory and applications*, 67–85.

Giorgetti, G., Cidronali, A., Gupta, S. K., and Manes, G. (2007). Exploiting low-cost directional antennas in 2.4 GHz IEEE 802.15. 4 wireless sensor networks. In *Wireless Technologies, 2007 European Conference on*, 217–220. IEEE.

Godha, S., Lachapelle, G., and Cannon, M. E. (2006). Integrated GPS/INS system for pedestrian navigation in a signal degraded environment. In *ION GNSS*, Volume 2006.

Gogolak, L., Pletl, S., and Kukolj, D. (2011). Indoor fingerprint localization in WSN environment based on neural network. In *Intelligent Systems and Informatics (SISY), 2011 IEEE 9th International Symposium on*, 293–296. IEEE.

GPSHistory (2016). Britain relies on rail. http://www.boatsafe.com/kids/navigation.htm. Accessed = 16-03-2016.

Grejner-Brzezinska, D., Hong, C.-K., Wielgosz, P., and Hothem, L. (2007). The impact of severe ionospheric conditions on the GPS hardware in the Southern Polar Region. In *Dynamic Planet*, 867–875. Springer.

Grewal, M. S. (2011). *Kalman filtering.* Springer.

GRIFFIN, G., KESSLER, R., ONDER, M., CHAR, R., and DONATH, M. (2006). Future Directions in RFID Applications and Research in Transportation. *Research Opportunities in Radio Frequency Identification Transportation Applications*, 32.

Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. *Technometrics, Taylor & Francis*, *11*(1), 1–21.

Gu, Y., Lo, A., and Niemegeers, I. (2009). A survey of indoor positioning systems for wireless personal networks. *Communications Surveys & Tutorials, IEEE*, *11*(1), 13–32.

Gustafsson, F., Gunnarsson, F., Bergman, N., Forssell, U., Jansson, J., Karlsson, R., and Nordlund, P. (2002). Particle filters for positioning, navigation, and tracking. In *IEEE Transactions on Signal Processing*, 425–437.

Güvenc, İ. (2003). *Enhancements to RSS based indoor tracking systems using Kalman filters.* Ph. D. thesis, University of New Mexico.

Hansen, P. A. (2001). Positive train control. *Trains*, *61*(1).

Harter, A., Hopper, A., Steggles, P., Ward, A., and Webster, P. (2002). The anatomy of a context-aware application. *Wireless Networks*, *8*(2/3), 187–197.

Heirich, O. (2016). Bayesian Train Localization with Particle Filter, Loosely Coupled GNSS, IMU, and a Track Map. *Journal of Sensors*, *2016*.

Heirich, O., Robertson, P., and Strang, T. (2013). RailSLAM-Localization of rail vehicles and mapping of geometric railway tracks. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, 5212–5219. IEEE.

Hill, R. (1981). Fail-safe speed sensor for London Transport automatic trains. In *IEE Proceedings B (Electric Power Applications)*, Volume 128, 277–284. IET.

Hill, R. J. and Weedon, D. N. (1990). Safety and reliability of synchronizable digital coding in railway track-circuits. *Reliability, IEEE Transactions on*, *39*(5), 581–591.

Hofestadt, H. (1995). GSM-R: global system for mobile radio communications for railways. In *Electric Railways in a United Europe, 1995., International Conference on*, 111–115. IET.

Howard, A. (2006). Multi-robot simultaneous localization and mapping using particle filters. *The International Journal of Robotics Research*, *25*(12), 1243–1256.

Hu, L. X., Wang, X. X., and Bian, X. B. (2012). A Research of the Positioning System for Rail Transit Based on Wireless Sensor Networks. In *Applied Mechanics and Materials*, Volume 229, 1276–1279. Trans Tech Publ.

Huang, X., Janaswamy, R., and Ganz, A. (2006). Scout: Outdoor localization using active RFID technology. In *Broadband Communications, Networks and Systems, 2006. BROADNETS 2006. 3rd International Conference on*, 1–10. IEEE.

Huang, Y. and Benesty, J. (2004). *Audio Signal Processing for Next-Generation Multimedia Communication Systems: For Next-Generation Multimedia Communication Systems*. Springer Science & Business Media.

Huffman, J. P., LaForest, J. P., Petit, W. A., and Smith, J. A. (1977). Alternating current track circuits. US Patent 4,065,081.

IEM-RM, S. E.-M. (2003). B-scan ultrasonic image analysis for internal rail defect detection.

Ikeda, M. (1993). Characteristic of position detection and method of position correction by rotating axle. *Railway Technical Research Institute, Quarterly Reports*, *34*(4).

Interactive, C. (2010). Firefly motion capture system.

J, N. D. and Faulkner, F. (1991). A Doppler Speed Sensor for Land Vehicles. In *Royal Institute of Navigation, International Conference on*.

Javed, A., Huang, Z., Zhang, H., and Deng, J. (2015). CAMS: Consensus-based Anchor node Management Scheme for train localization. In *Adhoc Networks and Wireless (ADHOC-NOW), 2015 International Conference on*. Springer.

Javed, A., Zhang, H., and Huang, Z. (2013). Performance Analysis of Duty-Cycling Wireless Sensor Network for Train Localization. In *Proceedings of Workshop on Machine Learning for Sensory Data Analysis*, 43. ACM.

Javed, A., Zhang, H., Huang, Z., and Deng, J. (2014). BWS: Beacon-driven wake-up scheme for train localization using wireless sensor networks. In *Communications (ICC), 2014 IEEE International Conference on*, 276–281.

Jimenez, A. R., Seco, F., Prieto, C., and Guevara, J. (2009). A comparison of pedestrian dead-reckoning algorithms using a low-cost MEMS IMU. In *Intelligent Signal Processing, 2009. WISP 2009. IEEE International Symposium on*, 37–42. IEEE.

Johnson, M. (2016). How track circuits detect and protect trains. http://greatergreaterwashington.org/post/5068/how-track-circuits-detect-and-protect-trains/. Accessed = 16-03-2016.

Jonsrud, G. (2008). Folded dipole antenna for CC2400, CC2420, CC2430, CC2431, and CC2480. *Texas Instruments, Application Note AN040*.

Kaemarungsi, K. and Krishnamurthy, P. (2004). Modeling of indoor positioning systems based on location fingerprinting. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, Volume 2, 1012–1022. IEEE.

Kaligineedi, P., Khabbazian, M., and Bhargava, V. K. (2008). Secure cooperative sensing techniques for cognitive radio systems. In *Communications, 2008. ICC'08. IEEE International Conference on*, 3406–3410. IEEE.

Kaligineedi, P., Khabbazian, M., and Bhargava, V. K. (2010). Malicious user detection in a cognitive radio cooperative sensing system. *Wireless Communications, IEEE Transactions on*, *9*(8), 2488–2497.

Kalis, A. and Dimitriou, T. (2005). Fast routing in wireless sensor networks using directional transmissions. *International Journal of Mobile Network Design and Innovation*, *1*(1), 63–69.

Khan, N. Y. (2014). *Self localisation in indoor environments using machine vision*. Ph. D. thesis, University of Otago.

Kim, J., On, J., Kim, S., and Lee, J. (2008). Performance evaluation of synchronous and asynchronous MAC protocols for wireless sensor networks. In *Sensor Technologies and Applications, 2008. SENSORCOMM'08. Second International Conference on*, 500–506. IEEE.

King, T., Kopf, S., Haenselmann, T., Lubberger, C., and Effelsberg, W. (2006). COMPASS: A probabilistic indoor positioning system based on 802.11 and digital compasses. In *Proceedings of the 1st international workshop on Wireless network testbeds, experimental evaluation & characterization*, 34–40. ACM.

Kishor, R. S., Ippar Umesh, L., Jethwani Amit, B., and Dusane, M. A. V. (2012). RAIL CRACK DETECTION SYSTEM USING IR MODULE.

Klepal, M., Pesch, D., *et al.* (2007a). A bayesian approach for RF-based indoor localisation. In *Wireless Communication Systems, 2007. ISWCS 2007. 4th International Symposium on*, 133–137. IEEE.

Klepal, M., Pesch, D., *et al.* (2007b). RF-Based Location System in Harsh Environment. In *Applied Wearable Computing (IFAWC), 2007 4th International Forum on*, 1–11. VDE.

Koch, J., Hillenbrand, C., and Berns, K. (2005). Inertial navigation for wheeled robots in outdoor terrain. In *5th IEEE Workshop on Robot Motion and Control (RoMoCo), Dymaczewo, Poland*.

Köpke, A., Swigulski, M., Wessel, K., Willkomm, D., Haneveld, P., Parker, T. E., Visser, O. W., Lichte, H. S., and Valentin, S. (2008). Simulating wireless and mobile networks in OMNeT++ the MiXiM vision. In *Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops*, 71. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Korve Engineering, I. (2007). *Audible Signals for Pedestrian Safety in LRT Environments*. Transportation Research Board.

Kuun, E. and Richard, W. (2004). Open Standards for CCTC and CCTV Radio-based Communication. *Alcatel telecommunications review*, (2), 243–252.

Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli, I., Fateh, M., Viola, E., and Pascale, G. (2005). Non-contact ultrasonic inspection of rails and signal processing for automatic defect detection and classification. *Insight-Non-Destructive Testing and Condition Monitoring*, *47*(6), 346–353.

Lardennois, R. (2003). Wireless communication for signaling in mass transit. *Siemens Transportation Systems*.

Leahy, F., Judd, M., and Shortis, M. (1993). Measurement of railway profiles using GPS integrated with other sensors. In *Vehicle Navigation and Information Systems Conference, 1993., Proceedings of the IEEE-IEE*, 706–709. IEEE.

Lee, S., Kim, E., Kim, C., and Kim, K. (2009). Localization with a mobile beacon based on geometric constraints in wireless sensor networks. *Wireless Communications, IEEE Transactions on*, *8*(12), 5801–5805.

Lemelson, J. H. and Pedersen, R. D. (1999). GPS vehicle collision avoidance warning and control system and method. US Patent 5,983,161.

Li, X. (2006). RSS-based location estimation with unknown pathloss model. *Wireless Communications, IEEE Transactions on*, *5*(12), 3626–3633.

Lin, T.-H., Chang, K.-h., Tian, J.-b., Chu, H.-h., and Huang, P. (2008). Modeling and simulation comparison of two time synchronization protocols. In *Proceedings of the 3nd ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks*, 117–123. ACM.

Liu, B.-C., Lin, K.-H., and Wu, J.-C. (2006). Analysis of hyperbolic and circular positioning algorithms using stationary signal-strength-difference measurements in wireless communications. *Vehicular Technology, IEEE Transactions on*, *55*(2), 499–509.

Liu, H., Darabi, H., Banerjee, P., and Liu, J. (2007). Survey of wireless indoor positioning techniques and systems. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, *37*(6), 1067–1080.

Liu, W., Zhao, D., and Zhu, G. (2014). End-to-end delay and packet drop rate performance for a wireless sensor network with a cluster-tree topology. *Wireless Communications and Mobile Computing*, *14*(7), 729–744.

Majumder, S., Durrant-Whyte, H., Thrun, S., and de Battista, M. (2000). An approximate Bayesian method for simultaneous localisation and mapping. *IEEE Trans. Automatic Control*, *45*(3), 477–482.

Mandal, A., Lopes, C. V., Givargis, T., Haghighat, A., Jurdak, R., and Baldi, P. (2005). Beep: 3D indoor positioning using audible sound. In *Consumer Communications and Networking Conference, 2005. CCNC. 2005 Second IEEE*, 348–353. IEEE.

Mao, G., Anderson, B., and Fidan, B. (2007). Path loss exponent estimation for wireless sensor network localization. *Computer Networks*, *51*(10), 2467–2483.

Mao, G., Fidan, B., and Anderson, B. D. (2007). Wireless sensor network localization techniques. *Computer networks*, *51*(10), 2529–2553.

Marti, S., Giuli, T. J., Lai, K., and Baker, M. (2000). Mitigating routing misbehavior in mobile ad hoc networks. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, 255–265. ACM.

Mayhew, G. L., Kivett, J., Himes, J. G., and Evans, J. A. (1994). Application, of radio navigation technology to advanced automatic train control. In *Position Location and Navigation Symposium, 1994., IEEE*, 217–224. IEEE.

Mazl, R. and Přeučil, L. (2003). Sensor data fusion for inertial navigation of trains in GPS-dark areas. In *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*, 345–350. IEEE.

Midya, S. and Thottappillil, R. (2008). An overview of electromagnetic compatibility challenges in European Rail Traffic Management System. *Transportation Research Part C: Emerging Technologies*, *16*(5), 515–534.

Mini, R. A., Loureiro, A. A., and Nath, B. (2004). The distinctive design characteristic of a wireless sensor network: the energy map. *Computer Communications*, *27*(10), 935–945.

Mishra, S. M., Sahai, A., and Brodersen, R. W. (2006). Cooperative sensing among cognitive radios. In *Communications, 2006. ICC'06. IEEE International Conference on*, Volume 4, 1658–1663. IEEE.

Misra, P., Jha, S., and Ostry, D. (2011). Improving the coverage range of ultrasound-based localization systems. In *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*, 605–610. IEEE.

Monte de Caparica, P. N. P. (2000). Track to Train Communications System.

Montemerlo, D., Thrun, S., and Whittaker, W. (2002). Conditional particle filters for simultaneous mobile robot localization and people-tracking. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, Volume 1, 695–701. IEEE.

Montemerlo, M., Thrun, S., Koller, D., and Wegbreit, B. (2003). FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges. In *International Joint Conference on Artificial Intelligence*, Volume 18, 1151–1156. Lawrence Erlbaum Associates LtD.

Montemerlo, M., Thrun, S., Koller, D., Wegbreit, B., *et al.* (2002). FastSLAM: A factored solution to the simultaneous localization and mapping problem. In *AAAI/IAAI*, 593–598.

Morales, Y. and Tsubouchi, T. (2007). DGPS, RTK-GPS and StarFire DGPS performance under tree shading environments. In *Integration Technology, 2007. ICIT'07. IEEE International Conference on*, 519–524. IEEE.

Morávek, P., Komosny, D., Sala, D. G., and Guillén, A. L. (2010). Received signal strength uncertainty in energy-aware localization in wireless sensor networks. In *Environment and Electrical Engineering (EEEIC), 2010 9th International Conference on*, 538–541. IEEE.

MTM (2012). *MTM Specifications.* MAXFOR Technology Inc. www.maxfor.co.kr/datasheet/MAXFOR_Digital_Brochure.pdf (accessed March 27, 2012).

Nagel, H. C. (1979). Power transfer apparatus for railroad track circuits. US Patent 4,171,788.

Navaraja, P. (2014). CRACK DETECTION SYSTEM FOR RAILWAY TRACK BY USING ULTRASONIC AND PIR SENSOR. , *5*.

NetworkRail (2016). Britain relies on rail. http://www.networkrail.co.uk/aspx/662.aspx. Accessed = 16-03-2016.

Ni, L. M., Liu, Y., Lau, Y. C., and Patil, A. P. (2004). LANDMARC: indoor location sensing using active RFID. *Wireless networks*, *10*(6), 701–710.

Oh, S., Park, S., and Lee, C. (2007). Vision based platform monitoring system for railway station safety. In *Telecommunications, 2007. ITST'07. 7th International Conference on ITS*, 1–5. IEEE.

OmegaUltraSound (2016). Ultrasound Controller prices. http://www.omega.com/pptst/LVCN210.html. Accessed = 09-03-2016.

Oukhellou, L., Come, E., Bouillaut, L., and Aknin, P. (2008). Combined use of sensor data and structural knowledge processed by Bayesian network: Application to a railway diagnosis aid scheme. *Transportation Research Part C: Emerging Technologies*, *16*(6), 755–767.

Pace, S., Frost, G., Lachow, I., Frelinger, D., and Fossum, D. (1995). The global positioning system: assessing national policies. Technical report, DTIC Document.

Pahl, P. and Radar, P. V. (2000). An In-Building RF-based User Location and Tracking System [C]. *IEEE Communications Socienties*, *2*, 775–784.

Patwari, N. and Hero III, A. O. (2003). Using proximity and quantized RSS for sensor localization in wireless networks. In *Proceedings of the 2nd ACM international conference on Wireless sensor networks and applications*, 20–29. ACM.

Peng, R. and Sichitiu, M. L. (2006). Angle of arrival localization for wireless sensor networks. In *Sensor and Ad Hoc Communications and Networks, 2006. SECON'06. 2006 3rd Annual IEEE Communications Society on*, Volume 1, 374–382. IEEE.

Pister, K. and Doherty, L. (2008). TSMP: Time synchronized mesh protocol. *IASTED Distributed Sensor Networks*, 391–398.

Polastre, J., Hill, J., and Culler, D. (2004). Versatile low power media access for wireless sensor networks. In *Proceedings of the 2nd international conference on Embedded networked sensor systems*, 95–107. ACM.

Priyantha, N. B., Chakraborty, A., and Balakrishnan, H. (2000). The CRICKET location-support system. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, 32–43. ACM.

Punetha, D., Tripathi, D. M., and Kumar, A. (2014). A Wireless Approach with Sensor Network for Real Time Railway Track Surveillance System. *International Journal of Engineering Trends and Technology (IJETT)*, *9*(9), 426–429.

Rados, J., Rados, B., and Kolar, V. (2007). European train control system. *Annals of DAAAM & Proceedings*, 623–625.

Railway-Technology (2016). Driverless train technology and the London Underground: the great debate. http://www.railway-technology.com/features/featuredriverless-train-technology/. Accessed = 16-03-2016.

Ramesh, S. and Gobinathan, S. (2012). Railway Faults Tolerance Techniques using Wireless Sensor Networks 1.

Receiver, U. (2016). Ultrasound Receiver prices. http://electricaltoolsdeal.com/products/ultrasonic-sensor-receiver. Accessed = 09-03-2016.

Ren, H. and Meng, M.-H. (2009). Power adaptive localization algorithm for wireless sensor networks using particle filter. *Vehicular Technology, IEEE Transactions on*, *58*(5), 2498–2508.

Robertson, P., Angermann, M., and Krach, B. (2009). Simultaneous localization and mapping for pedestrians using only foot-mounted inertial sensors. In *Proceedings of the 11th international conference on Ubiquitous computing*, 93–96. ACM.

Roller, D., Montemerlo, M., Thrun, S., and Wegbreit, B. (2003). Fastslam 2.0: an improved Particle Filtering algorithm for simultaneous localization and mapping that provably converges. In *Proceedings of the International Joint Conference on Artificial Intelligence*.

Royo, F., Olivares, T., and Orozco-Barbosa, L. (2009). A synchronous engine for wireless sensor networks. *Telecommunication Systems*, *40*(3-4), 151–159.

Rushant, K. and Spacek, L. (1997). An autonomous vehicle navigation system using panoramic machine vision techniques. Technical report, Technical report CSM-298.

Santos, A. J. D., Soares, A. R., Redondo, F. M. D. A., and Carvalho, N. B. (2005). Tracking trains via radio frequency systems. *Intelligent Transportation Systems, IEEE Transactions on*, *6*(2), 244–258.

Savage, P. G. (2013). Blazing gyros: The evolution of strapdown inertial navigation technology for aircraft. *Journal of Guidance, Control, and Dynamics*, *36*(3), 637–655.

Savic, V., Wymeersch, H., and Larsson, E. G. (2013). Simultaneous sensor localization and target tracking in mine tunnels. In *Information Fusion (FUSION), 2013 16th International Conference on*, 1427–1433. IEEE.

Saxena, M., Gupta, P., and Jain, B. N. (2008). Experimental analysis of RSSI-based location estimation in wireless sensor networks. In *Communication Systems Software and Middleware and Workshops, 2008. COMSWARE 2008. 3rd International Conference on*, 503–510. IEEE.

Scholten, H., Westenberg, R., and Schoemaker, M. (2009). Trainspotting, a WSN-based train integrity system. In *Networks, 2009. ICN'09. Eighth International Conference on*, 226–231. IEEE.

Schön, T., Gustafsson, F., and Nordlund, P.-J. (2005). Marginalized particle filters for mixed linear/nonlinear state-space models. *Signal Processing, IEEE Transactions on*, *53*(7), 2279–2289.

Schulz, D., Burgard, W., Fox, D., and Cremers, A. B. (2001). Tracking multiple moving targets with a mobile robot using particle filters and statistical data association. In *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, Volume 2, 1665–1670. IEEE.

Se, S., Lowe, D., and Little, J. (2001). Vision-based mobile robot localisation and mapping using scale-invariant features. In *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, Volume 2, 2051–2058. IEEE.

Seitz, P., Gale, M. T., Meier, H., Raynor, J. M., Wolff, P., and Hecht, M. (1990). Locometer: on-line inspection of locomotive wheel-to-rail movements using high-precision CCD metrology. In *The Hague'90, 12-16 April*, 39–45. International Society for Optics and Photonics.

Shen, X., Wang, Z., Jiang, P., Lin, R., and Sun, Y. (2005). Connectivity and RSSI based localization scheme for wireless sensor networks. *Advances in Intelligent Computing*, 578–587.

Skybrary (2009). Skybrary Repository. http://www.skybrary.aero/index.php/. Accessed = 11-02-2016.

Smeets, H., Shih, C.-Y., Zuniga, M., Hagemeier, T., and Marrón, P. J. (2013). TrainSense: a novel infrastructure to support mobility in wireless sensor networks. In *Wireless Sensor Networks*, 18–33. Springer.

Sommer, C., Joerer, S., and Dressler, F. (2012). On the applicability of Two-Ray path loss models for vehicular network simulation. *VNC*, *12*, 64–69.

Song, C. and Zhang, Q. (2008). Sliding-window algorithm for asynchronous cooperative sensing in wireless cognitive networks. In *Communications, 2008. ICC'08. IEEE International Conference on*, 3432–3436. IEEE.

Song, J., An, T., and Lee, D. (2012). Vision-Based Train Position and Movement Estimation Using a Fuzzy Classifier. *Digital Convergence Studies*, *10*(1), 365–369.

Song, Z., Jiang, G., and Huang, C. (2011). A survey on indoor positioning technologies. In *Theoretical and Mathematical Foundations of Computer Science*, 198–206. Springer.

Sonotronics (2016). Ultrasound receivers battery life. http://www.sonotronics.com. Accessed = 09-03-2016.

Srinivasa, S. and Haenggi, M. (2009). Path loss exponent estimation in large wireless networks. In *Information Theory and Applications Workshop, 2009*, 124–129. IEEE.

Srinivasan, A., Teitelbaum, J., and Wu, J. (2006). DRBTS: distributed reputation-based beacon trust system. In *Dependable, Autonomic and Secure Computing, 2nd IEEE International Symposium on*, 277–283. IEEE.

States, R. and Pappas, E. (2006). Precision and repeatability of the Optotrak 3020 motion measurement system. *Journal of medical engineering & technology*, *30*(1), 11–16.

Staton, G. (2005). INTEGRAIL Project Overview. *EGNOS/INTEGRAIL Demonstration, Guateng.*

Stewart, M. and Rizos, C. (2002). *GPS projects: some planning issues.* Taylor & Francis, London and New York.

Stoleru, R., He, T., Stankovic, J. A., and Luebke, D. (2005). A high-accuracy, low-cost localization system for wireless sensor networks. In *Proceedings of the 3rd international conference on Embedded networked sensor systems*, 13–26. ACM.

Stoyanova, T., Kerasiotis, F., Prayati, A., and Papadopoulos, G. (2009). A practical RF propagation model for wireless network sensors. In *Sensor Technologies and Applications, 2009. SENSORCOMM'09. Third International Conference on*, 194–199. IEEE.

Sun, Y., Du, S., Gurewitz, O., and Johnson, D. B. (2008). DW-MAC: a low latency, energy efficient demand-wakeup MAC protocol for wireless sensor networks. In *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, 53–62. ACM.

Sun, Y., Gurewitz, O., and Johnson, D. B. (2008). RI-MAC: a receiver-initiated asynchronous duty cycle MAC protocol for dynamic traffic loads in wireless sensor networks. In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, 1–14. ACM.

Tarrio, P., Bernardos, A. M., Casar, J. R., *et al.* (2008). A new positioning technique for RSS-Based localization based on a weighted least squares estimator. In *Wireless Communication Systems. 2008. ISWCS'08. IEEE International Symposium on*, 633–637. IEEE.

Texas Instruments (2003). *2.4 GHz IEEE 802.15.4/ZigBee-ready RF Transceiver*. Texas Instruments.

theGuardian (2016). Number of UK train journeys has doubled. http://www.theguardian.com/uk-news/2015/sep/14/train-journey-numbers-double-since-privatisation-railways-uk-report. Accessed = 16-03-2016.

TheRegister (2014). This 125mph train is fitted with LASERS. http://www.theregister.co.uk/2014/11/10/network_rail_new_measurement_train_ride/. Accessed = 01-09-2016.

Thevenot, V., Bruckmueller, T., Doederlein, C., Mattos, P., Sarfati, R., Lechner, W., and Tossaint, M. (2003). ECORAIL: A step towards safe railway Controlling systems based on satellite positioning. In *The European Navigation Conference*.

Thrun, S., Liu, Y., Koller, D., Ng, A. Y., Ghahramani, Z., and Durrant-Whyte, H. (2004). Simultaneous localization and mapping with sparse extended information filters. *The International Journal of Robotics Research*, *23*(7-8), 693–716.

Tian, Y., Chen, X., Huang, Q., Lv, P., Li, R., and Li, M. (2011). 3D Localization of Moving Object by High-Speed Four-Camera Vision System. In *Future Intelligent Information Systems*, 425–434. Springer.

uic (2016). Automatic train control. http://www.railway-energy.org/static/Automatic_train_control_79.php. Accessed = 16-03-2016.

Umiliacchi, P., Shingler, R., Langer, G., and Henning, U. (2006). A new approach to optimisation through intelligent integration of railway systems: the InteGRail project. In *Proceedings of the 7th WCRR, World Conference on Railway Research*.

Urech, A., Garcia Prieto, M., and Roberts, W. (2006). GRAIL, GNSS Introduction in the rail sector. In *Proceedings of ENC*.

Urech, A., Pérez Diestro, J., and Gonzalez, O. (2002). GADEROS, a GAlileo Demonstrator for Railway Operation System. In *Data Systems in Aerospace*, Volume 509, 46.

Van Dam, T. and Langendoen, K. (2003). An adaptive energy-efficient MAC protocol for wireless sensor networks. In *Proceedings of the 1st international conference on Embedded networked sensor systems*, 171–180. ACM.

Wang, J.-H. and Du, S.-S. (2010). Local area localization algorithm for underground tunnel based on WSN. *Meikuang Jixie(Coal Mine Machinery)*, *31*(12), 189–191.

Wang, W. and Zhu, Q. (2008). RSS-based Monte Carlo localisation for mobile sensor networks. *Communications, IET*, *2*(5), 673–681.

Want, R., Hopper, A., Falcao, V., and Gibbons, J. (1992). The active badge location system. *ACM Transactions on Information Systems (TOIS)*, *10*(1), 91–102.

WhereNet (2008). WhereNet Solutions. https://www.zebra.com/us/en/products/location-solutions/wherenet.html. Accessed = 11-04-2015.

Wiita, F. L. (1989). Railway anticollision apparatus and method. US Patent 4,864,306.

Wikil (2016). Track circuit failure. https://en.wikipedia.org/wiki/Track_circuit#Caused_by_track_circ Accessed = 16-03-2016.

Wong, Y., Ngoh, L., Wong, W., and Seah, W. (2007). A Combinatorics-Based Wakeup Scheme for Target Tracking in Wireless Sensor Networks. In *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE*, 3569–3574. IEEE.

Wu, M. and Pei, F. (2013). Improved Distributed Particle Filter for Simultaneous Localization and Mapping. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, *11*(12), 7617–7626.

Xu, J., Liu, W., Lang, F., Zhang, Y., and Wang, C. (2010). Distance measurement model based on RSSI in WSN. In *Wireless Sensor Network*, Volume 2, 606–611. Scientific Research Publishing.

Xu, K., Liu, D. W., Liu, H. L., and Gu, Z. R. (2013). Experimental Analysis of RSSI-Based Underground Tunnel Localization with Wireless Sensor Networks. In *Applied Mechanics and Materials*, Volume 385, 1756–1759. Trans Tech Publ.

Yang, J. and Chen, Y. (2009). Indoor localization using improved RSS-based lateration methods. In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, 1–6. IEEE.

Ye, W., Heidemann, J., and Estrin, D. (2004). Medium access control with coordinated adaptive sleeping for wireless sensor networks. *Networking, IEEE/ACM Transactions on*, *12*(3), 493–506.

Yu, M., Mokhtar, H., and Merabti, M. (2007). Fault management in wireless sensor networks. *Wireless Communications, IEEE*, *14*(6), 13–19.

Zhang, B., Yu, F., and Zhang, Z. (2009). A high energy efficient localization algorithm for wireless sensor networks using directional antenna. In *High Performance Computing and Communications, 2009. HPCC'09. 11th IEEE International Conference on*, 230–236. IEEE.

Zhang, T., Chen, Z., Ouyang, Y., Hao, J., and Xiong, Z. (2009). An improved RFID-based locating algorithm by eliminating diversity of active tags for indoor environment. *The Computer Journal*, *52*(8), 902–909.

Zhang, Z.-T., Xu, D., *et al.* (2009). High-speed vision system based on smart camera and its target tracking algorithm. *J. Robot.*, 229–234.

Zhao, Y., Liu, Y., and Ni, L. M. (2007). VIRE: Active RFID-based localization using virtual reference elimination. In *Parallel Processing, 2007. ICPP 2007. International Conference on*, 56–56. IEEE.

Zhao, Y. J., Govindan, R., and Estrin, D. (2002). Residual energy scan for monitoring sensor networks. In *Wireless Communications and Networking Conference, 2002. WCNC2002. 2002 IEEE*, Volume 1, 356–362. IEEE.

Zhou, X., Ma, J., Li, G. Y., Kwon, Y. H., and Soong, A. C. (2010). Probability-based combination for cooperative spectrum sensing. *IEEE Transactions on Communications*, *58*(2), 463–466.

Zhu, L., Yu, F. R., and Ning, B. (2010). Availability Improvement for WLAN-Based Train-Ground Communication Systems in Communication-Based Train Control (CBTC). In *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, 1–5. IEEE.