



THE UNIVERSITY OF QUEENSLAND A U S T R A L I A

Bachelor of Engineering
Undergraduate Thesis Proposal

Presence Detection/Location Tracking from WiFi Sniffing of Smart Phones

Chris Berry-Porter

42021395

christopher.berryporter@uqconnect.edu.au

28th August 2014

Under the supervision of

Mark Schulz

Abstract

Detecting the presence or location of smart devices (such as phones or tablets) has a wide range of uses in today's society. This is because tracking smart devices is synonymous with tracking users since users carry these devices on them most of the time. This device and user location information can be used for a variety of purposes including being used to track traffic flow on a highway, tracking shoppers in a store to gauge certain product interest or for control of lights, power or devices in a smart home or office.

The typical method of getting the location of an object is using GPS however this does not work reliably indoors. The solution proposed by this project is to locate smart devices using wireless sniffing. This technique allows wireless devices to be detected by sensor nodes placed throughout the area where the smart devices will be found. It is a passive technique and does not require the users to install any application at all.

This document outlines the features and functionality of the proposed solution as well as the background information required to understand how the system will operate.

Contents

Abstract	2
1. Introduction.....	4
1.1 Product Specification.....	5
2. Background.....	6
2.1 Wireless Localisation	6
2.2 Multilateration	6
2.3 Message Queue Telemetry Transfer	7
2.4 Existing Indoor Localisation Systems.....	8
2.4.1 Navizon	8
2.4.2 Insiteo	8
3. Project Plan	9
3.1 Gantt Chart	9
3.2 Milestones	10
3.2.1 Sensor Node Packet Capture.....	10
3.2.2 Sensor Node MQTT Publisher	10
3.2.3 Server MQTT Subscriber	11
3.2.4 Server Localisation.....	11
3.2.5 Server API	11
3.2.6 User Interface	11
3.3 Risk Assessment	12
3.3.1 OHS Risk Assessment.....	12
3.3.2 Project Risk Assessment	12
Appendix A – UQ Risk Rating Matrix	13
References.....	14

1. Introduction

One of the key benefits of applications involving mobile technology is the use of location awareness to provide a better experience for users. Location awareness allows an application to use the provided location information to perform additional tasks or provide extra functionality. This location information may not necessarily be in the form of latitude-longitude co-ordinates but could instead be a simple classification scheme such as indoors/outdoors or present/not present.

This location information can be put to use in a wide range of applications. In some cases the location information forms the core of the application functionality which is the case in most mapping, best route or guidance systems [1]. Other applications use this information to extract some secondary information such as tracking mobile devices to determine traffic flow on a highway [2], [3] or tracking user activity [4], [5].

While there are numerous localisation techniques, each of them has their own strengths and weaknesses. Localisation using the Global Positioning System (GPS), for example, is very accurate but only works reliably outdoors and consumes a lot of power. Localisation using mobile phone network transmission towers is another popular method [6], however not all devices use mobile phone networks and its performance and accuracy is highly reliant on the density of transmission towers in the area.

Localisation using wireless sniffing is a technique similar to localisation using mobile phone networks, however instead of using mobile phone network transmission towers, sensor nodes or access points can be used instead. The system works by using the sensor nodes to listen to the wireless transmissions from wireless devices. From this information, the location or presence of a wireless device can be calculated.

This project aims to create a modular and scalable system that can be used standalone or as a framework to track mobile devices using Institute of Electrical and Electronics Engineers (IEEE) 802.11 Wi-Fi modules. As shown in Figure 1 on the next page, the system will consist of four main parts: the wireless sensor nodes which will be placed throughout the area where the devices to be localised or detected will be present, a Message Queue Telemetry Transfer (MQTT) broker, an API server which will handle all client requests for data and perform any localisation computation, and a user friendly interface allowing the data collected from the sensor nodes to be visualised in a meaningful way. A database server could also optionally be added to allow for viewing of past data rather than just in real-time.

1.1 Product Specification

The final product shall:

- Detect the presence of wireless devices
- Operate passively without the need for users to install an application or modify their device
- Feature an opt-in scheme for privacy
- Provide a user friendly interface for visualisation of location data

The final product should:

- Allow extra sensor nodes to be added to the system without requiring server configuration
- Allow external programs access to collected data through both raw messages and an API

The final product may:

- Implement multilateration using the Received Signal Strength Indicator (RSSI) to improve location estimates
- Detect events of interest such as arrival or departure of a wireless device from an area
- Support viewing past data in addition to real-time data

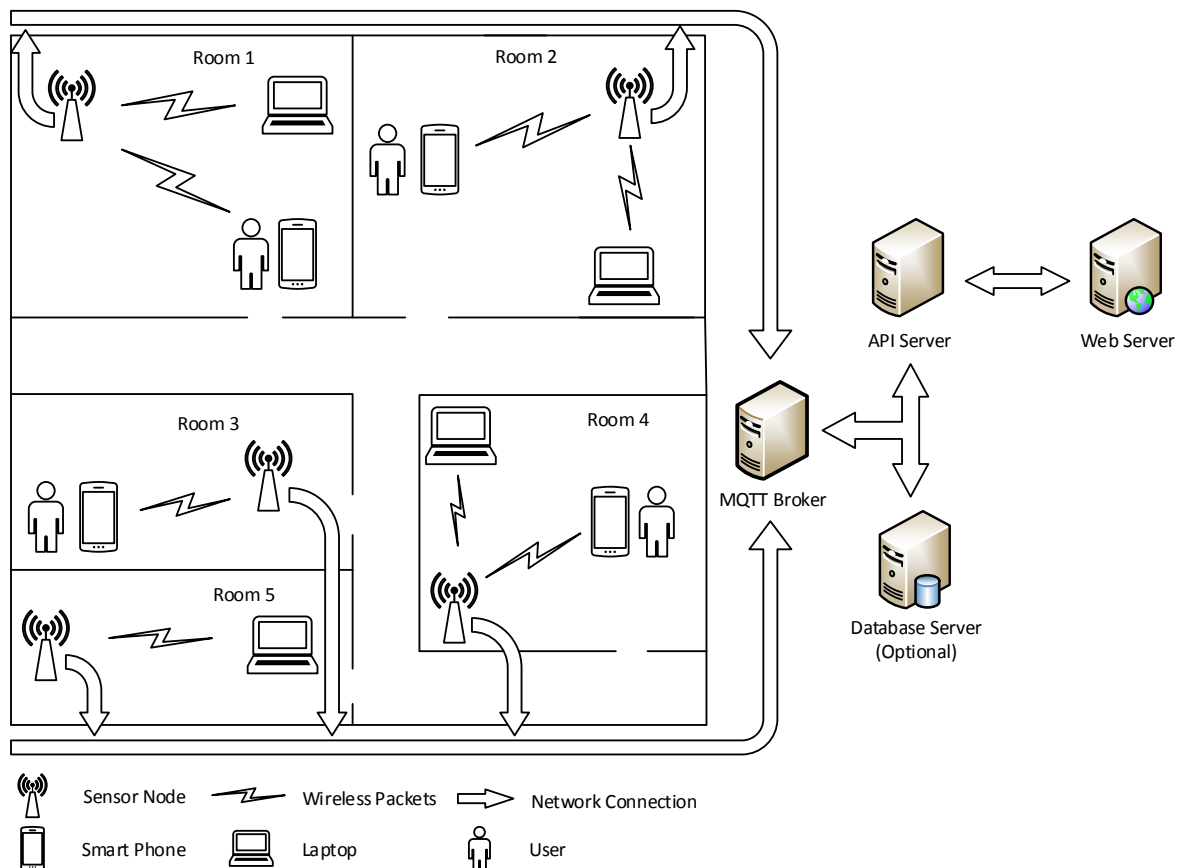


Figure 1: System Overview

2. Background

2.1 Wireless Localisation

Wireless localisation is one of many methods used for performing localisation of mobile devices. It is mainly used for indoor localisation due to the unreliability of GPS indoors and due to the need to place multiple sensor nodes in set spots throughout the area where the devices need to be tracked, which can be expensive or impractical for outdoor usage.

The way in which wireless localisation works is that all wireless devices communicate by sending packets of data to one another via radio transmission. All network devices identify themselves using a Media Access Control (MAC) address which is unique. This provides a reliable means for tracking wireless devices since within every packet, the MAC addresses of the device that sent the packet and the device the packet is destined for can be found [7]. Given this information, any wireless device can be tracked or detected by using sensor nodes or access points which listen for these data packets sent by the mobile devices.

In addition to the MAC addresses found in the wireless packets, many wireless receivers will also output the received signal strength of any wireless transmissions they receive [8]. Since signal strength decays as the distance between a radio transmitter and receiver increases, the received signal strength can be used as an indicator of how close a wireless device is to the receiver. This signal strength information can be used to improve location accuracy using multilateration.

2.2 Multilateration

Multilateration is a common method used to localise objects based on the distance between the object and many measurement points at set or known locations [9], [10], [11]. In the case of wireless localisation, the wireless device broadcasts to many sensor nodes at known locations. By using the received signal strength measured by each node, an estimate of the distance between the wireless device and the node can be calculated.

Having calculated the distance between each node and the wireless device, a circle around each node can be plotted where the wireless device must be at some point. Ideally, all circles will intersect at a single point where the wireless device is actually located, however it is common for the circles to not intersect at a single point or not at all due to many factors including multipath propagation, radio interference and other environmental factors affecting signal strength. These situations are illustrated below in Figure 2.

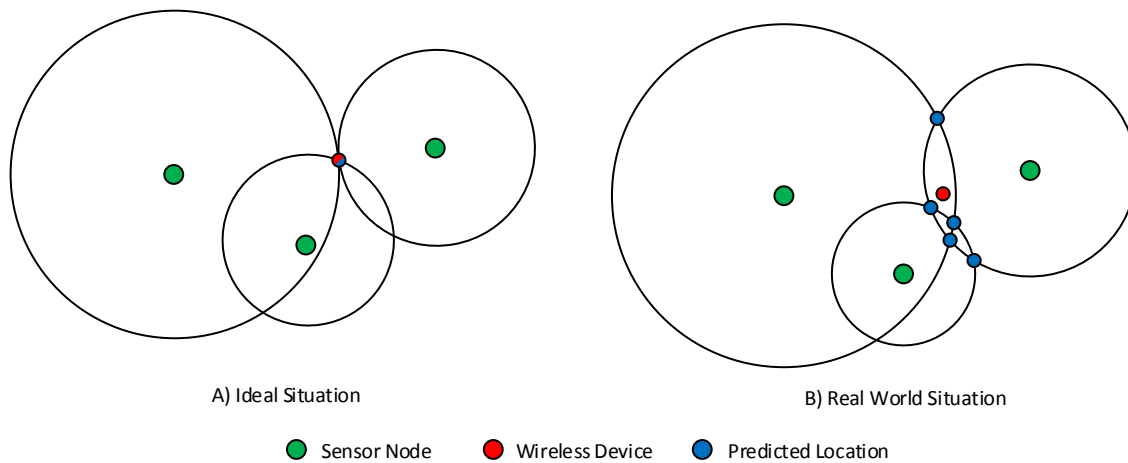


Figure 2: Multilateration example of a wireless device using three sensor nodes

In the real world case shown in Figure 2, since there are multiple predicted locations, the predicted locations would need an interpolation algorithm applied (averaging or least squares estimation) in order to better approximate the actual position of the wireless device.

2.3 Message Queue Telemetry Transfer

MQTT is a light-weight publish-subscribe protocol designed for the transfer of small messages over a network where network access is expensive or bandwidth may be limited [12]. The publish-subscribe messaging architecture works by having senders of messages, called publishers, send their messages without any knowledge of the receivers, called subscribers.

This works by having a central server called a broker which all publishers and subscribers connect to. This gives publishers a simple way to send data to multiple subscribers since the broker handles sending messages to individual subscribers. This system also allows subscribers to gather data from multiple publishers through a single connection to the broker.

Message streams are organised on the broker by topics which act similarly to modern file systems using forward slashes to separate levels [13]. This allows topics to contain a hierarchical structure so that subscribers can easily subscribe to topics and/or their subtopics using wildcards. MQTT defines two wildcard characters to be used when subscribing to topics. The hash (#) character matches all levels below the current one while the plus (+) character matches only one level below.

For example, the structure for data gathered by this project may be organised as: sensornodes/NODEX/raw and sensornodes/NODEX/json where the NODEX section would be replaced by the name or number of each sensor node publishing data. The raw topic for each node might contain data in a binary format while the json topic might contain the data in JavaScript Object Notation (JSON).

Using `sensornodes/+/json` as the subscription topic would allow a subscriber to receive data from all nodes in JSON format. Using `sensornodes/+/+` as the subscription topic however would be the same as using `sensornodes/#` and would allow the subscriber to receive data from all nodes in both JSON and binary formats. This is just a few examples of the flexibility offered by using the MQTT protocol.

2.4 Existing Indoor Localisation Systems

2.4.1 Navizon

Navizon is a company which provides an indoor localisation service which uses wireless sniffing. Navizon use a system very similar to the one described for this project [14]. They use many sensor nodes placed throughout a building in order to collect data about wireless devices, which is then sent over the internet to Navizon's own servers. From here, they provide an API for external applications to access or wireless device locations can be viewed directly via a web interface hosted by Navizon. There are no details as to whether they use multilateration or some other algorithm for calculating wireless device positions but this is a great example of an already working system claiming accuracy of 3 metres.




















2.4.2 Insiteo

Insiteo is another company which provides indoor localisation services, however they don't just rely on wireless sniffing. Insiteo use both wireless and Bluetooth sniffing in addition to tapping into the accelerometer, compass and barometer in smart phones and tablets in order to increase their location accuracy [15]. Insiteo also provides some additional features with their system including geo-fencing, proximity beacons and indoor mapping and guidance. This of course means that users must install an application on their device to use Insiteo's system, however it does offer a great deal more functionality than just simple location of devices.

3. Project Plan

Due to the limited time constraints placed upon this project, it is important that each of the tasks and milestones that make up the project have enough time allocated to them. Table 1 below outlines each of these tasks and milestones and the estimated time required to complete them.

Table 1: Project Tasks and Milestones

	 Task Mode ▾	Task Name ▾	Duration ▾	Start ▾	Finish ▾	Predecessors ▾
1		Thesis Project	231 days	Mon 28/07/14	Mon 15/06/15	
2		Background Research	25 days	Mon 28/07/14	Fri 29/08/14	
3		Sensor Node Packet Capture Development	15 days	Mon 1/09/14	Fri 19/09/14	2
4		Sensor Node MQTT Publisher Development	15 days	Mon 22/09/14	Fri 10/10/14	3
5		Server MQTT Subscriber Development	15 days	Mon 13/10/14	Fri 31/10/14	4
6		Exams and End of Year Break	85 days	Mon 3/11/14	Fri 27/02/15	5
7		Server Localisation Development	25 days	Mon 2/03/15	Fri 3/04/15	6
8		Server API Development	15 days	Mon 6/04/15	Fri 24/04/15	7
9		User Interface Development	20 days	Mon 6/04/15	Fri 1/05/15	7
10						
11		Thesis Milestones	215 days	Mon 18/08/14	Mon 15/06/15	
12		Project Proposal	8 days	Mon 18/08/14	Wed 27/08/14	
13		Project Proposal Submission	0 days	Thu 28/08/14	Thu 28/08/14	12
14		Progress Seminar Preparation	10 days	Mon 29/09/14	Fri 10/10/14	
15		Progress Seminar	0 days	Mon 13/10/14	Mon 13/10/14	14
16		Poster and Demonstration Preparation	15 days	Mon 4/05/15	Fri 22/05/15	
17		Poster and Demonstration	0 days	Mon 25/05/15	Mon 25/05/15	16
18		Thesis Report	14 days	Tue 26/05/15	Fri 12/06/15	
19		Thesis Report Submission	0 days	Mon 15/06/15	Mon 15/06/15	18

3.1 Gantt Chart

In order to better visualise the amount of time needed for the tasks and milestones listed above in Table 1, Figure 3 below shows two Gantt charts. The top Gantt chart displays all the development tasks required to complete the project while the bottom Gantt chart displays all the required deliverables such as reports and seminars.

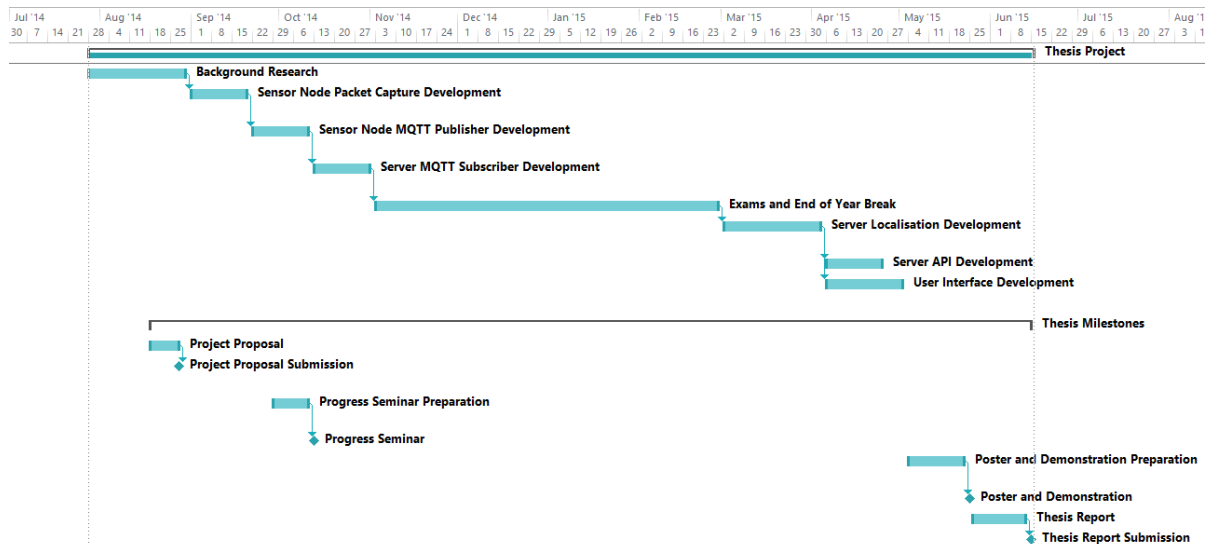


Figure 3: Project Gantt Chart

3.2 Milestones

The project has been broken up into several important milestones in order to help break the project down into manageable parts. The milestones listed below only include project implementation milestones, not general thesis deliverable milestones.

3.2.1 Sensor Node Packet Capture

The first milestone is to be able to listen and capture wireless packets being transmitted by wireless devices. In addition to capturing wireless packets, the data provided by each packet also needs to be extracted including received signal strength and sender/receiver MAC addresses. Development of this feature is expected to take approximately two weeks. This is a significant milestone since the rest of the project relies on the data collected through packet capture.

3.2.2 Sensor Node MQTT Publisher

Following on from the packet capture milestone, the data collected by each node needs to be transmitted to a server. The MQTT protocol will be used for all communication between the sensor nodes and the server so this milestone is about adding support for publishing MQTT messages to the sensor node software which captures the wireless packets. Many MQTT client libraries already exist to help make this task easier so it is expected to only require two weeks in development. No subscriber client will have been created at this point, so testing of this feature will be performed with existing MQTT tools which allow for displaying messages from topics as text.

3.2.3 Server MQTT Subscriber

With majority of the sensor node software complete, the focus will move onto the server. This milestone is to build the foundation of the server, allowing the server to subscribe to the data published by the sensor nodes. Again, there are many MQTT client libraries available to help reduce the complexity of this milestone. Testing of this milestone will be performed using several sensor nodes publishing to an MQTT broker with the server connected as a subscriber. The time allocated to completing this milestone is two weeks due to the use of an MQTT client library.

3.2.4 Server Localisation

The server localisation milestone will focus on implementing a localisation algorithm such as multilateration. This task will involve a large amount of testing in order to tune the localisation algorithm to work with a large range of wireless devices in different environments. Several algorithms may be implemented in order to compare which are the best in different situations. This is another significant milestone since it will largely determine how accurate the system is at calculating the location of a wireless device. Due to the required amount of testing, four weeks have been allocated for the completion of this milestone.

3.2.5 Server API

In order to allow external programs to be able to use the localisation data, an API will be implemented on the server. This will allow the server to perform all the calculations and localisation using the data from the sensor nodes itself and simply pass on locations of individual wireless devices. The API will also provide the positions of the sensor nodes themselves should this information be needed. Two weeks have been allocated for the development of the API which also overlaps with the time allocated for the user interface. This has been scheduled this way because the user interface will use the API and will help give a better idea of any extra features that might need to be added to the API.

3.2.6 User Interface

Finally, the system needs to be able to visualise the location data to end users. This milestone will be all about creating a user friendly interface capable of displaying the positions of detected wireless devices on a map. The user interface will use the server API in order to acquire the location data and will serve as an example of how an external application can use the API to get the location data. Three weeks have been allocated to this milestone however as mentioned above, this milestone will partly run in parallel with the server API milestone.

3.3 Risk Assessment

3.3.1 OHS Risk Assessment

Majority of the work for this project will be completed in a laboratory classified as low risk by the School of Information Technology and Electrical Engineering (ITEE). As such, general OHS laboratory safety rules will apply when working in such a laboratory.

3.3.2 Project Risk Assessment

The project risk assessment below has been created using the UQ risk rating matrix [16] which can be found in Appendix A.

Risk	Likelihood	Mitigation	New Likelihood
	Severity		New Severity
	Relative Risk		New Relative Risk
Risk of loss of code or software due to storage damage, corruption or accidental deletion.	Possible	Use of source code repository to keep a track of changes and to use as a backup. Additionally use cloud backup for non-source code related data.	Possible
	Moderate		Insignificant
	High		Low
Risk of damage to sensor node hardware causing inaccurate results or failure of the system.	Possible	Ensure equipment is placed in a secure environment where it is unlikely to be damaged.	Unlikely
	Moderate		Minor
	High		Moderate
Risk of thesis student falling ill or sustaining an injury preventing work from being completed and delaying the project.	Possible	Use of correct health and safety procedures will minimise the occurrence of this happening however this is generally uncontrollable.	Possible
	Moderate		Minor
	High		Moderate

Appendix A – UQ Risk Rating Matrix

Risk Rating Matrix									
			Consequence						
			People	Environment	Financial				
			No Injuries	Minor injury or First Aid Treatment Case.	Serious injury causing hospitalisation or multiple medical treatment cases.	Life threatening injury or multiple serious injuries causing hospitalisation.	Death or multiple life threatening injuries.		
				Spillage contained at site	Spillage contained but with outside help	Extensive spill	Toxic release of Chemicals		
			1% of Budget or <\$5K	2.5% of Budget or <\$50K	> 5% of Budget or <\$500K	> 10% of Budget or <\$5M	>25% of Budget or >\$5M		
			INSIGNIFICANT	MINOR	MODERATE	MAJOR	CATASTROPHIC		
			1	2	3	4	5		
Likelihood ↑	Is expected to occur in most circumstances	5	ALMOST CERTAIN	M (5)	H (10)	A (15)	A (20)	A (25)	
	Will probably occur	4	LIKELY	M (4)	M (8)	H (12)	A (16)	A (20)	
	Might occur at some time in the future	3	POSSIBLE	L (3)	M (6)	H (9)	H (12)	A (15)	
	Could occur but doubtful	2	UNLIKELY	L (2)	M (4)	M (6)	M (8)	H (10)	
	May occur but only in exceptional circumstances	1	RARE	L (1)	L (2)	L (3)	M (4)	M (5)	

Risk Category	Action	Tick
4 A: Acute	ACT NOW – Urgent – do something about the risk immediately. Requires immediate attention	
3 H: High	Highest management / Supervisor decision is required urgently	
2 M: Moderate	Follow Management Instructions	
1 L: Low	OK for now. Record and review if any equipment / people / materials / work process or procedures change	

References

- [1] GPS.gov, "GPS Applications," GPS.gov, 9 September 2013. [Online]. Available: <http://www.gps.gov/applications/>. [Accessed 23 August 2014].
- [2] A. Haghani, M. Hamed, K. F. Sadabadi, S. Young and P. Tarnoff, "Data Collection of Freeway Travel Time Ground Truth with Bluetooth Sensors," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2160, pp. 60-68, 2010.
- [3] A. Musa and J. Eriksson, "Tracking Unmodified Smartphones Using Wi-Fi Monitors," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, Toronto, Ontario, Canada, 2012.
- [4] B. Paramvir and V. N. Padmanabhan, "RADAR: An In-Building RF-based User Location and Tracking System," *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, pp. 775-784, 2000.
- [5] W. Qin, J. Zhang, B. Li and L. Sun, "Discovering Human Presence Activities with Smartphones Using Nonintrusive Wi-Fi Sniffer Sensors: The Big Data Prospective," *International Journal of Distributed Sensor Networks*, vol. 2013, p. 12, 2013.
- [6] V. Otsason, A. Varshavsky, A. LaMarca and E. de Lara, "Accurate GSM Indoor Localization," *UbiComp 2005: Ubiquitous Computing*, vol. 3660, pp. 141-158, 2005.
- [7] IEEE Standards Association, *IEEE Std. 802.11 - 2012*, 2012.
- [8] Radiotap, "Radiotap," Radiotap, 2 March 2011. [Online]. Available: <http://www.radiotap.org/>. [Accessed 23 August 2014].
- [9] B. Cook, G. Buckberry, I. Scowcroft, J. Mitchell and T. Allen, "Indoor Location Using Trilateration Characteristics," in *Proceedings London Communications Symposium*, London, UK, 2005.
- [10] "Dynamic Indoor Localization Using Multilateration With RSSI In Wireless Sensor Networks For Transport Logistics," *Procedia Engineering*, vol. 5, pp. 220-223, 2010.

- [11] M. Karagiannis, I. Chatzigiannakis and J. Rolim, "Multilateration: Methods For Clustering Intersection Points For Wireless Sensor Networks Localization With Distance Estimation Error," *CoRR*, vol. abs/1203.3704, 2012.
- [12] IBM, Eurotech, *MQTT V3.1 Protocol Specification*, 2010.
- [13] R. Light, "MQTT," Mosquitto, [Online]. Available: <http://mosquitto.org/man/mqtt-7.html>. [Accessed 23 August 2014].
- [14] Navizon, "Navizon ITS Fact Sheet," 2013. [Online]. Available: https://www.navizon.com/files/Navizon_ITS_Fact_Sheet.pdf. [Accessed 25 August 2014].
- [15] Insiteo, "Insiteo Platform," Insiteo, 2014. [Online]. Available: <http://www.indoor-gps.com/joomla/index.php/en/plateform>. [Accessed 27 August 2014].
- [16] UQ Health and Safety, "Risk Assessment Form," 11 December 2007. [Online]. Available: http://www.pf.uq.edu.au/pdf/SafetyForms/frm_PF388.pdf. [Accessed 27 August 2014].
- [17] I. Rose and M. Welsh, "Mapping The Urban Wireless Landscape With Argos," in *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, Zürich, Switzerland, 2010.
- [18] A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J. Tabert, P. Powledge, G. Borriello and B. Schilit, "Place Lab: Device Positioning Using Radio Beacons In The Wild," in *Proceedings of the Third international conference on Pervasive Computing*, Munich, Germany, 2005.