UNIVERSITI PUTRA MALAYSIA

KERNEL AND MULTI-CLASS CLASSIFIERS FOR MULTI-FLOOR WLAN LOCALISATION

MOHD AMIRUDDIN BIN ABD RAHMAN

FS 2016 52

KERNEL AND MULTI-CLASS CLASSIFIERS FOR MULTI-FLOOR WLAN LOCALISATION

By

MOHD AMIRUDDIN BIN ABD RAHMAN

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, & University of Sheffield in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

June 2016

Abstract of thesis presented to the Senate of Universiti Putra Malaysia and University of Sheffield in fulfilment of the requirement for the degree of Doctor of Philosophy

KERNEL AND MULTI-CLASS CLASSIFIERS FOR MULTI-FLOOR WLAN LOCALISATION

By

MOHD AMIRUDDIN BIN ABD RAHMAN

June 2016

Chair: Associate Professor Zulkifly Abbas, PhD

Faculty: Science

Indoor localisation techniques in multi-floor environments are emerging for location based service applications. Developing an accurate location determination and time-efficient technique is crucial for online location estimation of the multi-floor localisation system. The localisation accuracy and computational complexity of the localisation system mainly relies on the performance of the algorithms embedded with the system. Unfortunately, existing algorithms are either time-consuming or inaccurate for simultaneous determination of floor and horizontal locations in multi-floor environment. This thesis proposes an improved multi-floor localisation technique by integrating three important elements of the system; radio map fingerprint database optimisation, floor or vertical localisation, and horizontal localisation. The main focus of this work is to extend the kernel density approach and implement multi-class machine learning classifiers to improve the localisation accuracy and processing time of the each and overall elements of the proposed technique.

For fingerprint database optimisation, novel access point (AP) selection algorithms which are based on variant AP selection are investigated to improve computational accuracy compared to existing AP selection algorithms such as Max-Mean and InfoGain. The variant AP selection is further improved by grouping AP based on signal distribution. In this work, two AP selection algorithms are proposed which are Max Kernel and Kernel Logistic Discriminant that implement the knowledge of kernel density estimate and logistic regression machine learning classification.

For floor localisation, the strategy is based on developing the algorithm to determine the floor by utilising fingerprint clustering technique. The clustering method is based on simple signal strength clustering which sorts the signals of APs in each fingerprint according to the strongest value. Two new floor localisation algorithms namely Averaged Kernel Floor (AKF) and Kernel Logistic

Floor (KLF) are studied. The former is based on modification of univariate kernel algorithm which is proposed for single-floor localisation, while the latter applies the theory kernel logistic regression which is similar to AP selection approach but for classification purpose.

For horizontal localisation, different algorithm based on multi-class *k*-nearest neighbour (kNN) classifiers with optimisation parameter is presented. Unlike the classical *k*NN algorithm which is a regression type algorithm, the proposed localisation algorithms utilise machine learning classification for both linear and kernel types. The multi-class classification strategy is used to ensure quick estimation of the multi-class kNN algorithms.

The proposed algorithms are compared and analysed with existing algorithms to confirm reliability and robustness. Additionally, the algorithms are evaluated using six multi-floor and single-floor datasets to validate the proposed algorithms. In database optimisation, the proposed AP selection technique using Max Kernel could reduce as high as 77.8% APs compared to existing approaches while retaining similar accuracy as localisation algorithm utilising all APs in the database. In floor localisation, the proposed KLF algorithm at one time could demonstrate 93.4% correct determination of floor level based on the measured dataset. In horizontal localisation, the multi-class kNN classifier algorithm could improve 19.3% of accuracy within fingerprint spacing of 2 meters compared to existing algorithms.

All of the algorithms are later combined to provide device location estimation for multi-floor environment. Improvement of 43.5% of within 2 meters location accuracy and reduction of 15.2 times computational time are seen as compared to existing multi-floor localisation techniques by Gansemer and Marques. The improved accuracy is due to better performance of proposed floor and horizontal localisation algorithm while the computational time is reduced due to introduction of AP selection algorithm.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia dan University of Sheffield sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

KERNEL DAN KLASIFIKASI KELAS PELBAGAI UNTUK LOKALISASI BANGUNAN BERTINGKAT WLAN

Oleh

MOHD AMIRUDDIN BIN ABD RAHMAN

Jun 2016

Pengerusi: Profesor Madya Zulkifly Abbas, PhD

Fakulti: Sains

Teknik lokalisasi dalam bangunan untuk bangunan bertingkat sedang memuncul untuk aplikasi berdasarkan servis lokasi. Membangunkan sistem yang tepat dan cekap masa penting untuk anggaran lokasi semasa dalam bangunan bertingkat. Ketepatan dan kecekapan masa system bergantung terutamanya kepada prestasi algorithma yang terbenam dalam sistem. Namun, algoritma sedia ada adalah kurang cekap atau kurang tepat untuk dibangunkan dalam bangunan bertingkat. Tesis ini mencadangkan system lokalisasi bertingkat yang di tambah baik. Tiga elemen penting sistem iaitu pengoptimuman pangkalan data, lokalisasi lantai atau menegak, dan lokalisasi mendatar disepadukan dalam teknik lokalisasi bertingkat. Fokus utama kerja ini ialah untuk menambah baik ketepatan dan kecekapan pengiraan algoritma dengan mengambil kira setiap elemen tersebut.

Untuk pengoptimuman pangkalan data, teknik pemilihan titik akses (TA) yang baharu berdasarkan pemilihan TA berbeza dikaji untuk menambah baik ketepatan pengiraan berbanding teknik pemilihan TA yang lepas. Permilihan TA berbeza seterusnya ditambah baik dengan mengumpul TA berdasarkan ciri signal. Berdasarkan aspek ini, dua algoritma pemilihan TA dicadangkan iaitu algoritma Max Kernel dan Kernel Logistic Discriminant yang menggunapakai ilmu anggaran ketumpatan kernel dan klasifikasi pembelajaran mesin regresi logistik.

Untuk lokalisasi lantai, strategi adalah berdasarkan menggabungkan teknik pengklusteran fingerprint dengan algoritma lokalisasi lantai. Kaedah pengklusteran adalah berdasarkan pengklusteran kekuatan signal mudah dengan menyusun signal TA di setiap fingerprint berdasarkan nilai paling kuat. Dua algoritma lokalisasi lantai dinamakan algoritma Averaged Kernel Floor dan Kernel Logistic Floor dikaji. Algoritma pertama adalah berdasarkan pengubahsuaian algoritma kernel univariate yang digunakan untuk lokalisasi

satu aras. Algoritma kedua menggunakan teori kernel regresi logistic yang sama dengan teknik pemilihan TA tetapi untuk tujuan klasifikasi.

Untuk lokalisasi mendatar, algoritma lokalisasi berbeza berdasarkan algoritma klasifikasi kelas pelbagai k-nearest neighbour (kNN) dengan parameter pengoptimum dicadangkan. Tidak sama seperti algoritma kNN klasik yang merupakan algoritma jenis regresi, algoritma lokalisasi yang dicadangkan juga berdasarkan pengklasifikasi pembelajaran mesin. Algoritma tersebut dicadangkan dalam dua versi iaitu linear dan kernel. Strategi pelbagai-kelas untuk klasifikasi digunakan untuk memastikan anggaran pantas algoritma kNN pelbagai-kelas.

Kesemua algoritma yang dibangunkan dibandingkan secara kendiri dan dianalisa dengan algoritma terdahulu untuk mengesahkan kejituan dan kemapanan. Di samping itu, penilaian algoritma dibuat dengan pelbagai pangkalan data bertingkat dan satu aras untuk memastikan kebolehgunapakaian algoritma yang dicadangkan. Dalam pengoptimuman pangkalan data, algoritma pemilihan TA yang dicadangkan boleh mencapai sehingga 91.3% ketepatan di antara lokasi 2 meter dan pada masa yang sama menurunkan 17.7% kerumitan pengiraan. Dalam lokalisasi lantai, algoritma lantai yang dicadangkan menujukkan sehingga 96.8% ketepatan lantai. Dalam lokalisasi mendatar, algoritma yang dibangunkan mencapai sehingga 93.7% ketepatan di antara lokasi 2 meter.

Algoritma tersebut kemudian digabungkan untuk menganggarkan lokasi peranti untuk bangunan bertingkat. Keputusan purata 73.6% ketepatan di antara lokasi 2 meter dan 93.4% ketepatan lantai menunjukkan peningkatan berbanding teknik terdahulu oleh Gansemer dan Marques. Penambahbaikan kejituan disebabkan oleh prestasi algoritma lokalisasi lantai dan mendatar yang lebih baik manakala pengurangan kerumitan pengiraan disebabkan oleh pengenalan algoritma pemilihan TA.

ACKNOWLEDGEMENTS

Thanks to God! Many people have helped me in succession of this project by supporting me with different level of assistance. I by my deepest gratitude would like to thank them for their contributions.

First of all, I would like to thank my both main supervisors in UPM and University of Sheffield, Associate Professor Dr. Zulkifly Abbas and Professor Dr. Jie Zhang. Dr. Abbas has been excellent in helping me finishing this thesis in both technical and non-technical matters. Professor Zhang has given continuous support and guidance during this study. Also I would like to thank my supervisory committee, Dr. Xiaoli Chu (Sheffield) and Associate Professor Dr. Alyani Ismail (UPM) for their advices. My special gratitude goes to Dr. Marzieh Dashti (Nokia Bell Labs, Dublin) who has helped me in developing my technical skills in indoor localisation area. Also my appreciation goes to Dr. Holger Claussen, the manager in Nokia Bell Labs Dublin, and Dr. Fahim Kawsar, Director Internet of Things in Nokia Bell Labs Antwerp, for their trust in me to demonstrate the fingerprinting indoor positioning research work during my short period of time over there. Last but not least, my wonderful family for their continuous support from the beginning of this journey till now. Special gratitude goes to my wonderful wife, Siti Zeleha, for her never-ending love, patience, and encouragement in all these years.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Zulkifly Bin Abbas, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Jie Zhang, PhD
Professor
Department of Electronic and Electrical Engineering
University of Sheffield
(Member)

Alyani Binti Ismail, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Xiaoli Chu, PhD
Lecturer
Department of Electronic and Electrical Engineering
University of Sheffield
(Member)

BUJANG KIM HUAT, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly acknowledged;
- ownership of intellectual property from the thesis is as stipulated in the Memorandum of Agreement (MoA), or as according to the Universiti Putra Malaysia (Research) Rules 2012, in the event where the MoA is absent;
- permission from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) are required prior to publishing it (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: ________________________ Date: __________________

Name and Matric No.: Mohd Amiruddin Bin Abd Rahman (GS35311JD)

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: ______________________

Name of Chairman of Supervisory Committee: ______________________

Signature: ______________________

Name of Member of Supervisory Committee: **PROF DR JIE ZHANG**

Signature: ______________________

Name of Member of Supervisory Committee: ______________________

Signature: ______________________

Name of Member of Supervisory Committee: **DR XIAOLI CHU**

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIVIATIONS

AKF	Averaged Kernel Floor
ANN	Artificial Neural Network
AP	Access Point
CDF	Cumulative Density Function
FAF	Floor Attenuation Factor
GNSS	Global Navigation Satellite Service
IRLS	Iterative Reweighted Least Square
KDE	Kernel Density Estimate
KLF	Kernel Logistic Floor
KLPD	Kernel Logistic Pairwise Discriminant
KLR	Kernel Logistic Regression
*k*NN	k-Nearest Neighbour
LBS	Location-Based Service
LDA	Linear Discriminant Analysis
MAC	Media Access Control
MAP	*Maximum A Posteriori*
NLL	Negative Likelihood
NLOS	Non-Line of Sight
OVA	One-Versus-All
OVO	One-Versus-One
RBF	Radial Basis Function
RF	Radio Frequency
RSS	Received Signal Strength
SVM	Support Vector Machine
UPM	Universiti Putra Malaysia
WLAN	Wireless Local Area Network

AKF	Averaged Kernel Floor
ANN	Artificial Neural Network
AP	Access Point
CDF	Cumulative Density Function
FAF	Floor Attenuation Factor
GNSS	Global Navigation Satellite Service
IRLS	Iterative Reweighted Least Square
KDE	Kernel Density Estimate
KLF	Kernel Logistic Floor
KLPD	Kernel Logistic Pairwise Discriminant
KLR	Kernel Logistic Regression
*k*NN	k-Nearest Neighbour
LBS	Location-Based Service
LDA	Linear Discriminant Analysis
MAC	Media Access Control
MAP	*Maximum A Posteriori*
NLL	Negative Likelihood
NLOS	Non-Line of Sight
OVA	One-Versus-All
OVO	One-Versus-One
RBF	Radial Basis Function
RF	Radio Frequency

RSS	Received Signal Strength
SVM	Support Vector Machine
UPM	Universiti Putra Malaysia
WLAN	Wireless Local Area Network

LIST OF SYMBOLS

$\arg\max_{\xi} \mathrm{p}(.)$	*Maximum A Posteriori* estimate of ξ
$\mathbf{A}_v$	Pairwise APs signal vector of v-th combination in the radio map database
$\mathbf{A}_\beta$	Pairwise APs signal vector of β-th combination in the online sample
AP_l	l-th access point
$\hat{c}$	Estimated cluster/floor
$\boldsymbol{c}_f$	f-th cluster
D_i	Signal distance of i-th fingerprint
$D_{\beta i}$	Nearest signal distance of kNN classifier at i-th fingerprint location
$diag[.]$	Diagonal matrix function
$\mathbf{F}(\mathbf{p}_i)$	Fingerprint element at i-th location
h	Adjustable similarity parameter of Kernel Density Estimate
i	Fingerprint/Location index
k	Number of nearest estimated location
$\mathbf{K}(.)$	Kernel function
$NLL(.)$	Negative likelihood function of logistic regression
$\hat{\mathrm{p}}_{x\xi}$	Estimated x-coordinate location
$\hat{\mathrm{p}}_{y\xi}$	Estimated y-coordinate location
$\hat{\mathbf{p}}$	Estimated location
$\mathbf{p}_i$	Location of i-th fingerprint
$\mathrm{p}(.)$	Probability function
$\mathrm{P}_r(d_0)$	RSS at reference distance d_0 from an AP
$\mathrm{p}_{x\xi}$	Actual x-coordinate location
$\mathrm{p}_{y\xi}$	Actual y-coordinate location
$P_r(d)$	RSS at a fingerprint location of d distance from an AP
q	Location index of the AP l signal grouped in the cluster $\boldsymbol{c}_f$
$\boldsymbol{r}$	Vector of pairwise APs signal data
$RSS_{AP_l}(d_{AP_l})$	Received Signal Strength of AP_l at a distance d from the AP
$RSS_{AP_l}(d_0)$	Received Signal Strength of AP_l at a reference distance d_0
$S_{\beta i}$	Averaged of ϖ minimum $D_\beta(\psi_{\omega i})$ signal distance of i-th fingerprint location of kNN classifier
T	Transpose function
T	Total number of samples measured at one fingerprint location
U	Total number of points of signal range to determine the kernel in density estimate

$\boldsymbol{w}$	Model parameter vector of the logistic function
y	Class/Location of vector pairwise signal data $\boldsymbol{r}$
Ω_i	i-th class/location averaged $D_{\beta i}$or $S_{\beta i}$ signal distance of test sample AP combination
$\boldsymbol{\varphi}_i$	Offline signal vector of i-th fingerprint
Δf_{neigh}	Difference between two neighbouring kernel u and $u+1$
Ξ	Total number of location estimation
Φ	Total class labels or total number of vector signal data existed in both classes
ξ	Index of location estimation
η_{AP_l}	Path Loss Exponent referring to AP_l signals
η_{SF}	Similar floor path loss exponent
ρ	Number of class label from total of Φ
ϖ	Number of nearest minimum signal distance of each fingerprint location of kNN classifier
$\boldsymbol{\varphi}$	Online measured signal vector
γ	Adjustable similarity parameter of kernel function

CHAPTER 1

INTRODUCTION

1.1 Overview

Location is one of the most valuable information in mobile communication nowadays. Today's mobile devices are designed and programmed to have location features so it can be complemented with Location-Based Service (LBS) applications (Schiller and Voisard 2004). The location is important because it reflects interaction and context of the user based on the location of the device. In past times, location is mainly used to guide users to move from one place to another by giving the best possible route to reach the destination. However currently, the location information is used in much wider context. For example, by using smartphone a user can locate user's current position and share the location with his or her friends on the social network. Also, a user can book a taxi service or finding the nearest restaurants or cash machines by considering user's current location.

Unfortunately, all of these applications infer the device location mainly based on its position in outdoor environment which mainly depends on the information provided by Global Navigation Satellite System (GNSS) receiver integrated with the mobile device. However, with only information of outdoor location, further development or enhancement of LBS applications is restricted. In near future, LBS applications are designed and developed to work for indoor-based services. For example, to assist shoppers to find items that they want to purchase by locating the exact aisle of the item in a hypermarket, to help drivers find their car in a multi-story indoor airport car park, and to supply information for smart building administrators to monitor temperature, room availability, and lightings. Therefore it is a requirement to know accurate indoor location to achieve these objectives. GNSS receiver however is generally not suitable to provide indoor location due to blockage and attenuation of the signals by roofs, walls and other objects.

Researchers have been working to find alternative technologies to obtain accurate indoor location information. Some solutions includes Wireless Local Area Network (WLAN) infrastructure (Fang and Lin 2010, Prieto et al. 2012, Mirowski et al. 2014, Wang et al. 2015, Liang, Zhang, and Peng 2015), infrared (Petrellis, Konofaos, and Alexiou 2006, Tao et al. 2014), and Bluetooth (Hossain and Soh 2007, Jianyong et al. 2014, Gu and Ren 2015). Among the solutions, one of the most promising solutions is WLAN as its signal coverage is available almost anywhere in urban environments. Indoor localisation methods based on WLAN are largely documented in the literature and surpass any other indoor localisation technologies.

1.2 WLAN Indoor Localisation

Indoor localisation based on WLAN was pioneered by Bahl and Padmanabhan (2000). WLAN based localisation is established by associating received Radio Frequency (RF) signals with physical location. The received RF signals or also known as Received Signal Strength (RSS) could characterise different locations as the propagated signals are location dependent. To localise unique location, RSS is measured throughout the floor area as combination of multiple signals from multiple Access Points (APs).

The location estimation technique could be classified by two methods; radio propagation based model and fingerprinting method. In radio propagation based model, the location is estimated by triangulation where the location of three or more access points must be known and the path loss model such as log-distance, which described the environment dependent relationship of the distance between the transmitter (AP) and receiver (device) according to variation of signal strength value, of the APs are determined. During position request by the device, the signal vector measured by the device at an unknown location is used as the input to the path loss model to determine the distance of the device from the APs which translates the location of the device. On the other hand, the fingerprinting method first requires real surveying by collecting the signal signature at every unique physical location which is also called as fingerprint location. The collection of multiple signal signatures associated with the physical locations are stored in the database as radio map and during the location request by the device, the signal vector of the device is compared its similarity with the one in the database to determine the location.

Between the two methods, the latter technique, fingerprinting, is preferred. This is because higher positioning accuracy could be achieved compared to radio propagation based model. Radio propagation based model could not provide finer accuracy due to inability of the model to characterise complex multipath signals received at each specific locations. However, fingerprinting technique comes with the cost of high processing time of localisation algorithm due to large amount of signal signatures in the radio map. In today's application, indoor localisation system should be embedded in mobile device such as smartphone which has small computing capability, the localisation algorithm must be designed and developed to utilise as small processing power as possible and at the same time retain good positioning accuracy. Some examples of good and robust classical localisation algorithms are k-Nearest Neighbour (kNN) (Bahl and Padmanabhan 2000), univariate kernel (Roos et al. 2002), and multivariate kernel (Kushki et al. 2007). Additionally, the localisation algorithm must be designed so that it can work in multiple indoor environments especially in urban area where the application is demanded. Generally, these areas are occupied with various multi-floor constructions. Therefore, the indoor localisation system must be designed and developed for this kind of infrastructure.

1.3 Multi-Floor Localisation

Numerous studies can be found on development of indoor localisation system. However, majority of them are focussing on single-floor localisation (Wu et al. 2013, Sorour et al. 2015, Chen and Wang 2015). It is investigated that the research on multi-floor localisation receives less attention is mainly due to two reasons. First, large radio map datasets is required as multi-floor data must be collected e.g. large fingerprint dataset for fingerprinting method, or large AP location dataset for AP based method. Second, the perception that development of multi-floor localisation could be easily extended from single-floor localisation technique. However, multi-floor WLAN localisation is actually much more challenging compared to single floor localisation. Figure 1.1 illustrates the comparison between single-floor localisation and multi-floor localisation. It is understood that multi-floor localisation challenges comes additional floor environment which increases the complexity of localisation in multi-floor setting.

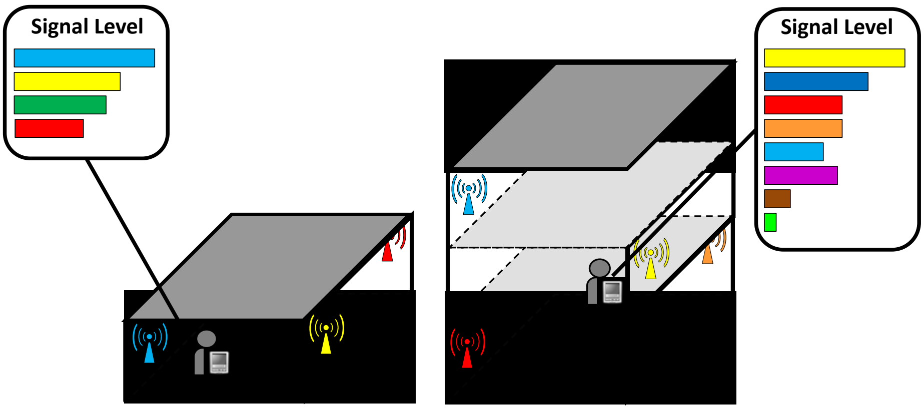

Figure 1.1. WLAN indoor localisation in single-floor and multi-floor settings

There are significant differences between multi-floor and single-floor WLAN localisation. First, the datasets of the collected signal data must be characterised by three-dimensional position, which includes floor level of the building and horizontal positions of the data compared to only horizontal positions required for single-floor localisation. As the amount of entries in the radio map varies according to the number of fingerprint locations, therefore the amount of entries of the datasets for multi-floor localisation is generally in multiplication of the number of fingerprint locations and number of floor level exists within the building. Second, the number of APs increases proportionally with the number of floor level in multi-floor building. This causes the dimensionality of AP during online phase increase. Additionally, during the signal measurement process, each signal vector for multi-floor localisation will be added with multiple APs signal from other floors in addition to signal from the APs on the floor itself compared to single-floor where the AP signals mostly come from the APs installed on the floor. The computation of the localisation algorithms depends on the three factors which are number of fingerprint locations, number of APs within the environment and the size of signal vector and these are the elements that are mentioned in first, second, and third comparisons. Therefore the computational complexity or processing time of the algorithm in multi-floor setting will be much higher compared to single-floor localisation. The fourth difference is the estimated locations in multi-floor environment require additional coordinate of the floor level or the z-coordinate compared to single-floor environment which is described by only x and y-coordinate. Lastly, the probability of error in multi-floor localisation is generally higher because of possibility that the estimated location is in different floor level than expected. All of these comparisons are summarised in Table 1.1.

Multi-floor localisation system could be divided into two main problems which are to locate the floor level of the device and to position the device on the chosen floor level which determines the horizontal location. The algorithms should be accurate and quickly processed the radio map database to give estimation of the location. Therefore, three important categories that should be investigated in order to produce an efficient multi-floor location are radio map fingerprint database optimisation, floor localisation algorithm, and horizontal localisation algorithm.

1.4 Kernel and Multi-Class Classifier Approach for Multi-Floor Localisation

In particular, the technique to estimate the location in multi-floor environment has been focusing on similar type of algorithm. The algorithm also is the extension of previously developed algorithm for single-floor localisation. Multi-floor localisation involves processing larger database compared to single-floor and therefore the extended algorithm is no longer suitable to be applied for multi-floor case. The usage of classical algorithms in multi-floor environment also leads to increasing computational complexity as related to increasing amount of database element. To solve the problem, this thesis proposes new multi-floor localisation technique based on kernel and multi-class classifier. The techniques implements kernel density estimate and multi-class classifier based on logistic regression as tools for AP selection, and floor localisation algorithm. The kNN multi-class classifier is applied for new horizontal localisation algorithm. Theoretical foundation and algorithm implementation of the technique is described in details in Chapter 3 and 4 respectively.

Table 1.1. Summary of differences between single-floor WLAN localisation versus multi-floor WLAN localisation

Issue	Single Floor Localisation	Multi-Floor localisation
• Radio map database	• Radio map database is for single floor and the quantity of the entries of the database depends on the number of fingerprint locations.	• Radio map database is in multiple number of floors exist in the building and the entries generally consists of multiplication of number of fingerprint locations and number of floor level.
• Dimensions of AP	• Dimensionality of AP depends on the number of APs installed in single floor.	• Dimensionality of APs increase with increasing number of APs installed on every floor level of the building.
• Fingerprint signal vector	• Signal vector at each fingerprint location is majorly from multiple APs that are installed on the floor level and the signal follows normal path loss model.	• Signal vector at each fingerprint location consists of signal from APs within similar floor level and also from other floor level within the building and the signal follows multi-floor path loss model.
• Estimated location coordinate dimension	• Estimated location is in 2 dimensional coordinate (x and y) which is the horizontal location.	• Estimated location is in 3 dimensional coordinate (x, y, and z) which includes floor location and horizontal location.
• Probability of location error	• The probability of location error is only within horizontal locations.	• The probability of location error may increase due to possibility of estimated location is located on different floor level than expected.

1.5 Problem Statement

To develop an efficient multi-floor localisation system, all of the elements of multi-floor localisation in Table 1.1 should be analysed according to categories mentioned above. There have been some developments of multi-floor WLAN localisation system and comprehensive review on the topic is written in Chapter 2 of this thesis. From the review, it is indicated that majority of multi-floor WLAN localisation algorithm are developed based on extension of single-floor localisation algorithms. Generally, the problem with this kind of system did not consider computational complexity of the algorithms implemented in multi-floor environment. Also, the algorithm is not optimised for simultaneous estimation of both floor and horizontal locations. Specifically, in order to develop an efficient and robust multi-floor indoor localisation system, this thesis investigates the following problems:

1. Majority of the proposed floor localisation algorithm is still based on classical similarity measure algorithm which is extended for multi-floor localisation. This means the developed floor localisation algorithm requires calculating every single entry of fingerprint to perform floor estimation. As discussed in Section 1.3, the multi-floor building problem involves the number of fingerprints is in the multiple of number of floor

level exists inside the building. Therefore the computational complexity increases as the number of floor increases.

2. Considering all APs for localisation may degrade the performance of the localisation system. The number of APs installed within building increases as the number of floor level increases so that the coverage of the signal is enough for localisation system to work. This leads to AP dimensionality in multi-floor environment is much higher compared to single floor.

3. The horizontal localisation algorithm in multi-floor setting is mainly implemented based on previously developed algorithm for single-floor problem. However in multi-floor building, additional AP signals are measured from other floor levels which degrade the performance of the algorithm. Consequently, the location estimation error in multi-floor location could not be minimised compared to single-floor location if similar algorithm is implemented for both environments.

4. The validity of some existing multi-floor localisation algorithm is questionable as the algorithms are only tested in limited testing area such as one or two buildings and the buildings are low rise which contains less than five floor levels, and it is not guaranteed that the proposed algorithm will produce the similar performance as in different environments.

5. The work on combining radio map database optimisation, floor localisation and horizontal localisation is not well studied to improve the multi-floor localisation system. Existing techniques are based solely on either improving floor localisation only or combination of database optimisation and floor localisation. The performance of combining all of the techniques is unknown.

1.6 Objective of the Research

The aim of this thesis is to develop a robust and efficient multi-floor localisation system emphasizing on: i) the accuracy of the localisation algorithms in both vertical and horizontal position which are characterised by estimated location error of the algorithms, and ii) the computational complexity of localisation algorithm which is to reduce the processing time of the developed algorithms. To achieve the aim, detail objectives are given as follows:

1. To optimise the radio map database by implementing AP selection technique to limit or reduce the number of required APs information to perform localisation and at the same time retain similar accuracy with using all APs information. Two novels AP selection algorithms are introduced to improve the selection of APs in optimising the database. The performances of the proposed AP selection schemes are compared with existing AP selection technique to evaluate the performance of the

proposed algorithms based on three classical localisation algorithms of kNN, univariate kernel, and multivariate kernel.

2. To reduce the processing complexity of floor localisation algorithm and at the same time to improve the accuracy of the estimated floor level. Proposed two new floor localisation algorithms based on clustered multi-floor radio map. The performances of the new floor localisation algorithms are compared with existing floor localisation algorithms to evaluate the effectiveness of the algorithm.

3. To improve the location estimation error of horizontal localisation algorithm by introducing novel localisation algorithms for both multi-floor and single-floor environments. The horizontal localisation algorithms are compared with classical single-floor localisation algorithms (kNN, univariate kernel, and multivariate kernel) and the performances of all algorithms are analysed and discussed.

4. To combine the proposed AP selection technique, floor localisation algorithm and horizontal localisation algorithms to evaluate performance in multi-floor environment. The test is in multiple multi-floor environments with different number of floor levels to verify the performance of the proposed algorithm. The performance is also compared with existing multi-floor localisation algorithm.

1.7 Organisation of the Thesis

In Chapter 2, review of the literature on the multi-floor localisation system is given. The review presents in depth study on problems exist in existing multi-floor localisation system which leads to the development of problem statement in Section 1.5. The review identifies the gaps in current research particularly in multiple scenarios of multi-floor e.g. validity of chosen environment for testing the multi-floor localisation system, radio map database optimisation techniques, floor localisation algorithms, and lastly the horizontal localisation algorithms.

Chapter 3 first presents the background on WLAN fingerprint localisation and theory that is related to localisation algorithm used in this thesis and second describes the novel algorithms developed for the multi-floor localisation. The state-of-the-art WLAN fingerprint localisation mechanism is introduced. The related improvement components of the multi-floor localisation are presented which involves the database optimisation, floor localisation, and horizontal localisation. The theory of three popular classical algorithms which are used as benchmark for proposed algorithms are also discussed. The theory on kernel density estimate and machine learning multiclass classification using kNN and logistic regression are explained. The theories are applied for the following proposed algorithms. The database optimisation algorithms implement Max Kernel and Kernel Logistic Pairwise Discriminant (KLPD). The floor localisation comprises of Averaged Kernel Floor and Kernel Logistic Floor algorithms. Normal and kernel multi-class kNN algorithms are used for horizontal localisation

Chapter 4 explains the measurement setup of collecting the RSS signal data. This includes the measurement tools, floor map, and specification of the measurement. Also the method to extract the path loss parameters of the measured signal data to be tested with propagation model is shown. The details of measured fingerprint database specification and evaluation of path loss model using extracted path loss parameter of the measured signal data from the database are described. The performance metrics to evaluate the developed algorithms are discussed. Additionally, the method to determined number of k used for kNN algorithm is presented.

Chapter 5 to 8 discuss the results related to the developed multi-floor system. The results explain the performance of the proposed AP Selection (Chapter 5), floor localisation (Chapter 6), and horizontal localisation algorithms (Chapter 7) of which results are evaluated and discussed. The results of combining the three proposed algorithms for multi-floor localisation are explained in Chapter 8. The results are mainly focusing on accuracy and computational complexity of the algorithms.

Lastly, Chapter 9 draws conclusion on the proposed multi-floor system. The work presented in the thesis is summarised. The contributions of the thesis are highlighted. Also, further research directions are suggested.

REFERENCES

Ahmad, U., A. Gavrilov, Lee Sungyoung, and Lee Young-Koo. 2006. "Modular Multilayer Perceptron for WLAN Based Localization." Neural Networks, 2006. IJCNN '06. International Joint Conference on, 0-0 0.

Ahmad, U., B. J. D'Auriol, Y.-K. Lee, and S. Lee. 2008. "Multi-Floor Semantically Meaningful Localization Using IEEE 802.11 Network Beacons." *IEICE Transactions on Communications* E91-B:3450-3460. doi: 10.1093/ietcom/e91-b.11.3450.

Akl, Robert G, Dinesh Tummala, and Xinrong Li. 2006. "Indoor propagation modeling at 2.4 GHz for IEEE 802.11 networks." The Sixth IASTED International Multi-Conference on Wireless and Optical Communications, July 3-5.

Al-Ahmadi, Abdullah Saad Mohammed, Abdusamea. I. A. Omer, Muhammad Ramlee Bin Kamarudin, and Tharek Abdul bin Abd Rahman. 2010. "MULTI-FLOOR INDOOR POSITIONING SYSTEM USING BAYESIAN GRAPHICAL MODELS." *Progress In Electromagnetics Research B* 25:241-259. doi: 10.2528/PIERB10081202.

Alsehly, Firas, Tughrul Arslan, and Zankar Sevak. 2011. "Indoor positioning with floor determination in multi story buildings." *2011 International Conference on Indoor Positioning and Indoor Navigation*:1-7. doi: 10.1109/IPIN.2011.6071945.

Arya, Azin, Philippe Godlewski, Marine Campedel, and Ghislain Du Chéne. 2013. "Radio database compression for accurate energy-efficient localization in fingerprinting systems." *IEEE Transactions on Knowledge and Data Engineering* 25:1368-1379. doi: 10.1109/TKDE.2011.241.

Bahl, P., and V.N. Padmanabhan. 2000. "RADAR: an in-building RF-based user location and tracking system." *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)* 2. doi: 10.1109/INFCOM.2000.832252.

Bhargava, Preeti, Shivsubramani Krishnamoorthy, and Aditya Karkada. 2013. "Locus: An Indoor Localization, Tracking and Navigation System for Multi-story Buildings Using Heuristics Derived from Wi-Fi Signal Strength." 120:212-223. doi: 10.1007/978-3-642-40238-8.

Bolliger, Philipp Lukas. 2011. "Robust indoor positioning through adaptive collaborative labeling of location fingerprints." Doctor of Sciences PhD Dissertation, ETH ZURICH (20109).

Campos, Rafael Saraiva, Lisandro Lovisolo, and Marcello Luiz R. de Campos. 2014. "Wi-Fi multi-floor indoor positioning considering architectural aspects and controlled computational complexity." *Expert Systems with Applications* 41:6211-6223. doi: 10.1016/j.eswa.2014.04.011.

Chen, Qiuyun, and Bang Wang. 2015. "FinCCM: Fingerprint Crowdsourcing, Clustering and Matching for Indoor Subarea Localization." *IEEE Wireless Communications Letters* 4:1-1. doi: 10.1109/LWC.2015.2482971.

Chen, Yiqiang, Qiang Yan, Jie Yin, and Xiaoyong Chai. 2006. "Power-efficient Access-Point selection for indoor location estimation." *IEEE Transactions on Knowledge and Data Engineering* 18:877-888. doi: 10.1109/TKDE.2006.112.

Cheng, Jiujun, Yueqiao Cai, Qingyang Zhang, Junlu Cheng, and Chendan Yan. 2014. "A New Three-Dimensional Indoor Positioning Mechanism Based on Wireless LAN." *Mathematical Problems in Engineering* 2014:1-7. doi: 10.1155/2014/862347.

Chintalapudi, Krishna, Anand Padmanabha Iyer, and Venkata N. Padmanabhan. 2010. "Indoor localization without the pain." *Proceedings of the sixteenth annual international conference on Mobile computing and networking - MobiCom '10*:173. doi: 10.1145/1859995.1860016.

Fang, Shih-Hau Fang Shih-Hau, and Tsung-Nan Lin Tsung-Nan Lin. 2010a. "Accurate Indoor Location Estimation by Incorporating the Importance of Access Points in Wireless Local Area Networks." *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*:1-5. doi: 10.1109/GLOCOM.2010.5684309.

Fang, Shih-hau, and Tsung-nan Lin. 2009. "Projection-Based Location System via Multiple Discriminant Analysis in Wireless Local Area Networks." *IEEE Transactions on Vehicular Technology* 58:5009-5019. doi: 10.1109/TVT.2009.2025134.

Fang, Shih-Hau, and Tsung-Nan Lin. 2010b. "A dynamic system approach for radio location fingerprinting in wireless local area networks." *IEEE Transactions on Communications* 58:1020-1025. doi: 10.1109/TCOMM.2010.04.090080.

Fang, Shih Hau, and Tsung-nan Lin. 2012. "Principal component localization in indoor wlan environments." *IEEE Transactions on Mobile Computing* 11:100-110. doi: 10.1109/TMC.2011.30.

Fang, Shih Hau, and Tsung Nan Lin. 2010c. "Cooperative multi-radio localization in heterogeneous wireless networks." *IEEE Transactions on Wireless Communications* 9:1547-1551. doi: 10.1109/TWC.2010.05.090061.

Farid, Zahid, Rosdiadee Nordin, and Mahamod Ismail. 2013. "Recent Advances in Wireless Indoor Localization Techniques and System." *Journal of Computer Networks and Communications* 2013:1-12. doi: 10.1155/2013/185138.

Gansemer, S., S. Hakobyan, S. Püschel, and U. Großmann. 2009. "3D WLAN indoor positioning in multi-storey buildings." *Proceedings of the 5th IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS'2009*:669-672. doi: 10.1109/IDAACS.2009.5342893.

Gansemer, Sebastian, Uwe Grossmann, and Syuzanna Hakobyan. 2010. *RSSI-based Euclidean Distance algorithm for indoor positioning adapted for the use in dynamically changing WLAN environments and multi-level buildings*: IEEE.

Genming, Ding, Tan Zhenhui, Zhang Jinbao, and Zhang Lingwen. 2013. "Fingerprinting localization based on affinity propagation clustering and artificial neural networks." Wireless Communications and Networking Conference (WCNC), 2013 IEEE, 7-10 April 2013.

Gu, Yu, and Fuji Ren. 2015. "Energy-efficient Indoor Localization of Smart Hand-held Devices Using Bluetooth." *IEEE Access* PP:1-1. doi: 10.1109/ACCESS.2015.2441694.

Gupta, Pankaj, Sachin Bharadwaj, Sthanunathan Ramakrishnan, and Jaiganesh Balakrishnan. 2014. "Robust floor determination for indoor positioning." *2014 Twentieth National Conference on Communications (NCC)*:1-6. doi: 10.1109/NCC.2014.6811285.

He, Suining, and S.-H. Gary Chan. 2015. "Wi-Fi Fingerprint-based Indoor Positioning: Recent Advances and Comparisons." *IEEE Communications Surveys & Tutorials*:1-1. doi: 10.1109/COMST.2015.2464084.

Honkavirta, Ville, Tommi Perala, Simo Ali-Loytty, and Robert Piche. 2009. *A comparative survey of WLAN location fingerprinting methods*: IEEE.

Hossain, A K M Mahtab, and Wee-seng Soh. 2007. *A Comprehensive Study of Bluetooth Signal Parameters for Localization*: IEEE.

Jianyong, Zhu, Luo Haiyong, Chen Zili, and Li Zhaohui. 2014. "RSSI based Bluetooth low energy indoor positioning."526-533. doi: 10.1109/IPIN.2014.7275525.

Khaoampai, Kornkanok, Kulit Na Nakorn, and Kultida Rojviboonchai. 2014. "Low complexity floor localization algorithm for mobile phone." *2014 11th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*:1-6. doi: 10.1109/ECTICon.2014.6839765.

King, Thomas, Stephan Kopf, Thomas Haenselmann, Christian Lubberger, and Wolfgang Effelsberg. 2006. "COMPASS: A probabilistic indoor positioning system based on 802.11 and digital compasses." *Online*:34-40. doi: 10.1145/1160987.1160995.

Kung, Sun Yuan. 2014. *Kernel methods and machine learning*: Cambridge University Press.

Kushki, Azadeh, Konstantinos N Plataniotis, A N Venetsanopoulos, and Senior Student Member. 2007. "Kernel-Based Positioning in Wireless Local Area Networks." *Mobile Computing, IEEE Transactions on* 6:689-705. doi: 10.1109/TMC.2007.1017.

Laoudias, C., P. Kemppi, and C. G. Panayiotou. 2009. "Localization Using Radial Basis Function Networks and Signal Strength Fingerprints in WLAN." Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE, Nov. 30 2009-Dec. 4 2009.

Laoudias, Christos, George Constantinou, Marios Constantinides, Silouanos Nicolaou, Demetrios Zeinalipour-Yazti, and Christos G Panayiotou. 2012. "The airplace indoor positioning platform for android smartphones." Mobile Data Management (MDM), 2012 IEEE 13th International Conference on.

Letchner, Julie, Dieter Fox, and Anthony LaMarca. 2005. "Large-scale localization from wireless signal strength." Proceedings of the national conference on artificial intelligence.

Liang, Dong, Zhaojing Zhang, and Mugen Peng. 2015. "Access Point Reselection and Adaptive Cluster Splitting-Based Indoor Localization in Wireless Local Area Networks." *IEEE Internet of Things Journal* 2:573-585. doi: 10.1109/JIOT.2015.2453419.

Lin, Tsung-nan, Senior Member, Shih-hau Fang, Wei-han Tseng, Chung-wei Lee, and Jeng-wei Hsieh. 2014. "A Group-Discrimination-Based Access Point Selection for WLAN Fingerprinting Localization." 63:3967-3976.

Liu, Hui, Houshang Darabi, Pat Banerjee, and Jing Liu. 2007. "Survey of wireless indoor positioning techniques and systems." *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews* 37:1067-1080. doi: 10.1109/TSMCC.2007.905750.

Liu, Hung-Huan, Chung-Wei Liao, and Wei-Hsiang Lo. 2015. "WiFi-based Floor Estimation for Multi-Storey Building: A Case Study for a Crowded AP Environment." Intelligent Systems and Applications: Proceedings of the International Computer Symposium (ICS) Held at Taichung, Taiwan, December 12–14, 2014.

Liu, Hung-Huan, and Yu-Non Yang. 2011. *WiFi-based indoor positioning for multi-floor Environment*.

Lohan, Elena Simona, Jukka Talvitie, Pedro Figueiredo e Silva, Henri Nurminen, Simo Ali-Loytty, and Robert Piche. 2015. *Received signal strength models for WLAN and BLE-based indoor positioning in multi-floor buildings*: IEEE.

Maneerat, Kriangkrai, and Chutima Prommak. 2014. "An Enhanced Floor Estimation Algorithm for Indoor Wireless Localization Systems Using Confidence Interval Approach."1058-1062.

Maneerat, Kriangkrai, Chutima Prommak, and Kamol Kaemarungsi. 2014. *Floor estimation algorithm for wireless indoor multi-story positioning systems*: IEEE.

Marques, Nelson, Filipe Meneses, and Adriano Moreira. 2012. *Combining similarity functions and majority rules for multi-building, multi-floor, WiFi positioning*: IEEE.

Marsland, Stephen. 2014. *Machine learning: an algorithmic perspective*: CRC press.

Martinez, Wendy L, and Angel R Martinez. 2007. *Computational statistics handbook with MATLAB*: CRC press.

Mirowski, Piotr, Dimitrios Milioris, Philip Whiting, and Tin Kam Ho. 2014. "Probabilistic Radio-Frequency Fingerprinting and Localization on the Run." *Bell Labs Technical Journal* 18:111-133. doi: 10.1002/bltj.21649.

Mu, Zhou, Zhang Qiao, Tian Zengshan, Qiu Feng, and Wu Qi. 2014. "Integrated location fingerprinting and physical neighborhood for WLAN probabilistic localization." Computing, Communication and Networking Technologies (ICCCNT), 2014 International Conference on, 11-13 July 2014.

Murphy, Kevin P. 2012. *Machine learning: a probabilistic perspective*: MIT press.

Park, Jun-geun, Ben Charrow, Dorothy Curtis, Jonathan Battat, Einat Minkov, Jamey Hicks, Seth Teller, and Jonathan Ledlie. 2010. "Growing an organic indoor location system." *Proceedings of the 8th international conference on Mobile systems applications and services MobiSys 10*:271. doi: 10.1145/1814433.1814461.

Petrellis, N, N Konofaos, and G P Alexiou. 2006. "Target localization utilizing the success rate in infrared pattern recognition." *Ieee Sensors Journal* 6:1355-1364. doi: Doi 10.1109/Jsen.2006.881342.

Prieto, J, S Mazuelas, A Bahillo, P Fernandez, R M Lorenzo, and E J Abril. 2012. "Adaptive Data Fusion for Wireless Localization in Harsh Environments." *IEEE Transactions on Signal Processing* 60:1585-1596. doi: 10.1109/TSP.2012.2183126.

Rappaport, Theodore S. 1996. *Wireless communications: principles and practice*. Vol. 2: prentice hall PTR New Jersey.

Razavi, Alireza, Mikko Valkama, and Elena-Simona Lohan. 2015. "K-Means Fingerprint Clustering for Low-Complexity Floor Estimation in Indoor Mobile Localization."

Redzic, Milan D., Conor Brennan, and Noel E. O'Connor. 2014. "SEAMLOC: Seamless indoor localization based on reduced number of calibration points." *IEEE Transactions on Mobile Computing* 13:1326-1337. doi: 10.1109/TMC.2013.107.

Roos, Teemu, Petri Myllyma, Henry Tirri, Pauli Misikangas, and Juha Sieva. 2002. "A Probabilistic Approach to WLAN User Location Estimation." 9.

Scott, David W. 2015. *Multivariate density estimation: theory, practice, and visualization*: John Wiley & Sons.

Seidel, Scott Y, and Theodore S Rappaport. 1992. "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings." *Antennas and Propagation, IEEE Transactions on* 40 (2):207-217.

Shih-Hau, Fang, and Lin Tsung-Nan. 2008. "Indoor Location System Based on Discriminant-Adaptive Neural Network in IEEE 802.11 Environments." *Neural Networks, IEEE Transactions on* 19 (11):1973-1978. doi: 10.1109/TNN.2008.2005494.

Snedecor, George W, and Witiiam G Cochran. 1989. "Statistical methods 8. ed." *Iowa State Univ*.

Sorour, Sameh, Yves Lostanlen, Shahrokh Valaee, and Khaqan Majeed. 2015. "Joint Indoor Localization and Radio Map Construction with Limited Deployment Load." *IEEE Transactions on Mobile Computing* 14:1031-1043. doi: 10.1109/TMC.2014.2343636.

Sun, Lin, Zengwei Zheng, Tao He, and Fei Li. 2015. "Multifloor Wi-Fi Localization System with Floor Identification." 2015.

Tao, Shuai, Mineichi Kudo, Bing-Nan Pei, Hidetoshi Nonaka, and Jun Toyama. 2014. "Multiperson Locating and Their Soft Tracking in a Binary Infrared Sensor Network." *IEEE Transactions on Human-Machine Systems* 45:1-12. doi: 10.1109/THMS.2014.2365466.

Thomas, King, Kopf Stephan, Haenselmann Thomas, Lubberger Christian, and Effelsberg Wolfgang. 2008. CRAWDAD dataset mannheim/compass (v. 2008-04-11).

Tran, D. A., and Nguyen Thinh. 2008. "Localization In Wireless Sensor Networks Based on Support Vector Machines." *Parallel and*

Distributed Systems, IEEE Transactions on 19 (7):981-994. doi: 10.1109/TPDS.2007.70800.

Wang, B, S Zhou, W Liu, and Y Mo. 2015. "Indoor Localization Based on Curve Fitting and Location Search Using Received Signal Strength." *Industrial Electronics, IEEE Transactions on* 62:572-582. doi: 10.1109/TIE.2014.2327595.

Wang, Hua-Yan, V.W. Zheng, Junhui Zhao, and Qiang Yang. 2010. *Indoor localization in multi-floor environments with reduced effort*: IEEE.

Wu, Chenshu, Zheng Yang, Yunhao Liu, and Wei Xi. 2013. "WILL: Wireless indoor localization without site survey." *IEEE Transactions on Parallel and Distributed Systems* 24:839-848. doi: 10.1109/TPDS.2012.179.

Yijun, Mo, Cao Zuo, and Wang Bang. 2012. "Occurrence-Based Fingerprint Clustering for Fast Pattern-Matching Location Determination." *Communications Letters, IEEE* 16 (12):2012-2015. doi: 10.1109/LCOMM.2012.111412.121909.

Youssef, M.a., a. Agrawala, and a. Udaya Shankar. 2003. "WLAN location determination via clustering and probability distributions." *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003. (PerCom 2003).*143-150. doi: 10.1109/PERCOM.2003.1192736.

Youssef, Moustafa. 2004. "HORUS: A WLAN-based indoor location determination system." Doctor of Philosophy, Department of Computer Science, University of Maryland.

Youssef, Moustafa A, Ashok Agrawala, A Udaya Shankar, and Sam H Noh. 2002. "A Probabilistic Clustering-Based Indoor Location Determination System Location Determination System Architecture The Joint Clustering Technique."

Zhu, Ji, and Trevor Hastie. 2005. "Kernel logistic regression and the import vector machine." *Journal of Computational and Graphical Statistics* 14 (1).