AN INTEGRATED MODEL USING IMPLICIT CONSTRAINT GENERATOR, FINGERPRINT BASED SIMULATOR AND MULTI OBJECTIVE OPTIMIZATION FOR INDOOR LOCALIZATION OF MOVING OBJECT

BY

YAZAN ALJERROUDI

A thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy Engineering

Kulliyyah of Engineering
International Islamic University Malaysia

JUNE 2022

ABSTRACT

Indoor localization is one of the most active research topics. It involves utilization of various sensing and technologies to accomplish global positioning system alternative solution for indoor localization. WiFi based indoor localization is regarded as the most promising sensing technology for non-invasive indoor localization with adequate accuracy. Typical approach of building WiFi based indoor localization is WiFi fingerprinting based on site survey and training machine learning to predict the location. This can be applied to any moving object as long as it is occupied with WiFi sensing and to any moving pedestrian accompanied with smartphone or tablets. Despite the significant research for developing WiFi based indoor localization, the literature is still yet to resolve various issues. Most importantly, multi-path and jumping behaviour, the dynamic aspect of navigation runs and topology optimization. This thesis tackles these aspects and aims at resolving them by proposing a model for WiFi based localization with an integration of various components. First, it develops a model for improving the accuracy of indoor localization based on the implicit constraints. Second, it develops an algorithm that simulates navigation behaviour in indoor navigation environment and use it for converting fingerprint to sequential navigation data. Third, it integrates implicit constraints model with online learning classifier for predicting the location based on fingerprint. Fourth, it develops a multi-objective optimization algorithm based on introducing the concept of crowding angle to optimize localization classifiers and integrate it with the localization algorithm.

The components of the model and the overall model were evaluated using state of the art approaches and benchmarks. The evaluation has included the verification of the superiority of the developed multi-objective optimization algorithm in the exploration and optimality. Furthermore, the present research evaluates the developed model with all its components generated navigation runs. Also, comparisons of its accuracy with both online sequential extreme learning machine (OSELM) and feature adaptive OSELM are conducted. Accomplished accuracy of the present model is around 95% with superiority over the benchmarks.

ملخص البحث

تحديد الموقع في البيئات الداخلية هو واحد من أهم الأبحاث الفعالة في أيامنا هذه. الهدف منه هو إيجاد بديل لتكنولوجيا نظام تحديد الموقع العالمي في البيئات الداخلية. نظام تحديد الموقع هذا بالاعتماد على شبكة الواي فاي هو ذو أداء واعد من ناحية الدقة والعملية. بينما تعتمد الطرق النموذجية على استخدام بصمة الواي فاي لتدريب الآلة لتتنبأ بالموقع، يسمح هذا البحث باستخدام الآلة المدربة في جهاز محمول للتنبؤ بالموقع وبدون أية قيود على طبيعة الجسم المتحرك ويعالج الأمور الآتية:

ظاهرة التعدد للمسارات والقفز التنبؤي بسبب عدم وجود خط رؤية في أغلب البيئات الداخلية. معالجة ظاهرة الديناмكية التي تؤثر على بصمة الواي فاي جزيئًا مع تقادم الزمن. أمثلة المتنبئ من وجهتي نظر الأولى البنية والثانية الدقة. إن نظامنا مكون من عدة بنى داخلية: الأولى هي بنية القيود الضمنية والتي تهدف إلى تقليص عدد خيارات التنبؤ باستبعاد ما يخالف القيود والثانية هي بنية توليد الملاحة الديناميكية ضمن البيئة الداخلية والتي توافق السلوك البشري بناء على تغيير بصمة الواي فاي إلى بصمة داخلية ويتم نظامنا بمكاملة البنيتين مع بنية الأمثلة ثنائية وجهات النظر. أما الاختبار فقد تم بفحص كل بنية على حده ومقارنة النظام مع آخر ما توصلت آليه الأبحاث ليصل الى 95% في تحديد الموقع في البيئات الداخلية.

APPROVAL PAGE

The thesis of Yazan Aljeroudi has been approved by the following:

Erwin Sulaeman
Supervisor

Mohamed Elsayed Okasha
Co-Supervisor

Salmiah Ahmad
Internal Examiner

Mohd Ezanee Rusli
External Examiner

Salwani Mohd Daud
External Examiner

Muhammad Ihsan Abdul Razak
Academic Secretariat

Radwan Jamal Elatrash
Chairman

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Yazan Aljeroudi

Signature Date 21/4/2022

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

An integrated model using implicit constraint generator, fingerprint based simulator and multi objective optimization for indoor localization of moving object

I declare that the copyright holders of this thesis are jointly owned by the student and IIUM.

Affirmed by Yazan Aljeroudi

	21/4/2022
Signature	Date

ACKNOWLEDGEMENTS

Firstly, it is my utmost pleasure to dedicate this work to my dear parents and my family, who granted me the gift of their unwavering belief in my ability to accomplish this goal: thank you for your support and patience.

I wish to express my appreciation and thanks to those who provided their time, effort and support for this project. To the members of my dissertation committee, thank you for sticking with me.

Finally, a special thanks to Associate Prof. Dr. Erwin Sulaeman for his continuous support, encouragement and leadership, and for that, I will be forever grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

AP	Access Point
AQCD	Angle Quantization Crowding Distance
CA	Crowding Angle
CAD-NSGA	Crowding Angle Distance Non-Dominated Sorting Genetic Algorithm
CA-NSGA	Crowding Angle Non-Dominated Sorting Genetic Algorithm
CD	Crowding Distance
CSI	Channel State Information
dBm	In Decibel Milliwatts
DGPS	Differential Global Positioning System
ELM	Extreme Learning Machine
FA-OELSM	Feature Adaptive Online Sequential Extreme Learning Machine
FBS	Fingerprint Based Simulator
GWO	Grey Wolf Optimization
ICG	Implicit Constraint Generator
ITM	Infinite Time Memory
KP-OLSM	Knowledge Preservation Extreme Learning Machine
MKT	Memory Knowledge Transfer
MOGA	Multi Objective Genetic Algorithm
MOO	Multi-Objective Optimization
mW	Milliwatts
NN	Neural Network
NPV	Negative Predictive Value
NSGA-II	Non-Dominated Sorting Genetic Optimization

PC	Probabilistic Classifier
PCA	Principle Component Analysis
PICP	Prediction Interval Coverage Probability
PINAW	Prediction Interval Normalized Average Width
RBF	Radial Basis Function
RN	Reference Nodes
RSS	Received Signal Strength
RSSI	Received Signal Strength Indicator
TNR	True Negative Rate

LIST OF FIGURES

LIST OF SYMBOLS

$H(x_i)$	The hidden output matrix
β	Weights between the hidden and output layer
$C(X_1, X_2)$	Set coverage or C metric
NDS	Number of non-dominated solution
P_s	Pareto front
TP	True positive
TN	True negative
FP	False positive
FN	False negative
ACC	The accuracy
y_i	The ground truth places
$\tilde{y}_i$	The predicted locations
d_f	The distance between the extreme solutions
d_I	The distance between the boundary solutions
d_i	The distance between the consecutive solutions
$\bar{d}$	The average of distances

CHAPTER ONE
INTRODUCTION

1.1 INTRODUCTION

The tremendous spread of smartphones, portable tablets, and other smart wireless devices in recent years has enabled the incorporation of various intelligent services and virtual personal assistants (Caporusso, Biasi et al. 2017, Voicu, Dobre et al. 2019, Yang and Lee 2019). Most intelligent services require the availability of numerous information from users. As part of its intelligent aspect, the service needs to obtain information implicitly with minimum interaction with the user using some models such as human activity recognition (Reyes-Ortiz, Oneto et al. 2016, Ronao and Cho 2016). One of the key information for any intelligence service to operate is the location of the user or the device (Huh and Kim 2019), location is subject to change with time, and many applications or services require precise identification of location (Abdullah, Rahman et al. 2018, Zhao, Huang et al. 2019). Location estimation or localization has been a research topic for decades (Horn and Schmidt 1995, Chintalapudi, Padmanabha Iyer et al. 2010). Estimating the location of the user or the device that is carried by the user requires exploiting a set of sensors that are carried by the user or embedded in the device. Typically, sensors that are embedded in daily usage devices are low cost and subject to errors. This motivates the work on developing intelligent algorithms that can handle the errors and provide a relatively accurate estimation of the location exploiting the processor that is deployed in any device and exploiting the remarkable development of intelligent and high-performance algorithms in general and artificial intelligence in particular.

Localization has been categorized under two categories: outdoor (Li, Xu et al. 2019, Nilwong, Hossain et al. 2019, Elawaad, Ezzeldin et al. 2020) and indoor (Chen, Zou et al. 2019, Zafari, Gkelias et al. 2019).). Outdoor is the less challenging category in general due to the availability of the Global Positioning System (GPS), which is considered the most common localization sensor. GPS has been considered the main solution for the majority of outdoor localization systems (Parkinson, Enge et al. 1996, Balico, Loureiro et al. 2018). This is because of the wide range of availability, the low cost of accessibility, and the adequate accuracy, which goes to as low as 10 cm for different variants of GPS or DGPS (Chew and Zakaria 2019). However, the most challenging category is indoor localization. This is due to the non-accessibility of GPS in the indoor environment and the non-availability of fully equivalent technology in the indoor environment. This has motivated researchers to put a significant amount of effort into the problem of indoor localization. Another motivation is the emerging need for deploying many intelligent services in the indoor environment that cannot operate unless a localization algorithm is capable of providing a relatively accurate estimation in an online mode.

WiFi is an available infrastructure that is almost available in all indoor environments. It is also easy to use WiFi fingerprints in order to predict people's locations after incorporating various models. In addition, WiFi sensing is available in all portable and smart devices. We add to that the absolute nature of the WiFi-based localization, which makes it not subject to drift. Such factors have made it a promising indoor localization technology and complementary sensing for GPS that works outdoor (Monica and Bergenti 2019, Zhang and Wang 2019).

This dissertation aims at developing an integrated model for the indoor localization of a moving object. The object can be a device carried by the user or by another device like a robot. Thus, estimating the location of the device implies estimating the location of the user or the moving vehicle that is carrying the device. This estimation has to be in an online mode in order to enable the functionality of the most intelligent indoor services. The remaining of this chapter is organized as follows. In section 2, we provide more information and data about the motivation of the dissertation. In section 3, we present the problem statement. Section 4 provides the objectives of the dissertation. The scope of the dissertation is provided in section 5. Finally, the outlines of the dissertation are given in section 6.

1.2 MOTIVATION

The literature on indoor localization is huge because it is an accumulated amount of research for nearly three decades. The researchers have tackled it from various perspectives: dealing with sensors errors, concentrating on the vision for the goal of localization, developing algorithms for specific applications or moving objects like robots or aerial vehicles, etc. However, less amount of work has been done on developing localization algorithms that can work on wireless portable smart devices that nearly exist in the hand of any person and can be carried or attached to any other object to enable various types of services. This fact has motivated me to conduct this dissertation on developing an integrated localization model for a moving object. Such a model, if developed, can be an essential part of a wide range of indoor intelligent services. We count some of them in this section.

1. Guide systems in airports: airports are considered one of the most crowded places, and the majority of travellers need a guiding service to assist them while navigating from one gate to another or from one terminal to another (Guerreiro, Ahmetovic et al. 2019). Unfortunately, there is no single reliable and consistent

service that serves this. In my opinion, the biggest challenge is the lacking of a reliable, real-time, and consistent indoor localization algorithm.

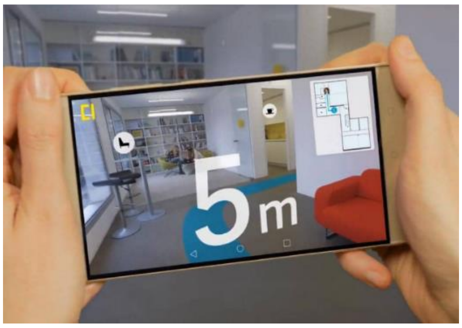
2. Rescuing systems in emergency cases: in various scenarios, when an emergency happens, like a fire inside a building or earthquake, many cases of losing people happen. Unfortunately, the primary level of current localization algorithms for the indoor environment or the non-feasibility of integrating or operating them is the obstacle in front of the availability of such services (Ferreira, Fernandes et al. 2017).
3. Management services in the industrial environment or construction resources (Won, Park et al. 2018) that has a big volume of labor. The management of labor is not an easy task. The existence of efficient services of labor requires the availability of indoor localization algorithms that has good reliability and are easy to be deployed in the industrial environment with less cost.
4. Sports assistance and training services needs also receiving information about the location of the users in real-time in order to guide them about the training practice (Umek, Tomažič et al. 2019).
5. Monitoring and surveillance services are essential in today's security applications. Indoor localization is also an essential input for such services (Kulshrestha, Saxena et al. 2019).
6. Marketing services are emerging, and companies are competing with each other in the style of marketing. Indoor localization provides a new way of marketing, like location-based marketing or promoting the product based on the action of passing by the store that has the product (Molitor, Spann et al. 2020).

7. Data gathering and analysis is becoming essential for the extraction of various types of information to the decision-makers. Indoor localization is also a part of such services.
8. Social services like notifying a person when one of his friends is nearby in the indoor environment. This also needs the availability of accurate and real-time indoor localization algorithms (Pellet, Shiaeles et al. 2019)) and social distancing applications for controlling pandemic spread (Gupta, Mehrotra et al. 2020).
9. Service of caring for elderly people requires real-time identification of their location while they are in their homes and their activities. Thus, an indoor localization algorithm is needed to enable such services (Kolakowski, Djaja-Josko et al. 2020).

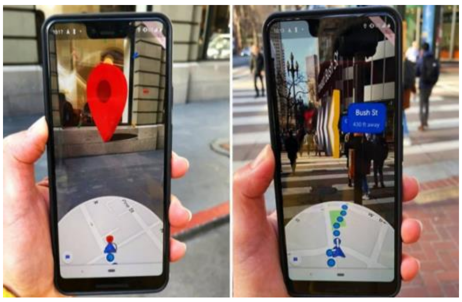
While they are in their homes and their activities. Thus, an indoor localization algorithm is needed to enable such services (Kolakowski, Djaja-Josko et al. 2020).

We counted some of the intelligent services that cannot operate unless the availability of indoor localization with certain attributes and requirements exists. Some examples of location-based services are shown in Figure 1.1.

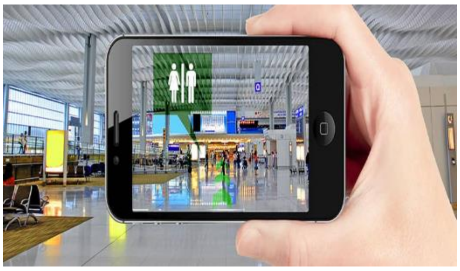
In the next section, we present the problem statement of the dissertation.

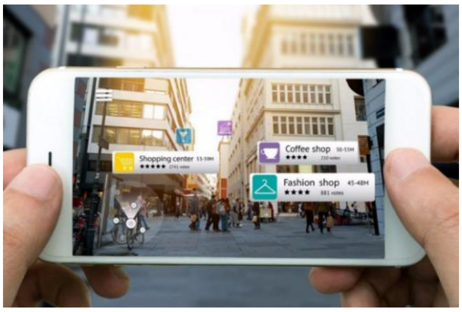


-a-



-b-





-c-　　　　　　　　　-d-

Figure 1.1 Map of a location-based services

1.3 PROBLEM STATEMENT

The literature contains numerous approaches and models for indoor localization of moving objects, e.g., (Sun, Xue et al. 2018, Zhang, Wen et al. 2018), etc. Neural network-based models have gained special attention, e.g. (Song, Fan et al. 2019), ,etc. Neural networks show a great capability in approximating the WiFi fingerprint in indoor localization and updating its approximation based on online learning, which makes it more attainable to predict location based on a WiFi sensor that exists in any portable electronic device today. Among neural network-based localization models, extreme learning machine ELM, a special type of neural network model, has recently gained interest in indoor localization (Al-Khaleefa, Ahmad et al. 2018, Zhang, Wen et al. 2018), etc. This is interpreted by the simple topology of ELM that consists of only one hidden layer. Such topology is adequate in the application of indoor localization that uses low dimensions of sensing compared with the vision that requires the deep type of models due to the high dimension of sensing data.

There are two aspects of tackling indoor localization in the literature; the first one is the model that aims at predicting the location, and the second one is the evaluation of the model.

From the model perspective, ELM models for indoor localization have been used in their typical way for location prediction (Gan, Khir et al. 2018) or it has been modified to handle various issues that arise when using WiFi for localization, such as feature adaptive model (Jiang, Liu et al. 2016), knowledge preservation (Al-Khaleefa, Ahmad et al. 2019), cyclic dynamic (Al-Khaleefa, Ahmad et al. 2019). Observing the various ELM-based developed models for WiFi localization, we see that the multi-path issue of WiFi signal in indoor localization has not been taken explicitly into consideration in ELM-based models. As a result, a jumping behaviour in the location prediction is observed. Another aspect of concern is missing an explicit consideration of the indoor localization geometry. As we know, indoor localization is characterized by straight-line segments (or corridors) connecting with rooms or other corridors. The last observation is the lacking of probabilistic prediction of the location, which degrades the performance in the case of connecting the model with other models for the overall sensing fusion localization model. Typically, the localization fusion models have to use the individual location prediction with a confidence level in order to assure better fusion performance. Thus, ELM-based localization models have to deliver location as well as confidence in the location.

From the evaluation approach, many researchers have built benchmark datasets and published them to enable other researchers to evaluate their models in an objective and comparative way. However, using a fingerprint dataset for evaluation is not adequate because the researchers, in most cases, divide the fingerprint dataset into training and testing, and they evaluate the model as a simple classifier which causes lacking evaluation of the model performance in a really dynamic way similar to actual navigation run. Building a simulator that approximates an actual navigation run and enables using fingerprint information in a time series way is crucial for a fair evaluation. In the work of (Al-Khaleefa, Ahmad et al. 2018), some efforts were made for building such a simulator; however, the simulator is restricted to cyclic dynamic behaviour, which is only one type of many navigation patterns in an indoor environment.

Another perspective of improving the performance is optimizing the random weights in the input-hidden layer as well as optimizing the topology of the classifier from the perspective of number of neurons. The two perspectives have their effects on two aspects of performance, namely, accuracy and time. Hence, the optimization can be seen as multi-objective optimization. In the literature, multi-objective optimization concerns generating a set of non-dominated solutions named as Pareto front. The typical way of performing multi-objective optimization is based on adopting various criteria of exploration, such as crowding distance (Deb, Agrawal et al. 2000), and angular quantization of solution space (Metiaf, Wu et al. 2019). The crowding distance concerns prioritizing furthest solutions, while angular quantization of solution space concerns prioritizing solutions in angular sectors that were not explored enough. The issue with angular quantization of solution space is its global view. This might lead to slow or non-convergence due to consuming the searching in sectors even if they do not contain parts of the true Pareto. Another alternative way is to replace it with local criteria of angle awareness.