Task-Driven Integrity Assessment and Control for Vehicular Hybrid Localization Systems

by

Nabil Drawil

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2012

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Throughout the last decade, vehicle localization has been attracting significant attention in a wide range of applications, including Navigation Systems, Road Tolling, Smart Parking, and Collision Avoidance. To deliver on their requirements, these applications need specific localization accuracy. However, current localization techniques lack the required accuracy, especially for mission critical applications. Although various approaches for improving localization accuracy have been reported in the literature, there is still a need for more efficient and more effective measures that can ascribe some level of accuracy to the localization process. These measures will enable localization systems to manage the localization process and resources so as to achieve the highest accuracy possible, and to mitigate the impact of inadequate accuracy on the target application.

In this thesis, a framework for fusing different localization techniques is introduced in order to estimate the location of a vehicle along with location integrity assessment that captures the impact of the measurement conditions on the localization quality. Knowledge about estimate integrity allows the system to plan the use of its localization resources so as to match the target accuracy of the application. The framework introduced provides the tools that would allow for modeling the impact of the operation conditions on estimate accuracy and integrity, as such it enables more robust system performance in three steps.

First, localization system parameters are utilized to contrive a feature space that constitutes probable accuracy classes. Due to the strong overlap among accuracy classes in the feature space, a hierarchical classification strategy is developed to address the class ambiguity problem via the class unfolding approach (HCCU). HCCU strategy is proven to be superior with respect to other hierarchical configuration. Furthermore, a Context Based Accuracy Classification (CBAC) algorithm is introduced to enhance the performance of the

classification process. In this algorithm, knowledge about the surrounding environment is utilized to optimize classification performance as a function of the observation conditions.

Second, a task-driven integrity (TDI) model is developed to enable the applications modules to be aware of the trust level of the localization output. Typically, this trust level functions in the measurement conditions; therefore, the TDI model monitors specific parameter(s) in the localization technique and, accordingly, infers the impact of the change in the environmental conditions on the quality of the localization process. A generalized TDI solution is also introduced to handle the cases where sufficient information about the sensing parameters is unavailable.

Finally, the produce of the employed localization techniques (i.e., location estimates, accuracy, and integrity level assessment) needs to be fused. Nevertheless, these techniques are hybrid and their pieces of information are conflicting in many situations. Therefore, a novel evidence structure model called Spatial Evidence Structure Model (SESM) is developed and used in constructing a frame of discernment comprising discretized spatial data. SESM-based fusion paradigms are capable of performing a fusion process using the information provided by the techniques employed. Both the location estimate accuracy and aggregated integrity resultant from the fusion process demonstrate superiority over the employing localization techniques. Furthermore, a context aware task-driven resource allocation mechanism is developed to manage the fusion process. The main objective of this mechanism is to optimize the usage of system resources and achieve a task-driven performance.

Extensive experimental work is conducted on real-life and simulated data to validate models developed in this thesis. It is evident from the experimental results that task-driven integrity assessment and control is applicable and effective on hybrid localization systems.

Acknowledgements

All praise is due to Allah for granting me the strength and knowledge to complete this work. I would like to express my deep and sincere gratitude and appreciation to my supervisor professor Otman Basir for his continuous encouragement, helpful discussion, and insightful guidance throughout the research presented in this thesis.

I would like to express my sincere gratitude to the members of my doctoral committee, Dr. Safieddin Safavi-Naeini, Dr. Steven Waslander, Dr. Dana Kulic, and Dr.Abdulmotaleb El Saddik for investing time in reading and providing many valuable comments on my thesis. Thanks are due to Margaret Frigula for her proofreading.

I would also like to thank my colleagues in the Pattern Analysis and Machine Intelligence (PAMI) research laboratory who created an excellent environment for the research work. Special thanks go to Akrem El-Gazal, Khaled Almahrough, Sayyid Anas Vaqar, Abduljalil Netfa, Andrew Rae, Hanan Saleet, Tarek Khalefa, Patrick Tsui, Moataz El Ayadi, Moataz Anazy, Haitham Amar, and Mostafa Faroque for their fruitful discussions and support.

The Libyan Ministry of Education is highly appreciated for sponsoring me during my study. I am also grateful to the University of Waterloo Faculty of Engineering, and Ontario Ministry of Training, Colleges and Universities for providing me with Engineering Graduate Scholarships and an Ontario Graduate Scholarship.

Thanks are due to Intelligent Mechatronic Systems (IMS) Inc. for providing the real-life measurements. I also would like to thank the Geospatial Centre at University of Waterloo for providing the Ontario digital map.

Last and by far not least, I am indebted to my parents, brothers, and sisters for their continuous support, patience, and prayers. I owe a great deal of gratitude to my wife and my children for their continuous understanding, unlimited encouragement, and unending love during the years of my study.

Dedication

To my mother and father,

my wife and my children

Fatima, Moaath, and Hajer ...

Table of Contents

List of Tables

List of Figures

Nomenclature

α	Time Window Memory Factor
δ_{eph}	Satellite Ephemeris Error
δ_{iono}	Satellite Ionospherical Error
δ_{mp}	Satellite Multipath Errors
δ_{trop}	Satellite Tropospherical Errors
$\epsilon(\hbar)$	Error Rate of Classifier $\hbar$
$\hbar$	Hierarchical Multi-Class Classifier
$\mathbb{T}$	GPS Receiver Clock Error
$\mu_{_{SNR}}$	Mean of SNR
$\overline{\pi}$	Limiting Probability
Ψ	Set of Localization Techniques
ρ_{code}	Satellite Pseudorange Code
ρ_t	IVCAL-TDI Measure at Time t

$\sigma_{_{SNR}}$	Standard Deviation of SNR
$\mathbf{P}$	Transition Probabilities Matrix
Θ	Frame of Discernment
υ	IVCAL Instant Integrity Estimate
Ξ_s	Target Localization QoS
ζ_{ω}^{q}	Similarity between Class q and class ω
Acc	Localization Accuracy Level
C	Set of Class Labels
h	Flat Classifier
P	Set of Patterns
p_k	Sample Pattern
P_{ij}	Transition Probability from S_i to S_j
r	Range from A Satellite to A Receiver
s	Speed of Light
T_i	Localization Technique i
v_{rcvr}	GPS Receiver Measurement noise Error
X_t	True location of a vehicle at time t
$Z_{i,t}$	Location Estimate by T_i at time t

A-GPS Assisted GPS

ANN Artificial Neural Network

AOA Angle of Arrival

BS Base Station

CBAC Context Based Accuracy Classification

CS Channel Selection

DGPS Differential GPS

DOP Dilution of Precision

DRS Dead Reckoning System

DSET Dempster-Shafer Evidence Theory

eCall Emergency Call

EFM Estimate Fusion and Management

EGNOS European Geostationary Navigation Overlay Service

EKF Extended Kalman Filter

FIS Fuzzy Inferencing System

FP Fingerprinting

GDOP Geometric Dilution of Precision

GIS Geographic Information System

GNSS Global Navigation Satellite Systems

GPS Global Positioning System

HCCU Hierarchical Classifier Using Class Unfolding

IMU Integrity Monitoring Unit

INS Inertial Navigation System

ITS Intelligent Transportation Systems

IVC Inter-Vehicle Communication

IVCAL Inter-Vehicle Communication Assisted Localization

KF Kalman Filter

LBS Location-Based Service

LEU Localization Enhancement Unit

LOS Line of Sight

MDU Multipath Detection Unit

MS Mobile Station

NLOS Non-Line of Sight

PCA Principle Component Analysis

PLU Primary Localization Unit

QoS Quality of Service

RSS Received Signal Strength

SNR Signal to Noise Ratio

SVC Support Vector Classification

TDI Task-Driven Integrity

TDOA Time Difference of Arrival

TOA Time of Arrival

UERE User Equivalent Range Error

VANET Vehicular Ad hoc Networks

WAAS Wide Area Augmentation System

WSN Wireless Sensor Networks

Chapter 1

Introduction

Over the last ten years, vehicle localization has been attracting attention in a wide range of applications. As reported in [2–17], a number of localization techniques have been developed to serve a variety of applications. Recently, the need for pinpointing a vehicle's location has become more important and in many cases a matter of safety; for instance, for emergency response systems, such as the eCall system, to deliver on their task they need reliable and accurate location estimates. Such expectation is now shared by other applications, including, accident avoidance and management, navigation systems, location sensitive billing systems, location based services.

In this chapter, a brief overview about vehicle localization techniques is presented. The motivation behind this work and its objectives are also provided. The chapter is concluded with an outline of the thesis.

1.1 An Overview of Vehicle Localization

Since Global Positioning System (GPS) is universally available and vehicles have enough power to operate its receivers, it has become a *de facto* standard for vehicle localization.

GPS was developed and operated by the U.S. Department of Defense [18]. GPS is based on a network of satellites that transmit continuous coded information that makes it possible to identify locations on Earth by measuring the distances from the satellites, and the GPS receiver who also has the ability to obtain information about its own velocity and direction. A GPS network consists of 24 satellites arranged in six orbital planes, as depicted in Figure 1.1, so that at any given time, a minimum of five satellites can be observed by a GPS receiver at any location in the world. According to the accuracy required, different types of GPS receivers have been developed.

Figure 1.1: GPS satellites arranged in six orbital planes.

Basic GPS receivers often have multiple radio channels so that the receiver can observe multiple GPS satellites at once and obtain a pseudo-range measurement from each satellite signal. Leva [19] and Hoshen [20] show two different techniques by means of which a GPS receiver can compute its location from four pseudo-range measurements, the minimum number of localization measurements required in three dimensions. However, if the signal from one of the GPS satellites is obstructed, it is difficult to identify the location of the

GPS receiver from only three measurements. Therefore, more advanced receivers with six or more radio channels have been developed. The extra channels provide a degree of redundancy that can be accessed in case one or more of the four required signals are lost. While this increase in the complexity of the hardware in the GPS receiver adds to the cost, it does not necessarily guarantee accurate location measurements. The pseudo-range measurement is still prone to many errors and hence has a typical accuracy range of 10 to 50 meters.

GPS errors can be categorized as either global or local. Global errors are produced when satellite signals originate and travel to the Earth, whereas local errors are dependent on the environment surrounding the receiver. The extent of the effect of global errors on receiver measurements varies from one geographical area to another due to ionospheric delays, tropospheric delays, ephemeris, and satellite clocks. Local errors may arise as a result of multipath effects, due to lack of a Line of Sight (LOS) or to the receiver hardware itself. More details about global and local GPS measurement errors are given in [21–25].

Most global errors can nevertheless be avoided, and accurate location measurements can be obtained using a Differential GPS receiver (DGPS). A DGPS consists of two receivers that observe the same GPS satellites. One of these receivers is stationary and the other, which is used to measure the locations, is roving. The stationary receiver resides at a known location and obtains the pseudo-range from the satellites' signals so that it can identify a global error by comparing the measurements with its location. It then transmits the global error correction to the roving receiver so that the roving one can correct its measurements. In the best cases, accuracy increases to the level of tens of centimeters when a DGPS is used. However, this type of receiver has some drawbacks, such as the cost of the communication channels between the stationary receiver and the roving receivers and the cost of the hardware. Moreover, DGPS receivers must be covered by the same GPS

satellites in order to be correlated (i.e., to have the same global error). This requirement thus ties up the roving receiver and allows it to move only in a bounded area. A significant disadvantage is that the DGPS can not correct local errors since multipath effects happen immediately around the roving receiver, and hardware errors are individual.

Another method for achieving improved localization is by means of a Dead Reckoning System (DRS), which has been adopted in some applications. With this method, the location estimation is computed based on how far the vehicle has moved from a known position given the directions and distances traveled over small periods of time. Since this method is simple and inexpensive, it has been utilized in a wide range of applications. A crucial disadvantage of this method is that errors in the direction and/or distance measurements affect the final location estimation. Furthermore, since measurement errors accumulate over a period of time, this method is recommended for use only over short periods of time.

A number of localization techniques that do not rely on GPS signals have been investigated [26, 27]. These techniques rely instead on an estimation of the distance between two objects (e.g., a sender and a receiver in a wireless network), a method inspired by the localization approaches used in cellular networks [28–31]. These approaches are based on radio-location methods, such as Time of Arrival (TOA), Angle of Arrival (AOA), Received Signal Strength (RSS), or Time Difference of Arrival (TDOA) [32]. Once the distances between an object and other reference points are obtained, trilateration can be performed in order to estimate the relative location of the object. However, radio-location methods still suffer from errors such as the multipath effect or Non-Line of Sight (NLOS), interference and hardware measurement errors.

1.2 Research Motivation

Vehicle location information is vital in many applications, including emergency response, location-based services, navigation, and assisted driving. Such emerging systems can not operate at the level required without reliable, accurate localization.

The eCall project, which facilitates automated emergency calls from cars in Europe, is one example of developing emergency systems. With this system, vehicles perform an automated call to the nearest emergency centre if a collision happens anywhere in Europe. Collision information, such as airbag deployment, measurements from impact sensors, and the vehicle's GPS coordinates, can be delivered through the automated call so that assistance can reach the victims. Nevertheless, the crucial piece of information, which is the exact location of the vehicle, is often not accurate when a GPS receiver is used in urban canyon environments [33].

Location-Based Services (LBS) are another example of emerging systems that rely mainly on vehicles' location information. According to [34, 35], in such systems, when the driver of a vehicle wishes to request a specific service from road-service providers, such as gas prices or restaurant menus, he or she can send a query over a Vehicular Ad hoc Network (VANET) to the provider's server. However, it is not possible to provide the desired services without knowledge of the accurate location of the vehicle requesting the service.

Despite the current advancements in localization techniques, more robust and widely available location information is required due to major developments that are emerging in Intelligent Transportation Systems (ITS), such as

- vehicle collision warning

- security distance warning
- cooperative cruise control
- automatic parking
- road pricing

1.3 Research Objectives

Of the vehicle localization techniques proposed in the literature, the most prevalent are those based on radio signals, such as GPS, DGPS, TOA, TDOA, and AOA. These methods are based on physical measurements related to the radio signal that travels from a sender to a receiver. Nevertheless, all these methods are extremely sensitive to multipath signals, which are produced in multipath environments where many high rise buildings and/or trees are present, such as in urban areas. In such environments, the sender radio-signals are reflected and/or diffracted before they reach the receiver. Vehicle localization in urban canyon environments is thus a rather challenging problem.

Fusion of more than one localization technique to achieve higher and more reliable localization accuracy has been reported in the literature. Therefore, incorporating independent sources of location information in the fusion process has been proposed by many researchers. For example, the Inter-Vehicle Communication Assisted Localization (IVCAL) incorporates a GPS measurement, an Inertial Navigation System (INS) measurement, and a relative distance measurement between the vehicle and a selected set of vehicles in its vicinity [13]. IVCAL is able to greatly diminish the multipath effect and other errors caused by the GPS receiver and the INS measurements.

Fusion, however, is itself a challenging issue. As Basir mentioned in [36], "A sensory system which uses more than one sensor to perform its sensory tasks is constantly confronted with two important questions. Firstly, how it should combine the information gathered by the different sensors, so as to obtain the finest possible sensory information? Secondly, how it should distribute the workload associated with the sensory tasks among its sensors so that the best performance in terms of speed and quality is achieved?" Along this view on fusion a reliable hybrid vehicle localization system is sought as the main objective which will be pursued as follows:

1. Conduct an extensive survey of the localization and tracking techniques that have been reported in the literature.

2. Classify the localization techniques based on the location information sources.

3. Identify performance criteria so that the outstanding issues of vehicle localization can be addressed in a practical context.

4. Develop a confidence measure that captures the integrity of the vehicle location estimate produced by localization techniques.

5. Study a variety of fusion techniques, especially those used at the decision level, such as ones based on the Bayesian, artificial intelligence, and evidential theories.

6. Design and develop a fusion technique that will integrate the different location estimates obtained by the different localization techniques. This fusion technique should be resilient to the stochastic behavior of the availability and uncertainty of the estimates produced by these techniques. Consequently, this fusion technique will be able to produce more accurate location estimates than those of the individual techniques.

7. Develop a framework that will provide vehicle location information with specific Quality of Service (QoS) that is required by a certain application or task. This framework is aimed to maintain the system aggregated integrity using the above developed components.

1.4 Thesis Outline

The thesis comprised of seven chapters:

Chapter 1 provides a brief overview on vehicle localization and its applications. The motivation behind this research work and its objectives are also introduced.

Chapter 2 contains background and a literature review that present a variety of fusion techniques in the context of the vehicle localization problem, in addition to highlighting important research gaps.

Chapter 3 presents a high level concept of a novel framework for fusing different localization techniques while maintaining a target level of quality set by a task or application.

Chapter 4 presents the development of two novel classification algorithms that are capable of determining the localization accuracy using information pertained to measurement conditions. The efficacy of the classification algorithms is shown in two case studies.

Chapter 5 discusses the development of an approach for capturing the impact of observation conditions on the localization performance and to consequently determine an integrity index with respect to the localization accuracy claimed by a technique. Experimental results for another two case studies are also provided.

Chapter 6 introduces the concept of location estimates meta-fusion scheme in light of the integrity assessment and aggregation of these estimates. A new evidence structure is developed in this chapter to handle the fusion of localization quality indexes. Furthermore,

task-driven resource allocation is introduces to the fusion process.

Chapter 7 contains conclusions derived from the research work and presents suggestions for future work.

Chapter 2

Background and Literature Review

2.1 Introduction

The focus of much recent research in localization systems has been to increase accuracy through the use of multiple localization modalities. This chapter provides a literature review of multi-modality based localization techniques and establishes a categorization of such techniques based on type of measurement and the strategy employed to fuse measurement from the various localization sources.

GPS as a source of localization measurements and the Kalman Filter (KF) as a tool for fusing measurements from multiple sources have been quite predominant in a large body of reported localization research work. A key feature that makes the Kalman Filter an attractive fusion tool is its ability to produce a minimum square error for linear measurements under zero mean Gaussian noise contamination [37].

2.2 Motion and GPS Measurement Data Fusion

DGPS and Assisted GPS (A-GPS) are two advanced types of GPS that provide a high level of accuracy and fast retrieving rate. Nevertheless, using a GPS receiver as the sole vehicle localization measurement source may turn to be unreliable, especially in urban canyons and other areas where the satellite signal can be distorted or lost. A number of solutions have been reported in the literature that proposed augmenting GPS measurements with information about the vehicle's motion in order to improve localization accuracy. In what follows I provide a summary of a number of such solutions.

2.2.1 DRS and GPS Integration

A DRS is a localization technique that estimates the next location of a mobile object over a series of short time intervals, given the object's direction, speed, and previous location. DRS is simple and known for producing incremental error and hence needs to be reset periodically. It is therefore suitable for use over short periods of time.

One approach to resetting the cumulative localization error is to combine DRS with GPS whereby GPS measurements are used to reduce the DRS cumulative error; when the GPS measurement is unavailable, the DRS estimates the location using sensors such as wheel odometers, a flux-gate compass, a gyroscope, and an accelerometer [38].

2.2.2 INS and GPS Fusion

INS employs a computing unit and motion sensors to estimate its location without relying on any external reference once it is initialized using for example a GPS measurement. To avoid the accumulated error caused by the measurements of internal sensors in INS, the INS

location estimate is fused with measurement data from other sources. As discussed in [39] fusing INS and GPS can take the form of a loosely or tightly coupled system architecture.

An example of a system that fuses INS and GPS is the real-time kinematic global positioning system which uses an Extended Kalman Filter (EKF) to fuse data [40]. In this system, GPS latency is defined as the time required for the satellite signals to travel to Earth and the time required for the computation of the location; GPS latency varies with the number of observed satellites. Therefore, the GPS latency is encapsulated in the EKF state so that the fusion of the INS and GPS data is synchronized with the readings of the sensors.

It is possible to fuse ordinary GPS INS by means of a KF [41]. In this case the computational complexity of the EKF can be reduced by preprocessing the INS measurements and inputting them into the KF as a linear component. However, preprocessing the INS measurement adds to the computational cost of the solution.

2.2.3 Other Motion Sensors and DGPS Fusion

Integrating the INS of a dynamic model with a differential GPS (DGPS) has also been investigated [42] as opposed to the kinematic models typically used in the literature. Due to the nonlinearity of the dynamic model, an EKF is used. In a regular INS, the motion sensors consist of an accelerometer and yaw rate. Due to the accelerometer noise and other motion sensors, such as six wheel-speed encoders, a steering angle encoder, and an optical yaw rate gyro, are used instead [42]. This location estimator maintains a localization accuracy of as low as 0.9 m for up to 100 m when the vehicle moves in a straight line and the system relies more on the dynamic model than on the GPS measurements. However, the localization errors become larger along curves as a result of the slow sampling frequency of

the GPS and the non-linearity approximation of the EKF. Furthermore, multipath effects are not addressed.

In [4] a method of positioning a vehicle on undulating ground by fusing DGPS data and motion sensor data is proposed. A fibre optic gyro, a roll pitch sensor, and wheel encoders are used as motion sensors. The positioning accuracy is improved by compensating for the error for each sensor. The error is determined by means of a KF, which is also utilized as a fusion unit.

In [43] an Artificial Neural Network (ANN) is chosen as a tool for detecting errors and noises in INS measurements using a DGPS as a guide to the true location of the vehicle during a training phase. The work reported in [43] is similar to that reported in [40] in that preprocessing operations are performed on the measurements before they are fused. An assumption that is made in this method is that the DGPS data is always either available or unavailable due to an outage in satellite signal. However, in urban areas, satellite signals are often available but quite often are contaminated by multipath noises, which effects the quality of the ANN learning.

2.3 Fusion of Landmark, DRS, and GPS Measurements

The detection and and recognition of landmarks provide spatial information related to the local environment. It is therefore possible to integrate spatial information with localization measurements from DRS and GPS in order to improve localization accuracy [11, 44–48]. Two approaches for detecting and augmenting landmarks for vehicle localization systems are presented next along with a localization technique that attempts to detect visible satellites (i.e., satellites with LOS) for use in the positioning process.

2.3.1 Laser Scanners, Digital Maps, and GPS/DRS

Due to the accumulated error caused by the long satellite outages in GPS/DRS localization systems, digital maps are utilized to perform localization during such outages [46]. A laser scanner mounted on a vehicle scans major objects in the vehicle's environment. The system matches these landmarks with other landmarks in the digital map that represent the region of interest. If there is a match, the vehicle location is estimated by correlating the identified landmarks.

However, segmentation is not a trivial job, especially in situations where landmarks are merged with background objects. Moreover, the system must be trained by having it traverse the regions of interest [47] to extract landmarks (features, such as traffic signs and the posts of traffic lights) that can later be used as reference points.

In [45], a vehicle equipped with an autonomous navigation system and a laser scanner is reported. The laser scanner is used to detect the edges of sidewalks and estimate the distance between the edge of the sidewalk and the vehicle. Distance measurements are utilized to improve the accuracy of a localization system that comprises GPS, DRS, and Geographic Information System (GIS). The GIS data contains digitized information such as abstract road maps, road edges, and other landmarks. Landmark information is created through a learning stage. During the testing stage, the EKF fusion technique produces an innovation value from which the system determines whether to accept the fusion location estimate. If the GPS data is corrupted by multipath signals or is unavailable, only the DRS location estimate is utilized. The vehicle location estimate is used to select the region of interest from the GIS database that contains the landmark information. To improve the vehicle location estimate, a matching scheme is performed to compare the GIS-extracted landmarks (i.e., sidewalk edges) with those extracted by the laser scanner,

and the estimated distances between the sidewalk edge and the vehicle are then used in fixing the vehicle location. Although the memory constraints are overcome by using the GIS, the accuracy of the estimate of the distances is not consistent due to occluding objects between the laser scanner and the edge of the sidewalk. The training phase required for any traversed region is also not insignificant.

2.3.2 Vision, Digital Maps, and GPS/DRS

Visual data is also utilized in localization techniques since digital images can provide a wide range of information about the surrounding environment. Due to the time required for image processing [11], only key images are maintained and linked to the GIS database [45]. Again, both GPS/DRS are used and the proximity of the vehicle location estimate to the roads in the GIS database is examined. The road segment closest to the location estimate is then selected, and key images of that road are extracted in order to compare their features with the features of the images taken during the navigation stage. The weakness of this strategy appears when the curvature of the vehicle's path is significant, especially when the vehicle turns in orthogonal intersections.

Visual features can, however, be blended with other location measurements, such as GPS and DRS data in the EKF formulation [48]. The main advantage of this strategy is that the uncertainty of all the information sources is kept local to the EKF, namely, in the error covariance matrix, which guarantees a minimum mean square error estimate. In [48], the EKF structure is derived and validated where the curvature of the roads is employed as a visual feature. It is shown that when the roads are curvy, the vehicle location estimate is dramatically improved. On the other hand, if the road traversed is not curved, then the accuracy of the location estimate remains the same as that produced by the GPS/DRS fusion localization technique.

2.3.3 Satellite Visibility and DGPS

In urban areas, GPS multipath signals cause unpredictable localization errors due to the NLOS satellite signals. An approach to mitigate the effect of NLOS signals is the localization system driven by tracking visible GPS satellites using an infrared camera. An omni-directional infrared camera mounted on the top of a vehicle is used to recognize obstacles and their height and to detect visible satellites by observing their positions with a satellite orbit simulator [49]. This method allows the system to exclude any radio waves emitted by invisible satellites to improve the localization accuracy.

The vehicle localization system used in this approach has high degree of accuracy since it employs a DGPS receiver. However, in high rise building areas, the availability of location estimates is low due to the lack of enough visible satellites, and even with enough visible satellites, the geometric configuration of the constellation may result in a high Dilution of Precision (DOP).

2.4 Cooperative Localization

Cooperative Localization is a recent location estimation approach that has been implemented in vehicular positioning and wireless communication systems. This localization scheme is suitable for scenarios which involve the coexistence of several entities that independently provide location information. The goal is to localize a mobile node or to enhance its location estimate given that it shares relative spatial information with nearby nodes (e.g., other vehicles or mobile network towers).

2.4.1 Radio Signal Measurement Data Fusion

Radio localization methods have been studied extensively for cellular networks in a wide range of applications (e.g., for CDMA networks see [28–32, 50–55] and for GSM networks see [56]). An example of these systems is a localization system that estimates the locations of emergency calls initiated by cellular phones. The system operates on the principle that measurements from different Base Stations (BS's) are combined in order to compute the location of a Mobile Station (MS). The base stations typically have different levels of uncertainty in their measurements, which are minimized as a result of the fusion process. The relative spatial information in this system is based on the measurements from radio signals, such as RSS, TOA, TDOA, and AOA. In some of these GPS-less approaches, a mix of two or more different types of radio signal measurements is utilized in order to relax constraints such as the synchronization of the BS's.

In the following subsections detailed models for some of these techniques are given. (x_m, y_m) signifies the MS location. The locations of n base stations: (BS_1, BS_2, $\mathrm{BS}_3, \ldots,$ BS_n) are denoted by $\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_n, y_n)\}$, respectively. For simplicity and without loss of generality, locations are represented by two coordinates, x and y, in the Cartesian coordinate system.

TOA Data Fusion

Time of arrival measurements are based on the time of flight of a signal as it travels between a source and a destination. Since the signal travels at the speed of light (c), it is possible to compute the distance between the two points as follows:

$$d_i = (t_i - t_m)c \tag{2.1}$$

where t_m signifies the signal sending time from the MS, t_i signifies the signal arrival time at the BS$_i$, and i signifies the BS's index (i.e., $i = \{1, 2, \ldots, n\}$).

According to [32], the TOA technique can be employed using three BS's, the minimum number of reference points in two dimensions (Figure 2.1), in order to estimate the MS location by computing the distances between each BS and the MS (i.e., d_1, d_2, d_3), as per Equation 2.1, and then formulating the following optimization problem:

$$\hat{x}_m, \hat{y}_m = arg \min_{x_m, y_m} \sum_{i=1}^{3} \left(d_i - \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2} \right)^2 \tag{2.2}$$

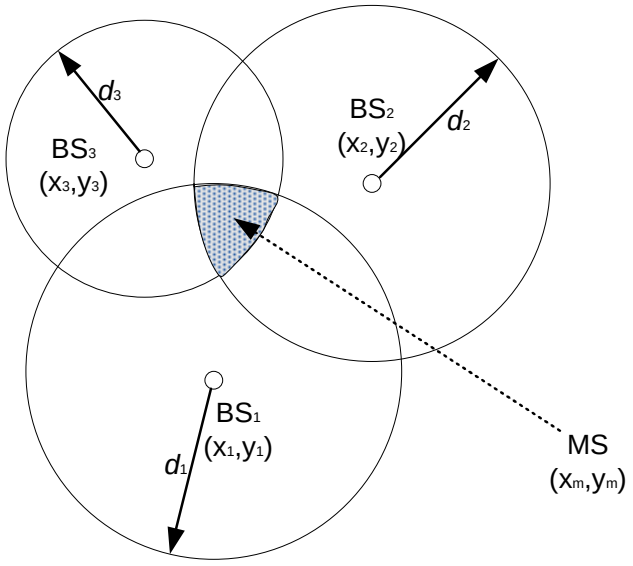


Figure 2.1: The TOA localization method.

Nevertheless, due to possible NLOS propagation conditions, the actual Euclidean dis-

tances between the MS and the BS_i is less than or equal to $(t_i - t_m)c$. This inequality creates more than one solution for the optimization problem in 2.2, all of which reside in a bounded area, as shown in Figure 2.1. A constrained version of the optimization problem in 2.2 is proposed in [28, 57] in order to increase the localization accuracy; however, the geometric arrangement of the BS's may produce poor location estimates due to the shape of the bounded area that contains the MS. This shortcoming might be avoided using more BS's. The next method, described below, utilizes more than three BS's in estimating the MS location so that ambiguity in the distance computation is reduced.

In [50,53] the Cartesian coordinate system is represented as follows. The location of one of the base stations is assumed to be the origin (e.g., BS_1 be the origin: $(x_1, y_1) = (0, 0)$) and the locations of the other objects in the network are computed with respect to the origin. Hence, the distances $(d_1, d_2, d_3 \ldots, d_n)$ can be used to estimate the location of the MS by solving the following set of equations:

$$\begin{aligned} d_1^2 &= x_m^2 + y_m^2 \\ d_2^2 &= (x_2 - x_m)^2 + (y_2 - y_m)^2 \\ d_3^2 &= (x_3 - x_m)^2 + (y_3 - y_m)^2 \\ &\vdots \\ d_n^2 &= (x_n - x_m)^2 + (y_n - y_m)^2 \end{aligned} \tag{2.3}$$

After rearranging terms, the above equations can be written as follows:

$$\begin{bmatrix} x_2 & y_2 \\ x_3 & y_3 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} \begin{bmatrix} x_m \\ y_m \end{bmatrix} = \frac{1}{2} \begin{bmatrix} k_2^2 - d_2^2 + d_1^2 \\ k_3^2 - d_3^2 + d_1^2 \\ \vdots \\ k_n^2 - d_n^2 + d_1^2 \end{bmatrix} \tag{2.4}$$

where $k_i^2 = x_i^2 + y_i^2$. Equation 2.4 can be expressed in a matrix form

$$\mathbf{Hx} = \mathbf{b} \tag{2.5}$$

where $\mathbf{H} = \begin{bmatrix} x_2 & y_2 \\ x_3 & y_3 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_m \\ y_m \end{bmatrix}$, and $\mathbf{b} = \frac{1}{2} \begin{bmatrix} k_2^2 - d_2^2 + d_1^2 \\ k_3^2 - d_3^2 + d_1^2 \\ \vdots \\ k_n^2 - d_n^2 + d_1^2 \end{bmatrix}$.

Equation 2.5 represents an overdetermined system (i.e., $n > 2$). Practically, such a system has no exact solution. Therefore a linear least squares method is used to estimate the location of the MS as follows:

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{b} \tag{2.6}$$

where $(.)^T$ signifies matrix transpose and $(.)^{-1}$ signifies matrix inverse.

Alternative techniques, such as the maximum likelihood are reported in [54,55]. In [58], terrestrial time of arrival (TOA) measurements are fused with the GPS localization using a weighted least square estimator to improve the final localization accuracy.

TDOA Data Fusion

TDOA is preferable to the TOA due to the fact that TDOA does not require synchronization between the MS and BS's, Figure 2.2. Instead, it takes advantage of the synchronization of the CDMA cellular network BS's to compute the difference between the time of arrivals of the MS signal at the BS_i and BS_1, where $i \in \{2, 3, \ldots, n\}$. The difference in the distance is therefore defined as follows:

$$\begin{aligned} d_{i1} &\equiv d_i - d_1 \\ &= (t_i - t_m)c - (t_1 - t_m)c \\ &= (t_i - t_1)c \end{aligned} \tag{2.7}$$

It can be seen that the difference is not affected by errors in the MS clock time t_m. Substituted Equation 2.7 in Equation 2.3, and then expanding and rearranging the terms

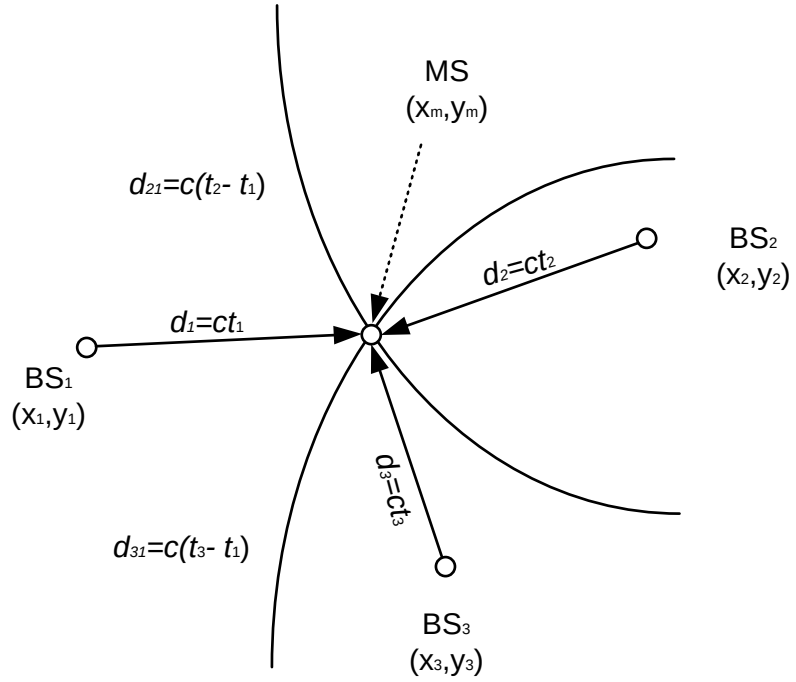


Figure 2.2: The TDOA localization method.

produce the following set of equations:

$$\begin{bmatrix} x_2 & y_2 \\ x_3 & y_3 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix} \begin{bmatrix} x_m \\ y_m \end{bmatrix} = d_1 \begin{bmatrix} -d_{21} \\ -d_{31} \\ \vdots \\ -d_{n1} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} k_2^2 - d_{21}^2 \\ k_3^2 - d_{31}^2 \\ \vdots \\ k_n^2 - d_{n1}^2 \end{bmatrix} \tag{2.8}$$

which can be expressed in a matrix form as follows:

$$\mathbf{Hx} = d_1\mathbf{c} + \mathbf{r} \tag{2.9}$$

where $\mathbf{H} = \begin{bmatrix} x_2 & y_2 \\ x_3 & y_3 \\ \vdots & \vdots \\ x_n & y_n \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} -d_{21} \\ -d_{31} \\ \vdots \\ -d_{n1} \end{bmatrix}$, and $\mathbf{r} = \frac{1}{2} \begin{bmatrix} k_2^2 - d_{21}^2 \\ k_3^2 - d_{31}^2 \\ \vdots \\ k_n^2 - d_{n1}^2 \end{bmatrix}$.

Similarly, Equation 2.9 can be solved using the following linear least squares formulation:

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1} \mathbf{H}^T \left(d_1\mathbf{c} + \mathbf{r}\right) \tag{2.10}$$

The solution of Equation 2.10 is determined in two steps. First, the estimate of the MS is determined in terms of d_1, which is substituted in the quadratic expression $d_1^2 = x_m^2 + y_m^2$ to compute d_1. Second, the value of d_1 is substituted back in Equation 2.10 to solve for $\hat{\mathbf{x}}$ [50].

AOA Data Fusion

AOA techniques estimate the location of an MS by measuring the angle of signal arrival from the MS at several BS's by means of an antenna array. The MS location is then estimated through the intersection of the straight paths leaving from at least two BS's, as depicted in Figure 2.3. However, combining only two AOA measurements may introduce a large amount of uncertainty with respect to the MS location estimate, especially when the MS is close to the line connecting the two BS's. Moreover, this localization method requires the MS to be in LOS with the participating BS's, since reflected or diffracted signals result in misleading information. For this reason, it is preferable for the AOA to be combined with another localization method, such as TOA or TDOA.

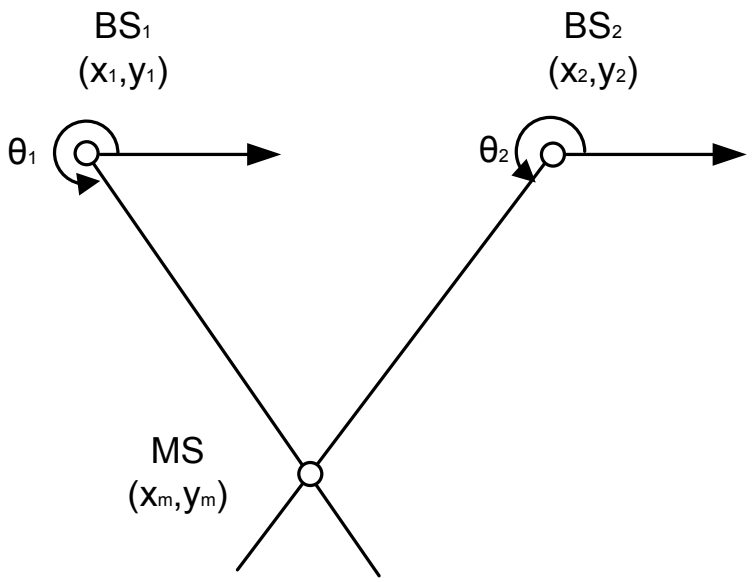


Figure 2.3: The AOA localization method.

RSS Data Fusion

RSS based localization is a method that employs mathematical models that describe the path loss as a function of distance. Since these models translate the received signal power into a distance between an MS and a BS, the MS must lie on a circle centred at the BS. Employing three or more BS's provides an estimate for the MS location.

RSS is well known for being drastically affected by multipath fading and shadowing (multipath signals). The error caused by multipath signals can be reduced by using prior knowledge available on the contours of the signal strength centred at the BS's [59]. However, such knowledge assumes a specific surrounding environment that can change due to

change in weather, moving objects, such as trucks, as well as new buildings and other barriers.

Fingerprinting (FP)

This localization method is a pattern recognition, or pattern matching, technique. The underlying concept of fingerprinting is that the radio signal propagation characteristics of an MS are unique in terms of TOA, AOA, and RSS when captured at different BS's [28,56]. Such characteristics can therefore be used as a signature to indicate the location of an MS [60]. The FP method has two phases: a training phase and localization phase [61–64]. In the training phase, a database is created to index the different patterns in the characteristics of the radio signal propagation. In the localization phase, the signature of the MS is matched with the patterns in the database. The challenging aspect of this method is assuring that the system can distinguish between similar patterns that represent different locations.

Of course, the more exhaustive training phase is (i.e., recording a signature for every small area in the environment), the more accurate the MS location estimate is. The main drawback of this method is the requirement to continually update the database as the configuration of the BS's changes when BS's are removed or new ones are added [65]. Nevertheless, this method is becoming more attractive for indoor applications because the database creation can be more comprehensive and manageable [65–67].

2.4.2 VANET Localization Using Relative Distances Measurements

This approach takes advantage of the emerging VANET environments. The distances between VANET nodes are estimated and exchanged among vehicles along with preliminary estimates of the vehicles' locations. Vehicles can then use this information to construct local

relative position maps that contain the vehicles and their neighbours. This strategy has initially emerged in Wireless Sensor Networks (WSN), but recently, a number of solutions have been proposed for use in VANET [13, 68, 69]

Vehicle Localization in VANET

A VANET based localization method was introduced in [68] for localizing vehicles with no GPS receivers, or those whose location can not be determined because satellite signals have been lost, for instance, in a tunnel. With this method, vehicles that are not equipped with GPS determine their own locations by relying on information they receive from vehicles that are equipped with GPS. Vehicles within transmission range can measure the distances between each other using one of the radio-location methods presented in [32]. By finding its closest three neighbours the unequipped vehicle can compute its position using trilateration.

Cooperative Vehicle Position Estimation

The work reported in [70] presents a method of distributed vehicle localization in VANET. The method utilizes RSS measurements to estimate the distances between one vehicle and others in its coverage area. It is assumed that vehicles initially estimate their own locations using a GPS receiver and then exchange their location information so that they can perform an optimization technique in order to improve their location estimates.

This technique demonstrated robustness of location estimates. However, it lacks the ability to detect and avoid the effect of multipath signals in the GPS measurements, which drastically degrades the localization accuracy in multipath environments (e.g., urban canyons).

In [71] an algorithm (i.e., IVCAL) has been proposed to mitigate the multipath effect in the location estimates of vehicles in VANET . A KF and an inter-vehicle communication system collaborate in order to increase the robustness and accuracy of the localization of every vehicle in the network. The two main components that allow the inter-vehicle communication system and the KF to interact are the Multipath Detection Unit (MDU), which detects the existence of a multipath effect in the output of the KF, and the Localization Enhancement Unit (LEU), which obtains the neighbours' information from the inter-vehicle communication system and feeds an optimized location estimate back to the KF (Figure 2.4).

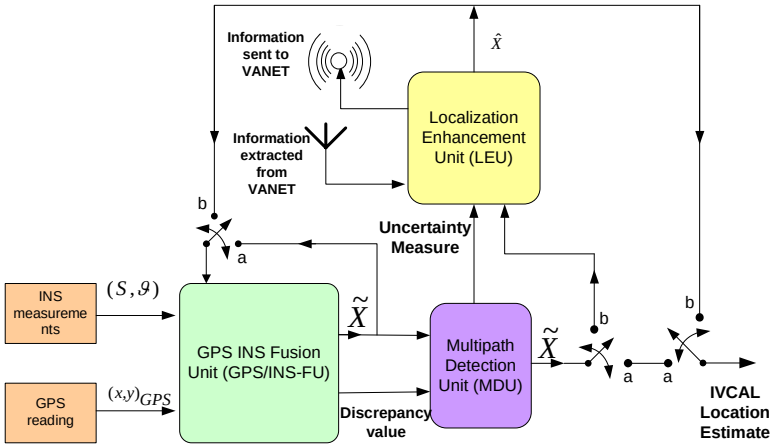


Figure 2.4: Block diagram of the inter-vehicle communication assisted localization technique.

As in [11] and [45], KF innovation is used as an indication of the contamination of the GPS measurement, and it has therefore been used as a learning pattern for the MDU in IVCAL. An uncertainty measure is also utilized in order to specify a subset of the most accurate network neighbours that can be used as anchors to enable vehicles to improve

their location estimates.

Lack of adequate location anchors and/or prolonged multipath conditions remain unsolved problems that continue to degrade localization accuracy.

2.5 Multi-Level Fusion Approaches

As it has been reported above, a verity of multi-modality localization methods have evolved in recent years. Typical modalities include satellite signals, VANET communication, vision features, laser rays, etc. This variety of information has motivated the concept of multi-level fusion.

For instance, in [72], a data-fusion model is proposed in the form of a three-level fusion localization system. In the first level, a variety of location information is gathered as row data and processed separately using local filters that are suitable for each type of location information. As with the system in [39], the second level combines the output of the first level and produces a better location estimate. In [72], the results are then fused in the third level based on contextual information (e.g., digital maps and traffic information). In this scheme, the final location estimate is fed back to the second level in order to improve future estimations.

Multi Level fusions aims to tackle data fusion as a hierarchical process so as to allow for combining measurements at various levels of abstraction in a simple manner. The estimates in the lowest-level filters are evaluated for reliability. The fusion of these estimates in higher-level filters will then be more robust.

2.6 Integrity of Localization Systems

Due to the inherent errors in the positioning information, a level of uncertainty in location estimates is inevitable. Therefore, it is essential to measure the confidence of the positioning information in order to identify any hidden anomalies. To achieve this task, a level of trust, integrity, in every estimate must be determined.

In the last two decades, a significant effort has been made in aviation to develop integrity monitoring systems [73, 74]. Integrity is defined as a measure of the trust which can be placed in the correctness of the information supplied by the total system; integrity includes the ability of a system to provide timely and valid measurements to users [75]. Three key components have been proposed for integrity monitoring: 1) fault detection, 2) fault isolation, and 3) removal of faulty measurement sources from the estimates [76]. The European Geostationary Navigation Overlay Service (EGNOS) and the Wide Area Augmentation System (WAAS), [74], are developed to form a redundant source of information for the Global Navigation Satellite Systems (GNSS) in order to perform integrity monitoring by providing correction information.

During the last decade, monitoring the integrity of land-vehicles' localization has attracted attention due to the increasing demand for highly reliable accurate location data. Since roving in dense urban environments may limit access to the signals from augmentation systems such as EGNOS or WAAS, other means of measuring integrity have been proposed [77–80].

For instance, [81] and [82] present a localization solution based on the fusion of GNSS and INS sensors. In this fusion process an interactive multimodel method is used. Different covariance matrices are used as a response to change in the noise behaviour. The proposed integrity measure is based on the covariance matrix of the EKF estimation error.

Relying on the error covariance matrix can be misleading especially when experiencing unmodeled environment noise. In other words, it is not possible in many cases to detect, isolate, and remove the estimation faults, let alone the unavoidable false alarms.

Also, in [80] a binary integrity decision-maker is proposed for a map-matching localization technique in which multihypothesis road-tracking method combines proprioceptive sensors (odometers and gyrometers) with GPS and map information. In this work, the integrity represents high or low confidence of the location estimate. The candidate tracks or roads are associated with a probability that is computed using the multihypothesis road-tracking method. If one credible road exists and the normalized innovation is below a prespecified threshold, the technique declares high confidence location estimate. However, the lack of granularity in the integrity measure limits the range of the integrity-level based application that can use this method.

Integrity monitoring of map-matching localization has also been proposed and tested in [77]. However, in this work three indicators has been monitored to achieve this task: distance residuals, heading residuals, and an indicator related to uncertainty of the map matched position. Due to the linguistic nature of these indicators, they have been combined using a fuzzy inference model to produce a value between 0 to 100 to indicate the integrity of the system. The derivation of the integrity threshold has been determined experimentally to be 70, where the type of the environment experienced during the experiment was not specified. The threshold value, nevertheless, might be different from one environmental condition to another, and dependent on the type of sensors used. More importantly, the integrity needed for localization is application/task dependent. For instance, the demand on navigation systems to produce accurate localization is more critical in urban areas, where road networks tend to be dense, than that on highways and rural areas.

Knowledge of the observation conditions that influence the accuracy performance of a

localization system is a key issue in evaluating the integrity of location estimation. In the following section, some operational parameters have been studied in the literature in terms of their ability to capture localization accuracy attributes.

2.6.1 Localization Accuracy Assessment

A fuzzy inference method to evaluate the accuracy of location estimate obtained by a GPS receiver is presented in [25, 83]. The decision about the accuracy of the location estimate is made based on the Geometric Dilution of Precision (GDOP) and the Signal to Noise Ratio (SNR). The fuzzy system output is defined as a reliability factor which indicates the amount of trust in the location estimate.

Chang *et al.* [84] use the number of visible satellites as a measurement conditions indicator. When the number of visible satellites drops below four, the pseudo-range and clock bias of some invisible satellites are estimated. In order to solve the GPS positioning problem, both the altitude and the speed of the vehicle are assumed to be constant. Of course, these two assumption are seldom valid, particularly in urban environments. Furthermore, the persistence of such environmental conditions for a long period of time degrades the estimation performance of the pseudo-range and clock bias, and consequently negatively affects the performance of the localization process. Thus, accuracy in this case cannot be taken for granted. Meguro *et al.* [49] attempt to mitigate the location estimate errors by eliminating the error source(s). To achieve this task, an IR camera is used to identify which satellite is introducing multipath signals so that its reading is discarded. Nevertheless, in high-rise building environments, where multipath effects are most prevalent, the number of satellites with LOS to the GPS receiver can be insufficient. Consequently, using only these satellites in such an environment introduces high DOP and therefore strongly diminish the accuracy of the available measurements [85, 86]. In general, eliminating the number

of observed satellites may lead to a situation in which the availability of highly accurate location estimates is reduced – a performance requirement that is crucial to mission critical applications [15, 23, 72].

Costa [87] proposed a simulation model to represent the channel of the GPS, upon which the path status (e.g., clear, shadowed,or blocked) and positioning error are estimated. The estimation process utilizes a digital elevation model, buildings' databases, and a vegetation model, in addition to number of observed satellites and their SNR. The estimations of the simulation model is comparable to the ground truth measurements; however, it can not be used during the absence of buildings' databases and the other prerequisites of the simulation model, which limits the mobility of such technique.

Tzoreff *et al.* [88] proposed a model that describes the vehicle path to achieve more accurate localization in urban environments. The model demonstrates superior performance even in the presence of one satellite with LOS. Nevertheless, satellites with NLOS – a common issue in urban areas [24, 71, 89] – are not considered.

2.7 Performance Criteria and Benchmarking

From the discussion above it is clear that vehicle localization is an increasingly growing area of research. Nevertheless, there is a number of outstanding issues that still need to be addressed. In order to put these outstanding issues in practical context I propose the following performance criteria.

2.7.1. *Accuracy:* *Accuracy of a vehicle location estimate is defined as the degree of closeness of a vehicle's location estimate to its actual (true) location.*

2.7.2. *Availability:* *Availability of a vehicle location estimate is defined as the ratio of the number of estimates produced to the number of estimates expected per one unit of time.*

2.7.3. *Response Time:* *Response time is the time required by a localization technique to produce a location estimate.*

2.7.4. *Integrity:* *Integrity is defined as the level of confidence that can be placed in the correctness of the location estimate* [75, 77, 90].

Based on the above performance criteria, a benchmark can be established in order to compare the performance of different localization techniques. Such comparison does not only consider reported best achievable accuracy localization performance, but also examines the integrity performance of these techniques. Table 2.1 provides a comparison summary of the modality used, best case accuracy, environmental constraints to indicate availability, need for synchronization where response time must be considered, and dependency on infrastructure.

Table 2.1: Specifications of Localization Techniques.

Modality(ies)	**Best Case Accuracy(m)**	**Availability**	**Synch.**	**Infr.str.**
GPS	10-20 [19, 20]	Out Door-Open Sky	Yes	No
DeadReckoning (DR)	Worsen with time [38]	Anywhere	No	No
DGPS with Visible Satellites	0.01-7.6 [49]	suburban-Open Sky	Yes	Yes
DGPS+DR+Map Matching	0.5-5 [91]	Out Door-Open Sky	N/A	Yes
GPS+Vision+Map Matching	0.5-1 [45, 92]	Out Door-Open Sky	No	Yes
Cellular Localization	90-250 [56]; 25-69 [28]	Under Network Coverage	Yes	Yes
Location Services	Submeter [93]	In Door	N/A	Yes
Relative Ad hoc Localization	2-7 (Simulation [13, 70])	Suburban	Yes	No

Table 2.2 reports emerging applications and their requirements with respect to localization accuracy only. It is evident from Tables 2.1 and 2.2 that current localization techniques do not live up to the required integrity.

As it is shown in Table 2.1, synchronization is a prerequisite in the fusion process of some multi-sensory systems. However, the synchronization issue is out of the scope of this thesis since it has been thoroughly investigated and covered in the literature (*cf.* [40,57,94]).

Table 2.2: Applications Requirement for Location Estimates [72].

Application	**Required Accuracy**		
	Low(10-20 m)	**Medium (1-5 m)**	**High (1 m or less)**
Message Routing (VANET)	X		
Data Dissemination	X		
Map Localization	X		
Road Pricing	X		
Coop. Cruise Control		X	
Coop. Intersection Safety		X	
Blind Crossing		X	
Platooning		X	
Collision Warning Sys.			X
Vision Enhancement			X
Automatic Parking			X

2.8 Task Driven Localization Integrity

From the above discussion it is obvious that for localization systems to meet the expectations of emerging applications it is imperative that they employ diverse location measurement sources and effective strategies to fuse these sources so as to achieve the QoS expected of them. Of course, this QoS is multi-dimensional as it pertains to expected accuracy, availability, response time and integrity. The QoS as a function of these performance criteria is application and task dependent. The more stringent the required QoS is with respect to a given performance criterion, the more resources are needed and the higher the computational cost is. This presents a challenge for the system as calls for effective use of resources to achieve the target QoS. For example, there are applications where accuracy can be traded for faster response time. On the other hand, there are applications where response time is not as important as accuracy (offline vehicle track mapping). There are also applications where both requirements, accuracy and response time, can not

be compromised for any other gain.

Indeed, task or goal driven localization is about effective allocating system resources and planning of localization tasks such that the system mission is achieved with maximum integrity possible. This strategy to performance is the core of this research work. In order for this strategy to work it is imperative that the impact of the environment is not ignored. Without modeling the impact of the environment on the system, the system can not be guaranteed to achieve its target performance, and even worse as it may falsely determine its task is accomplished. Thus, modeling the impact of the environmental conditions on the system is a central issue to this research work.

2.9 Summary

In this chapter, a variety of reported localization techniques are presented and classified based on the type of the measurement of the location information used.

Although, techniques that incorporate fusion of motion sensory data with GPS localization have demonstrated improvement in performance, there are still situations that can have a negative effect on their localization accuracy. Incremental localization errors in motion-sensor data and the multipath effect in urban canyon environments contribute significantly to such location estimate errors, which necessitates augmenting the initial location data with other sources of location information in order to overcome these shortcomings.

Digital maps and visual features enhance GPS-DRS localization by recognizing landmarks in the surrounding environment and matching them with others in a reference GIS map. A key problem associated with this scheme is that the landmark segmentation process is complex an ill conditioned process.

Range location methods offer localization performance that provides relatively accurate location estimates in urban areas. NLOS is often the main problem in such localization techniques. Lack of adequate location anchors and/or prolonged multipath conditions are challenges that continue to effect localization accuracy.

Multi-level fusion schemes are promising as they employ multiple location measurement phenomena. However, these schemes have given birth to new challenges in the localization problem in terms of resource management, and task driven performance.

In the next chapter, a proposed framework for vehicle localization is presented. The aim is to develop a vehicle localization system that can optimize and plan the use of its resources so as to achieve the performance requirements of the localization task or application.

Chapter 3

Task-Driven Localization Through Integrity Assessment and Control

3.1 Introduction

It is well understood that the techniques presented in Chapter 2 can estimate the location of vehicles relatively accurately in some situations if they are given adequate time to perform the task. However, they may not perform as well in other situations. The deficiencies of these localization techniques are uncorrelated as they are expected to be of diverse phenomena, and/or utilize different algorithmic paradigms. This motivates the development of systems that can take advantage of this diversity to achieve a reliable and accurate performance.

In this chapter, a high level concept of a novel framework for fusing different localization techniques is proposed. What distinguishes this framework from existing ones is its ability to take into account the impact of the measurement conditions on the individual techniques. Thus, it is able to optimize the fusion process so as to maximize the accuracy and integrity of the localization estimates.

Most of the fusion techniques reported in the literature employ an instance of KF or EKF to fuse location measurements [11, 40, 79, 95]. The integrity of the estimation is represented by the covariance matrix of the estimation error or the innovation. For example, in [11, 40, 95] the acceptance of the location estimate is evaluated through a binary decision maker that utilizes the innovation covariance matrix. In [79], an interactive multimodel filter (IMM-EKF) provides an estimate based on a weighted combination of these models. The weights are determined using a probabilistic model. The integrity of the location estimate is evaluated based on the error covariance matrix. It is important to note that KF and EKF assume low order nonlinearity of the measurement model and that the contaminant noise is of a zero mean Gaussian distribution; an assumption that does not hold in many real-life situations. Furthermore, it is often the case that the location measurement noise is time-varying, as such the error covariance matrices can not be relied on to provide insight into the measurement process integrity [96]. Time variant noise covariance matrices are proposed to tackle this issue; however, it is impossible to know a priori all the degrees of noise the system can experience. The objective of this research work is to develop a task-driven multi-modality localization system. Integrity in this system of location estimation is captured by means of a reliability model. The model accounts for the reliability of each modality as a function of the required integrity objectives as well as the impact of the measurement conditions on each modality [97, 98].

3.2 Problem Formulation

Consider a localization system that employs a set of N localization techniques, indexed by the set $\Psi = \{T_1, T_2, \ldots, T_N\}$. The system is to estimate the current location of a vehicle. Let X_t be the true location of the vehicle at time t. Each technique T_i uses its

measurement resources to compute an estimate $Z_{i,t+\Delta t_i}$. Δt_i signifies the time it takes technique T_i to compute the estimate. Each technique T_i declares a quality index ξ_i to its location estimate. It is conceivable that the quality indices of the different techniques to be unequal. For example, a GPS based technique may not produce location estimates with qualities equal to that of a map matching based localization technique. The goal of the system is to manage its set of techniques so as to compute a location estimate that meets a target quality Ξ_s.

Figure 3.1 depicts a block diagram of the system. The system consists of three logical layers: (1) Primary Localization layer which provides preliminary location estimates using the localization techniques in Ψ; (2) Integrity Monitoring layer which computes the reliability of the vehicle's location estimates produced by the Primary Localization layer- a process that captures the impact of measurement conditions; and (3) Estimate Fusion and Management layer which interacts with the application task to ensure that the task's expected localization accuracy and integrity are achieved by executing a proper fusion scheme. The following sections discuss these layers in more details.

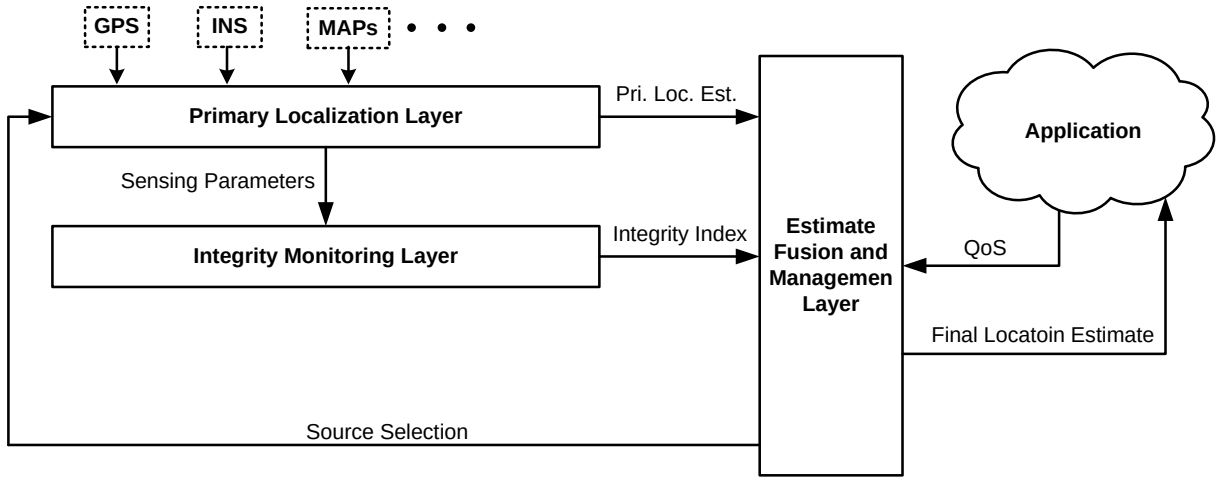


Figure 3.1: The structure of the proposed framework.

3.3 Primary Localization Layer

The primary localization layer contains system's localization techniques which are partitioned in the form of a set of Primary Localization Units (PLUs), as can be seen in Figure 3.2. Any localization technique, such as those mentioned in Chapter 2, can be used in any given PLU. These primary localization units receive localization requests from the Estimate Fusion and Management layer. PLUs are constructed from techniques that are based on different phenomena/algorithms to ensure minimum correlation. A primary localization unit can share its information sources with other units; it can constitute a single modality or multiple-modalities. An example of a single modality PLU is one that estimates the

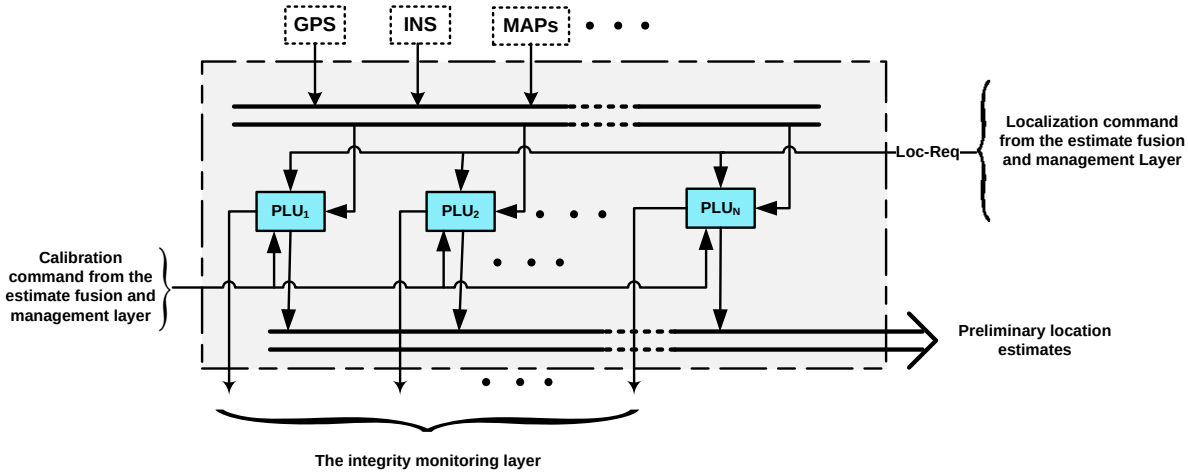


Figure 3.2: Primary localization layer.

vehicle location from a GPS information source. IVCAL is an example of a PLU that utilizes three modalities: GPS, INS, and Inter-Vehicle Communication (IVC) .

Central to the PLU functionality is determining the accuracy of the primary location estimate. Novel classification algorithms are introduced in Chapter 4 to assess the localization accuracy of PLUs.

3.4 Integrity Monitoring Layer

Central to the proposed framework is the integrity monitoring layer. Here, an Integrity Monitoring unit (IMU) is used to monitor the performance of a primary localization unit(Figure 3.3). The monitoring process takes into consideration the impact of the

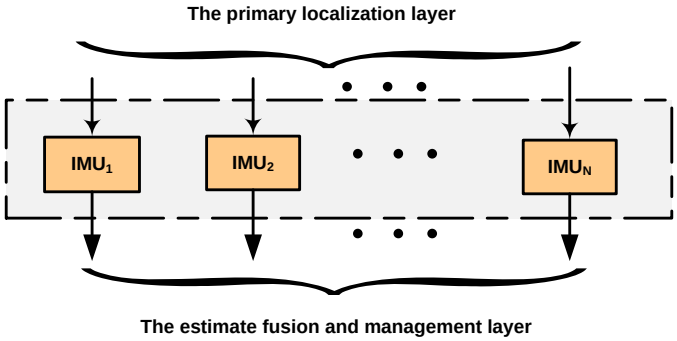


Figure 3.3: Integrity monitoring layer.

measurement conditions on the PLU, as well as its integrity historical profile. The theoretical formulation for the integrity monitoring units, as well as experimental results to demonstrate this concept, is presented in Chapter 5.

3.5 The Estimate Fusion and Management Layer

The Estimate Fusion and Management layer (EFM) is responsible for determining an effective integration (Meta-Fusion) strategy for fusing the estimates produced by the different primary localization units so as to achieve the required localization accuracy and integrity (Figure 3.4). The estimate fusion and management processes the location estimates produced by the different primary localization units in conjunction with their integrity assessments. Therefore, this layer employs a synchronization handler to manage timing issues

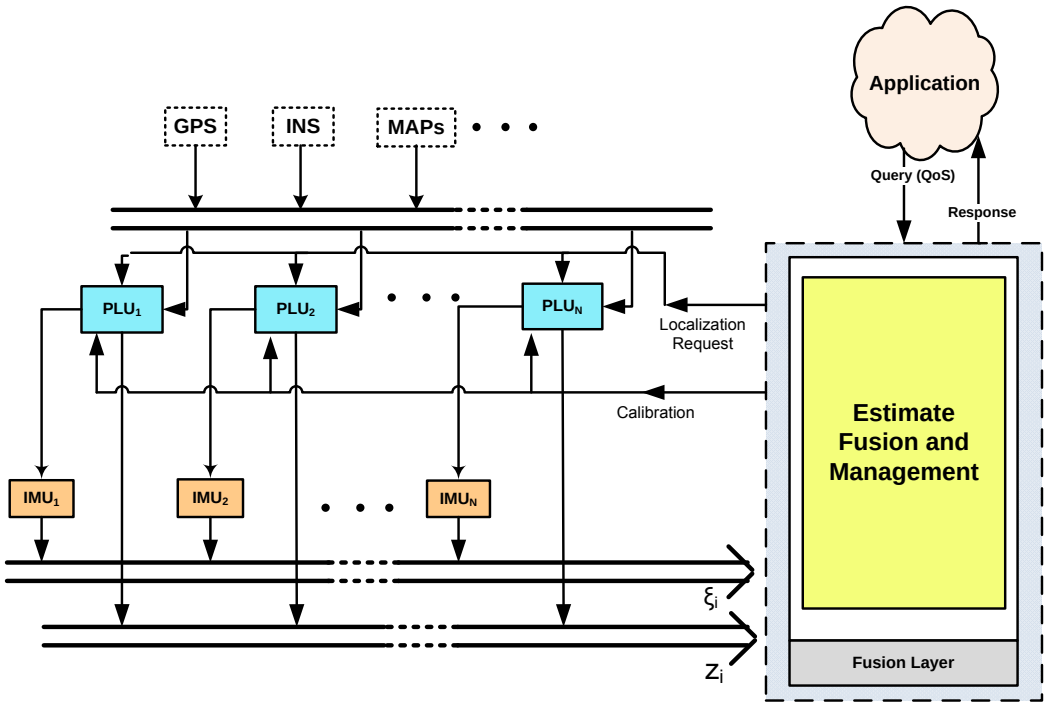


Figure 3.4: Estimate fusion and management layer.

among the different PLUs. Given the task's target accuracy and integrity, as well as that of the various PLUs, a measurement and fusion scheme is computed such that the scheme produces a location estimate that meets the task requirements. This layer will be introduced in Chapter 6 where localization accuracy, integrity, and consumed energy and cash are the QoS parameters considered.

3.6 Summary

In this chapter, a novel framework for vehicle location estimation is presented. The aim of this framework is to utilize localization technique diversity and localization source diversity to achieve robust localization performance that meets the application task's integrity and

accuracy constraints. It is expected that this framework will overcome some of the drawbacks of existing localization techniques. Central to the effectiveness of the framework is its ability to capture estimate reliability in computing the vehicle location. Furthermore, channel discounting, in the light of localization integrity assessment, will allow the framework to maximize the quality of estimate fusion as well as to address integrity aggregation among the different estimates. In the next two chapters, the development of the localization accuracy classification and the reliability assessment model are introduced along with testing and examples of their implementations.

Chapter 4

Localization Accuracy Classification: A Context Based Approach

4.1 Introduction

The localization accuracy of any system depends heavily on both the technique it uses to compute locations and the measurement conditions in its surroundings. However, while localization techniques have recently started to demonstrate significant improvement in localization performance (*cf.* [49,71,83,85,99,100]), they continue to be severely impacted by the measurement conditions in their environment. Indeed, the impact of the measurement conditions on the localization accuracy in itself is an ill conditioned problem due to the incongruent nature of the measurement process.

Reliable assessment of localization accuracy is a paramount prerequisite for any attempt that aims to improve localization process. Two main contributions are introduced in this chapter: 1) localization system parameters and location estimates are used to contrive a feature space representation of probable accuracy classes, and 2) two novel classification algorithms are proposed to classify the accuracy performance of a localization system based on a set of features that are computed from the system measurement process. The first

algorithm processes the feature space so as to identify class similarity, [101], among the different accuracy classes in the features space. A hierarchical classification algorithm is introduced to partition the feature space such that the classification process is performed in stages, over which the pattern space is merged and split according to the degree of similarity (i.e., overlap) among the different accuracy classes. Subsequently, the other classification algorithm efficiently employs contextual information about the measurement conditions to enhance the accuracy classification process.

The remainder of this chapter is organized as follows. Section 4.2 discusses the creation of feature space using measurements obtained during the localization process. Section 4.3 presents the formulation of the classification problem and introduced to novel accuracy classification algorithms. Section 4.4 and Section 4.5 explore two case studies in which the introduced classification algorithms are implemented on two different localization techniques. Experimental work and comparative results are provided in each case study. Section 4.6 summarizes the chapter and provides concluding remarks.

4.2 Feature Space Representation of Accuracy Classes

Accuracy is often determined as a real (i.e., continuous) figure. In practice, however, applications require specific margin/band(s) of accuracy. Therefore, the accuracy estimation problem is defined here as that of classification. Consider a localization technique, T_i, that has n operational parameters deemed to influence the accuracy performance. These parameters are used to construct a feature space where each pattern is represented by a point in $n-$dimensional space. Since the parameters are impacted by the measurement conditions, the feature space allows us to observe trends among the collected samples in this space so as to define predictive governing relationships between a measurement and

its accuracy index (class), Figure 4.1.

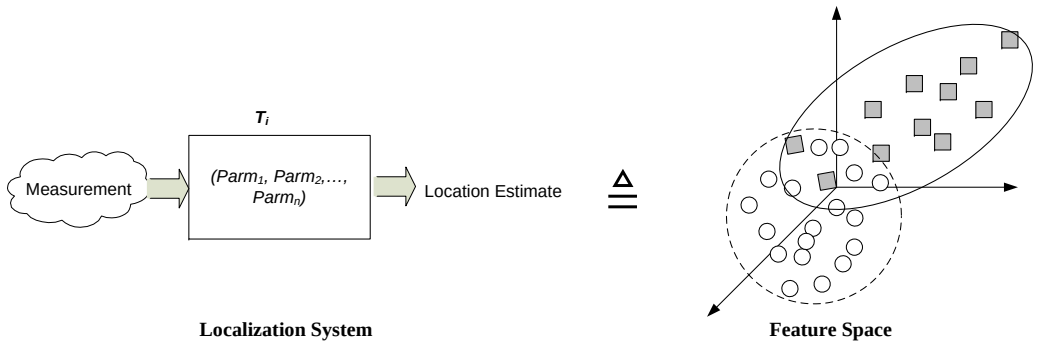


Figure 4.1: Mapping the location estimate process into feature space of accuracy classes.

Patterns of accuracy classes are known for being strongly overlapped (i.e., similar classes) to the extent that patterns can not be distinguished as to which class they belong to. Figure 4.2 shows an example of the classes overlap in the feature space of GPS accuracy classes. In the following section, the accuracy classes disambiguation and enhanced location accuracy classification algorithms are presented.

4.3 Context Driven Accuracy Classification

The localization accuracy classes exhibit a significant degree of overlap in the feature space. Thus the process of mapping a Localization measurement to an accuracy class using a naive classification strategy, such as a flat or non-similarity guided hierarchical classifiers, Figure 4.3, is error-prone. To address this issue we propose a context driven hierarchical classification strategy, whereby class unfolding and contextual information are used to regularize the classification process such that the classification error is minimized.

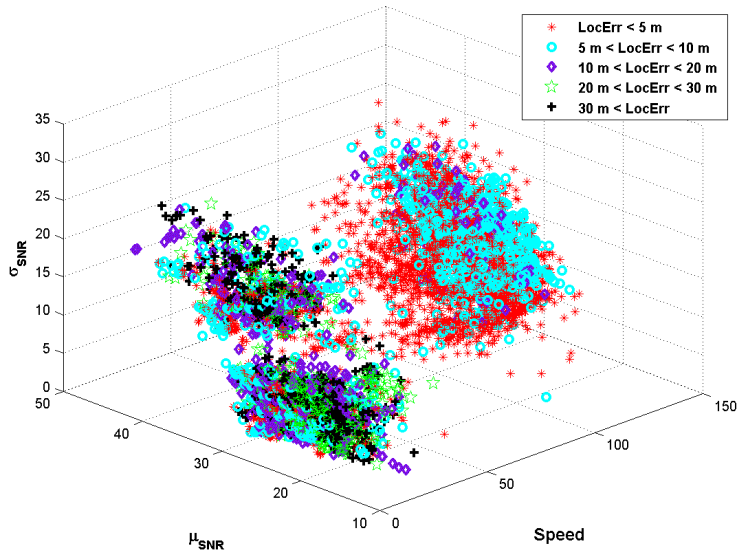
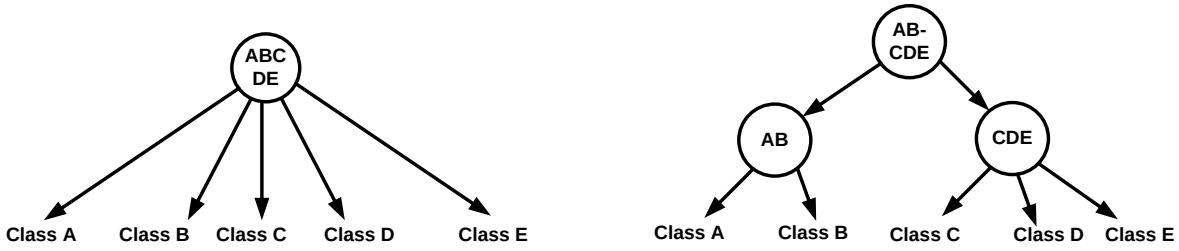


Figure 4.2: Example of accuracy classes' similarity (GPS measurements).



(a) Flat classifier (h).

(b) Hierarchical classifier ($\hbar$).

Figure 4.3: Multi-class flat classifier versus multi-class hierarchical classifier.

4.3.1 Hierarchical Classification using Class Unfolding

Consider a data set of patterns indexed by the set $P=\{p_1, \ldots, p_K\}$, where K represents the number of patterns in P. Let $C=\{c_1, \ldots, c_M\}$ be a set of class labels, such that $label(p_k)=q$, $q \in C$. Let G^q be the set of patterns $\{p_k\}$ such that $label(p_k)=q$, $q \in C$, for all $k = 1, \ldots, K$. The goal is to design a hierarchical multi-class classifier $\hbar$ that classifies each pattern p_k such that $\hbar(p_k) = q$ implies $p_k \in G^q$. The error rate, ϵ, of classifier $\hbar$ is defined as the

probability of misclassifying pattern p_k (Equation 4.1).

$$\epsilon(\hbar) = Pr(\hbar(p_k) = q | p_k \notin G^q) \tag{4.1}$$

The basic idea is to utilize a tree of binary base classifiers (h's) (e.g., linear classification, neural networks, or SVC), producing an equivalent of a multi-class classifier [102]. Each binary classifier constitutes a node in the binary tree and decides whether a pattern belongs to a specific class $q \in C$. In what follows, we explain the construction of the hierarchical classifier using class unfolding (HCCU) (Figure 4.4), that minimizes the error rate ϵ, regardless of the base classifier used.

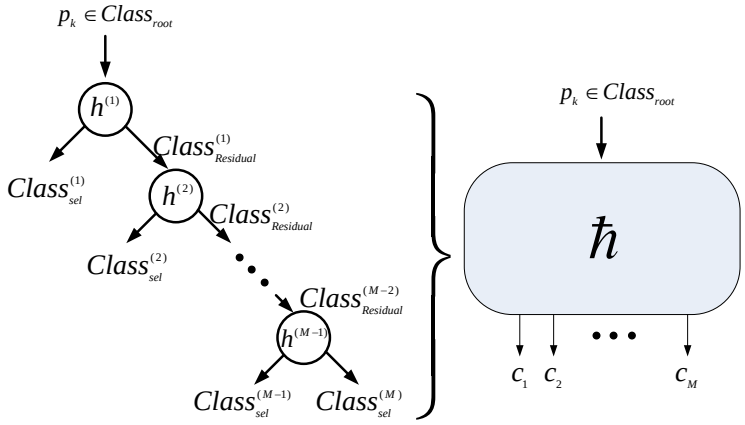


Figure 4.4: Hierarchical classifier using class unfolding.

Classifier Construction

As seen in Figure 4.2 it is possible for measurements from different classes to posses similar features. The extent of such similarity is class dependent. This similarity information plays an important role in the construction of the HCCU. The similarity between class q and class ω, ζ^q_ω, is defined as the number of patterns from class q that fall into the convex hull

of ω (Figure 4.5). Conversely, the similarity between class ω and class q, ζ_q^ω, is defined as the number of patterns from class ω that fall into the convex hull of class q. Generally speaking, $\zeta_q^\omega \neq \zeta_\omega^q$.

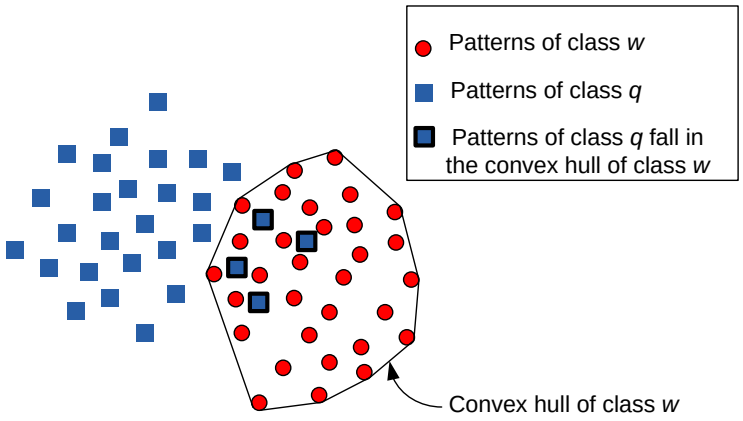


Figure 4.5: Classes similarity measure.

As a first step in the construction process, all measurements are collapsed into one class, i.e, $Class_{root}$. A binary decision process is employed to split the measurements into subclasses such that the similarity between these subclasses is minimized and the similarity between measurements of any given subclass is maximized. From $Class_{root}$ two classes are generated

$$Class_{Sel}^{(0)} = \phi$$
$$Class_{Residual}^{(0)} = C$$

At each stage i, $i = 1, \ldots, M-1$, $Class_{Sel}^{(i)}$ and $Class_{Residual}^{i}$ are computed as follows:

$$Class_{Sel}^{(i)} = \arg \min_{q \in C - S^{(i-1)}} \zeta_r^q \ , \ \forall r = (C - S^{(i-1)} - q) \tag{4.2}$$

$$Class_{Residual}^{(i)} = Class_{Residual}^{((i-1)} - Class_{Sel}^{(i)} \tag{4.3}$$

$$S^{(i)} = S^{((i-1)} \bigcup Class_{Sel}^{(i)} \tag{4.4}$$

where $S^{(i)}$ is a set of labels that have been selected in upper levels of $\hbar$ (e.g., at level i, $S^{(i)} = \{\phi, Class_{sel}^{(1)}, Class_{sel}^{(2)}, \dots \ Class_{sel}^{(i)}\}$). This completes the design of the HCCU.

Accuracy Classification

Classifying the accuracy of a new measurement x is a binary decision process that recursively traverses the tree starting from the root. At each level j the similarity between x and $class_{Sel}^{(j)}$ and that between x and $Class_{Residual}^{(j)}$ is tested. This process is iterated until a matching class is found (Algorithm 4.3.1).

Algorithm 4.3.1 : Accuracy Classification

```
Input: pattern x;
Output: Accuracy Class ŷ ;
for i = 1 : M − 1 do
    ŷ = h^(i)(x);
    if ŷ == Class_sel^(i) then
        return ŷ;
    end if
    /*x ∈ Class_Residual^(i)*/
end for
return Class_sel^(M);
```

There remain situations that give rise to degraded accuracy classification due to localization system performance sensitivity to objects in the surroundings. Generally speaking, the impact of environmental factors on the localization system tends to be stochastic in behaviour. This suggests the possibility of minimizing the impact of such influence by augmenting the measurement process by temporal contextual information.

4.3.2 Context Based Accuracy Classification (CBAC)

This algorithm is developed to enhance the performance of the proposed HCCU based on contextual information about the surrounding. The surrounding affects the localization system parameters, based on which the HCCU deduces a decision about the localization accuracy. Therefore, the deduced decision can be utilized as an index that is associated with certain *state*. A *state* can be defined as one that captures the effect of the surroundings. The contextual information is signified here by a set of states, each state represents environmental context in which a localization system produces a specific localization accuracy (i.e., c_1,..., c_M).

To capture the impact of the surrounding on the accuracy classification process a number of delayed signals ($\tilde{y}_k$'s) are introduced. These delayed signals are used to compute a state uncertainty index.

First, an HCCU classifier: $\hbar_1$, $\tilde{y}_k = \hbar_1(x_k)$, is used to determine the accuracy class of the localization measurement x_k. A delayed signal window $W_{k,L} = (\tilde{y}_k, \ldots, \tilde{y}_{k-L})$ of the last $L+1$ contextual states is constructed; $\tilde{y}_k$ is the accuracy class of the most recent localization measurement.

The extent of the impact of the environment on the localization accuracy will manifest itself as a degree of randomness in the output of the classifier $\hbar_1$. In order to quantify

the uncertainty in the decision made by classifier $\hbar_1$, this randomness is computed as the information entropy of the delayed signal window as follows:

$$H_{W_{k,L}} = -\sum_{i=1}^{M} Pr(c_i|W_{k,L}) \log_M Pr(c_i|W_{k,L}) \tag{4.5}$$

where $Pr(c_i|W_{k,L})$ is the probability that the current localization measurement, p_k, belongs to class c_i, given the delayed signal window $W_{k,L}$. $Pr(c_i|W_{k,L})$ is computed based on the frequencies of the classification decisions in $W_{k,L}$.

A zero value of $H_{W_{k,L}}$ signifies the fact that the classifier decision is consistent throughout the window, consequently, indicating classifier persistence. The higher is $H_{W_{k,L}}$ the less persistent is the classifier about its decision. This persistence factor is used to compute the state uncertainty index $f_{(k,L)}$ as follows:

$$\begin{aligned} A &= \arg\max_i Pr(c_i|W_{k,L}),\ \forall c_i \in C \\ B &= \arg\max_i Pr(c_i|W_{k,L}),\ \forall c_i \in C - c_A \\ f_{(k,L)} &= \begin{cases} A - \left(\frac{H_{W_{k,L}}}{2}\right) & when\ A \geqslant B \\ A + \left(\frac{H_{W_{k,L}}}{2}\right) & when\ A < B \end{cases} \end{aligned} \tag{4.6}$$

This index captures two important cues, namely, the contextual state, and the degree of uncertainty around the contextual state. For example, consider three accuracy classes: c_1 (*Accurate*), c_2 (*Marginally-Accurate*), and c_3 (*Inaccurate*). $f_{(k,L)} = 1$ signifies the fact that the classifier is quite certain that the environment is such that it coerces the system to produce location estimate with an accuracy class of type c_1. On the other hand, a value of 1.5 signifies the fact that the environment is such that it coerces the system to produce

a localization accuracy between type c_1 (*Accurate*) and c_2 (*Marginally-Accurate*). Thus, the accuracy feature space is augmented by this state uncertainty index so as to make the process of accuracy classification be a function of both the localization measurement x_k and the condition of the working environment around the localization system (indexed by $f_{(k,L)}$). This augmented feature space is used by another HCCU classifier: $\hbar_2$ to determine the accuracy of the GPS measurement x_k.

Figure 4.6 depicts CBAC where two HCCU classifiers ($\hbar_1$ and $\hbar_2$) collaborate to determine the accuracy class of localization measurements. The pseudo code for testing the proposed CBAC is depicted in Algorithm 4.3.2.

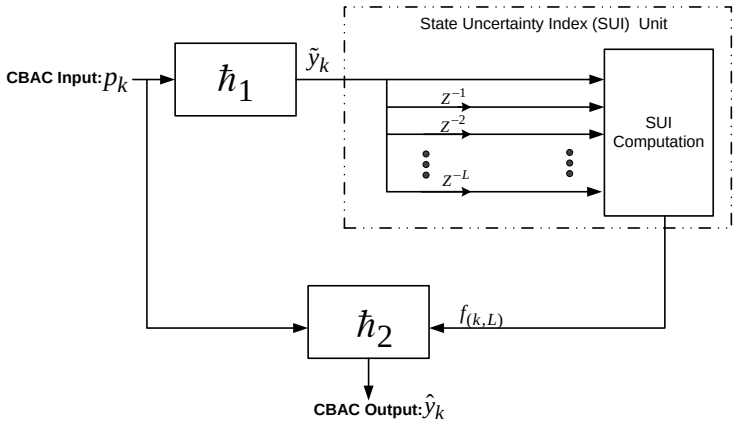


Figure 4.6: Context Based Accuracy Classification (CBAC) schema.

The validation of the introduced accuracy classification algorithms is conducted next, via two case studies. The first case study revolves around the dynamics of accuracy performance of a GPS receiver while the second one evaluates the accuracy classification of an FP localization technique.

Algorithm 4.3.2 : Pseudo Code for Testing CBAC

Input: Testing patterns x_k's;
Output: Accuracy class $\hat{y}_k$'s;
/*Note: $\tilde{y}_k$ is the output of $\hbar_1$ and $\hat{y}_k$ is the ultimate output of the CBAC*/
/*CBAC needs the first L measurements to initialize the delayed signal window using $\hbar_1$*/
for $k = 1$ to L **do**
    $\tilde{y}_k = \hbar_1(x_k)$;
    $\tilde{y}_k \rightarrow$ delayed signal window;
end for
/*Start feeding $\hbar_2$*/
loop
    $k++$;
    $\tilde{y}_k = \hbar_1(x_k)$;
    $\tilde{y}_k \rightarrow$ delayed signal window;
    /*Calculate $Pr(c_i|W_{k,L})'s$, $H_{W_{k,L}}$, and $f_{(k,L)}$ as in equations 4.5 and 4.6*/
    /*Construct pattern for $\hbar_2$*/
    $\hat{x}_k = (x_k, f_{(k,L)})$;
    $\hat{y}_k = \hbar_2(\hat{x}_k)$;
end loop

4.4 Case Study 1: GPS Accuracy Classification

The selection of GPS as the localization technique for this study is a result of the inherent status GPS enjoys as a basic building block in most of localization systems. Thus, degradation in GPS localization accuracy performance has the potential of negatively impacting the performance of the overall system.

The following section explains the GPS error components, and presents the features introduced in this work. These features are used collectively to infer the GPS receiver accuracy, Figure 4.7.

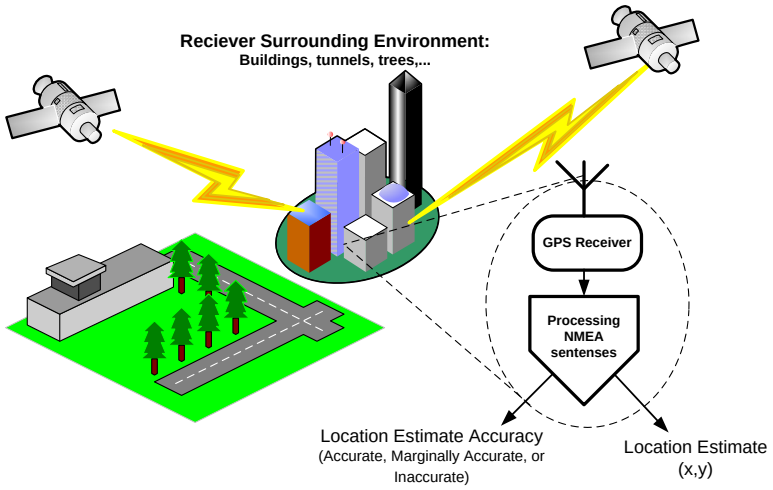


Figure 4.7: System Model: GPS receiver's surrounding environment has an impact on the location estimate accuracy.

4.4.1 Errors in GPS Location Estimate

As the signals of the GPS satellites travel to the GPS receiver, several error components accumulate in the pseudorange code ρ_{code}, and thus this code is a noisy estimate of the range, r, from the satellite to the receiver. These error components are summarized in the following equation

$$\rho_{code} = r + \delta_{eph} + \delta_{iono} + \delta_{trop} - s\mathbb{T} + \delta_{mp} + v_{rcvr} \tag{4.7}$$

where δ_{eph} signifies the satellite ephemeris error. The ephemeris error is caused by the difference between the actual GPS satellite position and the anticipated position. δ_{iono} and δ_{trop} signify the ionospherical and tropospherical errors, respectively; these spherical errors are caused by reduction in the speed of the satellite signal. The signal originally travels at the speed of light, s, but the speed is reduced when crossing the ionosphere and troposphere layers. The receiver clock error is signified by $\mathbb{T}$. The multipath error is

signified by δ_{mp}. The multipath error is caused by the delay of the signal arrival due to its reflection off building surfaces in the area. The error component v_{rcvr} signifies the receiver measurement noise which is device dependent.

As per [21, 22], the magnitude of δ_{eph} is about 3 m, δ_{iono} is about 5 m, and δ_{trop} is about 2 m. However, the range of the error caused by the receiver clock error depends on the value of T. In [103], it is reported that this error range can be reduced to less than 1 m. The Selective Availability (δ_{SA}) error, which previously was used to capture artificial falsification, was not incorporated. This error was eliminated in 2000 [104] and thus is no longer relevant for computing the ρ_{code}. In [21], it is reported that differential GPS can eliminate the common GPS error components (i.e., δ_{eph}, δ_{ion}, and δ_{trop}), while the multipath and device-noise error components remain dominant in ρ_{code}. Furthermore, the multipath component becomes severely pronounced in multipath environments. In particular, the multipath error has unknown distribution and bias in some situations [25]. Indeed, the GPS localization error ranges from a few meters in open sky environments to over 80 meters in urban canyons.

Data Set Creation

To accomplish the localization accuracy inference task, a total of 6520 GPS measurements have been collected as a data set using a vehicle equipped with a standard GPS receiver (*SiRFstarIII Evaluation Receiver*). The vehicle is driven in various environmental conditions so as to obtain a comprehensive insight into the behaviour of GPS localization error in a wide range of conditions. The data set consists of approximately two equal parts. The first part of the data set has been collected in an open sky environment over a segment of the highway 401 between the city of Waterloo and the city of Toronto, Ontario. The other part of the data set has been collected in environments with various degrees of multipath

signals (e.g., suburbs and downtown areas of the city of Toronto). This process will allow for investigating possible relationships between receiver parameters and the measurement conditions.

Parameters of interest are obtained from the GPS receiver measurements, using the National Marine Electronics Association (NMEA) sentences for every location estimate. The extracted parameters include: number of satellites used in the localization process, DOP, the mean and standard deviation of the SNR of the satellites used in the localization process, and the speed of the vehicle reported by the GPS receiver. Table 4.1 depicts the meta data of the extracted parameters.

Table 4.1: Meta Data of the gathered GPS parameters.

Parameter	Statistics		Range	
	Mean	STD	Min	Max
Speed (km)	62.22	53.48	0.02	135.00
σ_{SNR}(db)	13.18	5.36	1.79	31.04
μ_{SNR}(db)	26.67	5.35	11.30	44.40
No. of Sat.	9.80	2.13	4	12
DOP	1.41	0.94	0.80	11.95

For each measurement, the traced road segment is known, and thus is considered as a ground truth. The error in the location estimate is defined as the shortest distance between the reported GPS location estimate and the traced road segment (Figure 4.8).

This error computation ignores error components along the road direction. However, these components tend to be small compared to the amount of error nominally experienced orthogonal to the vehicle direction. This characteristic of the GPS error is due to the constellation of the GPS satellites and the high-rise buildings on the sides of the road segment [89], where satellite signals reach the receiver after bouncing off buildings in the surrounding. The error computed for each sample is mapped to an accuracy class (i.e,

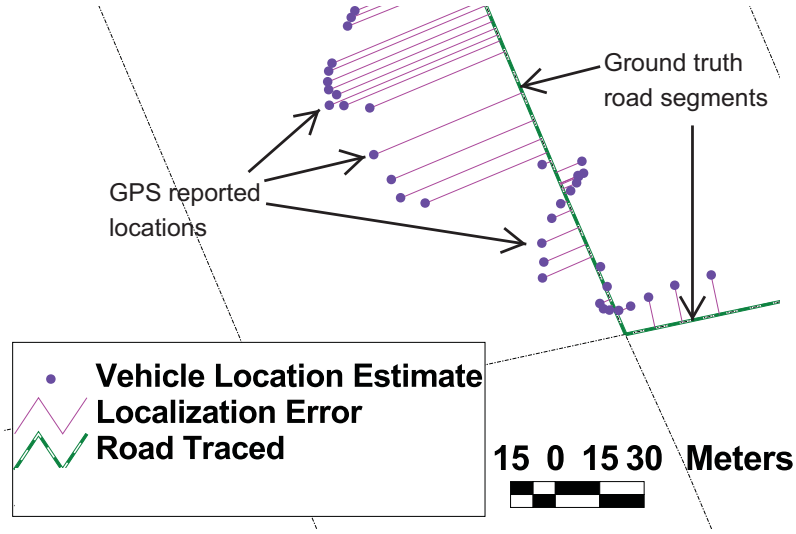


Figure 4.8: GPS Localization Error.

Accurate, *Marginally-Accurate*, and *Inaccurate*). In what follows we describe the approach we used to analyze the measurement samples collected, to construct the accuracy feature space.

GPS Localization Error Analysis

Principle Component Analysis (PCA) is an orthogonal transformation that converts a set of observations of possibly correlated variables into a set of values of uncorrelated variables called *principle components*. The first principal component is the one with maximum variance, and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to (uncorrelated with) the preceding components [105]. Based on the assumption that the location measurements are jointly normally distributed, PCA is applied to the measurements to determine the three most prominent GPS parameters. Table 4.2 depicts the GPS parameters and their significance after applying PCA. It can be seen that the most indicative parameters are speed, Signal to Noise Ratio mean

(i.e., μ_{SNR}) and Signal to Noise Ratio standard deviation (i.e, σ_{SNR}).

Table 4.2: The basis of the PCA space versus the GPS parameters' significance.

	PCA-1	PCA-2	PCA-3	PCA-4	PCA-5
Standard Deviation	53.60	6.10	2.94	1.58	0.56
Speed	0.998	-0.055	0.033	-0.028	0.002
σ_{SNR}	0.024	0.803	0.594	-0.031	0.006
μ_{SNR}	0.059	0.592	-0.803	-0.016	-0.006
No. of Sat.	0.028	0.031	0.009	0.923	-0.382
DOP	-0.009	-0.011	0.005	-0.382	-0.924

Figure 4.9 depicts the transformed GPS measurements in a space spanned by the first three principle components. It can be seen that the measurements constitute three groups in this three dimensional space.

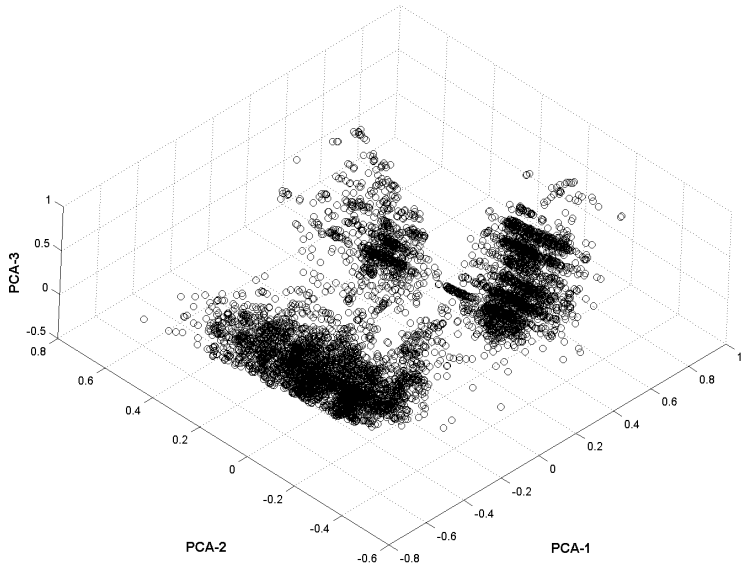


Figure 4.9: Data projection on the first three Principal Components.

A K-means clustering algorithm (K=3, since visual inspection of the measurements in the PCA space reveals a tendency to group into three clusters) [106] is used to segment the

PCA measurement space into three clusters. Figure 4.10 shows the three labelled clusters, where cluster-1 is marked by red 'o', cluster-2 is marked by blue 'x', and cluster-3 is marked by green '+'.

Without loss of generality, we consider the three GPS accuracy bands depicted in Table 4.3. The following analysis can be applied to finer accuracy bands.

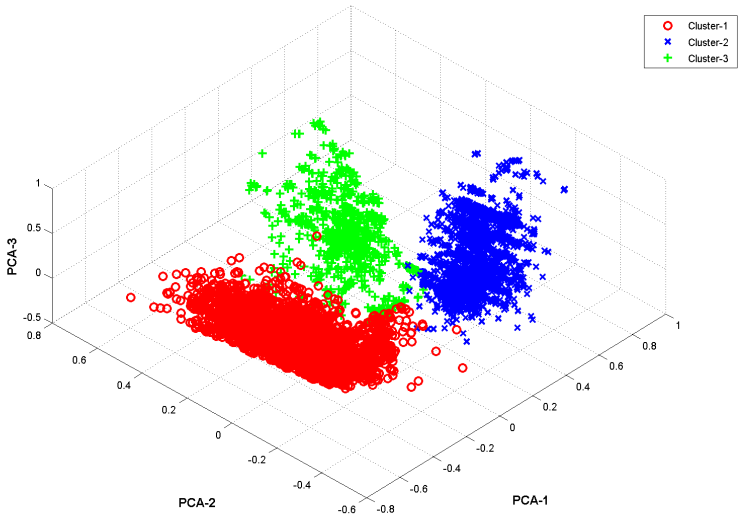


Figure 4.10: Clusters in the PCA feature space.

Table 4.3: Three classes of localization accuracy.

	Accurate	*Marginally-Accurate*	*Inaccurate*
Error Range	0-10 (m)	10-20 (m)	>20 (m)

As stated earlier, the objective is to develop a classification process that can map a GPS measurement to an accuracy class. In order to design such classifier we analyze the three PCA measurement clusters in the context of the desired accuracy bands. Ideally, we would like the accuracy of all measurements of any given cluster to strictly fall into the same accuracy band. Figure 4.11 and Table 4.4 depict the clusters composition of the

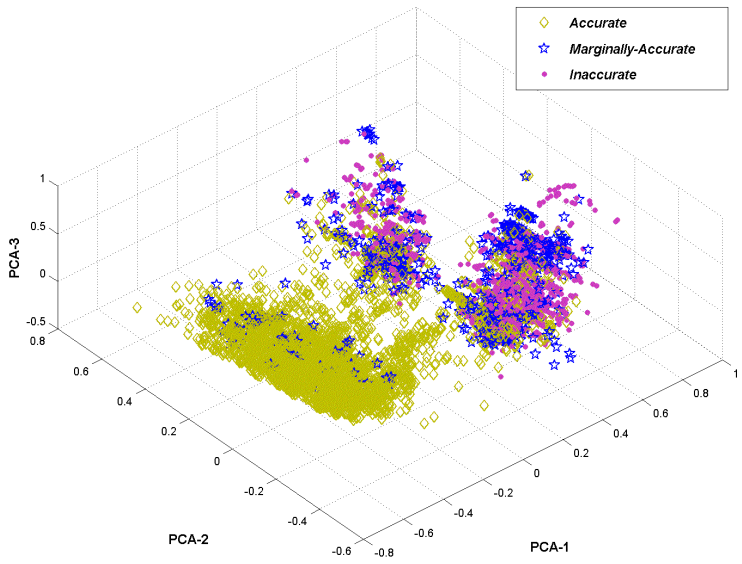


Figure 4.11: Associating accuracy indexes to the clusters in the PCA feature space.

Table 4.4: Associating accuracy indexes to the clusters in the PCA feature space.

	Accurate	*Marginally-Accurate*	*Inaccurate*
Cluster-1	96.91%	3.09%	0.00%
Cluster-2	29.66%	24.26%	46.08%
Cluster-3	52.70%	17.18%	30.12%

three accuracy bands. It can be seen that 96.91% of cluster-1 population belongs to the *Accurate* measurements class, 3.09% belongs to the *Marginally-Accurate* measurements class, and 0.0% of this cluster belongs to the *Inaccurate* measurements class. However, the population of both cluster-2 and cluster-3 consists of a mix of measurements belong to the three accuracy classes with various degrees. Therefore, given this distribution of the three accuracy bands across the three clusters, error in determining the accuracy class of a GPS measurement is expected to be high if the measurement happens to be from either cone of the two clusters – cluster-2 and cluster-3. Given such cluster decomposition, the class labels set is constructed as $C = \{c_1 : Accurate, c_2 : Marginally\text{-}Accurate, c_3 : Inaccurate\}$.

This overlap between the three accuracy bands (classes) in the parameters space (features) makes it impossible for any linear classifier to assign a GPS measurement to a unique accuracy band based on these five parameters, with acceptable confidence level, despite the fact that these parameters include the three most prominent ones. Kernel based classification methods such as Support Vector Classifiers (SVC) [101] are expected to provide decent performance in such space configuration; nevertheless, the computational requirement of such kernel based classifiers prevents them from being a viable choice in embedded environments such as the ones typically used in GPS systems.

4.4.2 HCCU and CBAC Validation on GPS Measurements

Both the flat and the introduced hierarchical classifiers are employed in this comparative study. First, the setup of the real-life experiment conducted is explained. The base classifier used in the flat and hierarchical classifiers in this study is then described. Comparative results and discussions are presented in the conclusion of this section.

Experiment Setup

A roving vehicle equipped with a standard GPS receiver (*SiRFstarIII Evaluation Receiver*) has been used to collect the measurements for this experiment. A Lenovo SL400 Laptop connected to the GPS receiver over a USB port has been used to store the NMEA sentences to be processed later (offline). The experiment has been carried out over an area that comprises various observation conditions. The GPS error classification under open sky environments has been tested over a segment on highway 401 between the city of Waterloo and the city of Toronto, Ontario. The suburbs and downtown areas of the city of Toronto have been used to test the GPS error classification under adverse environmental conditions.

Base Classifiers Implementation

The base classifiers used to facilitate the comparison among the classification algorithms introduced in this work are chosen to be FeedForward Neural Networks (FFNNet) with back-propagation learning algorithm [107]. FFNNet has the ability to classify linearly inseparable patterns, not to mention its efficiency in designing embedded systems due to its computational complexity. Each FFNNet classifier consists of three layers (i.e., input, hidden, and output layers). The number of nodes in the input layer is equal to the number of features in the pattern of interest. The size of the hidden layer has been determined based on minimizing the error rate using a cross validation method [101]. The size of the hidden layer, which provides the least error rate, is found to be between 20 and 25 nodes. The number of nodes in the output layer is equal to the number of classes of interest. The input layer employs a linear activation function; both the hidden and output layers employ hyperbolic tangent sigmoid activation functions (i.e., tansig function).

Comparative Results

The goal here is to classify the GPS measurements in the data set into different classes of localization accuracy using the two classification algorithms introduced in Section 4.3. To maintain a generalized independent data set evaluation, the training and testing procedures are performed using a tenfold cross-validation approach in which equal size of patterns from each class are randomly selected. The selected patterns are divided into ten partitions (i.e., folds); every time, one partition is used for testing and the remaining partitions are used for training. The results are added up, and then, the classification performances are evaluated. The comparisons are conducted over two *use cases*. In the first *use case*, GPS measurement accuracy classification is approached as a two-class problem, whereby a GPS measurement is to be classified as either highly accurate (i.e, *Accurate* band), or inaccurate

(in this case it views the two bands *Marginally-Accurate* and *Inaccurate* as one band). Such *use case* is relevant in applications that demand highly accurate location estimates, such as refining and updating digital maps and landmarks' localization. Since only two classes are considered in this *use case*, the performance of a flat classifier is compared with that of CBAC. In *Use Case*-2 the GPS accuracy classifications is approached as a three-class problem. The output of the classification process should be one of three classes: *Accurate*, *Marginally-Accurate* and *Inaccurate*. This *use case* is relevant in applications where *Marginally-Accurate* GPS measurements meet the required accuracy to the extent the system can achieve higher availability. Such achievement can significantly improve the fusion techniques' performance as presented in Chapter 6.

The evaluation of the classification performance is based on the classification Error Rate ϵ. Two types of error rate are used in the comparison: 1) Class Error Rate, and 2) Overall Error Rate. Class c_i Error Rate is defined as

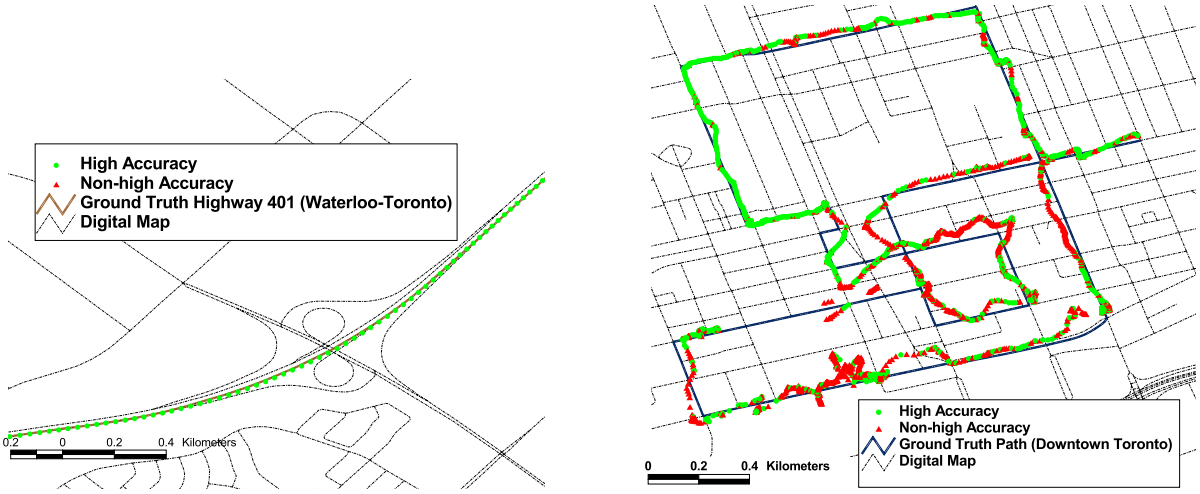
$$\epsilon_{c_i} = \frac{\#\text{ of misclassified patterns} \in c_i}{\#\text{ of patterns} \in c_i} \times 100$$

The overall Error Rate is defined as

$$\epsilon = \frac{\sum_{c_i \in C} \#\text{ of misclassified patterns} \in c_i}{\sum_{c_i \in C} \#\text{ of patterns} \in c_i} \times 100$$

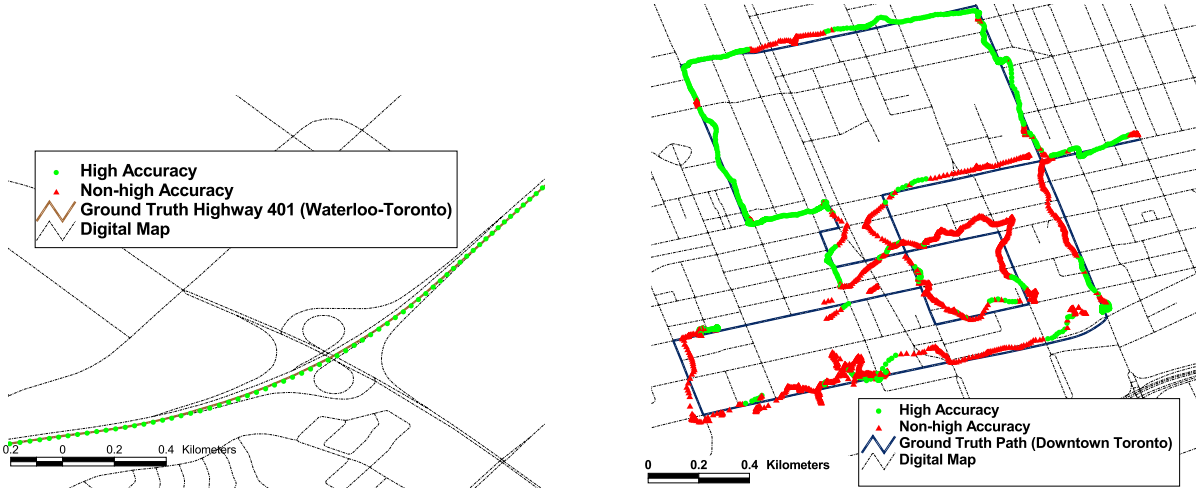
Results for *Use Case*-1

The performance of the flat classifier is shown in Figure 4.12 for various observation conditions: (a) Open Sky environment, and (b) Canyon environment. Figure 4.13 depicts the performance of CBAC for the two classes over the same observation conditions. It can be seen that both classifiers perform well in Open Sky environments (e.g, highway 401). On the other hand, CBAC performed consistently well even in canyon areas. This was not the case for the flat.



(a) (b)

Figure 4.12: Accurate and inaccurate location estimate classification performance using FFNNet. (a) Open Sky environment. (b) Urban canyon environment.



(a) (b)

Figure 4.13: Accurate and inaccurate location estimate classification performance using CBAC. (a) Open Sky environment. (b) Urban canyon environment.

The error rates are calculated using a tenfold cross-validation on 1000 measurements from each class. Table 4.5 shows a gain of 12% of the overall error rate by using CBAC. Table 4.6 presents the confusion matrix for *Use Case*-1. As can be seen, the error rate per class is improved for both classes.

Table 4.5: High accuracy versus non-high accuracy classification.

	Error Rate		
	$\epsilon_{Accurate}$	$\epsilon_{Inaccurate}$	ϵ
Flat	14.80%	9.30%	12.05%
CBAC	13.90%	7.30%	10.60%

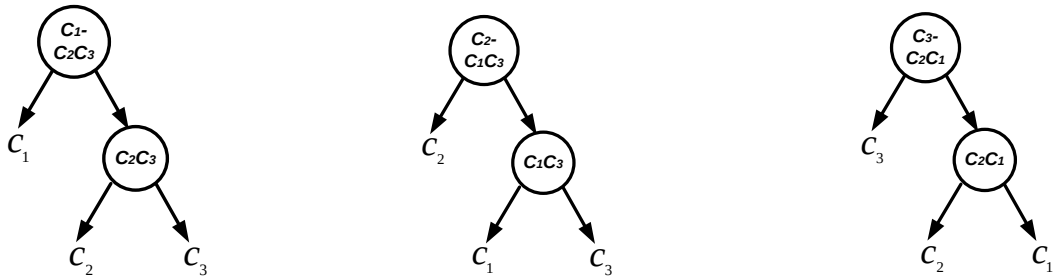
Table 4.6: Confusion matrix for *Use Case*-1.

		Classification	
Actual	Classifier	Accurate	Inaccurate
Accurate	Flat	852	148
	CBAC	861	139
Inaccurate	Flat	93	907
	CBAC	73	927

Results for *Use Case*-2

In this case, the classification task is more challenging as the patterns are to be classified into one of three classes, namely, *Accurate*, *Marginally-Accurate* and *Inaccurate* (i.e., c_1, c_2 and c_3, respectively). Due to the difficulty of obtaining multipath environment measurements, a tenfold cross-validation is applied on 800 measurements from each class.

The effectiveness of HCCU is first presented by comparing its error rate with that of a flat and other hierarchical classifiers. The hierarchical classifiers used in this comparison are shown in Figure 4.14. It can be seen in Table 4.7 that HCCU outperforms non-similarity guided classifiers. It is worth mentioning here that if a hierarchical classifier is



(a) Similarity guided ($\hbar_{1\_23}$). (b) Non-similarity guided ($\hbar_{2\_13}$). (c) Non-similarity guided ($\hbar_{3\_12}$).

Figure 4.14: Possible structures of a three classes hierarchical classifier.

Table 4.7: HCCU vs. non-similarity guided classifiers.

Overall Error Rate			
Flat	$\hbar_{2\_13}$	$\hbar_{3\_12}$	HCCU ($\hbar_{1\_23}$)
26.96%	28.04%	26.04%	25.33%

not structured carefully, its performance can be worse than that of a flat classifier (e.g., the performance of $\hbar_{2\_13}$).

Next comparison is between the performance of a flat classifier and that of the introduced classification algorithms. In the first implementation, the proposed hierarchical classifier is implemented without contextual information (i.e., HCCU). In the second implementation CBAC is augmented with contextual information.

The error rate of the classifiers used in *Use Case*-2 are also evaluated using a tenfold cross-validation. Table 4.8 depicts the class error rates and the overall error rate for the three classifiers (flat, HCCU, and CBAC), in addition to other conventional classifiers. It is clear that HCCU improves the classification performance compared with the flat classifier. Furthermore, CBAC performs the best where it achieves 17.77% performance improvement compared with the flat classifier.

The confusion matrix of the three classifiers is shown in Table 4.9. As to the confusion

Table 4.8: Three levels of accuracy performance classification.

Classifiers	**Error Rate**			
	ϵ_{c_1}	ϵ_{c_2}	ϵ_{c_3}	ϵ
RBF Network	26.88%	54.01%	23.00%	34.63%
Bayesian Network	22.12%	56.50%	24.13%	34.25%
Trees-J48	20.50%	37.90%	28.24%	28.88%
Flat	21.00%	38.38%	21.50%	26.96%
HCCU	19.63%	35.75%	20.63%	25.33%
CBAC	17.63%	32.63%	16.25%	22.17%

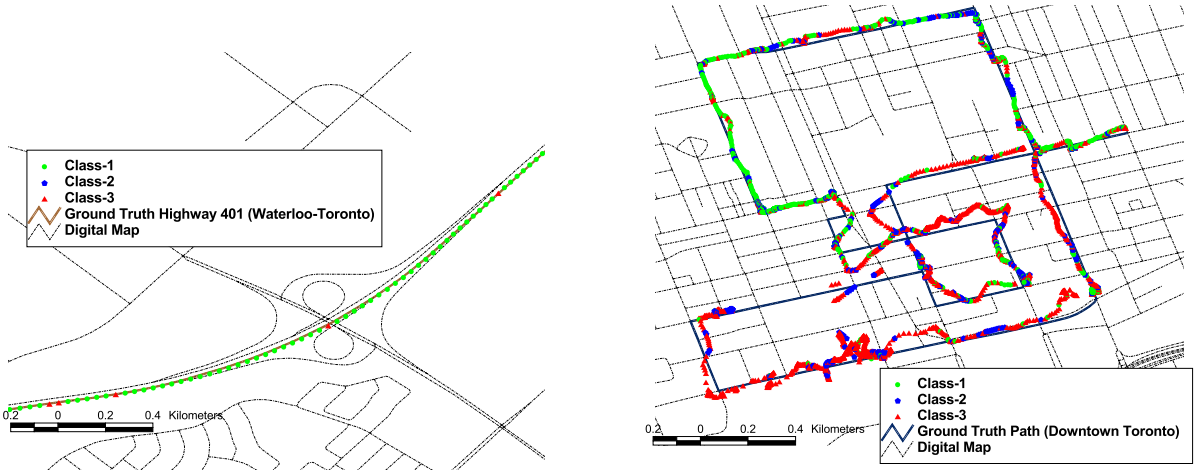
Table 4.9: Confusion matrix for *Use Case*-2.

		Classification		
Actual	Classifier	c_1	c_2	c_3
c_1	Flat	632	106	62
	HCCU	643	91	66
	CBAC	659	74	67
c_2	Flat	94	493	213
	HCCU	78	514	208
	CBAC	86	539	175
c_3	Flat	7	165	628
	HCCU	3	162	635
	CBAC	2	128	670

matrix, for c_1, CBAC fails to improve the misclassification of c_1 as c_3 by 8%. This shortcoming is recovered by the 34.45% improvement in classifying c_1 as c_1 or c_2. Classifying c_2 and c_3 is improved using CBAC in all aspects. This improvement comes at the cost of memory space, where CBAC needs $2 \times (n-1)$ multiples of a space needed by a flat classifier to classify n classes.

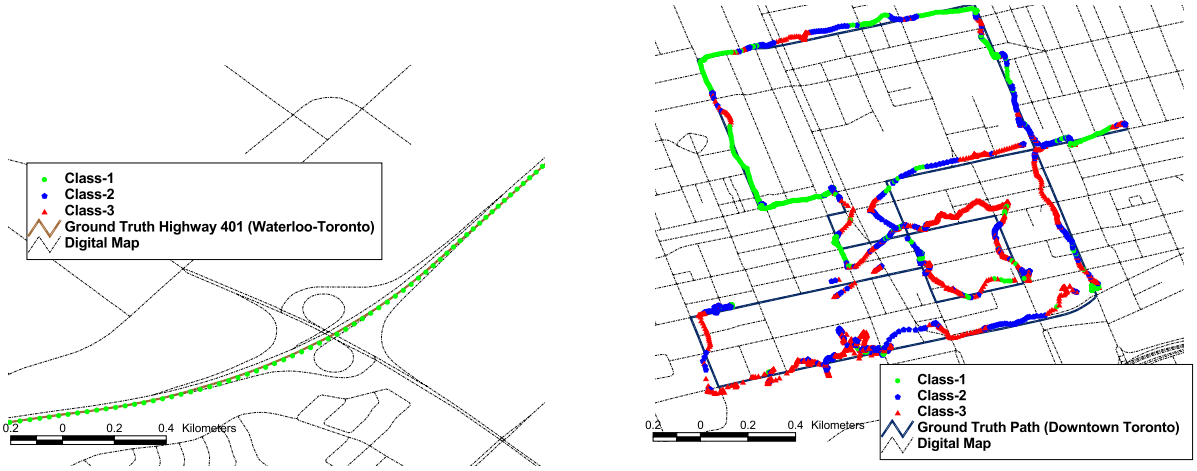
Figure 4.15 shows the performance of the HCCU, where Figure 4.16 depicts the per-

formance of CBAC under the same measurement conditions.



(a) (b)

Figure 4.15: Classifying the three levels of accuracy using HCCU.(a) Open Sky environment.(b) Urban canyon environment



(a) (b)

Figure 4.16: Classifying the three levels of accuracy using CBAC.(a) Open Sky environment.(b) Urban canyon environment

4.5 Case Study 2: FP Accuracy Classification

FP localization is selected in this study over other localization methods due to its cost effectiveness in terms of computational power and hardware requirements, in addition to its robustness to multipath environments [61,62,66,108]. Furthermore, FP is not subjected to limitations in either the magnitude of the localization error (e.g., trilateration-based localization), the accumulated error (e.g., DRS), or the signal availability (e.g., GPS).

The following section explains the affecting parameters on the FP localization errors. These parameters can be collectively used as features to infer FP localization accuracy.

4.5.1 Error in FP Location Estimate

FP localization is a pattern recognition/matching method in which parameters of interest, such as RSS, are measured in real-time and compared with similar type of measurements for pre-defined locations to estimate the location of a mobile object. The challenging aspect of this method is assuring that the system can distinguish between similar patterns that represent different locations. Such similar patterns, for example, typically exist when weak signals are received from remote BSs.

Typically, FP localization utilizes parameters obtained from cellular networks due to the high availability of their radio signals in indoor and outdoor environments. In cellular networks, the different types of towers and network architectures contribute the most to the discrepancy in degrees of localization accuracy. For example, macrocells of about 3 km radii are typically deployed in rural areas and over intercity highways, while microcells of about 300 m radii are commonly deployed in urban areas to serve more customers per unit area [57, 109]. As a result, urban cellular networks are characterized by their dense BSs

and strong RSS compared with those of rural and intercity areas – a condition that helps produce more accurate FP localization in urban areas.

4.5.2 Generation of FP Localization Data Set

In this case study, I maintain the hypothesis that the number of detected BSs and RSS can be used to infer FP localization accuracy. This hypothesis is validated here empirically by applying CBAC on a data set consisting of FP localization errors associated with measurements used in the localization process (i.e., number of BSs and RSS). This data set is created as a vehicle travels from the city of Waterloo to the downtown of the city of Toronto via a segment of highway 401. The simulation steps are depicted in Figure 4.17. The details of each step of the data set creation are presented next.

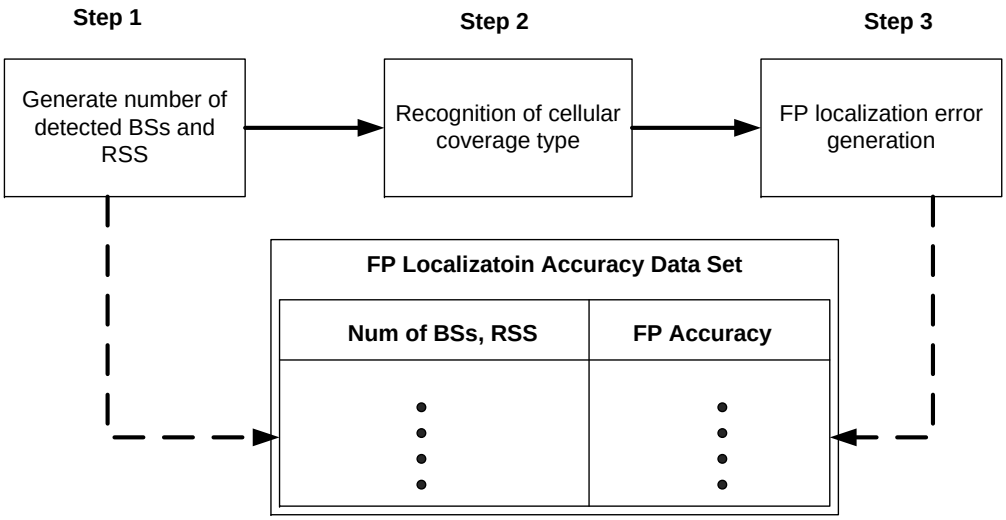


Figure 4.17: FP localization data set creation.

Step 1: Generation of Number of Detected BSs and RSS

The simulation scenario involves a vehicle travels through roads covered by cellular networks with various degrees of BSs' density. The density of the cellular networks is simulated based on Canadian cellular towers map provided in [1], Figure 4.18.

Figure 4.18: Canadian cellular towers map [1].

Towers of Rogers cellular network are considered in this simulation due to their high degree of availability, compared with other companies' networks, over the area of interest. The density of the cellular network is categorized into four categories, namely, macrocells density (i.e., over highway 401), microcell low-density (i.e., over small cities), microcell medium-density (i.e., over Toronto suburb areas), and microcell high-density (i.e., over downtown of the city of Toronto). The cellular network measurements, number of detected BSs and RSS, are generated randomly using the distributions depicted in Table 4.10, which are elicited from the information provided in [1] and [57].

Table 4.10: Probability distributions used in the simulation.

Category	Measurements	Distribution Type	Distribution Parameters
macrocells density	Num. of BSs	Discrete Uniform	n = 2, a = 1, b = 2
	RSS	Gaussian	μ = 5 db, σ = 10 db
microcell low-density	Num. of BSs	Discrete Uniform	n = 3, a = 2, b = 4
	RSS	Gaussian	μ = 15 db, σ = 10 db
microcell medium-density	Num. of BSs	Discrete Uniform	n = 3, a = 3, b = 5
	RSS	Gaussian	μ = 35 db, σ = 7 db
microcell high-density	Num. of BSs	Discrete Uniform	n = 5, a = 4, b = 8
	RSS	Gaussian	μ = 45 db, σ = 7 db

Step 2: Recognition of Cellular Coverage Type

This step simulates a typical behaviour of an FP localization technique. As the information about the number of detected BSs and RSS is captured by an MS, an FP localization technique is able to estimate the location of that MS. However, in simulation, the mapping process from network measurements to location estimate can be misleading. The reason is that the magnitude of the location estimate error is often based on the category of the covering cellular network. In other words, it is possible to map similar measurements from different network categories to contrasted localization errors [57]. This ambiguous situation results from the similarity in the signals' interference/attenuation levels affected by omni and directional antennas, receiving distance, and the existance of LOS/NLOS. Thus, the recognition of the cellular network category is pivotal in simulating realistic FP localization errors.

To handle such the linguistic ambiguous measurements (e.g., strong-week RSS, and low-medium-large number of BSs) a Fuzzy Inferencing System (FIS), Mamdani model [110], is used to produce a crisp value between 0 and 1. The crisp value is used to decide on

the network category from which the measurement is obtained. The decision is made by dividing the range between 0 and 1 into four pins where each pin represent specific category, 4.19. The pin where the crisp value falls is selected as the recognized network category.

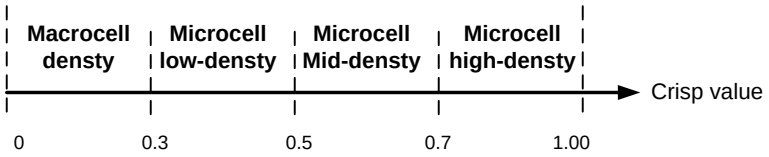


Figure 4.19: Network category decision making.

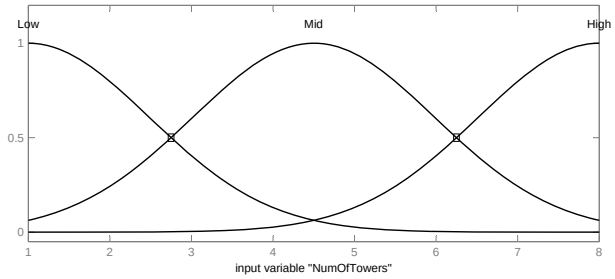
The FIS is designed so that the engineering knowledge of the measurement conditions is embedded in the FIS in a form of knowledge base. This knowledge base consists of a number of rules, such as the following

$$IF\ x\ is\ A_i\ AND\ y\ is\ B_i\ THEN\ z\ is\ C_i$$

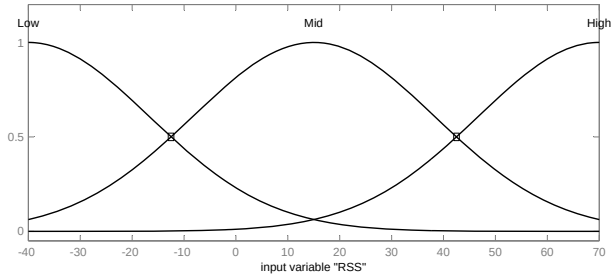
where x and y signifies fuzzified measurements of interest, and z signifies the output variable; A_i, B_i, and C_i are fuzzy sets introduced by Zadeh in [111] that governs the i^{th} rule. The fuzzification of the measurements and evaluation of the rules are performed using S-Norm operator of the FIS. The aggregation step is performed using summation method to maintain the contribution of each membership to the sets of output variable. The decision crisp value is computed by defuzzifying the aggregated output using mean of maximum method [112], which selects the set with the highest membership.

Figure 4.20 depicts the membership functions of fuzzy sets belonging to each input variable in the FIS, namely, Number of BSs and RSS. The membership functions of fuzzy sets belonging to the output of the FIS are shown in Figure 4.20. The rules embedded in the knowledge base are presented in Table 4.11. Furthermore, as an example, Figure 4.22 depicts a scenario of having 2 detected BSs and RSS equals to 20db. Accordingly, the

corresponding crisp value equals 0.26 which signifies macrocell density over highway.



(a) Number of detected BSs.



(b) Received signal strength in db.

Figure 4.20: Membership functions used in the FIS for the fuzzy sets of the input variables.

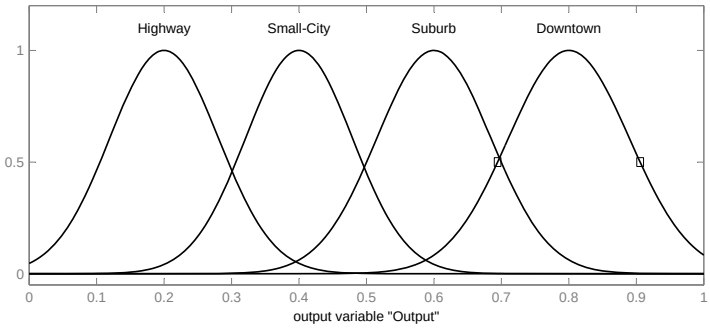


Figure 4.21: Membership functions used in the FIS for the sets of the output variable.

Table 4.11: FIS knowledge base rules.

RSS / Num of BSs	Low	Medium	High
Low	Highway	Highway	Suburb
Medium	Highway	Small city	Downtown
High	Small city	Suburb	Downtown

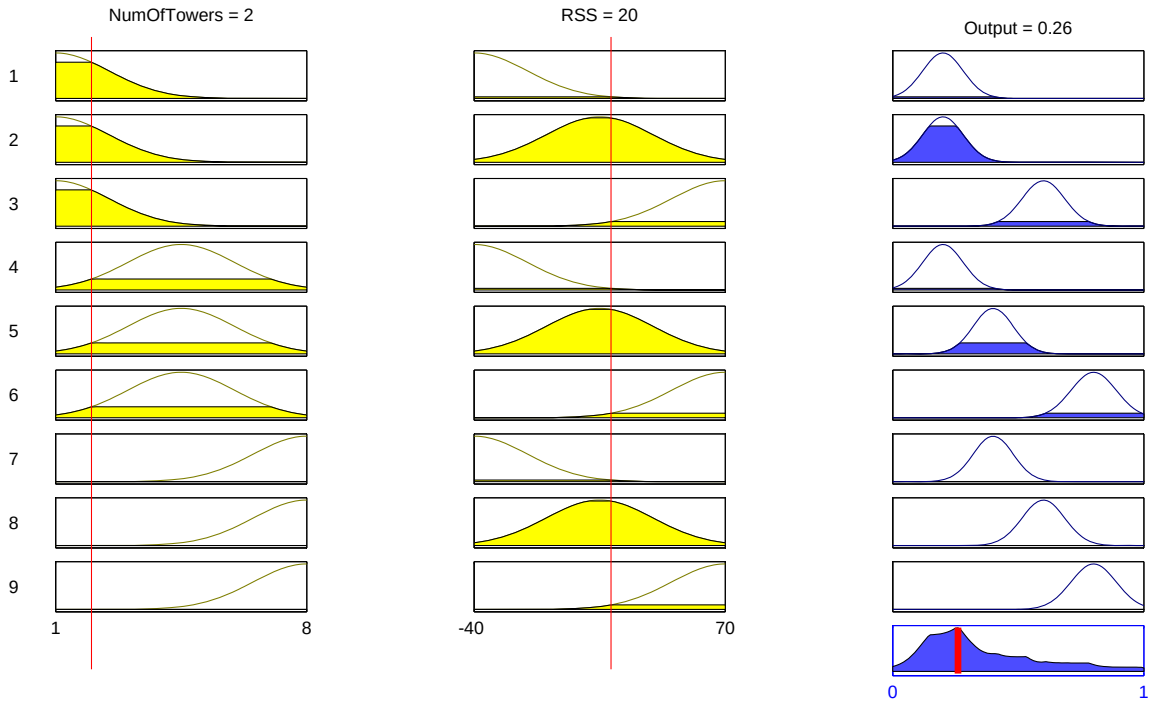


Figure 4.22: Example of producing crisp value for a network measurements.

Step 3: Generation of FP Localization Errors

The magnitude of the localization error is generated using independent and identical zero-mean-Gaussian distributed random variable for each network category. Figure 4.23 depicts the distributions used for each network category. Once the magnitude of the localization error, $|Loc_{err}|$, is generated then the location estimate of the FP,$Z_{FP,t}$, is computed by

adding $|Loc_{err}|$ to the actual location of the vehicle, X_t as in equation 4.8

$$Z_{FP,t} = X_t + \begin{bmatrix} |Loc_{err}|\cos(\gamma) \\ |Loc_{err}|\sin(\gamma) \end{bmatrix} \tag{4.8}$$

where γ is uniformly distributed between 0 and 2π. This completes the creation of the FP localization dataset.

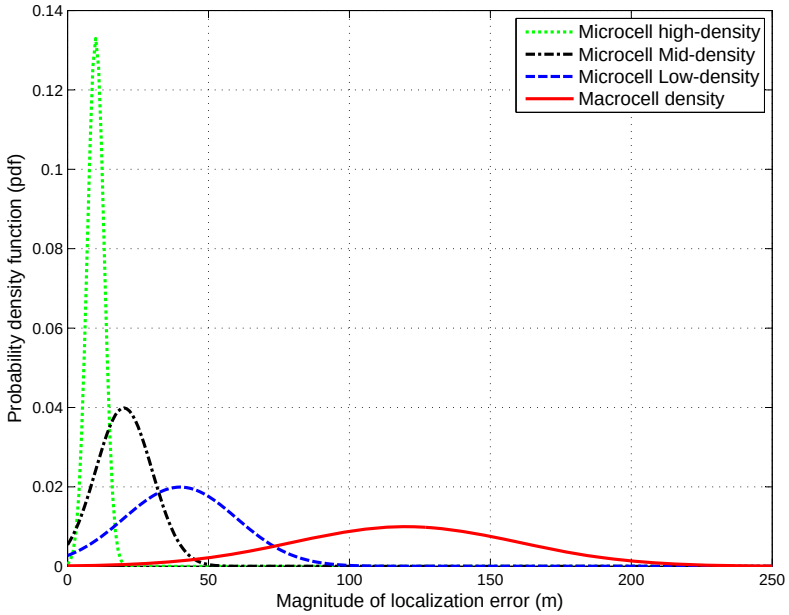
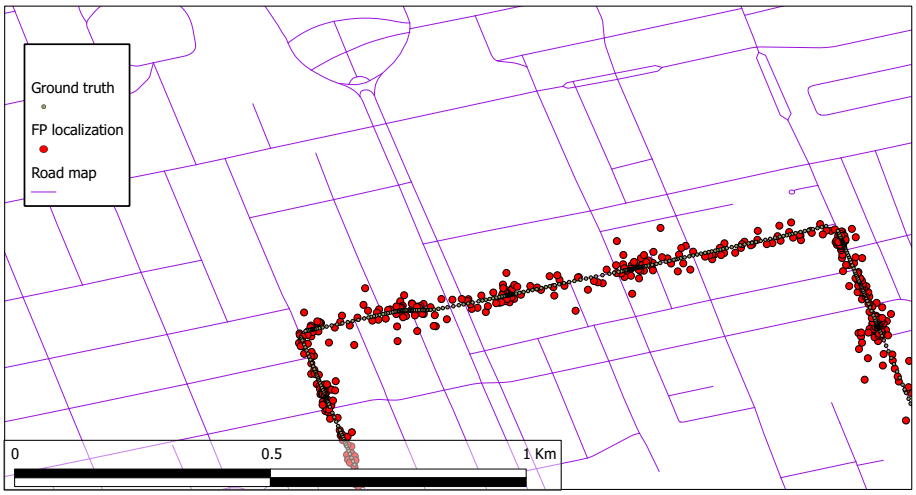
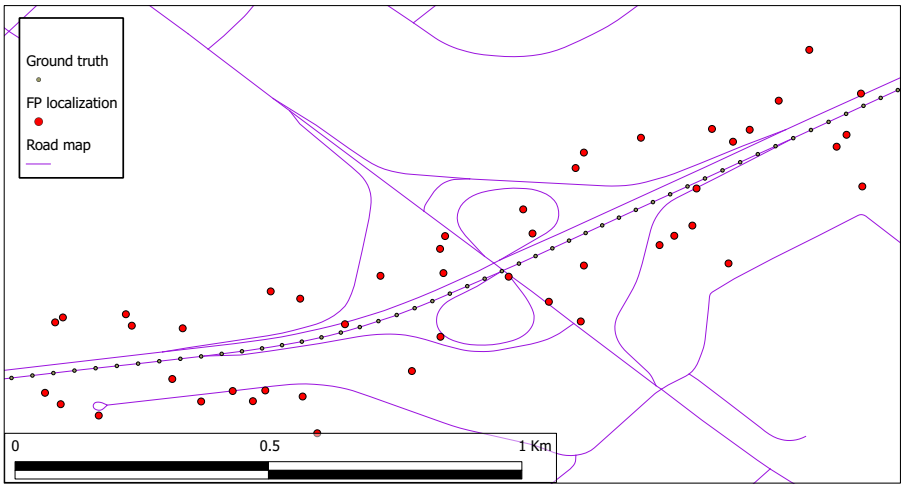


Figure 4.23: FP localization error distribution per network category.

The FP technique simulated above is executed to locate a vehicle using real traces over different areas with various measurement conditions. The location estimate is computed every 1 second providing an actual location for each estimate. Figure 4.24 shows the output of the location estimation in two different measurement conditions. In the following section, the classification algorithms developed in this chapter are implemented on the output of the FP localization to assess the localization accuracy.



(a)



(b)

Figure 4.24: FP localization simulation in various measurement conditions.(a) Downtown Toronto. (b) A segment of highway 401, Ontario

4.5.3 CBAC Validation on FP Measurements

As seen in the previous section, FP localization can relatively perform quite well in urban canyons since it accommodates well multipath environments [109, 113, 114]. Hence, FP

technique is a good candidate for a hybrid localization system that uses GPS technology. In this section, I will implement the CBAC algorithm on the FP data set to validate the efficacy of the algorithm on FP measurements. FP data set will be utilized, in the following chapters, as another source of location information for the hybrid vehicle localization system.

Experiment Setup

The traces of the vehicle used in Case Study 1 are used as ground truth. The FP localization simulator is used to generate location estimates with various accuracies in the different measurement conditions. Without loss of generality, two bands of localization accuracy are considered in this case study, namely, c_1 for errors less than or equal to 20 meters, and c_2 for errors greater than 20 meters. Figure 4.25 depicts the overlap (i.e., class similarity) between the two accuracy classes.

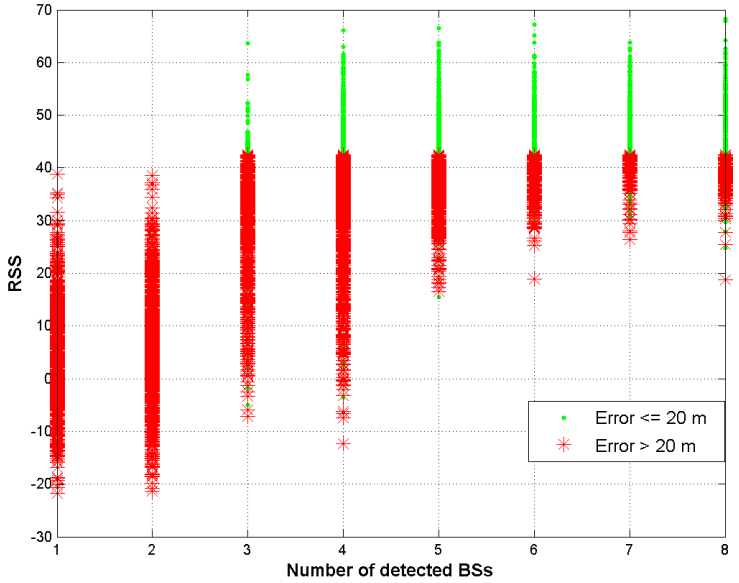


Figure 4.25: Classes overlap in the feature space of the FP measurements.

Comparative Results

FFNNet base classifiers are used in the flat and CBAC classifiers. The training and testing phases are conducted using a tenfold cross-validation method where ensemble size equals to 2500 pattern per each class. Three other conventional classification methods are used to classify the same patterns using 10-fold cross-validation. Table 4.12 shows the classification error rate of each classifier. It can be seen that CBAC outperforms the other classifiers. The confusion matrices of the flat classifier and CBAC, in Table 4.13, show slight improvement using CBAC; however, both have significantly improve the classification process compared with the other classification methods.

Table 4.12: FP accuracy classification performance.

Classifiers	**Error Rate**		
	ϵ_{c_1}	ϵ_{c_2}	ϵ
RBF Network	12.76%	13.04%	12.90%
Bayesian Network	11.76%	9.28%	10.52%
Trees-J48	13.60%	5.08%	9.34%
Flat	13.56%	4.88%	9.22%
CBAC	13.36%	5.04%	9.20%

Table 4.13: Confusion matrix of CBAC versus flat classifier.

		Classification	
Actual	Classifier	c_1	c_2
c_1	Flat	2161	339
	CBAC	2166	334
c_2	Flat	122	2378
	CBAC	126	2374

4.6 Summary

This chapter deals with localization accuracy classification. The objective here is to provide localization systems with a methodology for evaluating the accuracy of their location estimate so as to allow the target application to intelligently process such estimates in its decision making strategy. An analysis of the localization error of two typical localization techniques in various environments is provided. The relationship between the localization measurements and the error in its location estimation is investigated.

The chapter proposes a hierarchical classification strategy to address the class ambiguity problem via class unfolding approach (HCCU). The proposed HCCU strategy is proven to be superior with respect to other hierarchical configuration. Furthermore, in this chapter, a Context Based Accuracy Classification (CBAC) algorithm is proposed to enhance the performance of the classification process. In this algorithm, knowledge about the surrounding environment is utilized to optimize classification performance as a function of the observation conditions. The chapter presents experimental results to validate the effectiveness of the proposed accuracy classification algorithms.

The strong correlation and overlap between localization process measurements obtained under different measurement conditions make the accuracy classification an extremely hard problem and uncertain in many scenarios. It is pivotal for the target applications to obtain an integrity measure that assess the degree of trust placed in such a decision (i.e., location and accuracy estimates) in order to allow for intelligent fusion process. Next chapter will investigate the provision of localization integrity.

Chapter 5

Modeling the Impact of Observation Conditions on Localization Integrity

5.1 Introduction

It is typical in vehicle localization that location measurements obtained from multiple sensors (e.g, GPS, Vision, Inertial, etc) are combined together to compute accurate vehicle location. However, such improved accuracy can only be attained under nominal observation conditions. For example, some researchers have proposed augmenting GPS localization systems with vision sensors [115]. The vision sensors would use visual features to determine the location of the object. Nevertheless, a situation may ensue in which the vision sensors' localization accuracy deteriorates (e.g., under occlusion and/or poor lighting). Thus, it is obvious that for such multi-sensor system to achieve reliable performance consistently, it must determine the impact of the observation conditions on both sensors. To achieve this task, a degree of trust (i.e., integrity) in every estimate must be determined and utilized in real-time during the fusion process.

If one can deduce the observation conditions under which a technique's estimate was made, engineering knowledge about the technique in the context of this observation con-

dition can help in quantifying the integrity of the technique in producing the required accuracy [98, 116]. Such integrity deduction process, nevertheless, is itself uncertain due to the complexity and the inherent ill-posedness of the localization measurement process. I maintain that a multi-state stochastic model can be used to synthesize the dynamics of such uncertainty. The states of this model represent possible localization performance under certain observation conditions; the state transitions of this model represent the dynamics of the environment in changing from one state to another (i.e., from one condition to another). Therefore, a Markovian model constitutes a viable option for modeling this stochastic process, as the Markovian model has an inherent measure for quantifying the probability of being in a subset of its states [97].

In this chapter, a Markovian model is proposed to capture the impact of observation conditions on the localization performance and to consequently determine a reliability index with respect to the localization accuracy claimed by a technique. In this research work, the reliability index quantifies the integrity of a localization technique.

The remainder of this chapter is organized as follows. The derivation of the Markovian model is introduced in Section 5.2. In Section 5.3, the derived Markovian model is used to construct two instances of localization techniques. Localization quality indexes for the two techniques are computed and a variety of road segment scenario simulations are presented in Section 5.4. Section 5.5, presents a real-time adaptive Markovian model using non-parametric transition probability distributions. Finally, concluding remarks are provided in Section 5.6.

5.2 Task Driven Integrity

The proposed model captures the stochastic uncertainty in sensory measurements as it pertains to the observation conditions, and the impact of these conditions on the accuracy of location estimates. The point of interest here is the association between the observation conditions and the quality of decisions made based on any sensory measurements produced under such conditions. The hypothesis in this case is that since the vehicle is expected to be in similar environmental conditions frequently, the stochastic uncertainty of the localization process can be modeled as a set of states. The transitions among these states are governed by a probabilistic behaviour. This probabilistic behaviour can be learned and used to predict the accuracy of the system in localizing objects in its environment based on the likelihood of the environment being at a given state.

Markovian model analysis can yield a variety of useful performance measures that describe the operation of systems. These performance measures include, but are not limited to, system reliability, mean time to failure, mean time between failures, and probability of being in a given state at a given time.

A discrete Markovian model is therefore used in this work to provide a reliability measure for location estimation provided by vehicle localization systems. The derivation details of the proposed model and the assumptions made in this work are presented next.

5.2.1 Assumptions and Notation

Consider a localization technique that produces location estimates with accuracy Acc which takes value in $[Acc_1, Acc_2, \ldots, Acc_n]$, depending on the measurement conditions at which the estimate is made. The objective here is to model the localization accuracy of that technique as it changes from one level to another. A Markovian model with n states is

therefore used. Let state S_i, $\{S_i | i \in \{1, 2, \ldots, n\}\}$, of the Markovian model represent a localization accuracy Acc_i. To compute the reliability of a location estimate meeting a certain level of accuracy, one needs to know the average time, τ, during which the model does not visit states of lesser accuracy. In probability theory, this average time is referred to as the mean first-passage time [117]. The reciprocal of τ is known as the failure rate: a failure stat is a state that represents a level of accuracy below that required by the application task. Reliability can then be measured using the failure rate, as follows:

$$Reliability = 1 - failure\ rate \tag{5.1}$$

In the following we present the computation of the mean first-passage time which is used in Equation 5.1 to calculate the reliability measure of delivering accuracy Acc performance.

5.2.2 The Mean First-Passage Time Computation

Without loss of generality, let $\mathbf{S} = \{S_1, S_2, \ldots, S_n\}$ be an ordered list of a Markovian model states. Each state represents an accuracy level and the states are ordered in a descending order as a function of accuracy, i.e, best-to-worst: $\Theta = \{Acc_1, Acc_2, \ldots, Acc_n\}$, where $Acc_{i-1} \leqslant Acc_i$ (Figure 5.1). Let ϑ be the target accuracy of a location estimate. As depicted in Fig-

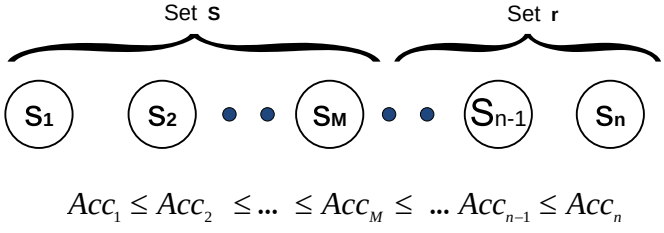


Figure 5.1: States of a Markovian model ordered according to the representation of accuracy from the best-to-the worst.

ure 5.1, the states are divided into two sets: $\mathbf{s} = \{S_i | Acc_i \leqslant \vartheta\}$ and $\mathbf{r} = \{S_i | Acc_i > \vartheta\}$.

Let S_M, $S_M \in \mathbf{s}$, be the state with the largest accuracy index (i.e., $Acc_M \geqslant Acc_i$, $\forall s_i \in \mathbf{s}$). The goal is then to compute the mean first-passage time for the set $\mathbf{r}$ (i.e., the mean return time to any state $S_i \in \mathbf{r}$). Therefore, a new Markovian model is generated from the original Markovian model as follows. The states $S_i \in \mathbf{r}$ are collapsed into one failure state S_F, The number of states in the resultant Markovian model is reduced to $M+1$: $\tilde{\mathbf{S}} = \{S_1, S_2, \ldots, S_M, S_{M+1} = S_F\}$. The transition probabilities of the new Markovian model $\tilde{\mathbf{S}}$ are computed as follows.

Definition 5.2.1. In a Markovian model, the limiting probability, $\overline{\pi}$, is the probability distribution of the model being in any state when it reaches statistical equilibrium, regardless of the initial state.

Definition 5.2.2. In a Markovian model, the transition probability, P_{ij}, is the probability that the next state is S_j, given that the current state is S_i (i.e., $P(S_j/S_i)$).

The limiting probability can be computed by solving the following linear system of equations [117]:

$$\overline{\pi} = \overline{\pi}\mathbf{P} \tag{5.2}$$

$$\sum_{i=1}^{n} \pi_i = 1 \tag{5.3}$$

where

$$\overline{\pi} = [\pi_1, \ldots, \pi_n]$$

and

$$\mathbf{P} = \begin{bmatrix} P_{11} & \ldots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \ldots & P_{nn} \end{bmatrix}$$

Once the values of $\overline{\pi}$ are computed, the limiting probability of state S_F, π_{M+1}, is

calculated as follows

$$\pi_{(M+1)} = \sum_{i=M+1}^{n} \pi_i \tag{5.4}$$

The limiting probabilities $(\pi_1, \ldots, \pi_M)$ do not change. Hence, the transition probabilities $P_{(M+1)i}$, $P_{(M+1)(M+1)}$, and $P_{i(M+1)}$ in the resultant Markovian model must be calculated as follows:

$$P_{(M+1)i} = \frac{\pi_i - \sum_{j=1}^{M} \pi_j P_{ji}}{\pi_{(M+1)}} \quad , where\ i = 1, \ldots,\ M \tag{5.5}$$

$$P_{(M+1)(M+1)} = 1 - \sum_{j=1}^{M} P_{(M+1)j} \tag{5.6}$$

$$P_{i(M+1)} = 1 - \sum_{j=1}^{M} P_{ij} \quad , where\ i = 1, \ldots,\ M \tag{5.7}$$

τ can be computed as the time the model spends in states $S_i \in \mathbf{s}$, before transitioning to state S_F

$$\tau = 1 + \sum_{j=1}^{M} P_{(M+1)j} T_{j(M+1)} \tag{5.8}$$

where $P_{(M+1)j}$ is the transition probability from state S_F to state $S_j \in \mathbf{s}$ and $T_{j(M+1)}$ is the mean visit time to transition from $S_j \in \mathbf{s}$ to S_F. This mean visit time $T_{j(M+1)}$ is a unique solution for the following set of linear system of equations:

$$T_{j(M+1)} = 1 + \sum_{i=1}^{M} P_{ji} T_{i(M+1)}, \quad j = 1, \ldots, M \tag{5.9}$$

which can be expressed in a matrix form as follows.

$$\mathbf{T} = \mathbf{e} + \mathbf{QT} \tag{5.10}$$

and consequently,

$$\mathbf{T} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{e} \tag{5.11}$$

where $\mathbf{I}$ is the identity matrix, $\mathbf{T} = \left[T_{1(M+1)}, T_{2(M+1)}, \ldots, T_{M(M+1)}\right]^T$, $\mathbf{e} = [\ 1,\ 1, \ldots,\ 1]^T$ with the same size of $\mathbf{T}$, and $\mathbf{Q}$ is the $M \times M$ sub-matrix of $\mathbf{P}$ ($q_{ij} = P_{ij}\ \forall i, j = 1, \ldots, M$).

These calculations complete the computation of the mean first-passage time and thus the failure rate and the reliability of providing accuracy ϑ. Since different levels of localization accuracy might be required by the various application tasks, the reliability of providing these levels can be evaluated similarly and the produce is what we call Task-Driven Integrity (TDI).

5.3 Two Case Studies

To implement the proposed TDI assessment model the parameters of the model need to be estimated under nominal and non-nominal observation conditions (*cf.* [97]). Two examples of the TDI model are constructed: a single-source localization technique (stand alone GPS system) and a multi-source localization technique (i.e., Inter-Vehicle Communication Assisted Localization (IVCAL) [71]).

5.3.1 Single Source Scenario: A GPS Receiver Integrity Assessment

In this section, a TDI model for a GPS receiver is constructed and compared with the baseline GPS localization quality index. Typically, the error in the location estimate produced by a GPS receiver is caused by various factors, including the satellite ephemeris error, the ionospherical and tropospherical errors, the receiver clock error, and the multipath error [71]. In GPS localization systems, the location estimate accuracy is calculated

based on two factors (Equation 5.12): 1) the Dilution of Precision (DOP), and 2) the User Equivalent Range Error (UERE) (*cf.* [22]).

$$Localization\ Accuracy = DOP \times UERE \tag{5.12}$$

The DOP is a function in the satellites' positions and therefore is computed based on the geometry of the constellation with respect to the receiver. The UERE is defined as the root mean square of the sum of the variances of the aforementioned error types. It is worth mentioning here is that the UERE is assumed to be the standard deviation of a zero mean Gaussian distribution. Therefore, if the value of the DOP is equal to 1, then the actual location is at most far from the estimate by a distance equal to the value of UERE with 68% confidence level. For the same estimate and the same DOP value, the confidence level can be as large as 95% if the localization error tolerance is increased to 2×UERE. This confidence level is utilized by many researchers as a degree of the trust placed in the estimate.

Nevertheless, the UERE measurement is not supported by all GPS receivers. Often, the UERE value is estimated statistically and used, as in Equation 5.12, under the assumption of a zero-mean Gaussian distribution. However, when a vehicle travels through different environmental conditions, the validity of this assumption is questionable due to the multimodal distribution of the localization error.

The model quantifies the integrity of producing two accuracy levels. To maintain a fair comparison between the two approaches of modeling the localization uncertainty, the TDI model will also utilize only the DOP value to compute the integrity of the system in estimating the GPS location with respect to two levels.

A TDI model with three states is constructed for a two level accuracy localization

application. Two states of the Markovian model represent the target accuracy levels, and the third one represents any accuracy value that is below the target accuracy performance, Figure 5.2. Hence, this Markovian model can capture the best to worst measurement conditions.

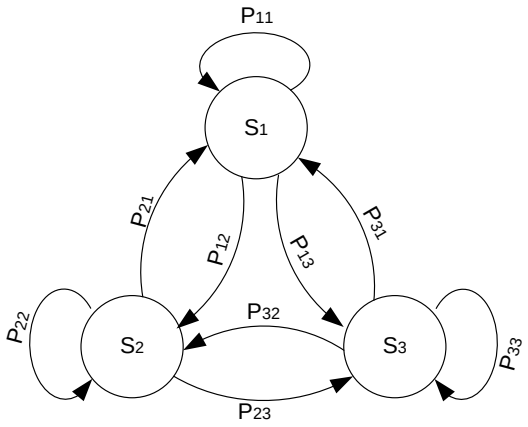


Figure 5.2: (a)Three states Markovian model for GPS-TDI assessment model.

The transition probabilities (i.e., P_{ij}) shown above are calculated based on the DOP value. Therefore, $P_{ij} = P(S_j/S_i, DOP)$; however, the GPS receiver calculates the consecutive estimates independently; Hence, $P_{ij} = P(S_j/DOP)$. The latter probability can be calculated using Bayes' rule as follows

$$P(S_j/DOP) = \frac{\phi(DOP/S_j)Pr(S_j)}{\sum_{k=1}^{3}\phi(DOP/S_k)Pr(S_k)} \tag{5.13}$$

where $\phi(DOP/S_i)$ signifies the probability density function of the likelihood of the DOP given the accuracy band S_i, and the $Pr(S_i)$ signifies the a priori probability of that accuracy band S_i. The a priori probabilities and the likelihoods functions of the DOP, given a specific accuracy band, can be determined statistically from measurements collected in different environmental conditions as presented in Section 5.4.1. This completes the design of the

GPS-TDI model.

5.3.2 IVCAL Integrity Assessment

In this section, the TDI assessment model is constructed for the IVCAL system. IVCAL is a multi-modal system that mitigates the effect of the multipath signals in the location estimation by utilizing a Multipath Detection Unit (MDU). Details of the components and operations of the IVCAL technique can be found in [71].

In IVCAL, two issues affect the accuracy of the localization: the observation conditions (nominal versus non-nominal), and errors in the MDU decisions. The KF innovation value is used in order to observe whether the satellite signals are contaminated by multipath signals. In practice, it is difficult to obtain sufficient patterns of different environments, such as clear sky, suburban landscape, and urban canyons, and use them to train the MDU of IVCAL. Instead, the two extremes, open sky and urban canyon environments, are used in the MDU training process as two classes of patterns that belong to two different environments. Accordingly, the MDU can classify most of the observations that belong to these two environmental conditions. This type of observation conditions is considered to be nominal observation conditions. In other environmental conditions, where the multipath signal is severe enough to affect the satellite signals, the MDU is more susceptible to making incorrect decisions; hence this type of observation conditions is considered as non-nominal observation conditions. Table 5.1 lists possible observation conditions for IVCAL as well as the MDU decisions and the Markovian states for each case.

As depicted in Table 5.1, IVCAL uses a KF to determine a vehicle's location when the MDU detects no multipath signals, and IVC when multipath signals are detected. It can therefore be concluded that a TDI assessment model of IVCAL has four states with four different levels of localization accuracy (Figure 5.3).

Table 5.1: IVCAL observation conditions and multipath detector decisions.

Observation conditions	Nominal		Non-Nominal	
Environment state	Clear	Multipath	Clear	Multipath
MDU decision	Clear	Multipath	Multipath	Clear
positioning approach	KF	Inter-Vehicle Comm.	Inter-Vehicle Comm.	KF
Markovian states	S_0	S_2	S_3	S_1

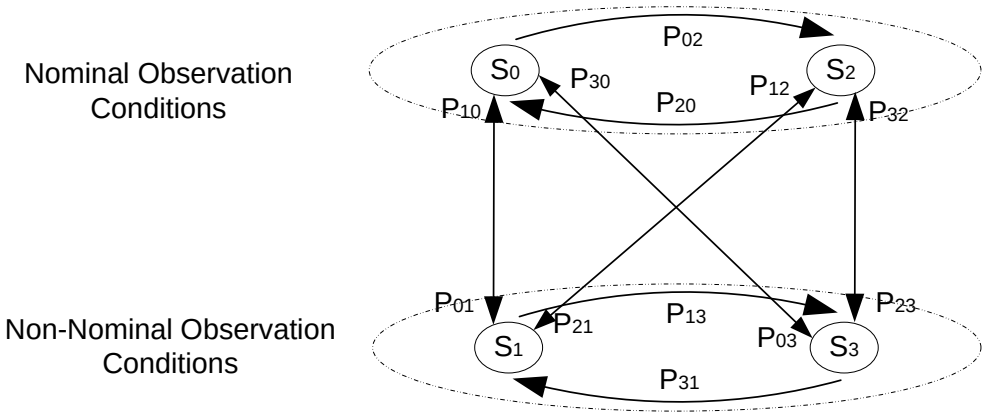


Figure 5.3: The Markovian model for the TDI assessment model of IVCAL.

The next step is to calculate the value of the transition probability of the Markovian model, P_{ij}. To calculate the value of P_{ij}, the KF innovation value is used as an indicator of the observation conditions, and the MDU decision error is used as an indicator of the correctness of the decisions made. The two indicators are considered to be independent random events. Consequently, the four states of the Markovian model are divided into two groups: one represents the level of accuracies in the nominal observation conditions, and the other represents the level of accuracies in the non-nominal observation conditions. Each group contains two states: multipath and non-multipath. This approach simplifies the calculation of the original transition probabilities.

In the following subsection, the calculation of the probability of transition between the

nominal and non-nominal groups of states and between the multipath and non-multipath decision states is explained.

Nominal/Non-Nominal State Transitions

The transition probability, depicted in Figure 5.4, depends on the IVCAL's KF innovation value $\overrightarrow{\lambda}$ and how this value deviates from the distribution of the innovation value in open sky environments or multipath environments (i.e., the two extremes). Therefore, the non-

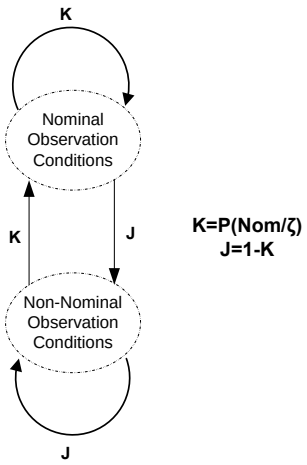


Figure 5.4: Transition between two groups of states that signify nominal and non-nominal observation conditions.

nominal observations are produced by innovation values lying in the midway between the two extremes of the KF innovations. The norm of the KF innovation (i.e., $\lambda = \| \overrightarrow{\lambda} \|$) can be thus considered as a random variable that has different ranges, as depicted in Figure 5.5 according to the measurement conditions.

It can be seen that the values of λ in nominal measurement conditions are separated by other values of λ in non-nominal observation conditions. This configuration of λ values makes the process of calculating a transition probability to a nominal condition state a

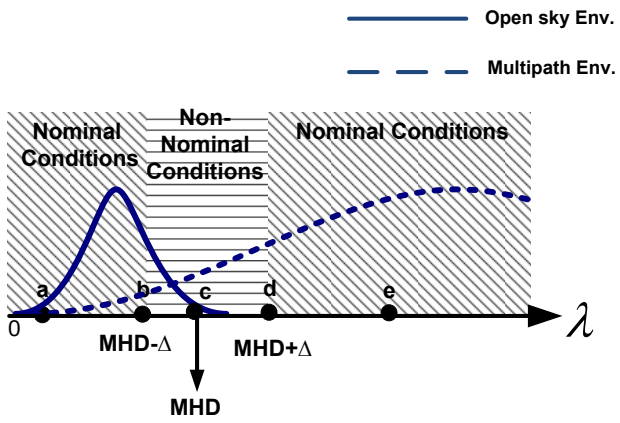


Figure 5.5: Kalman filter innovation values and the corresponding measurement conditions.

difficult task. If we can segregate the nominal and the non-nominal observation conditions given λ, the calculation of the transition probability is much easier. Therefore, another random variable ζ, which is a function of λ, will be used rather than λ:

$$\zeta = \mid \lambda - MHD \mid$$

where MHD is the innovation value that has equal Mahalanobis distance to the two extremes of the nominal observation distributions. In general the Mahalanobis distance is computed as follows:

$$\text{Mahalanobis Distance} = \sqrt{(\mathbf{y} - \bar{\mathbf{y}})^T \Sigma^{-1} (\mathbf{y} - \bar{\mathbf{y}})}$$

where $\mathbf{y}$ is a multivariate vector, $\bar{\mathbf{y}}$ is the means vector, and Σ is the covariance matrix of the target distribution.

Figure 5.6 depicts the value range of ζ with respect to the different measurement conditions. It can be seen that the nominal and non-nominal observation conditions are well segregated.

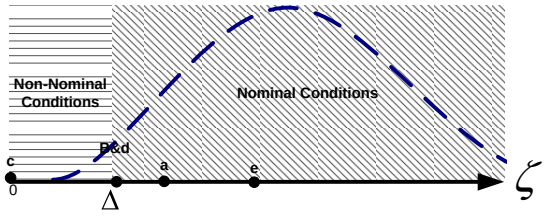


Figure 5.6: The separation of the observation conditions.

In Figure 5.5, points (b) and (d) are the most uncertain points with respect to the observation conditions depicted in the figure. The probability of being in nominal observation conditions should therefore be 0.5. Points (a) and (e) should have a high probability of being in nominal observation conditions. Point (c) is the most certain point with respect to the system being in non-nominal observation conditions, and as a result, the probability of going into a nominal observation condition state must be very small.

For the purpose of computing such a probability, a cumulative probability distribution, φ, for a normal probability density function, ϕ, is assumed (Figure 5.7). The mean and

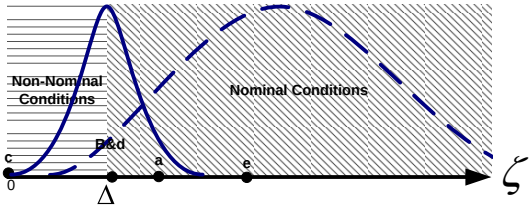


Figure 5.7: Probability distribution to compute $P(Nom/\zeta)$.

standard deviation (STD) of ϕ are $\mu_\phi = \Delta \leq \sigma_\lambda/2$ and $\sigma_\phi = 0.34 \times \Delta$, respectively, where σ_λ is the STD of the less spread nominal distribution. Accordingly, the value of the

probability being in a nominal observation condition, denoted by $K = P(Nom/\zeta)$:

$$\begin{aligned} K &= \varphi(\zeta) \\ &= \int_{-\infty}^{\zeta} \phi(\nu) d\nu \\ &= \int_{-\infty}^{\zeta} \frac{1}{\sqrt{2\pi}\sigma_{\phi}} e^{-\frac{(\nu-\mu_{\phi})^2}{2\sigma_{\phi}^2}} d\nu \end{aligned} \tag{5.14}$$

Intuitively, the probability of transition going into a non-nominal observation condition state is equal to $J = 1 - K$.

Table 5.2 presents the mapping of some values of λ into ζ and the respective values of the transition probability $P(Nom/\zeta)$. Figure 5.8 further shows the change in the transition probability $P(Nom/\zeta)$ with respect to ζ.

Table 5.2: Mapping from λ to ζ and calculating $P(Nom/\zeta)$ when $\Delta = 10$ m.

Observation Conditions	λ	ζ	$P(Nom/\zeta)$
(a)Nominal	5	15	0.92
(b)Uncertain	10	10	0.50
(c)Non-Nominal	20	0	2.13×10^{-3}
(d)Uncertain	30	10	0.50
(e)Nominal	40	20	0.99

Transition Between Multipath and Non-Multipath States

The transition probability required in this subsection is the transit probability between two states that signify the accuracy in a multipath (S_2, or S_3) and non-multipath (S_0, or S_1) environment, as shown in Figure 5.3. These two states reside in either the nominal observation conditions group or in the non-nominal observation group of states. The

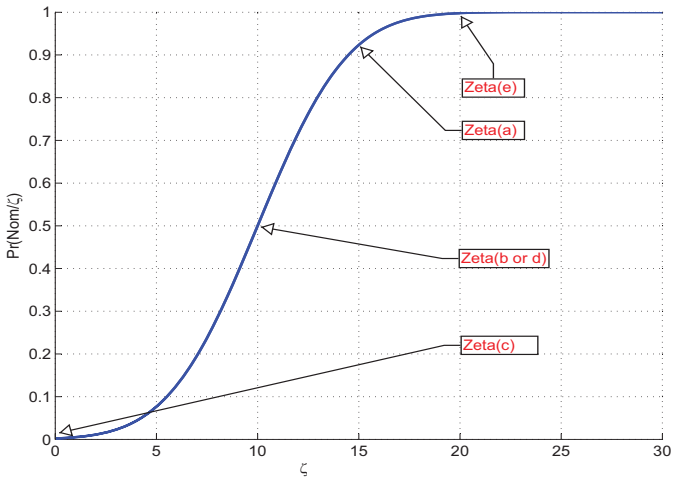


Figure 5.8: $P(Nom/\zeta)$ distribution.

transition probability is based mainly on the correctness of the decision made by the MDU, which classifies the environment as multipath or non-multipath.

The MDU, internally, has two parameters: γ_m which provides a figure that signifies the amount of the multipath effect, and γ_c which provides a figure that signifies the likelihood of being in a clear sky environment. The MDU compares the values of γ_m and γ_c and selects the largest for making its decision. Using the values γ_m and γ_c, it is possible to calculate the probability of the correctness of the decision as follows:

$$N = P(Multipath/\gamma_m, \gamma_c) = \frac{\gamma_m}{\gamma_m + \gamma_c} \tag{5.15}$$

$$L = P(Clear/\gamma_m, \gamma_c) = \frac{\gamma_c}{\gamma_m + \gamma_c} \tag{5.16}$$

It should be noted that the two probabilities sum to one and that they can build the two-state Markovian model depicted in Figure 5.9.

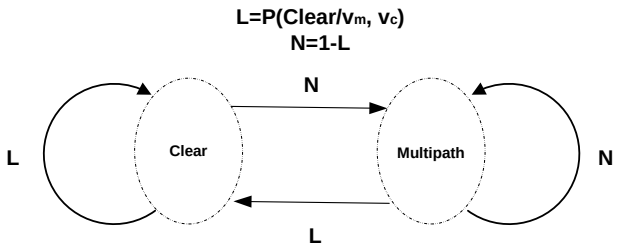


Figure 5.9: Two-state Markovian model that signifies the level of accuracy in multipath and non-multipath environments.

Transition Probabilities of the IVCAL Markovian Model

The two events described in the two previous subsections are assumed to be independent since they are generated from two independent entities in IVCAL. It is therefore possible to group the two events in one Markovian model that contains the four states shown in Figure 5.3. The transition probability distribution of this Markovian model is the joint probability of the transition probabilities depicted in Figures 5.4 and 5.9. Table 5.3 shows the values of the transition probabilities for the four-state Markovian model.

Table 5.3: Transition probability distribution for the four-state Markovian model of IVCAL.

$P_{00} = LK$	$P_{01} = LJ$	$P_{02} = NK$	$P_{03} = NJ$
$P_{10} = LK$	$P_{11} = LJ$	$P_{12} = NK$	$P_{13} = NJ$
$P_{20} = LK$	$P_{21} = LJ$	$P_{22} = NK$	$P_{23} = NJ$
$P_{30} = LK$	$P_{31} = LJ$	$P_{32} = NK$	$P_{33} = NJ$

Now, due to the sequential filtering of IVCAL, The TDI measure (ρ_t) at time t must be influenced by the previous location estimate integrity (ρ_{t-1}) and the instant integrity estimate υ produced by the proposed TDI model. The following expression shows how ρ_t

is calculated:

$$\rho_t = \upsilon \times \rho_{t-1}^{\alpha}, \qquad 0 < \alpha < 1 \tag{5.17}$$

where α signifies a memory factor maintains a time window that makes ρ_t influenced only by the recent integrity measurements.

5.4 Task-Driven Integrity Implementation

In the following subsections, the implementation of the TDI assessment model is described in order to show the viability of the instances constructed above. The model is applied on measurements of GPS receiver to compare it with the baseline GPS quality index used in the community. The model is also applied on more complicated localization technique (i.e., IVCAL technique) that provides a location estimate with different qualities depending on the surrounding environment.

5.4.1 The GPS-TDI Model Implementation

The distribution of the DOP value is assumed to be Gaussian over the different accuracy bands. The moments of the likelihood functions and the a priori probabilities are computed out of the real-life data collected in downtown Toronto (Ontario - Canada) and over a segment of the highway 401 between the city of Waterloo (Ontario - Canada) and the city of Toronto.

The GPS-TDI Model Quality indices

The accuracy bands of interest are assumed to be 10 m, 20 m, and greater than 20 m. The first band is equivalent to one UERE value, which normally used for a standalone

standard GPS receiver. The second band is equivalent to 2× UERE value. However, the integrity produced by the GPS-TDI of these accuracy levels is dynamic as the measurement conditions change.

The distributions of the DOP value over these bands is studied by collecting measurements from different environmental conditions (e.g., open sky, suburban, and urban areas). The sample size used in this study is equal to 4300 which comprises of 2445 measurements in band S_1, 662 measurements in band S_2, and 1193 measurements in band S_3. From each band, half of the measurements is randomly selected to calculate the statistics of the DOP value for each band. The remainder of the measurements are used in the testing of the TDI model (i.e., Holdout Cross-Validation testing method). The random selections and testing phases are repeated 100 times, and the model outputs are averaged to validate the model consistency. Table 5.4, depicts the statistics used in the GPS-TDI model for one of the 100 iterations.

Table 5.4: A GPS-TDI quality index and respective DOP statistics.

Index	**State**	$Pr(S_i)$	**The DOP distributions' statistics(m)**
10 m	S_1	0.5683	$\mu = 1.2043$ $\sigma = 0.9$
Between 10 *and* 20 m	S_2	0.1543	$\mu = 1.5838$ $\sigma = 1.2347$
Greater than 20 m	S_3	0.2774	$\mu = 2.3129$ $\sigma = 1.1473$

Results and Comparison

The performance of the GPS-TDI assessment model for subset of the testing data is shown in Figure 5.10. The magnitude of the localization error is depicted in (a) for the three

bands. In Figure 5.10(b), it can be seen that the integrity index of the relaxed accuracy (i.e., 20 m) is always above the stricter one (i.e., 10 m). The integrity measure appears to be relatively high and stable for the errors below 10 m. When the localization error becomes larger than 10 m, the indices become heavily fluctuating and on average the tight index reads between 0 and 0.4, where the relaxed accuracy index reads between 0.4 and 0.9.

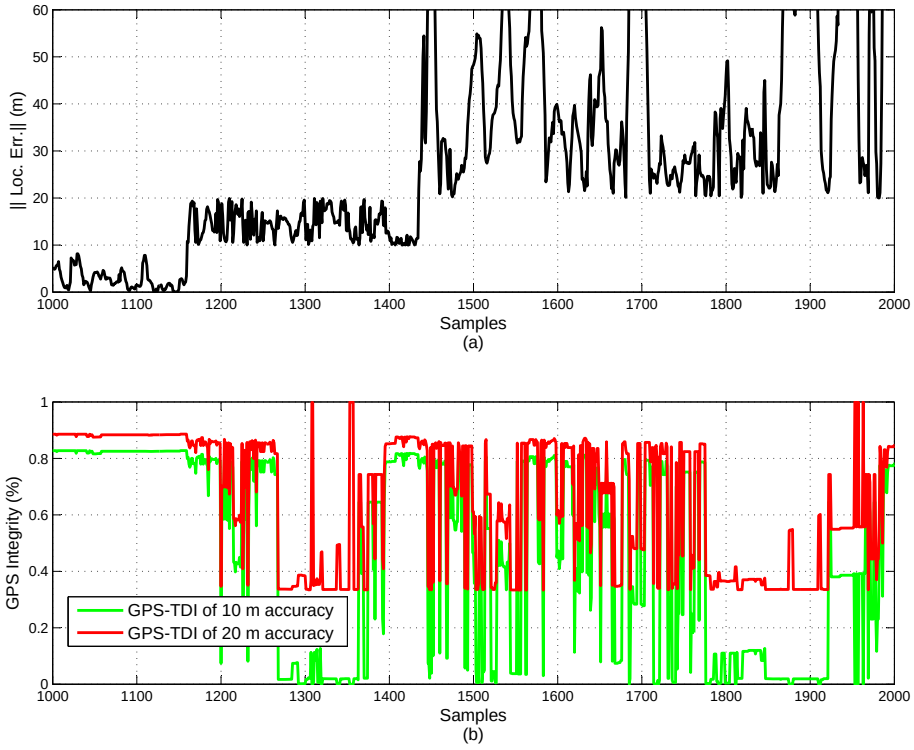


Figure 5.10: The TDI measure for two quality levels of GPS receiver: (a) The magnitude of the localization error. (b) The GPS-TDI measure for the quality indices.

To validate the TDI measure in this experiment we show the consistency of the integrity measure. This is conducted by extracting the data points indicated by certain level of integrity and calculate the percentage of points that have a localization error falling in the

claimed accuracy band. This computation is done for each level of integrity between 0.5 and 1.0 with step equal to 0.05. These percentages are averaged over 100 experiments. The standard deviation of each integrity level is also computed to examine the uncertainty of the integrity assessment. The resultant statistics are shown in Figure 5.11 where the integrity assessment is signified by the curves and the uncertainty is signified by the vertical bars. It can be seen that as the GPS-TDI model output increases, its value becomes closer to the statistical percentage of localization errors fall within the accuracy claimed. Furthermore, as the GPS-TDI measure increases the indices become more certain about their readings.

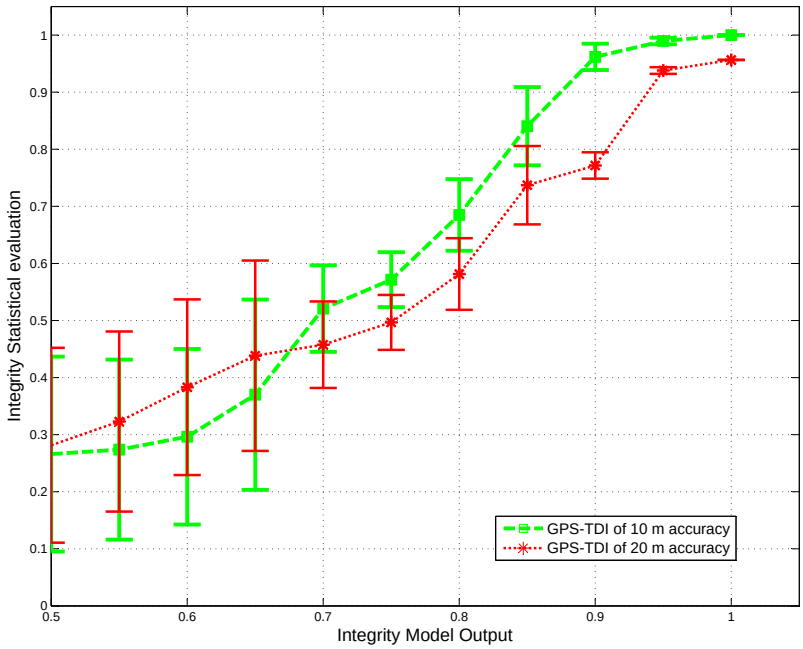


Figure 5.11: The GPS-TDI assessment model validation.

Figure 5.12 shows the inconsistency of the confidence measure when the accuracy computation is relying on in Equation 5.12. The vertical axes signifies the percentage of the data points that fall within the accuracy claimed, and the horizontal axis signifies the different values of the parameter UERE. The green line signifies the performance when

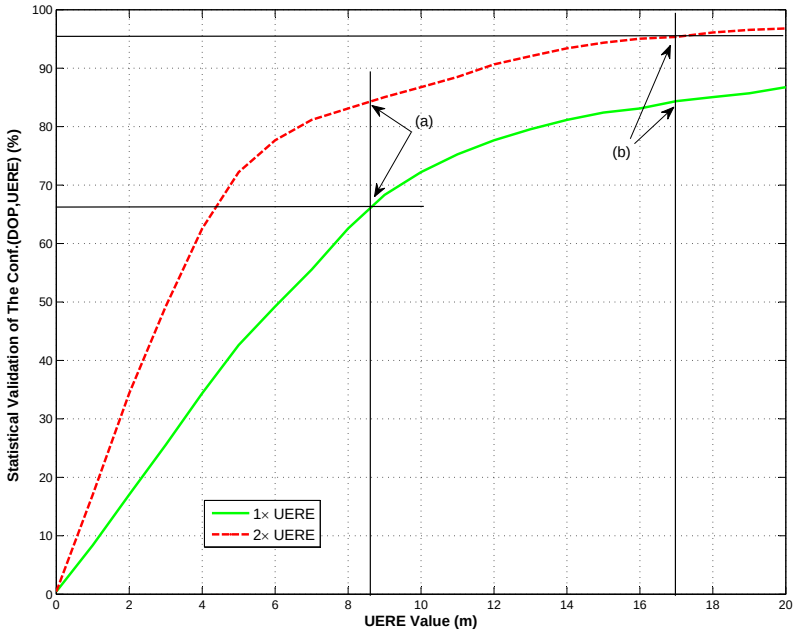


Figure 5.12: Uncertainty modeling using the DOP and UERE values in stochastic environmental conditions.

1× UERE is used, and the red dashed-line signifies the performance when 2 × UERE is used. As it can be seen, regardless of the index of UERE used, there is no value for the UERE where the confidence theory can be validated (i.e., 1 × UERE ≡ 68% confidence and 2 × UERE ≡ 95% confidence). Obviously, the reason for this violation is the dynamic environment that causes stochastic measurement conditions.

GPS-TDI Performance Sensitivity to GPS Parameters

The performance of GPS-TDI is shown above while only DOP is used to assess the system integrity. However, other GPS parameters, sensitive to the measurement conditions, can be used to improve the GPS-TDI performace. Therefore, different combinations of μ_{SNR}, σ_{SNR}, and DOP, as a multivariate vector, are used in Equation 5.13 and the GPS-TDI is

evaluated respectively. Figure 5.13 shows the performance of the GPS-TDI when different

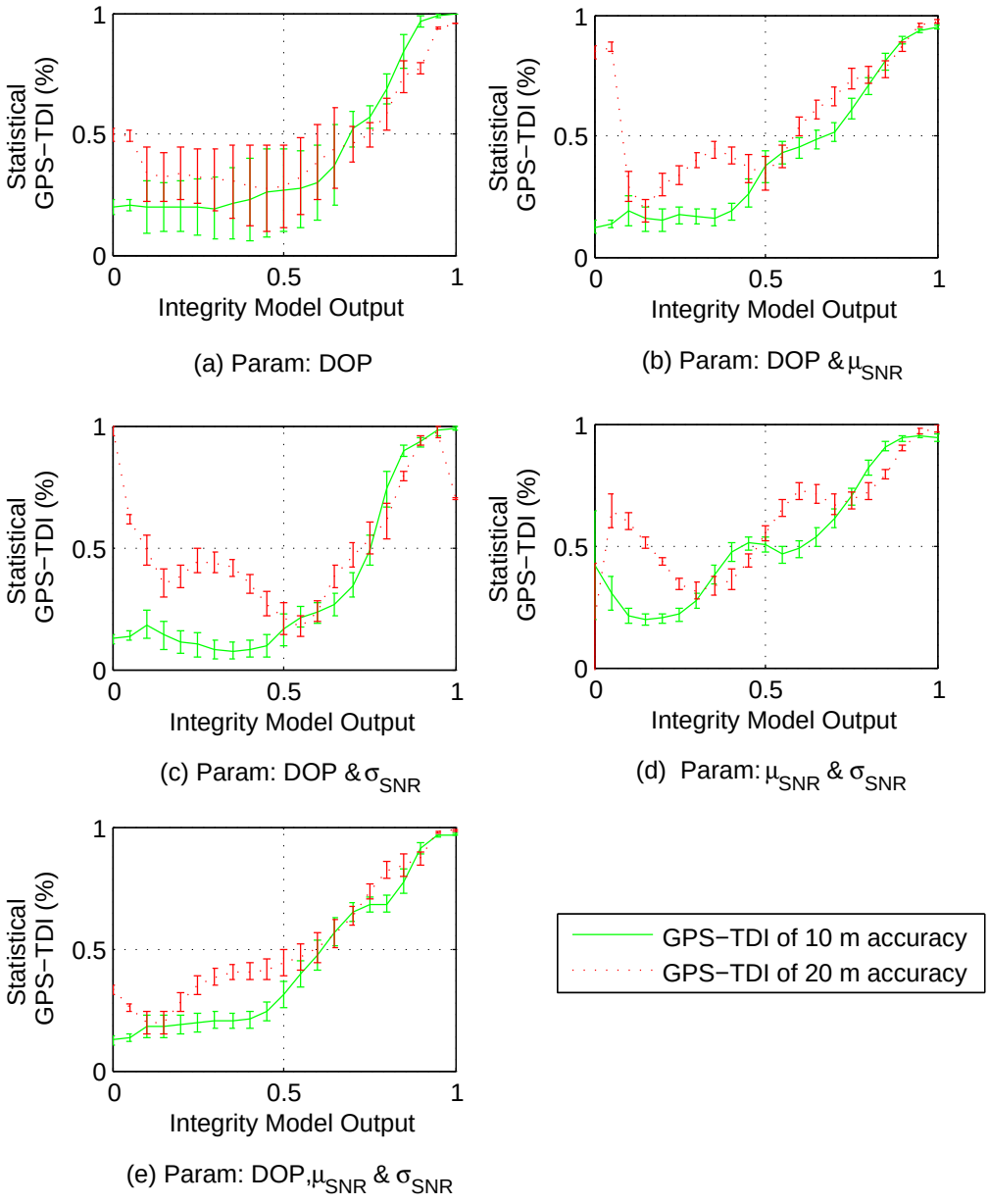


Figure 5.13: The sensitivity of the GPS-TDI performance to the GPS parameters used in the integrity assessment process.

GPS parameters are used. As seen, regardless of the GPS parameters used, the certainty of the integrity assessment increases with the increase in the level of integrity. Using DOP only results in highly uncertain integrity assessment when the level of integrity is below 0.7. The μ_{SNR} and σ_{SNR} introduce higher degrees of precision to the GPS-TDI measure; however, they could not maintain better accuracy. Obviously, using the three parameters (i.e., DOP, μ_{SNR}, and σ_{SNR}) all together produces the best performance in terms of precession and accuracy.

Numerically, the performance sensitivity to the parameters used is evaluated by computing the root mean square distance ($RMSD$) between the integrity statistical evaluation and the integrity model output.

$$RMSD = \sqrt{\frac{\|\overline{TDI}_{Stat} - \overline{TDI}_{model}\|^2}{n}} \tag{5.18}$$

where n is the number of pins used to descretize the model output. The smaller the $RMSD$ is the superior the integrity assessment performance is. Further numerical details about the performance sensitivity are listed in Table 5.5.

Table 5.5: GPS-TDI performance sensitivity to the GPS parameters.

Params used	$RMSD$	**Uncertainty**		
		Min	Average	Max
DOP	0.38	0.00	0.10	0.18
DOP and μ_{SNR}	0.39	0.01	0.04	0.07
DOP and σ_{SNR}	0.56	0.00	0.04	0.07
μ_{SNR} and σ_{SNR}	0.33	0.01	0.04	0.07
DOP, μ_{SNR}, and σ_{SNR}	0.21	0.01	0.04	0.06

5.4.2 The IVCAL-TDI Model Implementation

For the IVCAL-TDI model, the qualities of localization were first determined through a number of experiments in which the environmental conditions are simulated. The different quality levels are represented by a set, which is called quality index: $\xi_{IVCAL} = \{\kappa_1, \kappa_2, \kappa_3, \kappa_4\}$. The IVCAL-TDI model was then tested in a variety of simulated road segment scenarios. The results and comparisons are also provided in this section.

The Quality Index of IVCAL-TDI Model

The elements of the IVCAL quality index, κ_i, were computed and represented by a table in which rows signify the levels of quality. Each row contains the Markovian state associated with the respective quality of the localization. Prior to providing the IVCAL quality index, we first present the experiments performed to obtain this index.

The Simulation of the Observation Conditions:

The environmental conditions depicted in Table 5.1 were simulated in order to obtain localization error statistics. These statistics constitute the IVCAL quality index. In the simulations, the GPS measurement error is modeled as a random variable e_{GPS}, where

$$e_{GPS} \sim \left\{ \begin{array}{ll} N(0 \text{ m}, 15^2) & \text{Open area environments} \\ N(0 \text{ m}, 50^2) & \text{Slight multipath might not be detected by the MDU} \\ N(0 \text{ m}, 80^2) & \text{Multipath might be detected by the MDU} \\ N(0 \text{ m}, 150^2) & \text{Severe multipath environments} \end{array} \right\} \tag{5.19}$$

Each vehicle in the simulations starts with a location estimate that contains an error modeled by a zero mean Gaussian random variable with a 5 m STD.

Experiment 1

The first experiment was performed in order to determine the localization error while the IVCAL system is in state S_0, Figure 5.14, during which no high buildings or obstacles exist in the environment surrounding the vehicle, and the MDU of IVCAL detects no multipath signals. Consequently, IVCAL uses only the KF location estimate.

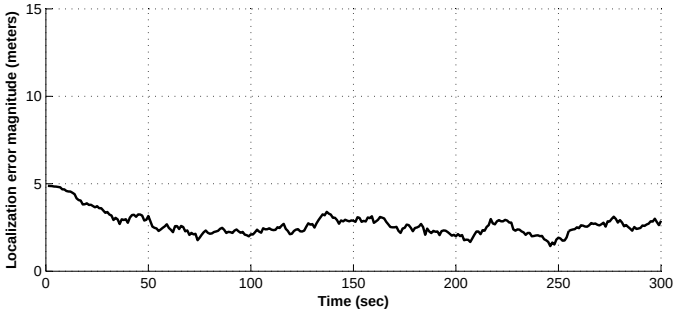


Figure 5.14: Vehicle localization error in an open area using a KF.

Experiment 2

The second experiment simulates the environmental conditions of state S_1 in the Markovian model. In this environment, the MDU might not detect the multipath effect where $e_{GPS} \sim N(0\text{m}, 50^2)$, and hence, localization is performed using only the KF. During the simulations, the vehicles pass through 500 m of open area, followed by 500 m of a multipath region, and then enter another 500 m of an open area, and so on. An example of such an environment is a town centre full of buildings that are not very high (four to seven stories). The magnitude of the computed localization error in this environment is depicted in Figure 5.15. It can be seen that the error increases whenever the vehicle passes through a multipath region due to the misrepresentation of the observation noise in the KF.

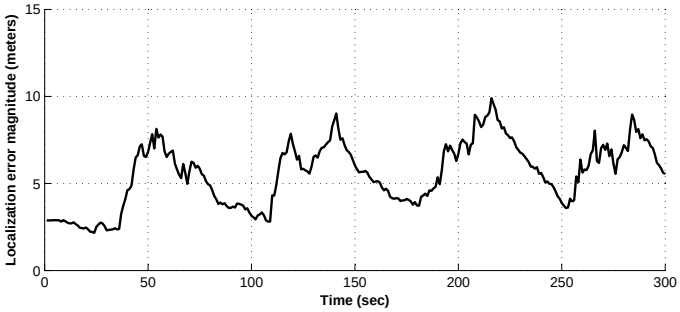


Figure 5.15: Vehicle localization error in a medium height building environment using a KF.

Experiment 3

To simulate a mild multipath that can be detected by the MDU, the GPS error is modeled as $e_{GPS} \sim N(0\text{m}, 80^2)$, and IVC is used when a vehicle goes through multipath regions. As can be seen in Figure 5.16, some spikes appear in the localization error when IVC is used, due to the bad alignment of the anchors used in the localization process. The average localization error is nevertheless better than that in the previous experiment because of the ability of IVCAL to improve the localization process even in open areas. This environment is an example of a typical suburban region where high buildings are not consolidated in one area.

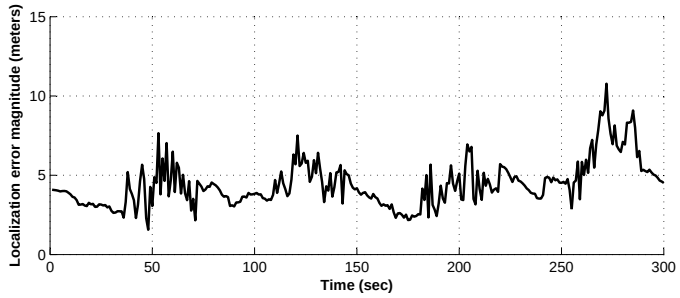


Figure 5.16: Vehicle localization error using IVC in regions with scattered high buildings.

Experiment 4

In this experiment, the environmental conditions of state S_2 were simulated: an area with a strong multipath effect. To model such an environment, the following scenario is implemented. Vehicles pass through 500 m of road segments that resemble regions with high building where the STD of the GPS error is 150 m. These regions are interlaced with 500 m of open areas in order to maintain number of anchors for the other vehicles in the multipath regions. Figure 5.17 shows the magnitude of the localization error during 300 seconds of the total time of the experiment while a vehicle was passing through interlaced open-sky and multipath regions.

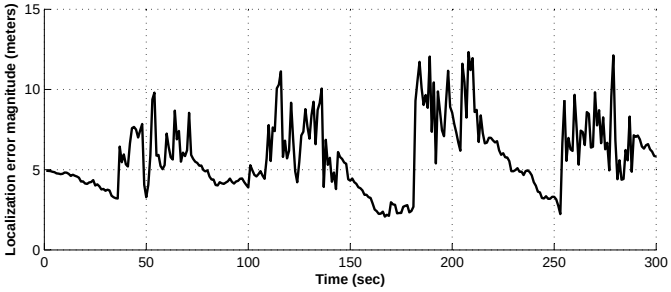


Figure 5.17: Vehicle localization error in severe multipath regions using IVC.

IVCAL Quality Index Values:

The statistics in the previous experiments are calculated in terms of the magnitude of the localization error. They include the maximum value obtained, the mean, and the STD. Table 5.6 presents the states of the IVCAL Markovian model and their respective statistics. The states are arranged from best to worst based on the localization performance represented by the statistics, assuming that the value representing the application requirements, ϑ, is based only on the average error.

Table 5.6: IVCAL quality index.

Index	State	Localization error statistics(m)		
κ_1	S_0	Max	=	9.30
		μ	=	2.19
		σ	=	1.42
κ_2	S_3	Max	=	25.25
		μ	=	4.40
		σ	=	3.31
κ_3	S_1	Max	=	20.00
		μ	=	5.39
		σ	=	3.30
κ_4	S_2	Max	=	32.10
		μ	=	5.86
		σ	=	4.53

IVCAL-TDI Testing

To test the IVCAL-TDI model, a comprehensive experiment that includes a variety of environmental conditions was performed and is explained in this section. The following assumptions were used for the simulations of IVCAL: 1) vehicles are using IVCAL to estimate their locations; 2) IVCAL is working over an infrastructure called VANET that facilitates wireless communication among vehicles; 3) vehicles are equipped with a standard GPS receiver, an inertial navigation system, and a radio transceiver; 4) the radio coverage is 500 m in open areas.

Road Segments' Scenarios:

The simulation covers a 5 km portion of a straight road. Vehicles travelling this road experience different local environments, such as an open area with no multipath effect, an inner-city area where high buildings cause a severe multipath effect, and a suburban area that causes a mild multipath effect. These different environments are depicted in Figure

5.18, which shows that vehicles travel from left to right, passing through five open sky regions, three mild multipath regions, and two severe multipath regions. Based on statistics

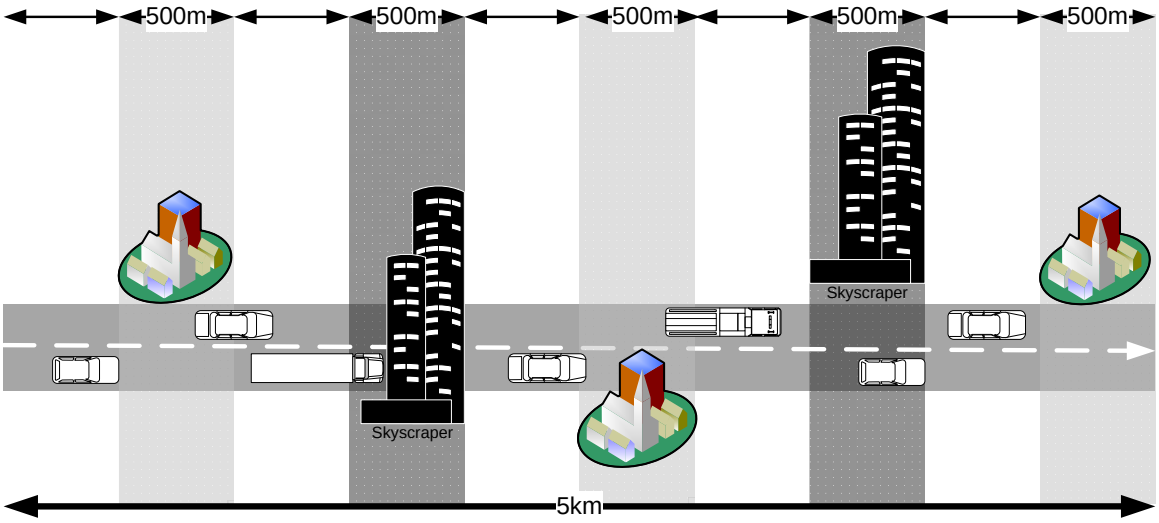


Figure 5.18: Road segment scenarios for testing the integrity model.

obtained from the real-life measurements the GPS error used in the severe multipath regions follows the distribution $e_{GPS} \sim N(0 \text{ m}, 150^2)$, in the mild multipath regions follows the distribution $e_{GPS} \sim N(0 \text{ m}, 65^2)$, and in no-multipath regions follows the distribution $e_{GPS} \sim N(0 \text{ m}, 15^2)$.

The road consists of two lanes in one direction for vehicles with different speeds. The right lane contains vehicles travelling at 50 km/h, and vehicles in the left lane are travelling at a higher speed: 60 km/h. The simulation uses 100 vehicles uniformly distributed on each lane to comprise VANET. The simulation period is equal to the time required for a vehicle to pass through 5 km at a speed of 50 km/h: 360 seconds.

IVCAL-TDI Results and comparisons:

The localization error was calculated for a vehicle travelling through the road depicted in Figure 5.18 from left to right in the right lane. Each time the IVCAL technique produces a location estimate, the integrity of that estimate is evaluated in terms of the quality indices. The memory factor (α) was set to 0.7 for IVCAL.

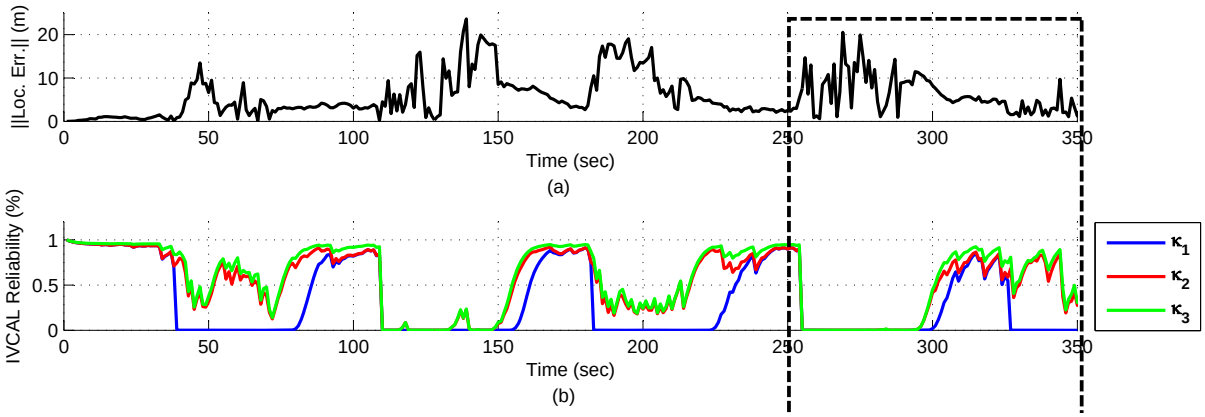


Figure 5.19: The TDI measure for three quality levels of IVCAL: (a) The magnitude of the localization error. (b) The TDI measure for the location estimate using IVCAL.

Figure 5.19 (a), shows the magnitude of the vehicle localization error using IVCAL. It can be seen that the error changes with time, and the change differs depending on the environment surrounding the vehicle. In multipath environments, the change is abrupt, and the error is large. However, the change is smooth in open sky environments since the KF estimate is selected by the MDU.

Figure 5.19 (b) shows the IVCAL-TDI assessment of the vehicle location estimate, ranging from 0 to 1, plotted against time. Since IVCAL provides four levels of accuracy (in terms of the average localization error), the best three levels (i.e., κ_1, κ_2, and κ_3 in Table 5.6) were selected to be assessed here. κ_1 is the most restricted accuracy, which represents

an average location error (ALE) $\leqslant$ 2.19 m. The integrity of κ_1 is represented by the blue line in the figure. The red line in the figure represents the integrity of κ_2, which is less accurate: $ALE \leqslant$ 4.40 m. The green line in the figure represents the integrity of κ_3, which has the worst accuracy of the three levels: $ALE \leqslant$ 5.39 m. The fourth accuracy (i.e., κ_4) is not considered here since, in all the experiments, it is guaranteed that IVCAL can not make the accuracy worse than this level ($ALE \leqslant$ 5.86 m) on average. It can be seen how the integrity values for the location estimates are affected by the environment: the greater the multipath effect is, the less reliable the estimate is. It can also be seen that the more restricted the level of accuracy, the less reliable the location estimate is for that level of accuracy.

Figure 5.20 shows an enlarged version of a portion of Figure 5.19. In part (b) of this figure, it is clear that the IVCAL-TDI measure needs some time to recover from the multipath effect meanwhile IVCAL was producing high accuracy localization. In other words, the IVCAL-TDI measure sometimes underestimates the quality of the localization technique. Therefore, the affect of the memory factor on the IVCAL-TDI measure performance has been examined through the following experiment.

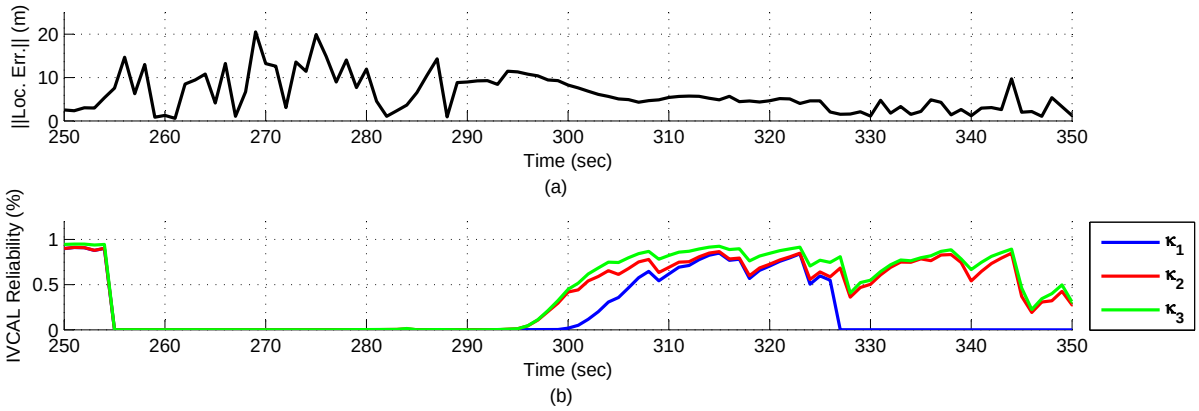


Figure 5.20: The impact of sequential localization on the integrity measure.

To examine the estimate of the IVCAL-TDI measure for the system integrity, the range of the measure (i.e., 0 to 1) has been divided into 10 equal pins. For each pin, the following statistics is calculated for every level of quality κ_i:

- Number of Estimates ($NoE_{l,i}$)= number of estimates produced with TDI value falls in the pin l for κ_i.
- Miss Detection ($MD_{l,i}$)= number of estimates that has errors exceeds κ_i and falls in pin l.
- Statistical Integrity Measure ($SIM_{l,i}$)= 1 - $\frac{MD_{l,i}}{NoE_{l,i}}$, where $l = 1, \ldots, 10$.

The IVCAL-TDI is evaluated by comparing it to the SIM. The above experiment has been repeated many times with different values of α in order to examine its effect on the IVCAL-TDI measure. The values of α used in the experiment ranges from 0.01 to 0.99 with increment of 0.01. Figure 5.21 shows averaged results of this experiment. It can be concluded from this figure that the IVCAL-TDI measure becomes more indicative as the value for the memory factor is in the range between 0.75 and 0.85. Yet, it can be seen that

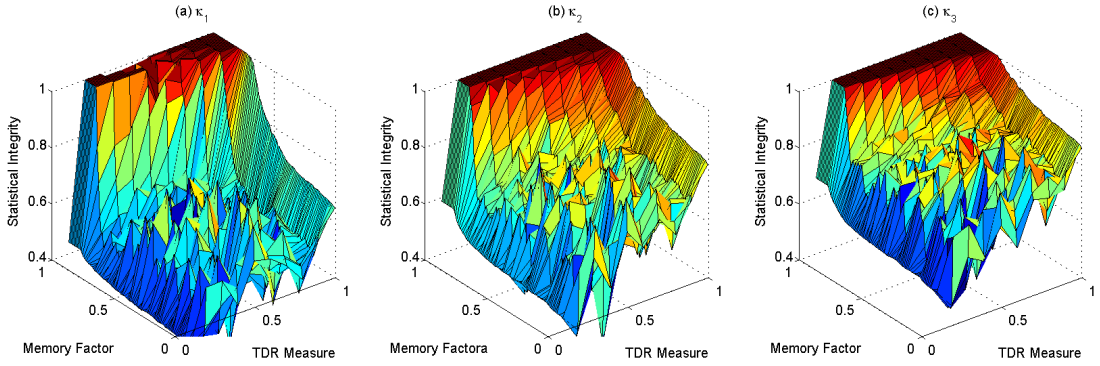


Figure 5.21: The impact of the memory factor on the TDI.

the IVCAL-TDI measure on average is conservative since most of the time it is less than the SIM when α is between 0.75 and 0.85.

5.5 Generalized Task-Driven Integrity Implementation

In the case studies presented above, quite a few assumptions that do not hold in many real-life situations have been made. For example, a governing normal distribution is assumed for the sensing parameters used in the integrity assessment. It is often the case that the location measurement noise is time variant, as such the error distributions can not be relied on to provide insight into the measurement process uncertainty. Furthermore, the construction of the TDI model is not a trivial task as the engineering knowledge of the key environmental impacts on the system performance need to be known a priori. This task is more challenging when the environmental events are large in number and correlated. In sequential localization techniques, the choice of the memory factor constitutes another burden in the design of the TDI model.

To relax the aforementioned assumptions, a generalized construction of the TDI model is proposed and implemented here. The basic idea revolves around maintaining a dynamic Markovian model that keeps track of the recent recognized accuracies. In order to accomplish this task an accuracy classifier, such as those presented in Chapter 4, is used to determine the instantaneous localization accuracy. The distribution of the model's transition probabilities is estimated based on the historical profile of the events determined by the accuracy classifier. Thus, the resultant distribution is non-parametric distribution.

5.5.1 Non-Parametric Transition Probability Distribution Estimation

In this section, the transition probabilities' computation is described using the recent recognized localization accuracies. A historical time window is used to keep k values, signifying the recent sequence of transitions from one state (i.e., accuracy level) to another. For example, the following window W_t is captured at time t and it contains k consecutive transitions

$$W_t = \left[S_3^1, S_3^2, S_1^3, S_2^4, \ldots, S_4^{k-1}, S_1^k\right]$$

where the superscript signifies the time index and the subscript signifies the state index. The size of the window (i.e., k) should be neither too large to reflect the impact of the current environment conditions, nor small to facilitate statistical computations.

Accordingly, the probability P_{ij}^t that the localization accuracy (i.e., state) will change from S_i to S_j is evaluated according to the observed transition frequency in the historical window:

$$P_{ij}^t = \varrho_{ij}^t / \varrho_i^t \tag{5.20}$$

where ϱ_{ij}^t signifies the number of transitions from S_i to S_j in W_t, and ϱ_i^t signifies the number of transitions from S_i regardless the next state in W_t.

Since

$$\varrho_{ij}^t \leqslant \varrho_i^t, \text{ and } \sum_j \varrho_{ij}^t = \varrho_i^t$$

then

$$P_{ij}^t \leqslant 1, \text{ and } \sum_j P_{ij}^t = 1$$

which satisfy the stochastic matrix conditions.

The caveat of this computation is that the transition matrix can easily lose its *ergodicity*,

[117], due to the lack of transition information about a subset of the accuracy states. Typically this situation happens when the system is stuck in a subset of the Markovian states which increases the sparsity of the transition matrix. Therefore, two adjustments are performed on the transition matrix to hinder the development of non-ergodic Markovian model.

1. The first adjustment is performed on rows with some entries equal to zero. To avoid having zeros in row i we substitute the respected entries with a very small positive value ($0 < \Delta \ll 1$) as below.

$$\widetilde{P_{ij}} = \frac{\Delta}{\#\text{ of zeros in row } i}, \forall j \text{ where } P_{ij} = 0 \tag{5.21}$$

and the non-zero entries of row i are adjusted as follows

$$\widetilde{P_{ij}} = P_{ij} - \frac{\Delta}{\#\text{ of non-zeros entries in row } i}, \forall j \text{ where } P_{ij} \neq 0 \tag{5.22}$$

2. The second adjustment is performed on rows containing nothing but zeros, which signify isolated states. Such situation happens, for example, when a vehicle travels on a highway or open-sky environments while using GPS for localization. Nevertheless, in real-life, there is always the possibility of visiting the isolated states. This possibility is maintained by Equation 5.21. At the same time those isolated states should maintain transition probabilities to the states appeared in W_t. The value of this transition probabilities, in row i, must be set according to the states' incoming probabilities, which are computed as follows.

$$P_j = \sum_i P_{ij}, \forall j = 1, \dots n \tag{5.23}$$

where n is the number of states in the Markovian model.Two sets of indices must be

maintained to perform the adjustment: the first set indicates the states with some incoming probability

$$J = \{j | P_j \neq 0\}$$

while the second set indicates the states with zero incoming probability,

$$J' = \{j | P_j = 0\}$$

Now the transition probabilities of row i can be set as follows.

$$\widetilde{P_{ij}} = \frac{P_j}{\sum_{k \in J} P_k} \times (1 - \Delta), \forall j \in J \tag{5.24}$$

$$\widetilde{P_{ij}} = \frac{\Delta}{\text{Cardinality of } J'}, \forall j \in J' \tag{5.25}$$

An example of a Markovian model with five states is presented in Figure 5.22 in which the adjusted Markovian model is shown where $\Delta = 10^{-3}$.

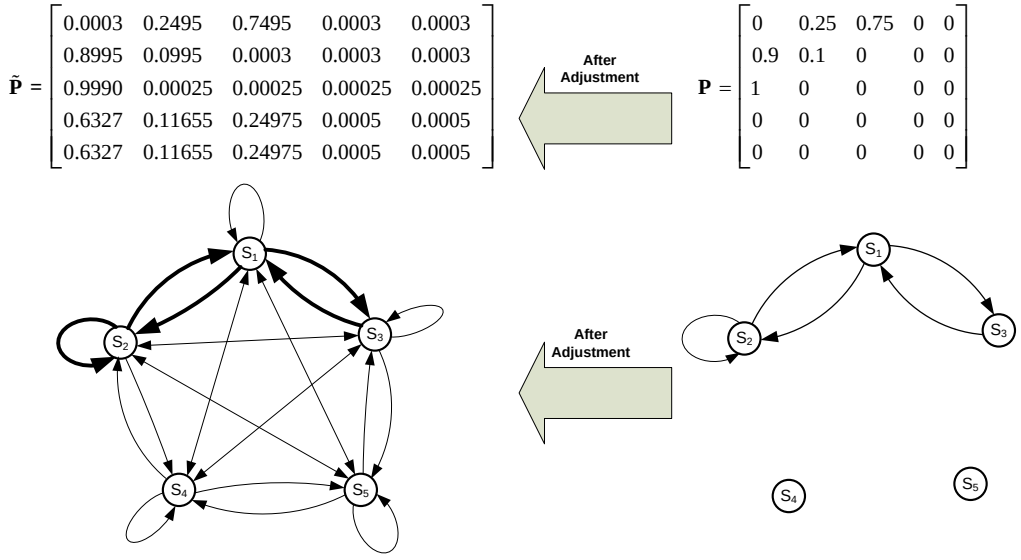


Figure 5.22: Example of adjusting non-informative Markovian model.

5.5.2 Parametric versus Non-parametric Distributions TDI Performance

The TDI model proposed in this section is tested on the GPS data set depicted in Section 4.4.1. CBAC is used here for GPS accuracy classification – the output of which is utilized to estimate the transition probabilities using non-parametric distribution. The size of the historical window ranges from 10 seconds to 180 seconds with 5 second steps. The validation of the TDI model is averaged over the different window sizes used. Figure 5.23.(a) shows the performance of the proposed TDI model using non-parametric distribution for two accuracy levels: 10 m and 20 m; where the curves depicts the average value and the bars signify the uncertainty using the STD among the different window sizes. It can be seen that the TDI measure becomes independent on the window size as the integrity level increases. Despite the lack of any information about the underlying distribution of the

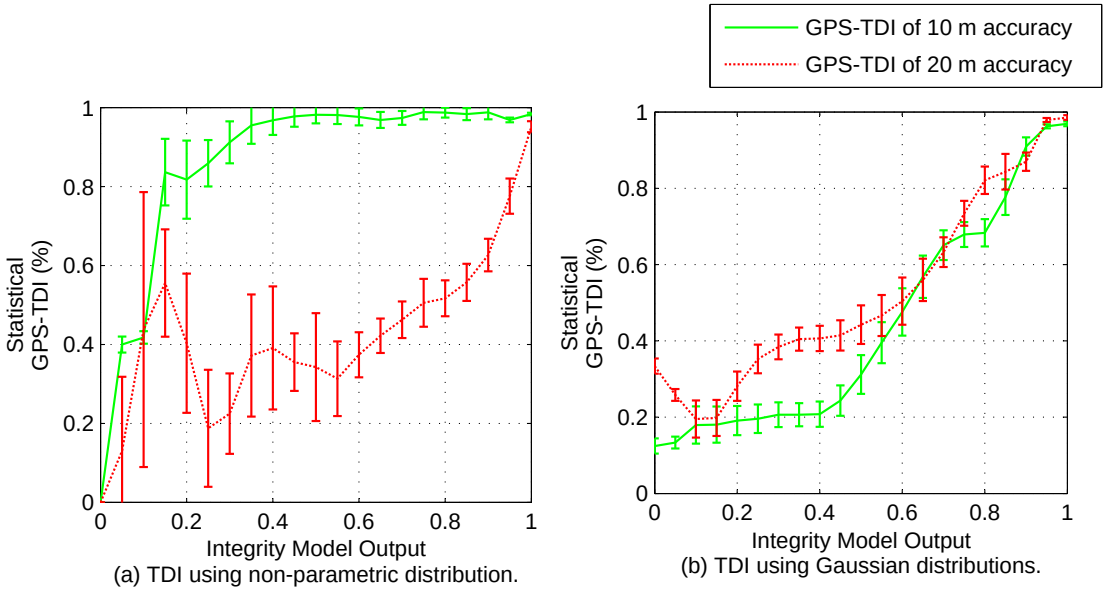


Figure 5.23: GPS-TDI model performance for different approaches of computing transition matrices.

GPS parameters, the TDI measure shows a decent degree of ability to assess the integrity of the localization, especially at the accuracy level of 20 m.

Figure 5.23.(b) shows the performance of the TDI model when Gaussian distribution is used for the sensing parameters (i.e., GPS, μ_{SNR}, and σ_{SNR}). As seen, the assessment using non-parametric distribution is very conservative at the accuracy level of 10 m, that is to say when the integrity level is above 0.4 the actual integrity is way above 0.9. On the other hand, the assessment of the 20 m accuracy level is comparable between the two approaches when both are above 0.6. These observations must be taken in consideration and used as rules to recover the shortcomings of the usage of non-parametric distribution in estimating transition probabilities in the IMUs.

5.6 Summary

Localization process has always a chance to fail to provide the required level of performance with respect to the applications, as the environmental conditions unexpectedly change. Therefore, a task-driven integrity (TDI) model is introduced in this chapter. The TDI model enables the applications modules to be aware of the trust level of the localization output. TDI model monitors specific parameter(s) in the localization technique and, accordingly, infers the impact of the change in the environmental conditions on the quality of the localization process. A generalized TDI solution has also been introduced in this chapter for the cases where information about the sensing parameters is unavailable or insufficient.

TDI model is implemented and tested on two different localization techniques. It is evident from the experimental results presented in this chapter that the proposed IMUs are capable of capturing the stochastic uncertainty in the location estimates and indicating

its integrity as to specific quality requirements. I consider TDI measure as a key for further research in many interesting areas, such as fusion dynamic uncertain information sources, like the ones that will be introduced in Chapter 6.

Chapter 6

Estimates Fusion and Management

6.1 Introduction

This chapter introduces the estimate fusion and management layer. The goal of this layer is to process, in the light of QoS required, the location and accuracy estimates obtained from the primary localization layer as well as the integrity level produced by the integrity monitoring layer. Based on the classification algorithms developed in Chapter 4 and integrity assessment model developed in Chapter 5, the information obtained from two location sources is assessed and used in the fusion process.

Effective fusion performance can only be achieved if adequate and appropriate a priori knowledge is available. Although, at least in some situations, assumptions can be made with respect to a priori and a posteriori probabilities (*cf.* Section 5.3), these assumptions can turn out to be unreasonable in many other situations, especially if we are to allow for non-probabilistic estimators in the PLU layer (*cf.* Section 4.5). Therefore, the focus in this chapter will be on Dempster-Shafer evidence theory (DSET) as an extension to the Bayes theory. One important aspect of DSET is that reasoning or decision making can be carried out with incomplete and/or conflicting pieces of evidence – a reality that is quite

common in localization problems. A new evidence structure, derived from DSET, called Spatial Evidence Structure Model, SESM for short, is introduced in this chapter to handle the fusion of localization quality indexes.

The remainder of this chapter is organized as follows. Section 6.2 provides an overview on the mathematical formulation of the DSET. Section 6.3 presents the development of the location information based evidence structure. In Section 6.4, three approaches for the fusion of location estimates are explained. Section 6.5, introduces the concept of task-driven resource allocation to the fusion process. Section 6.6 summaries the chapter and provides concluding remarks.

6.2 Dempster-Shafer Evidence Reasoning

Dempster-Shafer Theory is a mathematical theory of evidence. This theory is introduced by Dempster, [118], and extended later by Shafer [119]. DSET manages uncertainty while dealing with conflict through three levels: 1) representing evidence by focal elements and masses (basic probability assignments); 2) combing evidence by the Dempster's rule of combination; and 3) making decisions by certain decision rules.

Let $\Theta = \{\theta_1, \ldots, \theta_n\}$ be a frame of discernment. 2^{Θ} is the power set, denoted by Ω composed of all possible subsets of Θ. For example

$$\Theta = \{y_1, y_2, y_3\}$$

then

$$\Omega = \{\phi, \{y_1\}, \{y_2\}, \{y_3\}, \{y_1, y_2\}, \{y_1, y_3\}, \{y_2, y_3\}, \{y_1, y_2, y_3\}\}$$

where ϕ is the empty set.

Elements in Θ are assumed to be mutually exclusive and exhaustive. The mass function m is defined as a mapping of the power set elements to a number between 0 and 1, as follows

$$m : \Omega \rightarrow [0, 1] \tag{6.1}$$

$$\sum_{A \subseteq \Omega} m(A) = 1 \tag{6.2}$$

$$m(\phi) = 0 \tag{6.3}$$

Subset A with non-zero mass is called a *focal element* or focal for short. Focal elements and their masses constitute an *evidence structure*, expressed in form:

$$\{(A, m(A)) | A \subseteq \Omega, m(A) > 0\}.$$

Thus, $(A, m(A))$ is signifies a piece of evidence and the value of $m(A)$ represents the degree of evidential support with which a specific element of Ω belongs to the exact set A, not to subsets of A. The belief function Bel is defined as

$$Bel: \ \Omega \rightarrow [0, 1] \text{ and } Bel(A) = \sum_{B \subseteq A} m(B) \tag{6.4}$$

where $A \subseteq \Theta$, $B \subseteq \Theta$, and $A \neq \phi$. The plausibility function Pls is defined as

$$Pls: \ \Omega \rightarrow [0, 1] \text{ and } Pls(A) = 1 - Bel(\bar{A}) = \sum_{B \cap A \neq \phi} m(B) \tag{6.5}$$

The belief function $Bel(A)$ measures the total amount of probability that must be distributed among the elements of A, and it constitutes a lower limit function on the probability of A. On the other hand, the plausibility function $Pls(A)$ measures the maximal amount of probability that can be distributed among the elements in A, and it constitutes an upper limit function on the probability of A. Therefore, the interval span $[Bel(A), Pls(A)]$ reflects uncertainty, which describes the unknown with respect to A. In addition to quan-

tifying uncertainty, belief can be used to make decisions at the *pingnestic level* in which a probability distribution, called pignistic probability, $BetP$, is transferred from an evidence structure [120, 121] as follows.

$$BetP: \ \Omega \rightarrow [0,1] \text{ and } BetP(A) = \sum_{B \subseteq \Omega} \frac{|A \cap B|}{|B|} \times m(B) \tag{6.6}$$

where $|\cdot|$ denotes the cardinality. Note that the pignistic probability is normalized to one (i.e., $\sum_{A \subseteq \Omega} BetP(A) = 1$).

The Dempster's rule combines independent evidence by

$$\oplus_{i=1}^{n} m_i(A) = \frac{1}{1-K} \sum_{B_1 \cap \cdots \cap B_n = A} m_1(B_1) \cdot \ldots m_n(B_n) \tag{6.7}$$

where $B_1, \ldots, B_n \subseteq \Theta$ and

$$K = \sum_{B_1 \cap \cdots \cap B_n = \phi} m_1(B_1) \ldots m_n(B_n)$$

K signifies a basic probability mass associated with conflicts among the source of evidence. The above rule of evidence combination satisfies the commutative and associative properties:

$$m_1 \oplus m_2 = m_2 \oplus m_1 \tag{6.8}$$

$$m_1 \oplus (m_2 \oplus m_3) = (m_1 \oplus m_2) \oplus m_3 \tag{6.9}$$

To handle conflict of information sources, a discounting scheme has been introduced in DSET such that $Bel(A) = (1-\alpha) \times Bel(A)$, $\forall A \subset \Theta$, and $Bel(\Theta) = 1$. Accordingly, mass functions are usually modified in the following manner [119, 120, 122, 123]:

$$m^{\alpha}(A) = \begin{cases} (1-\alpha) \times m(A), \text{ if } A \subset \Theta \\ \alpha + (1-\alpha) \times m(\Theta), \text{ if } A = \Theta \end{cases} \tag{6.10}$$

where $\alpha \in [0,1]$ signifies a discounting factor, and $m^{\alpha}(A)$ signifies the discounted mass of $m(A)$. The larger α is, the more discounted a non frame of discernment focal is while the more mass is assigned to the frame of discernment.

Remark 6.2.1. Dempster's rule of combination strongly emphasizes the agreement between multiple sources and ignores all the conflicting evidence through the normalization factor K. Therefore, the use of the Dempster rule has come under serious criticism when significant conflict in the information is encountered [124, 125]. Consequently, researchers have developed modified Dempster rules that attempt to represent the degree of conflict in the final result [126]. For example, Yager proposed in [125] an alternative evidence combination rule that allocates conflict to the frame of discernment (Θ) instead of to the empty set (ϕ). Thus, mass associated with conflict is interpreted as the degree of ignorance. Without loss of generality, the modified combination rules for two sources of information become as follows.

$$\begin{aligned} K &= \textstyle\sum_{B\cap D=\phi} m_1(B)\cdot m_2(D), \\ m(\phi) &= 0, \\ m(A) &= \textstyle\sum_{B\cap D=A} m_1(B)\cdot m_2(D),\ A\neq\phi,\Theta, \\ m(\Theta) &= (\textstyle\sum_{B\cap D=\Theta} m_1(B)\cdot m_2(D)) + K \end{aligned} \tag{6.11}$$

where A, B, and $D \subseteq \Theta$.

6.3 Spatial Evidence Structure Model (SESM)

A new evidence structure model is proposed under the assumption that each location information source provides three pieces of information: location estimate denoted by Z_i, localization accuracy denoted be Acc_i, and integrity level denoted by ρ_i (see Figure 6.1).

The location estimate and the accuracy are determined by the respective PLUs, while the integrity level is specified by the corresponding IMUs. It is worth mentioning here is that a vehicle location estimate can be confined to any type of spatial data, such as points of road-segments' intersections, lines signifying road segments, or areas covering places of interest. Without loss of generality and to avoid errors caused by map-matching processing, I will consider the location estimate to be within an area of interest. The three pieces (i.e., Z_i, Acc_i, and ρ_i) of information are utilized in building the spatial evidence structure as explained below.

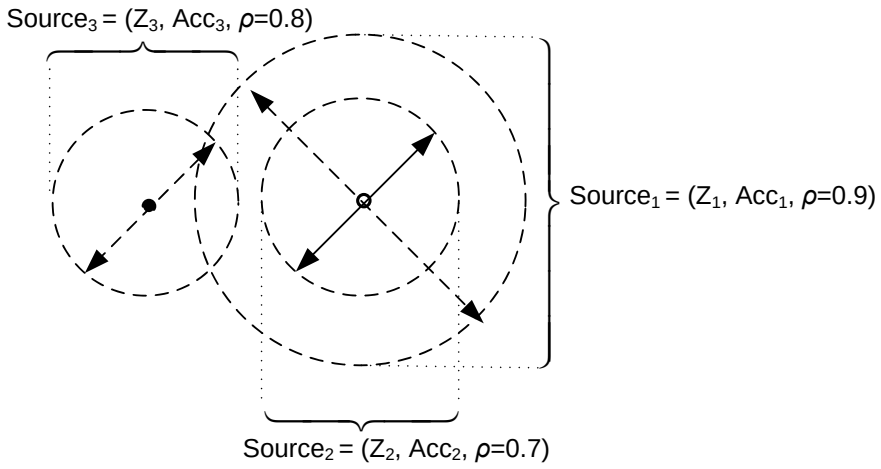


Figure 6.1: Data provided by each location information source.

6.3.1 Frame of Discernment in SESM

Since the final result of any localization process is a *location*, then the frame of discernment must encompasses all possible locations that the object of interest (e.g., vehicle) can be in. However, spatial data is continuous which makes it unsuitable for DSET. Therefore, the domain of spatial data of interest must be discretized. Fortunately, the localization

accuracy provided by a source of information can be used to split the spatial data domain into two areas. As in Figure 6.2, A_1 is a circle centred at the location estimate, Z_1, with radius equal to the accuracy value, Acc_1, and $\bar{A}_1$ signifies the remaining area of the area of interest. Hence, the actual location of the vehicle can be in one of these areas. Accordingly, $\Theta_1 = \{A_1, \bar{A}_1\}$, and $\Omega_1 = \{\phi, \{A_1\}, \{\bar{A}_1\}, \{A_1, \bar{A}_1\}\}$.

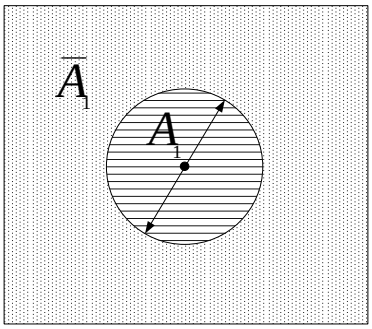


Figure 6.2: Discretizing the area of interest using localization accuracy.

Consider another source of information that provides A_2 which is centred at location estimate, Z_2, with radius equal to Acc_2 where $Acc_2 \neq Acc_1$. Therefore, $\Theta_2 = \{A_2, \bar{A}_2\}$, and $\Omega_2 = \{\phi, \{A_2\}, \{\bar{A}_2\}, \{A_2, \bar{A}_2\}\}$. Since $\Omega_1 \neq \Omega_2$, it is impossible to combine the data provided by the two sources. Hence, one must create a universe of discourse that covers the area of interest to facilitate the combination process.

To accomplish this task, one needs to discretize the area of interest into smaller areas called cells denoted by $\chi_{(j)}$, where j is a unique identification number. The shape of these cells must be selected so that they become mutually exclusive and exhaustively cover the area of interest. For simplicity I selected the cells' shape to be a square. Each cell is represented by its centre; however, the vehicle is considered to be anywhere within a cell. Therefore, the size of the cells is chosen as small to reflect high accuracy. On the other hand, as the cell size becomes smaller, the number of these cells increases dramatically

which raises the computation complexity. Thus, the cell size is selected to be comparable with the best accuracy required by the application (i.e., ϑ). Figure 6.3 shows the minimal cell size when the accuracy required is at least 10 m. The resultant frame of discernment depicted in Figure 6.4, $\Theta = \{\chi_{(1)}, \ldots, \chi_{(L^2)}\}$, is independent of the location accuracy provided by the PLUs.

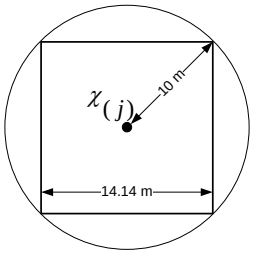


Figure 6.3: Example of square cell dimension when 10 m accuracy is needed.

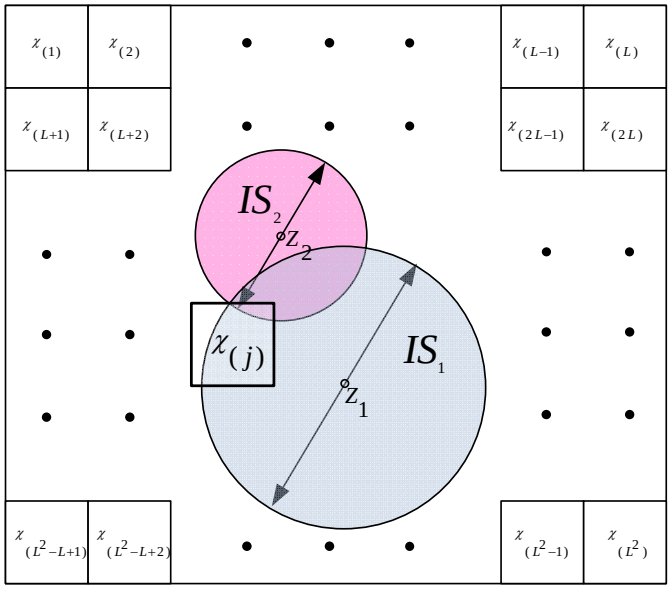


Figure 6.4: Discretizing the area of interest using mutual exclusive and exhaustive cells.

6.3.2 Evidence Construction in SESM

The next aspect of the spatial evidence structure model is the formation of the evidences provided by the sources of information in terms of focal elements, subsets of Θ, and their respective mass function, $m(\cdot)$. Both, the integrity level and localization accuracy provided by the sources of information are utilized here to accomplish this task.

Consider a source of information, IS_i, claims location estimate equals Z_i, localization accuracy equals Acc_i, and integrity level equals ρ_i. The Euclidean distance between the centre of $\chi_{(j)}$ and Z_i is denoted by $d_{j,i}$. For each IS_i, let the cells $\chi_{(j)}$ that have $d_{ij} \leq Acc_i$ be collectively supported by the integrity level ρ_i. Therefore, a *simple support function* [119] is used to set the mass function. A simple support function is a special case of a mass function, where the evidence only supports a certain subset E of Θ, and zero mass is assigned to all subsets of Θ other than E (i.e., the focal elements using a simple support function are $\{\{E\}, \{\Theta\}\}$). So, for any $D \subseteq \Theta$ the mass function is as follows.

$$m_i(D) = \begin{cases} \rho_i \text{ if } D = E, \text{ where } E = \left\{\chi_{(j)} | d_{ij} \leq Acc_i\right\} \\ 1 - \rho_i \text{ if } D = \Theta \\ 0 \text{ otherwise} \end{cases} \tag{6.12}$$

As an example, in Figure 6.4 where we have two location sources, IS_1 and IS_2, it can be seen that $\chi_{(j)} \in E_1$ while it is not the case for IS_2 (i.e., $\chi_{(j)} \notin E_2$).

6.3.3 SESM Combination Rule

As seen above, each source, IS_i, has two focal elements, $\{\{E_i\}, \{\Theta\}\}$. Typically in Dempster-Shafer combination rule, conflicts among sources of information are completely ignored, which leads to counterintuitive results in some situations when significant conflict exists. R. Yager's modified combination rule depicted in Equation 6.11 is adopted here to

avoid such issue. Yager's combination rule for n sources is shown below.

$$K = \sum_{\cap_{i=1}^{n} A_{ih} = \phi} m_1(A_{1h}) \cdot \cdots \cdot m_n(A_{nh}), \tag{6.13}$$

$$m^y(\phi) = 0, \tag{6.14}$$

$$m^y(D_r) = \sum_{\cap_{i=1}^{n} A_{ih} = D_r} m_1(A_{1h}) \cdot \cdots \cdot m_n(A_{nh}),\ r = 1, \ldots, R \tag{6.15}$$

$$m^y(\Theta) = \left(\sum_{\cap_{i=1}^{n} A_{ih} = \Theta} m_1(A_{1h}) \cdot \cdots \cdot m_n(A_{nh}) \right) + K \tag{6.16}$$

where $A_{ih} \in \{\{E_i\}, \{\Theta\}\}$, and $D_r \subset \Theta$. This combination rule produces $R + 1$ focal elements including Θ. The pignistic probabilities, $BetP(D_r)$, of the resultant focal elements are computed as in Equation 6.6 to make a decision about the fused location estimate, accuracy, and aggregated integrity:

- The focal with the highest $BetP$ is selected as a winning focal, D^*

$$D^* = \arg\max_{D_r} BetP(D_r), \text{ where } r = 1, \ldots, R + 1$$

- The centroid of the cells in the winning focal is considered to be the fused location estimate

$$Z_f = \frac{\sum_{\chi_k \in D^*} \text{centre of } \chi_k}{|D^*|}$$

- The accuracy of Z_f is considered to be equal to the distance between Z_f and the farthest $\chi_k \in D^*$

- The aggregated integrity is set to be equal to $BetP(D^*)$ since the pignistic probability reflects the belief transferred from the contributing location sources [120].

This completes the SESM fusion that yields the three pieces of information we are interested in, namely, location estimate, localization accuracy, and aggregated integrity.

6.4 Location Estimates Fusion Paradigms

In this section different hybrid vehicle localization systems are considered for the real-life vehicle traces presented in the two previous chapters. The sources of location information used in providing the preliminary localization are a GPS receiver and the simulated FP measurements presented in Chapter 4. The localization accuracy is determined using CBAC algorithm while the localization integrity is assessed using both the TDI model and the generalized TDI model presented in Chapter 5. The accuracy bands evaluated by the integrity models are 10 m and 20 m for the GPS measurements, and 20 m for the FP measurements. The following subsections are focused on various approaches to the fusion process within the fusion layer. Comparative results are also provided.

6.4.1 SESM Based Fusion

The fusion process here is fed with the location estimates and the integrity levels for different accuracy bands as in Figure 6.5. Since the IMU_{GPS} is evaluating two nested accuracy bands for the same PLU_{GPS} location estimate, GPS data is considered to be coming from two dependent location information sources. Due to the dependency among the two sources, each source is fused independently with the data produced by the PLU_{FP}. The fusion result with the least uncertainty is selected.

The pseudo code 6.4.1 shows the fusion paradigm steps where $\oplus_{SESM}(\xi_i, \xi_j)$ means building the spatial evidence structure model and performing the combination rule depicted in section 6.3.3 for the two quality indexes ξ_i and ξ_j.

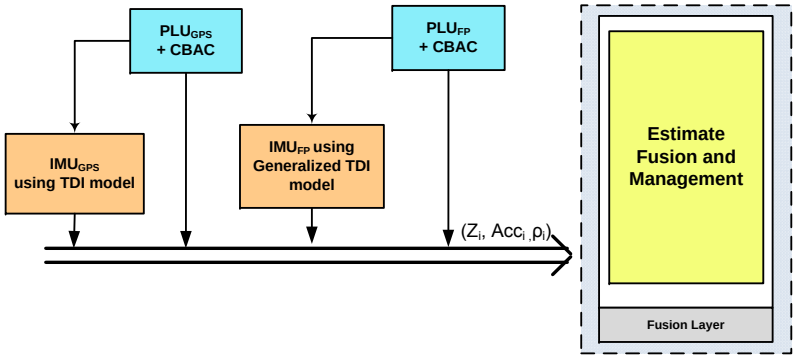


Figure 6.5: SESM fusion paradigm using GPS and FP measurements.

Algorithm 6.4.1 : Pseudo Code for Fusing GPS and FP measurement using SESM

Input: ξ_1 =(Z_{GPS}, Acc_{10m}, ρ_{10m}), ξ_2 =(Z_{GPS}, Acc_{20m}, ρ_{20m}), and ξ_3 =(Z_{FP}, Acc_{20m}, ρ_{20m});
Output: Z_f, Acc_f, and ρ_f;
/*Fuse FP and GPS$_{10m}$:*/
$[Z_1, Acc_1, \rho_1, m^y_{10}(\cdot)] = \oplus_{SESM}(\xi_1, \xi_3)$;
/*Compute $Bel\,(D^*_1)$ and $Pl\,(D^*_1)$ using $m^y_{10}(\cdot)$*/
$Uncertainty_{(1)} = Pl\,(D^*_1) - Bel\,(D^*_1)$;
/*Fuse FP and GPS$_{20m}$:*/
$[Z_2, Acc_2, \rho_2, m^y_{20}(\cdot)] = \oplus_{SESM}(\xi_2, \xi_3)$;
/*Compute $Bel\,(D^*_2)$ and $Pl\,(D^*_2)$ using $m^y_{20}(\cdot)$*/;
$Uncertainty_{(2)} = Pl\,(D^*_2) - Bel\,(D^*_2)$;
$f = \arg\ \min_{i=1,2} Uncertainty_{(i)}$;

Moreover, this paradigm is performed on the hybrid vehicle localization system while the generalized TDI (GTDI) is used in the IMU$_{GPS}$. Figure 6.6 shows the performance of SESM in terms of the cumulative distribution function, CDF, of the localization error. As seen, the CDF of the localization error is computed for the GPS localization, FP localization, average between the GPS and FP location estimates, the GPS and FP SESM fusion using TDI for the GPS accuracy, and the GPS and FP SESM fusion using GTDI for both the GPS and FP accuracies. The average between GPS and FP localization produced

intuitive result that falls between the performance of the GPS and FP. SESM fusion, while using TDI to attribute the GPS accuracy, outperforms the GPS-only localization especially for errors larger than 10 m. When GTDI is used in SESM fusion the performance degraded below that of GPS-only localization. This degradation is due to the increase in the uncertainty of evaluating the localization integrity when GTDI is used (*cf.* Figure 5.23). It is worth mentioning here that the usage of GTDI is more practical in many situations indicated in the previous chapter. It is mentioned in Chapter 5 that such degradation can be handled by enforcing a number of rules that guide the fusion process. The tuning of the fusion process is tackled in the following section.

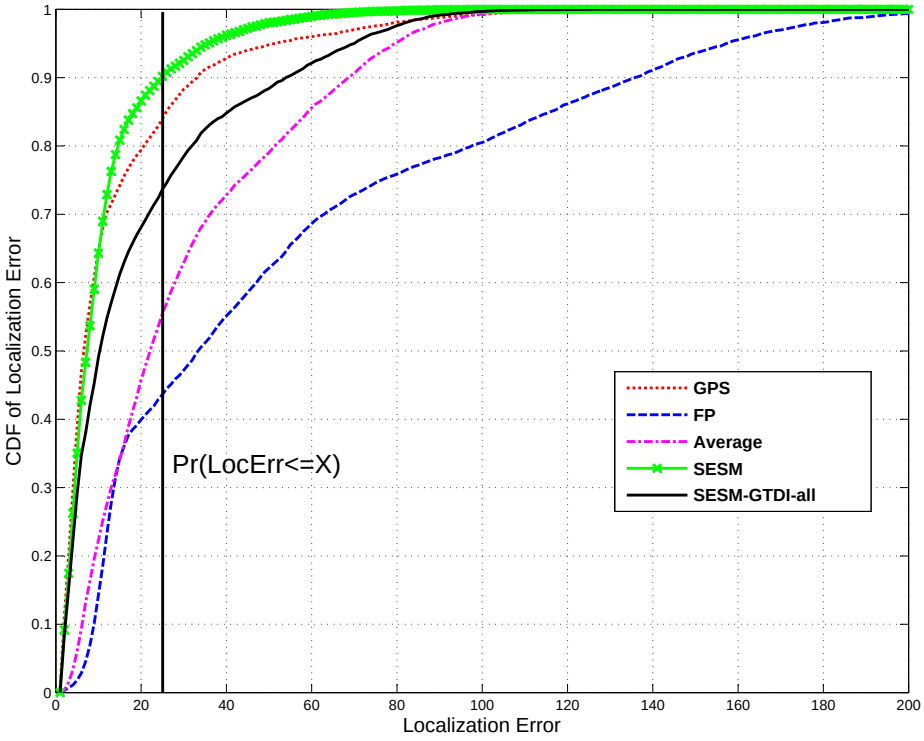


Figure 6.6: Localization error performance for SESM paradigm.

6.4.2 Channel Selection Guided SESM Fusion

The inaccurate information provided by the GTDI about certain source of location information causes a significant degradation in the fusion process performance. Therefore, inaccurate sources must be excluded in the fusion process. As a first step to improve the fusion process, a Channel Selection (CS) mechanism is utilized here. In GTDI, the significant effort spent on classifying the localization accuracy is not perfectly utilized. Thus, the CS mechanism is used to select the source with better accuracy based on the CBAC output. In the case where sources provide comparable accuracies, fusion is performed. The fusion is performed one time by simply linearly averaging the location estimates, and another time by implementing SESM based fusion but this time there would not be dependent sources since they are already filtered out using CBAC.

Table 6.1 depicts the set of rules enforced in the CS mechanism. These rules are extracted based on the observations made on the performance of the GTDI shown in Figure 5.23. For example, GPS localization has a very high integrity level when its accuracy is declared as 10 m; however, the GPS integrity level is not as high when accuracy is declared as 20 m. Consequently, GPS channel is selected whenever 10 m accuracy is declared while fusion between GPS and FP estimates is performed when both declare 20 m accuracy – the remaining situations are listed in the table below.

Table 6.1: CS rules for GPS-FP fusion when GTDI is used.

FP bands \ GPS bands	LocErr $\leq$ 10 m	10 m $<$ LocErr $\leq$ 20 m	LocErr $>$ 20 m
LocErr $\leq$ 20 m	ξ_{GPS}	$Fuse(\xi_{GPS}, \xi_{FP})$	ξ_{FP}
LocErr $>$ 20 m	ξ_{GPS}	ξ_{GPS}	$Fuse(\xi_{GPS}, \xi_{FP})$

The CDFs of the CS guided fusion are depicted in Figure 6.7. The improvement in the localization error using CS guided fusion is significant with respect to the GPS localization

and the paradigm of SESM with only GTDI models. It is noticed that CS guided SESM fusion is less accurate than CS guided linear mixed fusion (denoted by CS-LinMix), yet it outperforms the SESM fusion with TDI model of GPS.

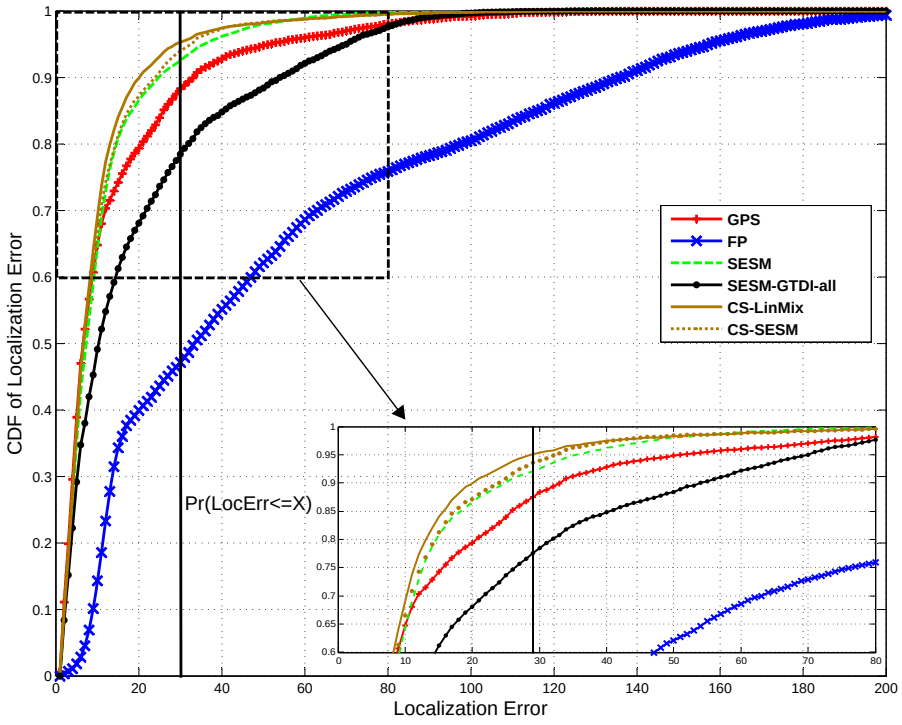


Figure 6.7: Channel selection guided fusion versus pure SESM performances.

6.4.3 Source Discounting for CS Guided SESM Fusion

The relative degradation in the fusion performance, when SESM fusion is used, provides irrefutable evidence on the existence of discredited sources of information. The fact of having discredited sources of information is a common issue in data fusion paradigms. Fortunately, DSET is able to handle this issue using source discounting mechanism shown in Equation 6.10. Since we have two sources of information (i.e., IS_{GPS} and IS_{FP}), two

discounting factors must be determined, namely, α_{GPS} and α_{FP}, as such the overall localization performance gets improved. To accomplish this task, an optimization problem is formulated as follows.

$$[\alpha_{_{GPS}}, \alpha_{_{FP}}] = \arg \max_{\alpha_{GPS}, \alpha_{FP}} \int_0^{\widehat{LE}} CDF_{csf}(\widehat{le})\ d\widehat{le} \tag{6.17}$$

$$s.t.$$

$$0 \leq \alpha_{_{GPS}}, \alpha_{_{FP}} \leq 1,$$

$$CDF_{csf}(\widehat{le}) = fun(\oplus_{csf})|_{m^{y}_{\alpha_{GPS}}(\cdot), m^{y}_{\alpha_{FP}}(\cdot)}$$

where $\widehat{le}$ signifies the resultant localization error after performing the fusion process, $\widehat{LE}$ signifies an upper bound for the variable $\widehat{le}$, $CDF_{csf}\left(\widehat{le}\right)$ signifies the cumulative distribution function of the localization error when CS guided SESM fusion (i.e., $\oplus_{csf}$) is used, and $m^{y}_{\alpha_{i}}(\cdot)$ signifies the Yager's modified mass function when α_{i} is used as a discounting factor.

The aim of Problem 6.17 is to maximize the probability of obtaining smaller error than certain value (e.g., $\widehat{LE}$). The objective function is complex, non-smooth, and may have many local maxima due to the different combination rules shown in Table 6.1 as well as the mass evaluation rules depicted in equations 6.14 to 6.16. Hence, derivative (i.e., traditional) based optimization methods will not work in solving this problem. In such cases, especially when the search space is spanned by large number of variables, global optimization methods [127], such as 1) stochastic methods: simulated annealing, direct Monte-Carlo sampling, etc., and 2) heuristics/metaheuristics: evolutionary algorithms, swarm-based optimization algorithms, to name few, are expected to produce optimal solution. Fortunately, the search space of this optimization problem is constrained and spanned by only two variables: α_{GPS} and α_{FP}. Therefore, the search space is sampled and the objective function is evaluated for

each sample to investigate the surface of the objective function. As depicted in Figure 6.8, the objective function is maximized when α_{GPS} is greater than 0.8 and α_{FP} is less than 0.2. In other words, the GTDI assessment of the GPS localization is much less credited that the GTDI assessment of the FP localization when the performance accuracy is classified as between 10 m and 20 m. It is also observed that the objective function becomes insensitive to the change in both α_{GPS} and α_{FP} when they are greater than 0.5 and 0.4, respectively. This insensitivity to the evidence provided by the IS_i is a result of discrediting these sources to the extent that Θ, the set signifying complete ignorance, becomes always the winner set ($D^* = \Theta$) so that the result of the fusion process is the same regardless of the change in the tuning parameters: α_{GPS} and α_{FP}.

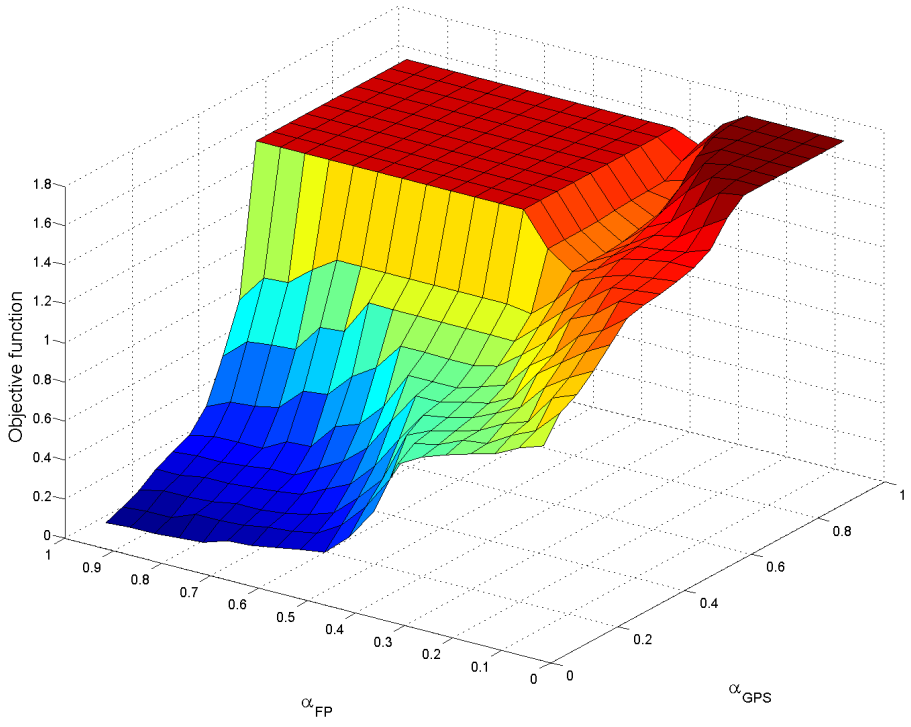


Figure 6.8: Objective function surface as a function in α_{GPS} and α_{FP}.

Thus, the discounting factors are set to the values 0.8 and 0.2 for α_{GPS} and α_{FP}, respectively. The CS guided SESM fusion is implemented while using these discounting factors, denoted by $CS - alpha - SESM$, and then compared with the previous fusion paradigm. Figure 6.9 shows the comparison between the fusion processes developed in this chapter in terms of their CDF of localization error. As seen, the $CS-alpha-SESM$ fusion significantly outperforms all the implemented fusion methods when errors are larger than 12 m and that is because source discounting factors are more effective when the accuracy is worse than 10 m.

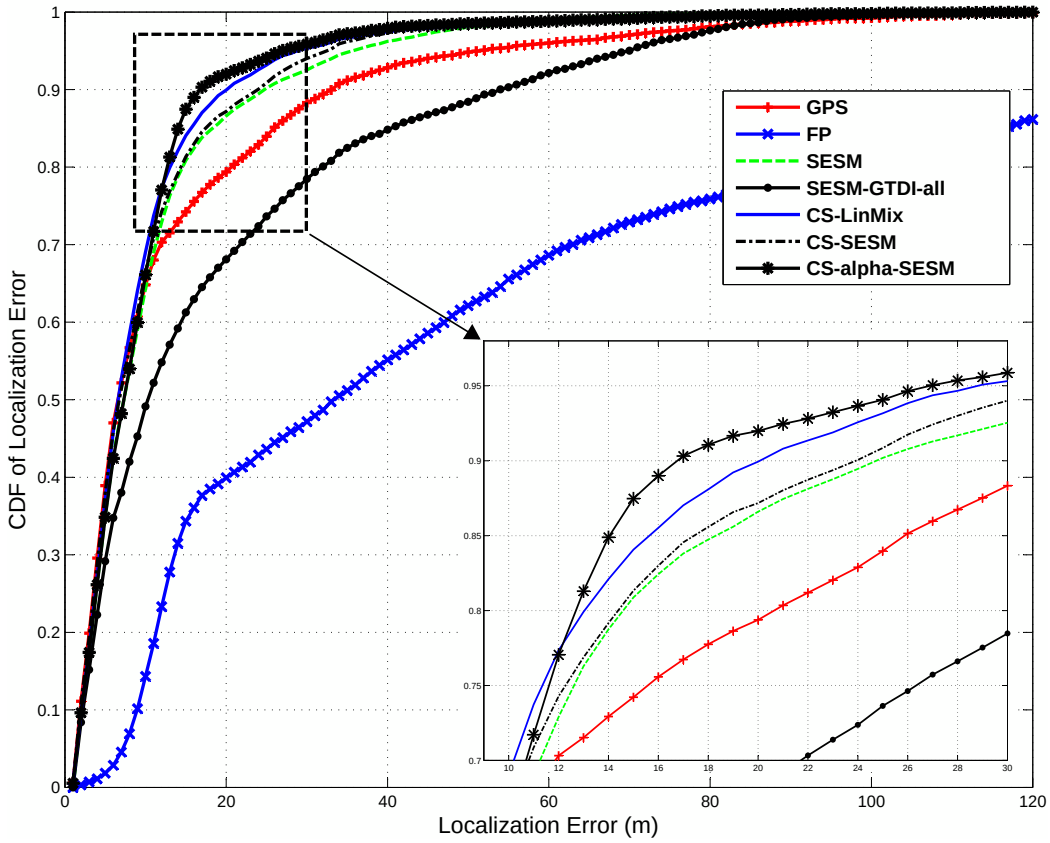


Figure 6.9: The effect of source discounting on the fusion process performance in terms of CDFs of localization error.

6.4.4 Comparative Results

Several comparisons are conducted in this section among the localization methods mentioned above. The comparisons are categorized into two categories that present the improvement in the localization error and the improvement in the claimed accuracy and integrity levels.

Improvement in Location Accuracy Bands:

The performance of the implemented fusion methods are further investigated from a closer point of view. The percentage of improvement per localization error bands depicted in Table 4.3 is calculated. Since FP localization error is much worse than that of the GPS, only GPS error bands are compared with the fusion results. Error band covering errors less than 10 m is denoted by b_1, error band covering errors between 10 m and of 20 m is denoted by b_2, and error band covering errors larger than 20 m is denoted by b_3.

In terms of location error minimization, the percentage of the errors shifted from b_3 to both b_2 and b_1 is computed. The same computation is done to the errors in b_2. Table 6.2 depicts the figures of these computations per each fusion paradigm. It can be seen that SESM fusion outperforms SESM-GTDI fusion due to the extra information provided. Nevertheless, CS fusion versions recover from the degradation caused by SESM-GTDI fusion, especially when source discounting is used.

Table 6.2: Improvement in the localization bands after the fusion process.

Error bands	SESM	SESM-GTDI	CS-LinMix	CS-SESM	CS-alpha-SESM
$b_3 \longrightarrow b_2$	38.53%	27.06%	38.38%	24.54%	38.00%
$b_3 \longrightarrow b_1$	25.08%	18.96%	18.96%	18.43%	32.11%
$b_2 \longrightarrow b_1$	39.95%	30.21%	30.21%	29.35%	51.16%

Another statistic called Gain-b_i is calculated, Equation 6.18. Gain-b_i signifies the percentage of the change in the number of measurements falling in band b_i after performing the fusion process.

$$\text{Gain-}b_i = \frac{\text{size of band } b_i \text{ using GPS} - \text{size of band } b_i \text{ after fusion}}{\text{size of band } b_i \text{ using GPS}} \times 100 \qquad (6.18)$$

Figure 6.10 shows the performance of fusion methods in terms of Gain-b_i per each band. Positive gains indicate increasing the size (number of measurements) of the band; while negative gains indicate shrinking the size (number of measurements) of the band. Few concluding remarks can be obtained from this chart:

- the significant improvement in the localization process happens in decreasing the huge errors in b_3 to b_2,
- part of the measurements in b_3 and b_2 have been gained in b_1 – a gain that does not seem large due to the extremely large number of samples in b_1–,

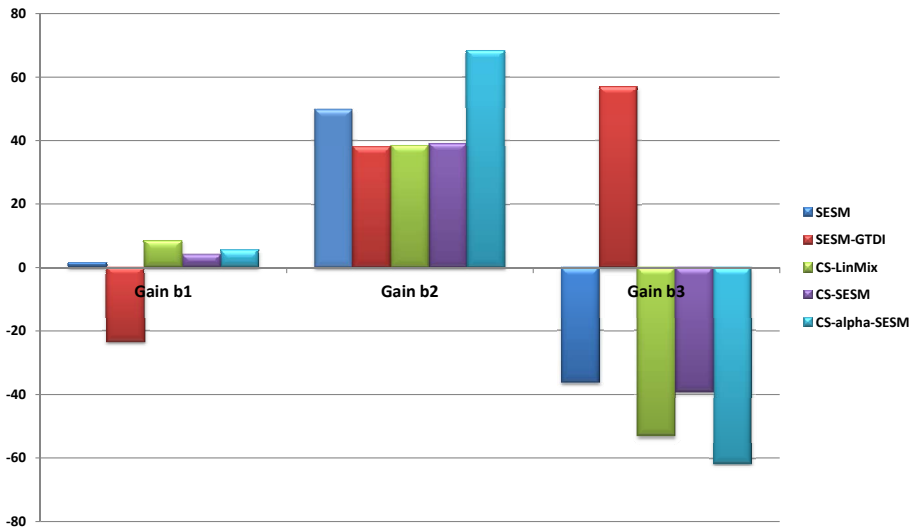


Figure 6.10: The statistics of band gains after performing estimate fusion.

- SESM-GTDI fusion, which is using GTDI for GPS integrity assessment, significantly degrades the performance in b_1 and b_3 as a result of giving stronger evidence to FP localization which is typically much less accurate than GPS localization, and
- CS fusion versions clearly recover from the aforementioned degradation in all bands, though CS versions are still using GTDI for GPS integrity assessment.

Improvement in Integrity Assessment and Aggregation:

The fusion process produces two pieces of information which are the claimed localization accuracy and the integrity aggregated. In this section, comparisons between the GPS performance and the fusion paradigms are presented. First, localization accuracy is considered. Three accuracy indexes are used in this comparison: $Acc_1 \leq 10m$, $Acc_2 \leq 20m$, and $Acc_3 > 20m$. The performance metric used here is the number of false alarms produced by a technique when claiming specific accuracy index. A false alarm is counted if the actual location is outside the accuracy band specified by the claimed index. Table 6.3 shows the

Table 6.3: Accuracy false alarm comparison.

Technique	**Statistic**	**Accuracy Index**		
		Acc_1	Acc_2	Acc_3
GPS	No. of Claim	2356	3477	824
	False alarm	39	780	260
CS-LinMix	No. of Claim	2356	4109	192
	False alarm	39	492	72
CS-SESM	No. of Claim	2356	3902	399
	False alarm	39	558	160
CS-alpha-SESM	No. of Claim	2356	4002	299
	False alarm	39	381	178

statistics collected from the experimental work. The techniques have similar performance in terms of Acc_1; however, it can be seen that the fusion techniques achieved a significant improvement in Acc_2 and Acc_3 in terms in lessening the false alarms. Furthermore, the increase in claiming Acc_2 instead of claiming Acc_3, when estimate fusion paradigms are used, confirms the localization improvement presented in the previous section. Amongst the estimate fusion paradigms, CS-LinMix is the best in diminishing the false alarm rate of claiming specific accuracy band.

Second, improvement in estimation integrity level is investigated. Intuitively, integrity is aggregated via fusion and becomes higher in level than that of the individual sources. This statement is verified by showing the instantaneous integrity level of the accuracies claimed by the PLU_{GPS} and the fusion paradigms. Figure 6.11 depicts six snapshots of integrity assessment provided by the different techniques. The snapshots are taken while a vehicle is travelling through different environment conditions. The GPS-GTDI assessment is represented by the solid red line in the figure while the integrity claimed by the fusion paradigms is represented by the blue-marked solid line, dotted-dashed line, and dashed line. It can be clearly seen that fusion paradigms produce higher level of integrity, when they report their location estimate and accuracy, compared with the GPS-GTDI integrity levels in almost all snapshots regardless of the measurement conditions. The integrity of CS-LinMix fusion paradigm achieves the highest level – a performance that supports the results of the accuracy false alarm rates presented above. The integrity levels reported by the CS-SESM and CS-alpha-SESM fusion paradigms are higher than the GPS-GTDI integrity and below that of the CS-LinMix fusion paradigm in most of the cases. It is worth mentioning here is that despite the lower integrity level reported by CS-alpha-SESM fusion paradigm with respect to CS-LinMix fusion paradigm the former paradigm achieved better localization error performance as per Figure 6.9. This observation reveals the fact that the

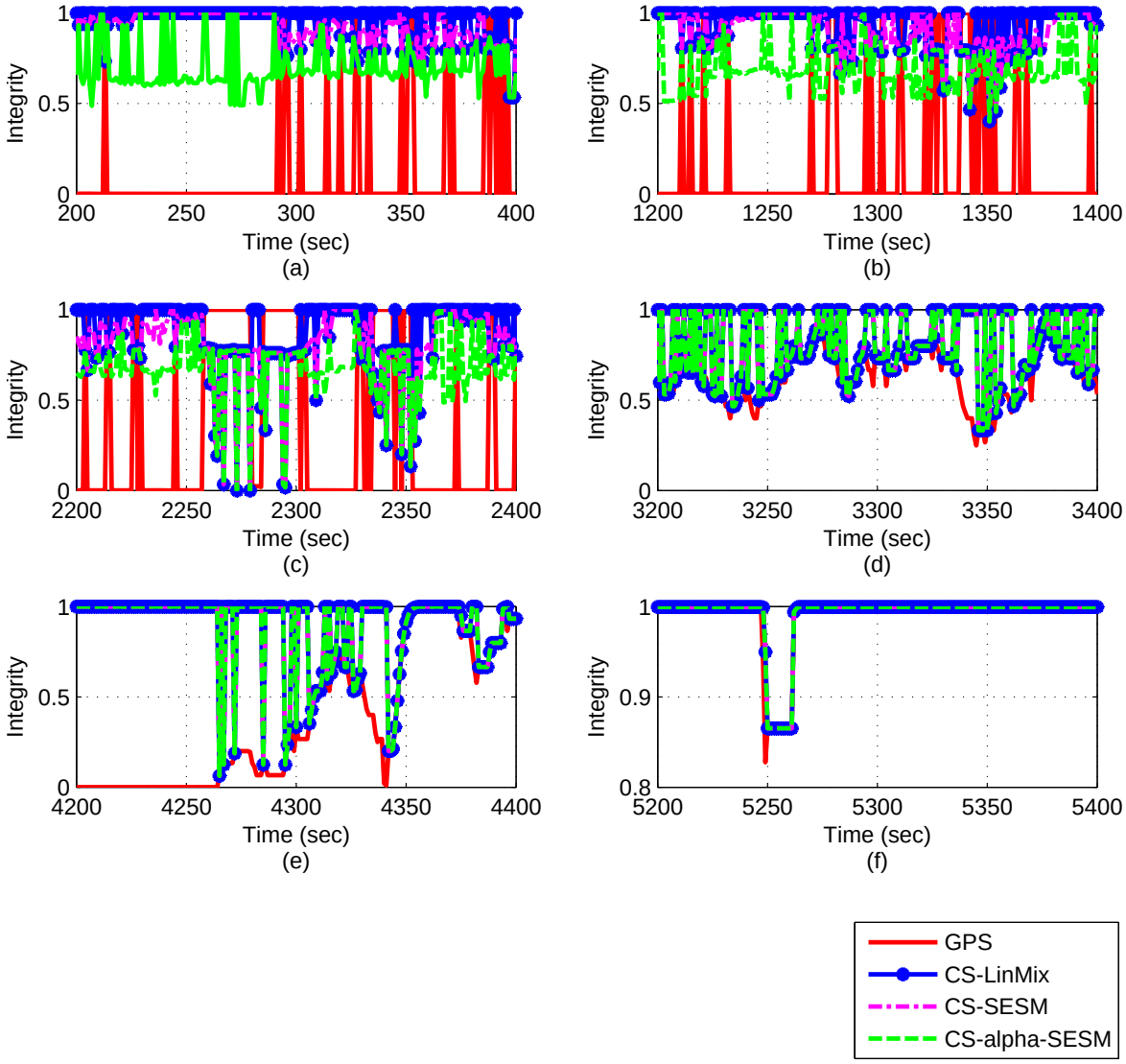


Figure 6.11: Integrity aggregation compared with GPS integrity level in different measurement conditions.

different fusion paradigms can be used to achieve different goals. For example, if the target is to minimize the localization error as much as possible and the application can tolerate some false alarms, such as in navigation systems, then CS-alpha-SESM fusion paradigm

is the best choice. On the other hand, if the target is to minimize the accuracy false alarms and increase the estimate trust level, such as in vehicular critical applications, then CS-LinMix is the best choice among the developed paradigms.

6.5 Task-Driven Resource Allocation

The aim of the localization approaches investigated up to this section is to achieve the best localization in terms of accuracy and integrity. In practice, other factors must be considered in the evaluation of the localization performance, such as the power consumed to process the data, the service cost of accessing cellular networks, the speed of processing the data, etc. We call these factors performance criteria, which are function in the localization technique at hand and the state of nature (i.e., position and measurement conditions). Most of the current localization techniques, single or multi-modality systems, consider one performance criterion at a time using fusion and/or channel selection paradigms. Consequently, one might ask which is the best modality or technique's assignment to improve specific performance criterion. The issue becomes more challenging when more than one conflicting criterion is considered simultaneously. A multicriteria formulation is introduced in the next section to answer the aforementioned question.

6.5.1 A Multicriteria Formulation

Consider a localization system that employs a set of N localization techniques available under the disposal of the estimate fusion and management layer. These techniques are indexed by the set $\Psi = \{T_1, \ldots, T_n, \ldots, T_N\}$ and estimating the true location X_t at time t. Each technique T_i uses its measurement resources to compute an estimate $Z_{i,t}$ and provide a quality index $\xi_{i,t}$ which is in terms of different criteria β_m (i.e., $\xi_{i,t} = (\beta_1, \ldots, \beta_m, \ldots, \beta_M)$).

A binary selection vector, $V_t = (v_1, \ldots, v_N)$ where $v_n \in \{0, 1\}$ and $\sum_n v_n = 1$, is used to select T_i that optimize $\xi_{i,t}$. Therefore, each criterion β_m is function in X_t and V_t as follows.

$$\beta_m = f(X_t, V_t),\ m = 1, \ldots, M \tag{6.19}$$

Nevertheless, the solution of optimizing the group of criteria β_m simultaneously may do not exist due to conflicts among the criteria. It is also possible to quickly consume the resources of the system while providing quality very much beyond the demand of the system. A key for intelligence of such system is its capability to utilize its different resources according to the different needs of the task at hand. This is what we call *Task-Driven Resource Allocation*. In order to tackle the two aforementioned issues, it is imperative to combine the criteria functions in one problem and incorporate the demands of the task at hand with respect to each criterion so that a *Pareto-optimal* solution can be obtained.

6.5.1. *Pareto-optimal solution:* *A vector V_t^* is said to be a Pareto-optimal for the multicriteria problem in* 6.19 *if and only if there is no vector V_t* **that can improve any** *β_m, $m = \{1, \ldots, M\}$, without worsening at least one β_k, $k = \{1, \ldots, M\} - \{m\}$.*

Without loss of generality, the demands of the localization tasks considered here constitute the target performance needed by a task/application. Basir and Shen investigated in [128] other types of demands, such as soft and hard task demands. The task target is signified by the vector $\Xi_t = (\eta_1, \ldots, \eta_M)$ where η_m signifies an assumed demand on the m_{th} performance criterion. Thus, the objective function becomes as follows.

$$\left[\sum_{m=1}^{M} \left(\beta_m \left(X_t, V_t\right) - \eta_m\right)^d\right]^{\frac{1}{d}} \tag{6.20}$$

where $1 \leq d \leq \infty$ and the the value of d specifies the distance type between the task demand levels represented by Ξ_t and the functional-efficient solutions. For example, if d is

set to 2, the objective function is simply the Euclidean distance metric.

In some situations the task may further provide some important information structure on the set of criteria functions. This means that the criteria functions do not possess the same level of importance. In such situation, the objective function is updated to accommodate a weighting scheme for all criteria functions:

$$\left[\sum_{m=1}^{M}\left(W_m\left(\beta_m\left(X_t, V_t\right)-\eta_m\right)\right)^d\right]^{\frac{1}{d}} \tag{6.21}$$

where $W_1, \ldots, W_M$ are chosen to satisfy the importance structure induced by the problem on the set of criteria functions. The more important is a criterion function the heavier weight it is assigned. It is often assumed that $0 < W_m < 1$ and $\sum_m W_m = 1$. Consequently, the Pareto-optimal solution V_t^* for the multicriteria optimization problem can be stated as follows.

$$\begin{aligned} V_t^* = \arg \min_{V_t} & \left[\sum_{m=1}^{M}\left(W_m\left(\beta_m\left(X_t, V_t\right)-\eta_m\right)\right)^d\right]^{\frac{1}{d}} \\ & s.t. \\ & V_t = (v_1, \ldots, v_n, \ldots, v_N), \text{ where } v_n \in \{0,1\}, \\ & \sum_n v_n = 1, \text{ and} \\ & \sum_m W_m = 1, \text{ wher } 0 < W_m < 1 \text{ and } m = 1, \ldots, M \end{aligned} \tag{6.22}$$

It is possible to generate Pareto-optima for the criteria functions by varying the weights W_m in the optimization problem of Equation 6.22.

6.5.2 Task-Driven Resource Allocation Implementation

The effectiveness of task-driven resource allocation is presented here using four single/multi modality localization techniques, namely, GPS, FP, CS-LinMix fusion, and CS-alpha-SESM fusion. Using such techniques in one system yields contrasted power consumption due to their different computational complexity, and incurs some financial cost in accessing cellular networks. The criteria functions chosen in this implementation are in terms of accuracy, integrity, power consumption, and financial cost per every location estimate. Table 6.4 depicts the assumed power consumption and financial cost per every location estimate for the selected techniques, as reported in [129–131]

Table 6.4: Power and financial cost of the chosen techniques.

Technique	**Power Consumption (mW)**	**Cash Cost (¢)**
GPS	370	free
FP	500	0.5
CS-LinMix	920	0.5
CS-alpha-SESM	1070	0.5

Both CS-LinMix and CS-alpha-SESM utilize FP technique which scans the cellular network towers and accesses data bases to obtain a preliminary location estimate; hence, they incur the same assumed cash cost. The 0.5 ¢ in the three localization techniques represents the service cost of the localization. Moreover, the two fusion paradigms utilize both GPS and FP data in order to produce a location estimate; therefore, their power consumption is equivalent to the summation of the consumption of the GPS, FP, and the consumption resultant from the computation process. Due to constructing the frame of discernment and performing all the SESM-related operations, the consumption of the computation process in CS-alpha-SESM is considered to be equivalent to that of CS-LinMix (50 mW) multiplied by three.

The application querying the vehicle location performs different tasks upon which specific demands are set. Given these demands, the pre-knowledge about the power consumption, and cash cost, as well as the real-time accuracy and integrity assessment, the estimate fusion and management is able to assign the adequate technique(s) and plane to achieve the target demand, Figure 6.12.

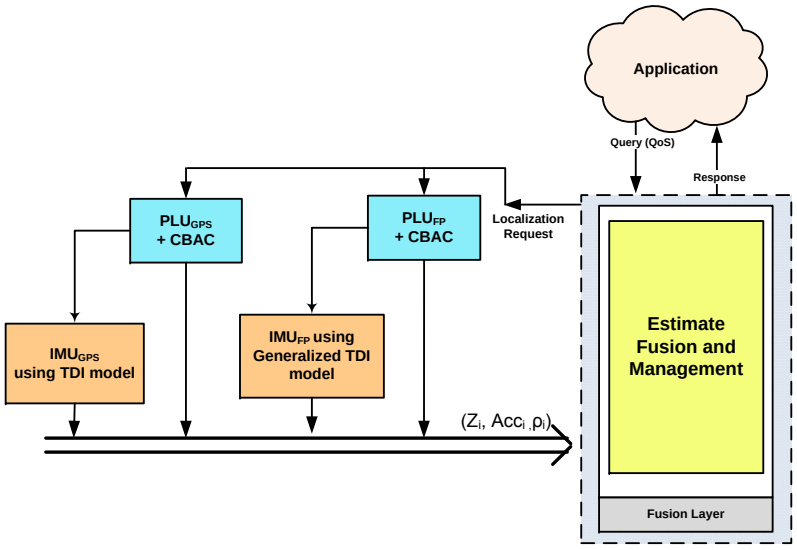


Figure 6.12: Task-driven integrity assessment and control for hybrid vehicle localization system.

The efficacy of the task-driven integrity assessment and control localization is examined over four *use cases*. Each *use case* presents certain task/application's target demands for a vehicle travelling first in the downtown of the city of Toronto and then leaves to the city of Waterloo through highway 401. Table 6.5 depicts the demands set by the *use cases* in terms of the performance criteria. Without loss of generality, the entry ∞ signifies no constraints on the performance criterion. Also, the value zero signifies spend/obtain the minimal possible value of the designated performance criterion. As explained in the previous section, the criteria functions may possess different levels of importance. These

Table 6.5: Localization demands for different *use cases*.

Criteria / *Use Cases*	Cash	Power	Accuracy	Integrity
Super performance	∞	∞	0.00 meters	100%
Navigation system	0.00 ¢/service	∞	20.00 meters	90%
Limited resources	0.25 ¢/service	0.00 mW	15.00 meters	70%
Running out of power	∞	0.00 mW	10.0 meters	100%

levels of importance are maintained using the weighting vector shown in Equation 6.22. The results of these *use cases* are then presented in terms of the CDF of the localization error, the consumed energy, and the cash spent. The techniques chosen to achieve the target performance are also depicted for each *use case*.

***Use Case* 1- Super Performance:**

In this *use case*, there are no constraints on the cash or power consumption, and the localization performance is needed to be the best. Since the information provided about accuracy and integrity has some level of uncertainty and we know that fusion paradigms produce better localization, more weight is given to the cash and power consumption than is given to the accuracy and integrity criteria functions: $W = [0.1, 0.1, 0.4, 0.4]$.

The CDF of the localization error of the task-driven localization as well as the other four techniques is shown in Figure 6.13. It can be seen that the task-driven fusion for this *use case* outperforms the CS-alpha-SESM in the error band less than 12 m, and outperforms the CS-LinMix in the error band greater than 12 m. In other words, task-driven fusion produces a logical balance between the two fusion paradigms.

Figure 6.14 shows a comparison among the different localization techniques in terms of the energy consumed and the money spent in sub-figures (a) and (b), respectively, while the sub-figure depicts the selection of the localization technique over the time of the experiment.

Since no constraints are enforced in this experiment on the cash or the power consumption, the cellular network is accessed all the time to improve the localization process, and the energy consumed is just below that of the CS-alpha-SESM due to the utilization of CS-LinMix in some situations which require less power consumption. In general, both CS-alpha-SESM and CS-LinMix alternate the localization task in the multipath area (i.e., downtown Toronto) based on the accuracy and integrity assessment. In the open sky environment the CS-alpha-SESM is chosen to produce the best localization possible.

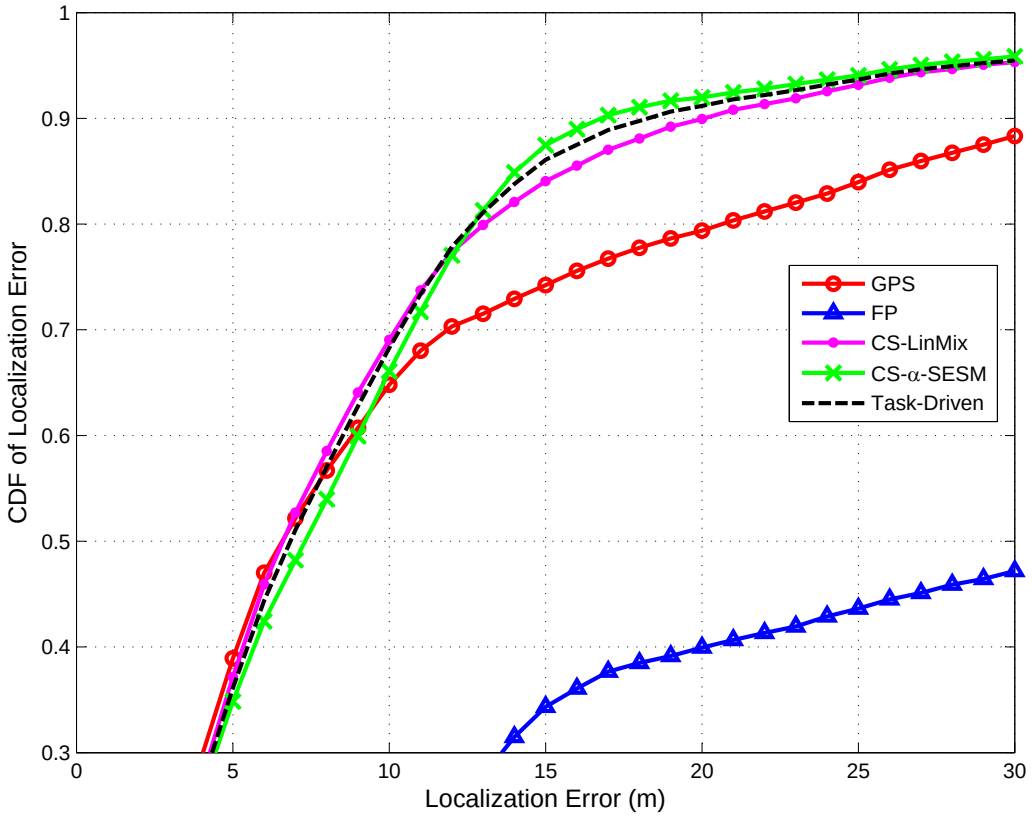


Figure 6.13: Comparison between the CDF of localization error of *use case* 1 and that of the other localization techniques.

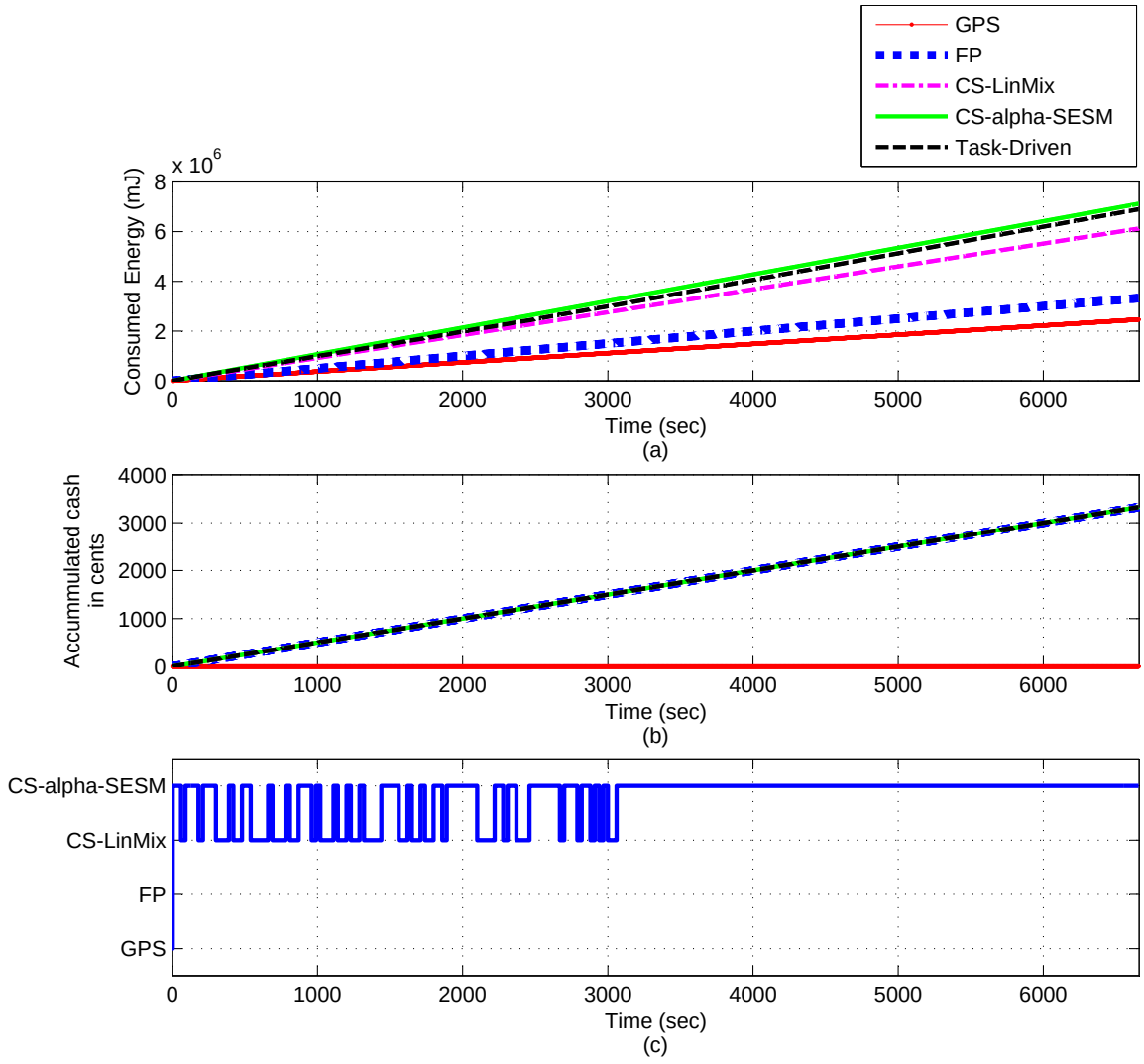


Figure 6.14: Task-driven localization performance in *use case* 1 with respect to the other localization techniques in terms of (a) consumed energy, and (b) cash spent. (c) selection of the localization techniques made by the Pareto-optimal solution.

Use Case 2- Navigation System:

In vehicles' navigation systems the power consumption does not constitute any constraint since the car battery covers the power needed. The localization accuracy does not need to

be very accurate, but the integrity needs to be relatively high. Often, there is no interest in spending cash in such *use case* except in multipath environment where GPS coverage and integrity become very weak. Therefore, more weight is assigned to the power and cash criteria and higher weight than in *use case* 1 is given to the accuracy and integrity: $W = [0.2, 0.2, 0.35, 0.25]$. Figure 6.15 depicts the performance in terms of CDF of the localization error. As seen, the performance of *use case* 2 is extremely good when errors are less than 12 m. For larger errors, task-driven localization significantly improves the GPS performance.

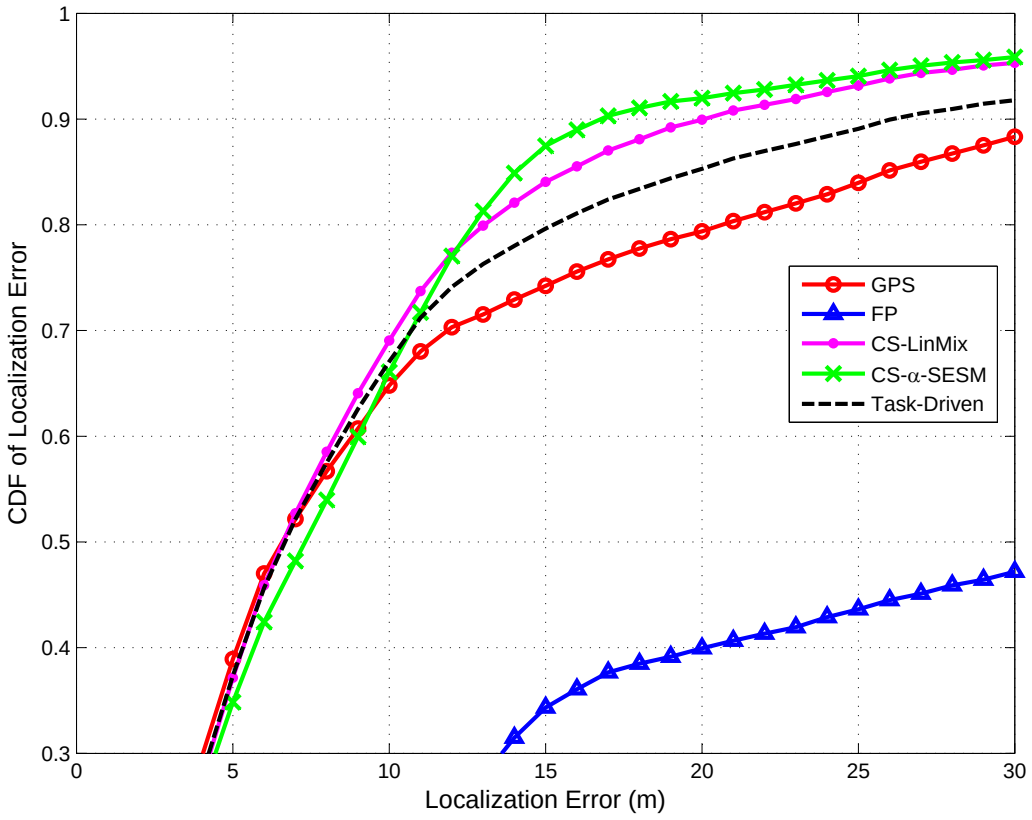


Figure 6.15: Comparison between the CDF of localization error of *use case* 2 and that of the other localization techniques.

Figure 6.16 shows the dynamics of the task driven localization as to the change in the criteria functions. This dynamic behaviour can be seen in the non-linear increase in the consumed energy and cash. It is obvious that the consumption rate is high in the

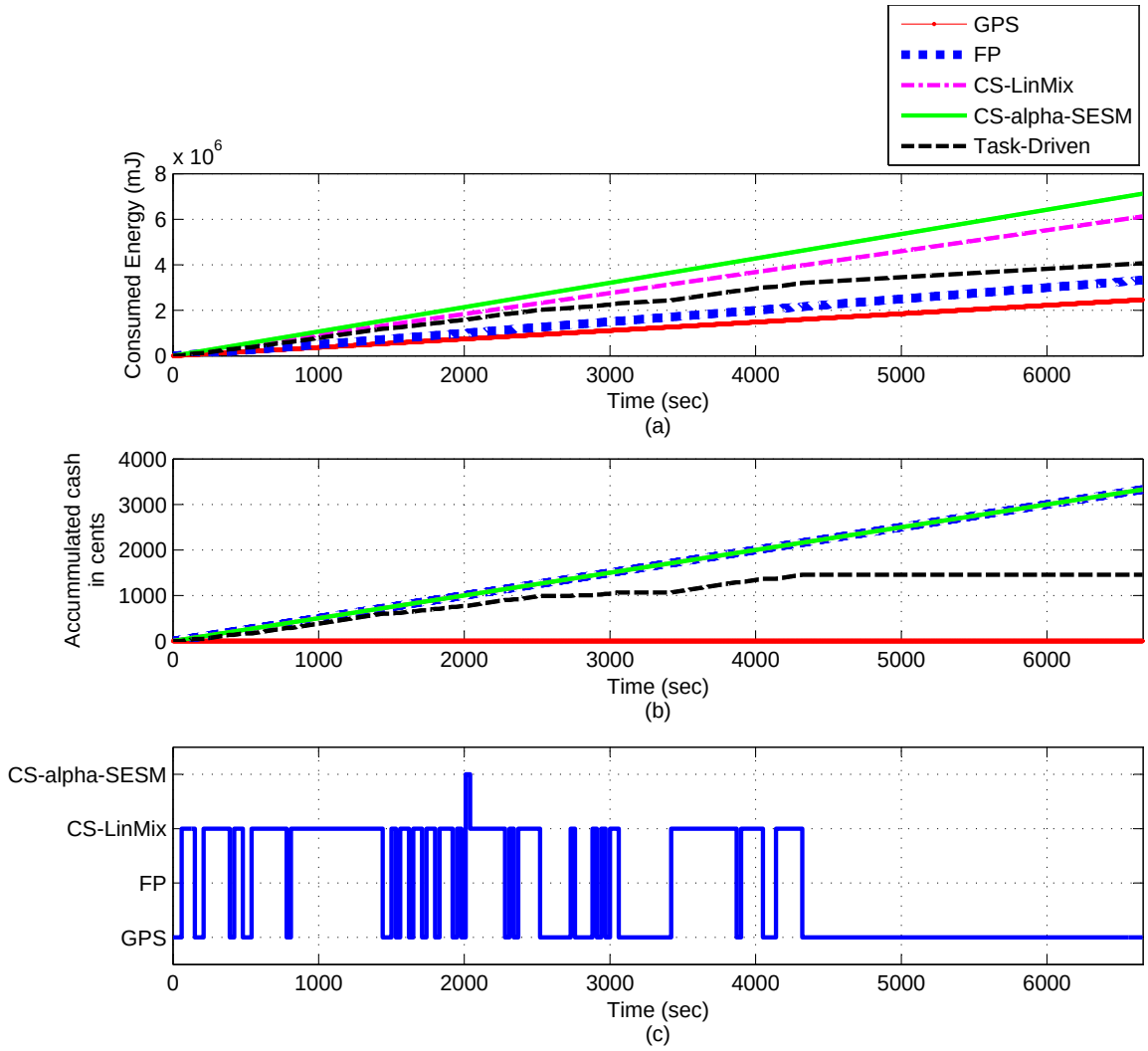


Figure 6.16: Task-driven localization performance in *use case* 2 with respect to the other localization techniques in terms of (a) consumed energy, and (b) cash spent. (c) selection of the localization techniques made by the Pareto-optimal solution.

multipath environment where task-driven localization is mainly switching between GPS and CS-LinMix fusion. FP and CS-alpha-SESM are not selected due to their relative high cash cost and energy consumption, respectively. On the highway, the task-driven localization chooses only GPS localization since it surpasses the target performance.

Use Case 3- Limited Resources

This scenario exemplifies the utilization of smartphones which are equipped with a GPS or A-GPS receiver and have an access to the cellular networks. Such smartphones provide location based services to car drivers or pedestrians. In such cases, localization accuracy is relatively relaxed to the range of 20 m but with good level of integrity. Power consumption is an issue and, therefore, it is strongly constrained. Cash is also spent with caution, so relatively more weight is assigned to the cash spending and power consumption: $W = [0.2, 0.2, 0.25, 0.35]$.

It is expected in such *use case* that the task-driven localization will focus on the utilization of GPS and FP localization and avoid the fusion paradigms due to their high consumed energy and cash cost. Figure 6.17 shows a comparison between the different localization techniques in terms of the CDF of the localization error. It can be seen that the performance of the task-driven localization is way better than the FP localization. GPS localization is, also, significantly improved in the large error band – such error band is commonly experienced in urban canyons.

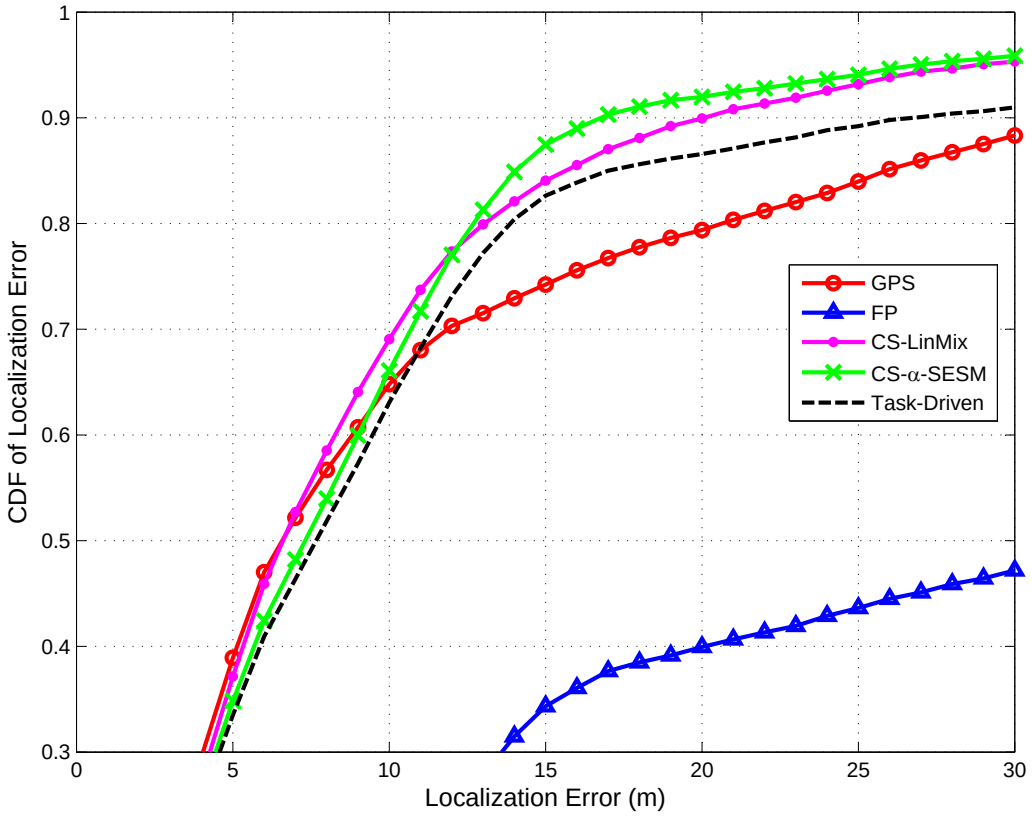


Figure 6.17: Comparison between the CDF of localization error of *use case* 3 and that of the other localization techniques.

It is evident from Figure 6.18 that task-driven localization consumes less energy that FP localization especially when the multipath effect becomes lesser. In terms of cash spending, the task-driven localization spends less than half of what the other techniques spend using cellular network assess. In sub-figure (c), it is clear that task-driven relies only on GPS and FP localization in multipath environments. In open sky environment, during this *use case*, the sole localization technique used is GPS .

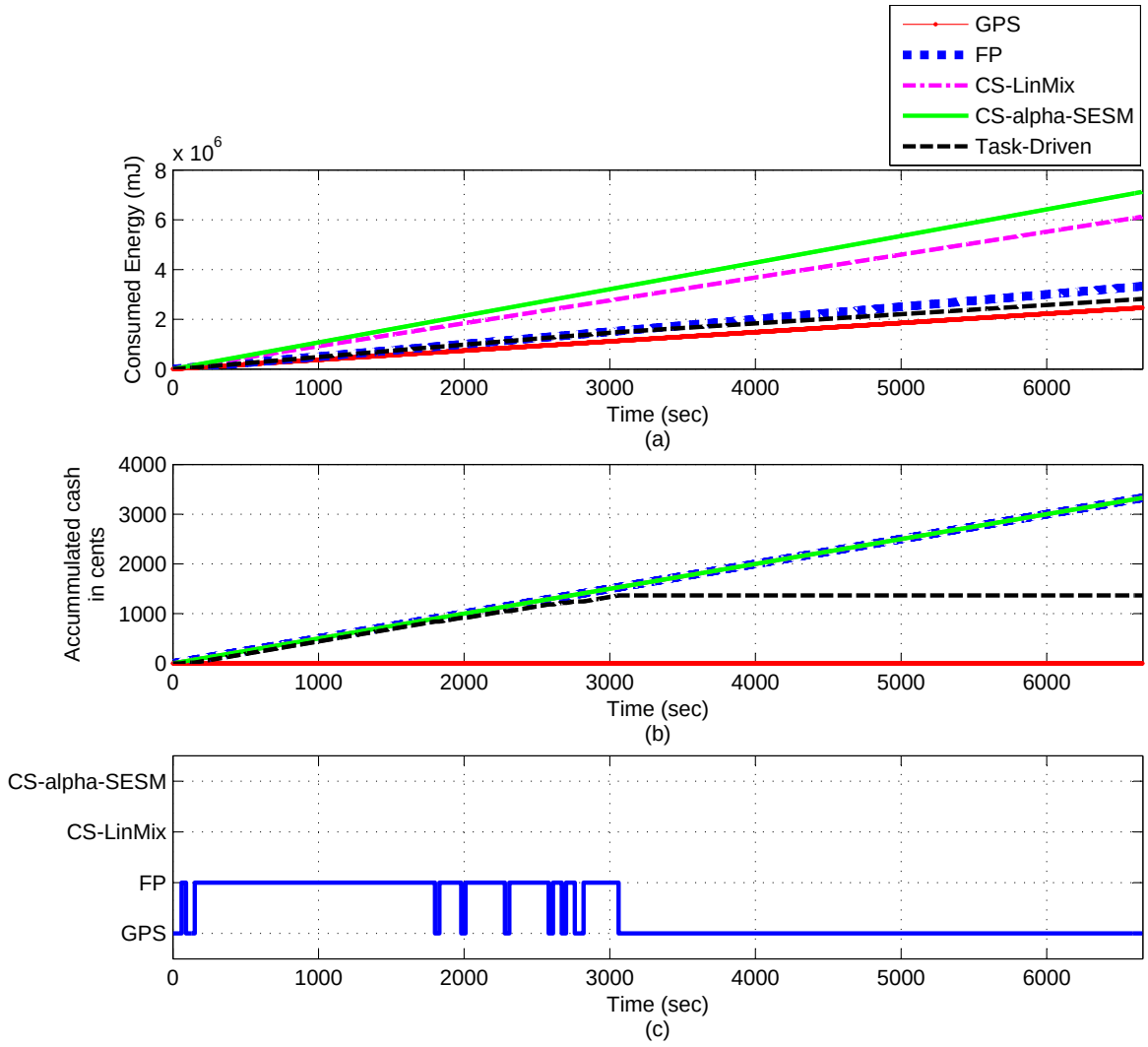


Figure 6.18: Task-driven localization performance in *use case* 3 with respect to the other localization techniques in terms of (a) consumed energy, and (b) cash spent. (c) shows the selection of the localization techniques made by the Pareto-optimal solution.

Use Case 4- Running Out of Power

Here I present an adverse scenario in which the conflict among the target criteria is increased. Assume that cash spending is allowed to obtain the best accuracy and in-

tegrity level, while the system is running out of power. The question, then, is how to maintain the highest possible quality of localization while keeping the power consumption low. The weights assigned to the criteria functions are equal in this *use case* (i.e., $W = [0.25, 0.25, 0.25, 0.25]$) so that decisions are based only on the distance between the target performances and the criteria functions.

Figure 6.19 shows the performance of the different localization techniques in terms of the CDF of the localization error. The balance between utilizing the resources in task-driven localization can be seen in producing accuracies similar to those of the fusion paradigms

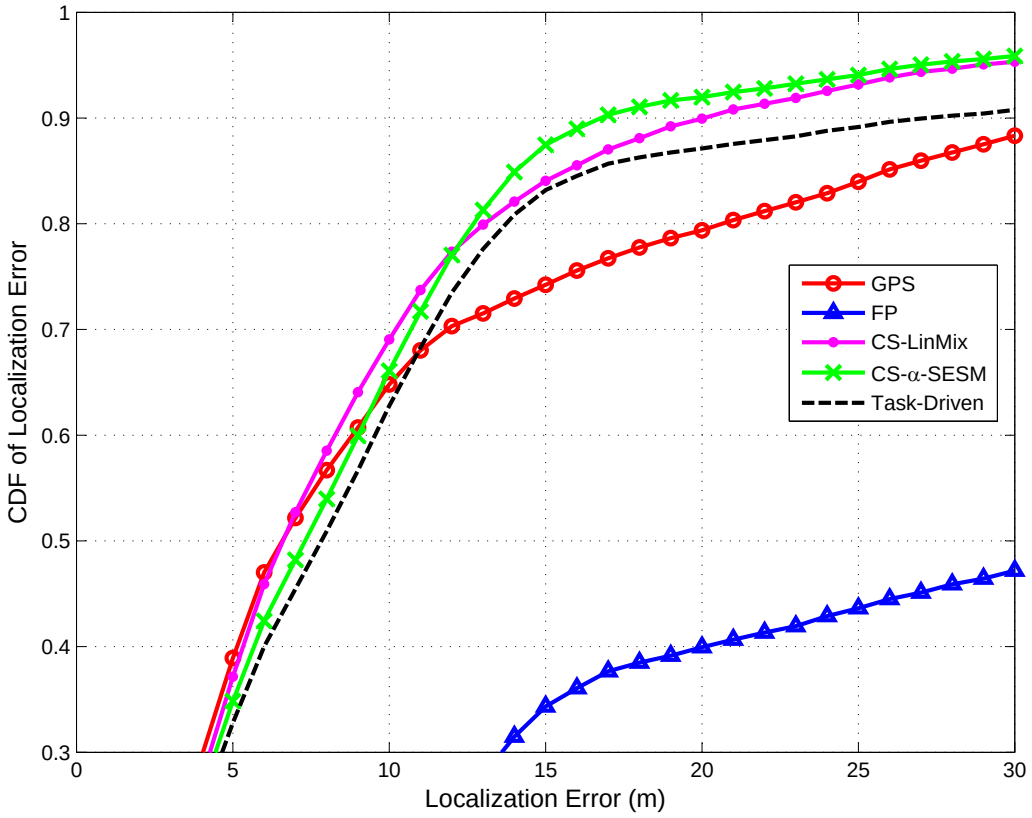


Figure 6.19: Comparison between the CDF of localization error of *use case* 4 and that of the other localization techniques.

in the low error band, and significantly improving the localization in the large error band.

As shown in Figure 6.20, the power consumption is kept as low as that of the FP localization in the multipath environment , where on the highway CS-LinMix is used due to its low power consumption with respect to CS-alpha-SESM and its high accuracy/integrity

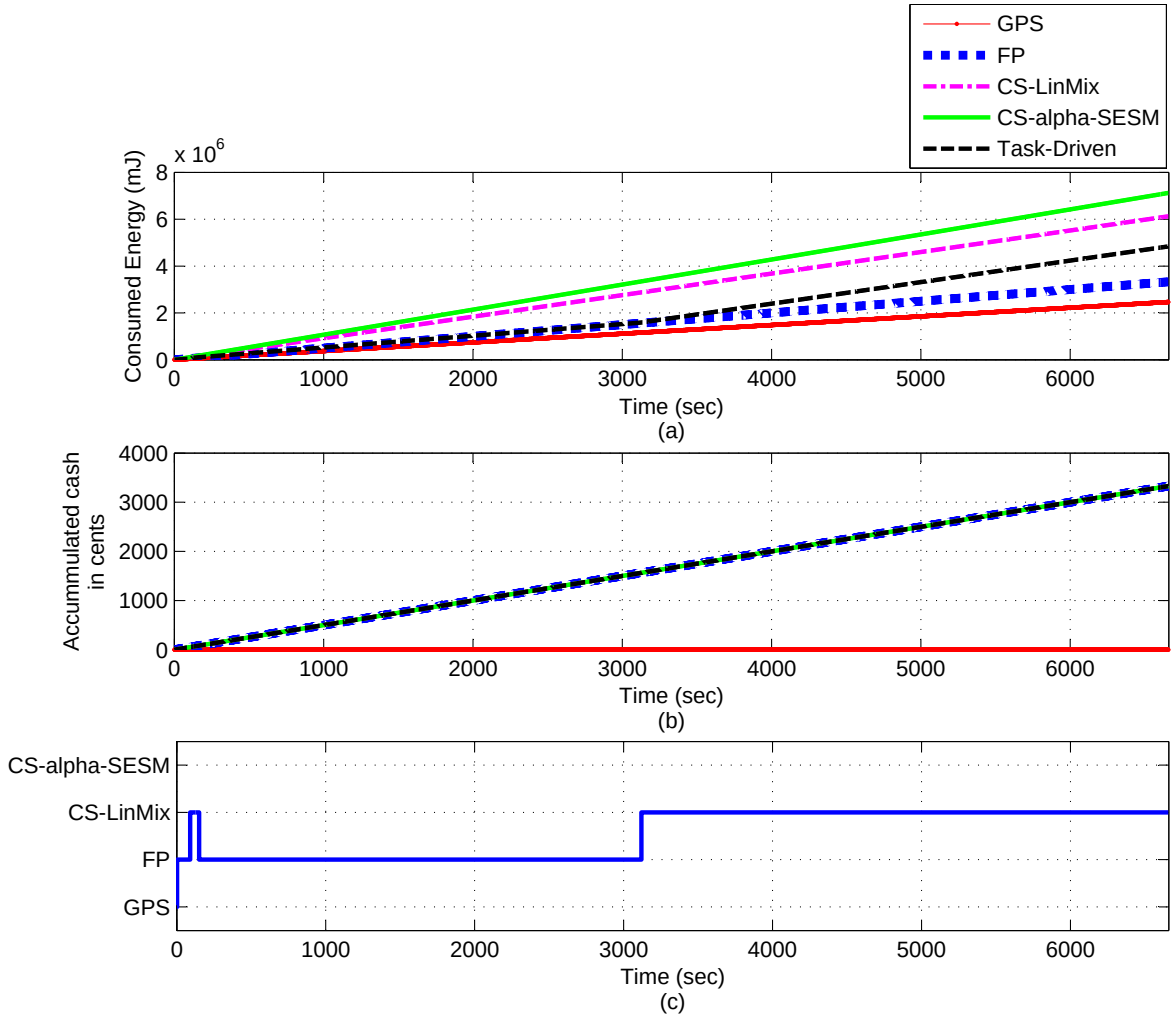


Figure 6.20: Task-driven localization performance in *use case* 4 with respect to the other localization techniques in terms of (a) consumed energy, and (b) cash spent. (c) shows the selection of the localization techniques made by the Pareto-optimal solution.

with respect to GPS localization. Due to the relaxed constraint on the cash spending, the task-driven localization utilizes the cellular network access in all localization requests.

6.6 Summary

This chapter presented the details of the proposed estimates fusion and management layer. The lack of complete information about the techniques' performance, in addition to the possible conflicts among the evidences, constitute the main reason of employing Dempster-Shafer Evidence Theory (DSET) in the fusion process. A novel evidence structure model called Spatial Evidence Structure Model (SESM) is developed in this chapter to deal with spatial data. SESM-based fusion paradigms are also introduced. These paradigms are capable of utilizing the information provided by the primary localization units and integrity monitoring units in the fusion process. Both the location estimate accuracy and aggregated integrity of the fusion process demonstrate superiority over the individual primary localization units.

The developed fusion paradigms extend the degree-of-freedom of managing the fusion process to the extent of controlling the system resources so as to achieve a task-driven performance. It is evident from the experimental results that task-driven integrity assessment and control is applicable and effective on hybrid localization systems.

Chapter 7

Conclusions and Future Directions

Vehicle localization has recently attracted significant attention in a wide range of applications, which can not deliver on their requirements without a specific level of localization accuracy. Commonly used localization techniques may not provide the accuracy level required by many of such applications. This goal, however, may seem tractable in some situations through fusion techniques, but at the cost of fast resource depletion.

Therefore, the main objective of the research work conducted in this thesis is to develop a novel hybrid vehicle localization system. This hybrid system is capable of combining vehicle location information, gathered by different sources, so as to achieve the target localization quality needed for a certain task/application while optimizing the assignment of the system resources. In the following sections, I summarize the major contributions in this research work, and I, then, provide some suggestions for future research directions.

7.1 Major Contributions

1. A variety of reported localization techniques have been presented and classified based on the type of the source of the location information used, such as radio, motion,

and vision sensors. A set of performance criteria has been introduced to highlight the research gaps in the reported localization techniques.

2. A novel framework for vehicle location estimation has been presented. The aim of this framework is to utilize localization technique diversity and localization source diversity to achieve robust localization performance that meets the application task's integrity and accuracy constraints. Central to the effectiveness of the framework is its ability to capture estimate integrity in computing the vehicle location. Furthermore, channel discounting in the light of localization integrity assessment allows the framework to maximize the quality of estimate fusion, as well as to address integrity aggregation out of different sources.

3. A methodology for evaluating the accuracy of location estimate has been provided so as to allow the target application to intelligently process such estimates in its decision making strategy. First, localization system parameters and location estimates have been used to contrive a feature space representation of probable accuracy classes. This feature space resembles the relationship between the localization measurements and the error in its location estimation. Second, a hierarchical classification strategy to address the class ambiguity problem via class unfolding approach (HCCU) has been developed. HCCU strategy is proven to be superior with respect to other hierarchical configuration. Furthermore, a Context Based Accuracy Classification (CBAC) algorithm has been introduced to enhance the performance of the classification process. In this algorithm, knowledge about the surrounding environment is utilized to optimize classification performance as a function of the observation conditions.

4. A task-driven integrity (TDI) model has been introduced. The TDI model enables the applications modules to be aware of the trust level of the localization output. TDI

model monitors specific parameter(s) in the localization technique and, accordingly, infers the impact of the change in the environmental conditions on the quality of the localization process. A generalized TDI solution has also been developed for the cases where information about the sensing parameters is unavailable or insufficient.

5. A novel evidence structure model called Spatial Evidence Structure Model (SESM) has been introduced to deal with evidences in form of spatial data with specific degree of quality. SESM-based fusion paradigms have been developed with the capability of performing fusion process on the information provided by the primary localization layer and integrity monitoring layer.

6. A context aware task-driven resource allocation mechanism has been developed to manage the fusion process. The main objective of this mechanism is to optimize the usage of system resources and achieve a task-driven performance. Several practical use cases have been presented and tested to show the efficacy of a context aware task-driven resource allocation mechanism. It is evident from the experimental results that task-driven integrity assessment and control is applicable and effective on hybrid localization systems.

7.2 Future Research Directions

The research work presented in this thesis has tackled the main aspects declared herein as research gaps, and has also demonstrated the capabilities of the developed framework in solving these aspects. Furthermore, this work uncovers other issues that, I believe, deserve further research work.

7.2.1 Sensitivity of HCCU/CBAC to Base Classifier Selection

The Base classifiers used in the HCCU and CBAC are FeedForward Neural Networks (FFNNet) with back-propagation learning algorithm. FFNNet has the ability to classify linearly inseparable patterns and also they are known for their efficient computational complexity, especially for embedded systems. Other types of classifiers, such as KNN, have the potential to be a good candidate as a base classifier. The performance sensitivity of the upper-level classifiers needs to be investigated as to the choice of the type of base classifier.

7.2.2 Ambiguous Classes Identification

In the classification problem, the target is to find a classifier h, which geometrically signifies a hyperplane in the feature space. The classifier h is optimized so that it provides the minimum classification error,$\epsilon(h)$, while splitting the feature space into number of volumes equal to the number of classes of interest.

$$h = \arg\min_{h} \epsilon(y = h(x))$$

where $y \in \{c_1, \ldots, c_M\}$. However, as it is noticed in the localization accuracy classification, the strong correlation and overlap between localization process measurements obtained under different measurement conditions make the accuracy classification an extremely difficult problem and uncertain in many scenarios.

Identifying volumes in the feature space in which the classification task is highly uncertain and declaring new classes called *Ambiguous Classes* ($\{c_{ambg_1}, \ldots, c_{ambg_k}\}$) for these volumes will definitely improve the total classification certainty. But this suggestion does not completely solve the issue since it is always possible to classify mea-

surements as belonging to the *Ambiguous Classes*. It is pivotal for such set of classes ($\{c_1, \ldots, c_M, c_{ambg_1}, \ldots, c_{ambg_k}\}$) to incorporate linguistic (i.e., fuzzy) frames of discernment and investigate the characteristics of the elements of these frames of discernment.

7.2.3 Task-Driven Integrity Model Extension

The construction of the Task-Driven Integrity (TDI) model is not a trivial task as the engineering knowledge of the key environmental impacts on the system performance need to be known a priori. This task continues to be more challenging in the case where the system's random events happen to be dependent and/or correlated. Furthermore, the construction complexity of the TDI model exponentially increases as the number of random events increases. Tentative approach to tend to these issues revolves around the concept of "divide and conquer". First, dependency among random events must be investigated. Second, the impact of the each set of dependent events need to be determined collectively. A respective integrity module is constructed. Finally, the integrity modules are combined with the independent events in one TDI model that captures the impact of the modules and the remaining events on the performance of the system.

7.2.4 Integrity Assessment and Control on VANETs

With the advent of Vehicular Ad hoc Wireless Networks (VANETs), a tremendous amount of applications, which literally rely on location information, have been proposed in the research community. The vast majority of the reported work on VANETs' applications assume accurate vehicle localization or at most assume small margin of error, such that of GPS in open sky [132,133] – an assumption that does not hold in many practical situations.

The integrity assessment and control framework developed in this thesis can handle

the issue of localizing a vehicle in different measurement conditions. Since it is able to exchange information among vehicles, VANET can be used to improve the localization process all over the network, in different measurement conditions, using distributed algorithms, as in [71]. Augmenting such distributed localization techniques with accuracy and integrity assessment mechanisms will not only significantly improve the localization process in VANET but also outperform the robustness of the individual-vehicle-based integrity assessment and control.

APPENDICES

Appendix A

Publication Related to This Thesis

Book chapter

1. Nabil Drawil and Otman Basir (2012). Emerging New Trends in Hybrid Vehicle Localization Systems, Global Navigation Satellite Systems: Signal, Theory and Applications, Shuanggen Jin (Ed.), ISBN: 978-953-307-843-4, InTech, PP: 279-298.

Journal Papers

2. Nabil Drawil, Otman Basir, Modeling the Impact of Observation Conditions on Localization Systems, Information Fusion, Available online 13 October 2012, ISSN 1566-2535, 10.1016/j.inffus.2012.09.002.

3. Drawil, N. M.; Amar, H. M.; Basir, O. A.; , "GPS Localization Accuracy Classification: A Context-Based Approach," Intelligent Transportation Systems, IEEE Transactions on , vol.PP, no.99, pp.1-12, 0 doi: 10.1109/TITS.2012.2213815

4. Nabil Drawil, Otman Basir, Intervehicle-Communication-Assisted Localization, *IEEE*

Transactions on Intelligent Transportation Systems, vol.11, no.3, pp.678-691, Sept. 2010 doi: 10.1109/TITS.2010.2048562.

5. Haitham Amar, Nabil Drawil, Otman Basir, Traveler Centric Trip planning: A Situation-Aware System, To be submitted to IEEE Systems Journal 2012

6. Nabil Drawil, Otman Basir, A Spatial Evidence Structure Model for Location Estimate Fusion, to be submitted to *IEEE Transactions on Intelligent Transportation Systems* 2012.

7. Nabil Drawil, Otman Basir, Task-Driven Location Estimates Fusion and Management, to be submitted to *IEEE Transactions on Intelligent Transportation Systems* 2012.

Conference Papers

8. Nabil Drawil, Haitham Amar, Otman Basir, A Solution to the Ill-Conditioned GPS Accuracy Classification Problem: Context Based Classifier, *IEEE GLOBECOM 2011Joint Workshop of SCPA and SaCoNAS*, Decemper 5-9, 2011, Houston, Texas USA, PP 1117-1122.

9. Nabil Drawil, Otman Basir, Toward Increasing the Localization Accuracy of Vehicles in VANET, *The fifth IEEE International Conference on Vehicular Electronics and Safety*, ICVES 2009 Nov. 10-12, , Pune, India, PP 13-18.

10. Nabil Drawil, Otman Basir, Vehicular Collaborative Technique for Location Estimate Correction, *International IEEE 68th Vehicular Technology Conference:* VTC2008-Fall, Sept. 21-24, Calgary, Alberta, PP 1-5.

11. Nabil Drawil, Otman Basir, Information Theoretic Approach for Fault Detection in Wireless Sensor Networks, *The Regional UW and IEEE Kitchener-Waterloo Section Joint Workshop on Multimedia Mining and Knowledge Discovery (MMKD)*, 17-18 October 2007, University of Waterloo, Ontario. PP: 57-62.

References

[1] Canadian Cellular Towers Map. `http://www.ertyu.org/steven_nikkel/cancellsites.html`. Accessed: 27/Jul/2012.

[2] Y. Nishimura, I. Tanahashi, S. Taniguchi, N. Matsumoto, and K. Nakamura. A New Concept for Vehicle Localization of Road Debiting System. *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 93–98, Sep. 1996.

[3] M. Cramer. GPS/INS Integration. 97:3–12, 1997.

[4] T. Aono, K. Fujii, S. Hatsumoto, and T. Kamiya. Positioning of Vehicle on Undulating Ground using GPS and Dead Reckoning. *IEEE International Conference on Robotics and Automation*, 4:3443–3448, May 1998.

[5] A. Stockus, A. Bouju, F. Bertrand, and P. Boursier. Web-Based Vehicle Localization. *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 436–441, 2000.

[6] A. Bouju, A. Stockus, R. Bertrand, and P. Boursier. Location-Based Spatial Data Management in Navigation Systems. *IEEE Intelligent Vehicle Symposium*, 1:172–177, Jun. 2002.

[7] D. Dao, C. Rizos, and J. Wang. Location-Based Services: Technical and Business Issues. *GPS Solutions*, 6(3):169–178, 2002.

[8] C. Lai and W. Tsai. Location Estimation and Trajectory Prediction of Moving Lateral Vehicle using Two Wheel Shapes Information in 2-D Lateral Vehicle Images by 3-D Computer Vision Techniques. *IEEE International Conference on Robotics and Automation*, 1:881–886, Sep. 2003.

[9] A. Benslimane. Optimized Dissemination of Alarm Messages in Vehicular Ad-hoc Networks (VANET). *Lecture Notes in Ccomputer Science.*, pages 655–666, 2004.

[10] O. Al-Bayari and B. Sadoun. New Centralized Automatic Vehicle Location Communications Software System Under GIS Environment. *INTERNATIONAL JOURNAL OF COMMUNICATION SYSTEMS*, 18(9):833, 2005.

[11] M. Jabbour, V. Cherfaoui, and P. Bonnifait. Management of Landmarks in a GIS for an Enhanced Localisation in Urban Areas. *Intelligent Vehicles Symposium, 2006 IEEE*, pages 50–57, 2006.

[12] M.K. Sliety. Impact of Vehicle Platform on Global Positioning System Performance in Intelligent Transportation. *Intelligent Transport Systems, IET*, 1(4):241–248, Dec. 2007.

[13] N. Drawil and O. Basir. Vehicular Collaborative Technique for Location Estimate Correction. *IEEE 68th Vehicular Technology Conference*, pages 1–5, Sep. 2008.

[14] S.A. Vaqar and O. Basir. Traffic Pattern Detection in a Partially Deployed Vehicular Ad Hoc Network of Vehicles. *IEEE Wireless Communications*, 16(6):40 –46, Dec. 2009.

[15] R. Lu, X. Lin, X. Liang, and X. Shen. A Dynamic Privacy-Preserving Key Management Scheme for Location-Based Services in VANETs. *Intelligent Transportation Systems, IEEE Transactions on*, 13(1):127 –139, march 2012.

[16] W. Lee, S. Tseng, J. Shieh, and H. Chen. Discovering Traffic Bottlenecks in an Urban Network by Spatiotemporal Data Mining on Location-Based Services. *Intelligent Transportation Systems, IEEE Transactions on*, 12(4):1047 –1056, dec. 2011.

[17] S. Hensel, C. Hasberg, and C. Stiller. Probabilistic Rail Vehicle Localization With Eddy Current Sensors in Topological Maps. *Intelligent Transportation Systems, IEEE Transactions on*, 12(4):1525 –1536, dec. 2011.

[18] www.trimble.com. `http://www.trimble.com/gps_tutorial/whygps.aspx`. Accessed: 27/Jul/2012.

[19] J. L. Leva. An Alternative Closed-Form Solution to the GPS Pseudo-Range Equations. *Aerospace and Electronic Systems, IEEE Transactions*, 32(4):1430 – 1439, 1996.

[20] J. Hoshen. The GPS Equations and the Problem of Apollonius. *Aerospace and Electronic Systems, IEEE Transactions*, 32(3):1116 – 1124, 1996.

[21] J. Rankin. An Error Model for Sensor Simulation GPS and Differential GPSg. In *Position Location and Navigation Symposium, 1994., IEEE*, pages 260–266. IEEE, 1994.

[22] M.S. Grewal, L.R. Weill, and A.P. Andrews. *Global Positioning Systems, Inertial Navigation, and Integration*, volume 416. Wiley Online Library, 2001.

[23] J. Cosmen-Schortmann, M. Azaola-Sáenz, MA Martinez-Olague, and M. Toledo-López. Integrity in urban and road environments and its use in Liability Critical applications. In *Position, Location and Navigation Symposium, 2008 IEEE/ION*, pages 972–983. IEEE, 2008.

[24] J. Rife, S.P.P. Enge, and B. Pervan. Paired Overbounding for Nonideal LAAS and WAAS Error Distributions. *Aerospace and Electronic Systems, IEEE Transactions on*, 42(4):1386–1395, 2006.

[25] S. Syed and ME Cannon. Fuzzy Logic-based Map Matching Algorithm for Vehicle Navigation System in Urban Canyons. In *ION National Technical Meeting, San Diego, CA*, volume 1, pages 26–28, 2004.

[26] Y. Zhang and L. Cheng. PLACE: Protocol for Location And Coordinate Estimation—a wireless sensor network approach. *Computer Networks*, 46(5):679–693, 2004.

[27] S. Čapkun, M. Hamdi, and J.P. Hubaux. GPS-free Positioning in Mobile Ad Hoc Networks. *Cluster Computing*, 5(2):157–167, 2002.

[28] M. Porretta, P. Nepa, G. Manara, and F. Giannetti. Location, Location, Location. *Vehicular Technology Magazine, IEEE*, 3(2):20–29, Jun. 2008.

[29] S. Venkatraman, Jr. Caffery, J., and H.-R. You. Location using LOS Range Estimation in NLOS Environments. *Vehicular Technology Conference, IEEE 55th*, 2:856–860, 2002.

[30] B.L. Le, K. Ahmed, and H. Tsuji. Mobile Location Estimator With NLOS Mitigation using Kalman Filtering. *Wireless Communications and Networking, IEEE*, 3:1969–1973, Mar. 2003.

[31] M.P. Wylie and J. Holtzman. The Non-Line of Sight Problem in Mobile Location Estimation. *5th IEEE International Conference on Universal Personal Communications*, 2:827–831, Sep. 1996.

[32] J.J. Caffery and G.L. Stuber. Overview of Radiolocation in CDMA Cellular Systems. *IEEE Communications Magazine*, 36(4):38 – 45, apr 1998.

[33] T.E. Melgard, G. Lachapelle, and H. Gehue. GPS Signal Availability in an Urban Area-Receiver Performance Analysis. *IEEE Position Location and Navigation Symposium*, pages 487–493, 1994.

[34] M.D. Dikaiakos, A. Florides, T. Nadeem, and L. Iftode. Location-Aware Services over Vehicular Ad-Hoc Networks using Car-to-Car Communication. *Selected Areas in Communications, IEEE Journal on*, 25(8):1590–1602, 2007.

[35] P. Zhou, T. Nadeem, P. Kang, C. Borcea, and L. Iftode. EZCab: A Cab Booking Application Using Short-Range Wireless Communication. *Third IEEE International Conference on Pervasive Computing and Communications*, pages 27–38, Mar. 2005.

[36] O. Basir. *Sensory Data Fusion and Integration: A Team Consensus Approach*. PhD thesis, University of Waterloo, 1993.

[37] R.G. Brown and Hwang P.Y.C. *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley & Sons, third edition, 1997.

[38] W.W. Kao. Integration of GPS and Dead-Reckoning Navigation Systems. *Vehicle Navigation and Information Systems Conference, 1991*, 2:635–643, Oct. 1991.

[39] I. Skog and P. Handel. In-Car Positioning and Navigation TechnologiesA Survey. *IEEE Transactions on Intelligent Transportation Systems*, 10(1):4–21, Mar. 2009.

[40] D. Bouvet and G. Garcia. Improving the Accuracy of Dynamic Localization Systems using RTK GPS by Identifying the GPS Latency. *IEEE International Conference on Robotics and Automation*, 3:2525–2530 vol.3, 2000.

[41] Qi. Honghui and J. B. Moore. Direct Kalman Filtering Approach for GPS/INS Integration. *IEEE Transactions on Aerospace and Electronic Systems*, 38(2):687 – 693, 2002.

[42] S. Rezaei and R. Sengupta. Kalman Filter Based Integration of DGPS and Vehicle Sensors for Localization. *Mechatronics and Automation, IEEE International Conference*, 1:455–460, Jul. 2005.

[43] R. Sharaf, A. Noureldin, A. Osman, and N. El-Sheimy. Online INS/GPS Integration with A Radial Basis Function Neural Network. *IEEE Aerospace and Electronic Systems Magazine*, 20(3):8–14, 2005.

[44] J. Rife. Collaborative vision-integrated pseudorange error removal: Team-estimated differential gnss corrections with no stationary reference receiver. *Intelligent Transportation Systems, IEEE Transactions on*, 13(1):15 –24, march 2012.

[45] M. Jabbour, P. Bonnifait, and V. Cherfaoui. Enhanced Local Maps in a GIS for a Precise Localisation in Urban Areas. *IEEE Intelligent Transportation Systems Conference*, pages 468–473, Sep. 2006.

[46] T. Weiss, N. Kaempchen, and K. Dietmayer. Precise ego-Localization in Urban Areas using Laserscanner and High Accuracy Feature Maps. *Proceedings of IEEE Intelligent Vehicles Symposium*, pages 284–289, Jun. 2005.

[47] K.Ch. Fuerstenberg and T. Weiss. Feature-Level Map Building and Object Recognition for Intersection Safety Applications. *Proceedings of IEEE Intelligent Vehicles Symposium*, pages 490–495, Jun. 2005.

[48] A. Rae and O. Basir. A Framework for Visual Position Estimation for Motor Vehicles. *4th Workshop on Positioning, Navigation and Communication*, pages 223–228, Mar. 2007.

[49] J.-i. Meguro, T. Murata, J.-i. Takiguchi, Y. Amano, and T. Hashizume. GPS Multipath Mitigation for Urban Area Using Omnidirectional Infrared Camera. *IEEE Transactions on Intelligent Transportation Systems*, 10(1):22–30, Mar. 2009.

[50] A.H. Sayed, A. Tarighat, and N. Khajehnouri. Network-Based Wireless Location: Challenges Faced in Developing Techniques for Accurate Wireless Location Information. *Signal Processing Magazine, IEEE*, 22(4):24–40, Jul. 2005.

[51] S. Al-Jazzar and Jr. Caffery, J. NLOS Mitigation Method for Urban Environments. *IEEE 60th Vehicular Technology Conference*, 7:5112–5115, Sep. 2004.

[52] J.J. Caffery and G.L. Stuber. Vehicle Location and Tracking for IVHS in CDMA Microcells. *5th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 4:1227–1231, Sep. 1994.

[53] Jr. Caffery, J.J. A New Approach to the Geometry of TOA Location. *IEEE 52nd Vehicular Technology Conference*, 4:1943–1949, 2000.

[54] M. McGuire, K.N. Plataniotis, and A.N. Venetsanopoulos. Location of Mobile Terminals using time Measurements and Survey Points. *IEEE Transactions on Vehicular Technology Conference*, 52(4):999–1011, Jul. 2003.

[55] X. Wang, Z. Wang, and B. O'Dea. A TOA-Based Location Algorithm Reducing the Errors due to Non-Line-of-Sight (NLOS) Propagation. *IEEE Transactions on Vehicular Technology Conference*, 52(1):112–116, Jan. 2003.

[56] M.Y. Chen, T. Sohn, D. Chmelev, D. Haehnel, J. Hightower, J. Hughes, A. LaMarca, F. Potter, I. Smith, and A. Varshavsky. Practical Metropolitan-scale Positioning for GSM Phones. *Lecture Notes in Computer Science*, 4206:225, 2006.

[57] J.J. Caffery. *Wireless Location in Cdma Cellular Radio Systems*. Kluwer Academic Pub, 1999.

[58] R.W. Ouyang, A.K.-S. Wong, and Kam Tim Woo. GPS Localization Accuracy Improvement by Fusing Terrestrial TOA Measurements. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1 –5, may 2010.

[59] Jr. Smith, W.W. Passive Location of Mobile Cellular Telephone Terminals. *IEEE International Carnahan Conference on Security Technology*, pages 221–225, Oct. 1991.

[60] R. Llorente, M. Morant, J.F. Puche, J. Romme, N. Amiot, B. Uguen, and J. Duplicy. Localization and Fingerprint of Radio Signals Employing a Multichannel Photonic Analog-to-Digital Converter. *Microwave Theory and Techniques, IEEE Transactions on*, 58(11):3304 –3311, nov. 2010.

[61] Y. Jin, W.-S. Soh, and W.-C. Wong. Error Analysis for Fingerprint-Based Localization. *Communications Letters, IEEE*, 14(5):393 –395, may 2010.

[62] C. Takenga and K. Kyamakya. A Low-cost Fingerprint Positioning System in Cellular Networks. In *Communications and Networking in China, 2007. CHINACOM '07. Second International Conference on*, pages 915 –920, aug. 2007.

[63] X.-C. Liu, S. Zhang, Q.-Y. Zhao, and X.-K. Lin. A Real-Time Algorithm for Fingerprint Localization Based on Clustering and Spatial Diversity. In *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2010 International Congress on*, pages 74 –81, oct. 2010.

[64] T.D. Le, H.M. Le, N.Q.T. Nguyen, D. Tran, and N.T. Nguyen. Convert Wi-Fi Signals for Fingerprint Localization Algorithm. In *Wireless Communications, Networking and Mobile Computing (WiCOM), 2011 7th International Conference on*, pages 1 –5, sept. 2011.

[65] A. Papapostolou, W. Xiao, and H. Chaouchi. Cooperative Fingerprint-Based Indoor Localization using Self-Organizing Maps. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International*, pages 1814 –1819, july 2011.

[66] L. Gogolak, S. Pletl, and D. Kukolj. Indoor Fingerprint Localization in WSN Environment Based on Neural Network. In *Intelligent Systems and Informatics (SISY), 2011 IEEE 9th International Symposium on*, pages 293 –296, sept. 2011.

[67] D.J. Suroso, P. Cherntanomwong, P. Sooraksa, and J. Takada. Location Fingerprint Technique using Fuzzy C-Means Clustering Algorithm for Indoor Localization. In *TENCON 2011 - 2011 IEEE Region 10 Conference*, pages 88 –92, nov. 2011.

[68] A. Benslimane. Localization in Vehicular Ad hoc Networks. *Systems Communications. Proceedings*, pages 19–25, Aug. 2005.

[69] R. Parker and S. Valaee. Cooperative Vehicle Position Estimation. *IEEE International Conference on Communications*, pages 5837–5842, Jun. 2007.

[70] R. Parker and S. Valaee. Vehicle Localization in Vehicular Networks. *IEEE 64th Vehicular Technology Conference*, pages 1–5, Sept. 2006.

[71] N.M. Drawil and O. Basir. Intervehicle-Communication-Assisted Localization. *Intelligent Transportation Systems, IEEE Transactions on*, 11(3):678–691, 2010.

[72] A. Boukerche, H.A.B.F. Oliveira, E.F. Nakamura, and A.A.F. Loureiro. Vehicular Ad hoc Networks: A New Challenge for Localization-Based Systems. *Computer Communications*, 31(12):2838–2849, 2008.

[73] T. Walter and P. Enge. Weighted RAIM for precision approach. In *PROCEEDINGS OF ION GPS*, volume 8, pages 1995–2004. Citeseer, 1995.

[74] S. Hewitson. GNSS receiver autonomous integrity monitoring: A separability analysis. *Proc. ION GPS*, pages 1502–1509, 2003.

[75] ESA. Making EGNOS Work for You,CD-ROM.

[76] S. Hewitson, H.K. Lee, and J. Wang. Localizability Analysis for GPS/Galileo Receiver Autonomous Integrity Monitoring. *The Journal of Navigation*, 57(02):245–259, 2004.

[77] M.A. Quddus. *High integrity map matching algorithms for advanced transport telematics applications*. PhD thesis, Citeseer, 2006.

[78] M. Schlingelhof, D. Betaille, P. Bonnifait, and K. Demaseure. Advanced Positioning Technologies for Co-operative Systems. *Intelligent Transport Systems, IET*, 2(2):81–91, Jun. 2008.

[79] R. Toledo-Moreo, M.A. Zamora-Izquierdo, B. Ubeda-Miarro, and A.F. Gomez-Skarmeta. High-Integrity IMM-EKF-Based Road Vehicle Navigation With Low-Cost GPS/SBAS/INS. *Intelligent Transportation Systems, IEEE Transactions on*, 8(3):491 –511, sept. 2007.

[80] M. Jabbour, P. Bonnifait, and V. Cherfaoui. Map-Matching Integrity using Multi-Sensor Fusion and Multi-Hypothesis Road Tracking. *Journal of Intelligent Transportation Systems Technology Planning and Operations*, 12(4):189–201, 2008.

[81] K. Jo, K. Chu, and M. Sunwoo. Interacting Multiple Model Filter-Based Sensor Fusion of GPS With In-Vehicle Sensors for Real-Time Vehicle Positioning. *Intelligent Transportation Systems, IEEE Transactions on*, 13(1):329 –343, march 2012.

[82] R. Toledo-Moreo, M.A. Zamora-Izquierdo, and A.F. Gomez-Skarmeta. A Novel Design of a High Integrity Low Cost Navigation Unit for Road Vehicle Applications. In *Intelligent Vehicles Symposium, 2006 IEEE*, pages 577–582, Jun. 2006.

[83] M.R. Mosavi, K. Mohammadi, and M.H. Refan. Fuzzy Processing on GPS Data to Improve Positioning Accuracy, before and after S/A is Turned Off. In *The Asian GPS Conference*, pages 117–120, 2002.

[84] T.-H. Chang, L.-S. Wang, and F.-R. Chang. A solution to the ill-conditioned gps positioning problem in an urban environment. *Intelligent Transportation Systems, IEEE Transactions on*, 10(1):135 –145, Mar. 2009.

[85] S. Wu, J. Li, and S. Liu. An Improved Reference Selection Method in Linear Least Squares Localization for LOS and NLOS. In *Vehicular Technology Conference, 2011. VTC 2011-Fall. IEEE 74th*, Sept. 2011.

[86] M.A. Quddus, W.Y. Ochieng, and R.B. Noland. Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transportation Research Part C: Emerging Technologies*, 15(5):312–328, 2007.

[87] E. Costa. Simulation of the Effects of Different Urban Environments on GPS Performance Using Digital Elevation Models and Building Databases. *Intelligent Transportation Systems, IEEE Transactions on*, 12(3):819 –829, sept. 2011.

[88] E. Tzoreff and B.-Z. Bobrovsky. A Novel Approach for Modeling Land Vehicle Kinematics to Improve GPS Performance Under Urban Environment Conditions. *Intelligent Transportation Systems, IEEE Transactions on*, 13(1):344 –353, march 2012.

[89] M. Modsching, R. Kramer, and K. ten Hagen. Field trial on GPS Accuracy in a medium size city: The influence of built-up. In *3rd Workshop on Positioning, Navigation and Communication (WPNC'06), Hannover*, Mar. 2006.

[90] B. Bakhache and I. Nikiforov. Reliable detection of faults in measurement systems. *International Journal of Adaptive Control and Signal Processing*, 14(7):683–700, 2000.

[91] A. Lahrech, C. Boucher, and J.-C. Noyer. Accurate Vehicle Positioning in Urban Areas. *31st Annual Conference of IEEE Industrial Electronics Society*, page 5 pp, Nov. 2005.

[92] F. Chausse, J. Laneurit, and R. Chapuis. Vehicle Localization on a Digital Map using Particles Fltering. pages 243–248, Jun. 2005.

[93] G. Zhang, S. Krishnan, F. Chin, and C.C. Ko. UWB Multicell Indoor Localization Experiment System with Adaptive TDOA Combination. pages 1–5, Sept. 2008.

[94] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins. *Global Positioning System: Theory and Practice*. Progress in Astronautics and Aeronautics. Springer-Verlag, 2001.

[95] S. Sukkarieh, E.M. Nebot, and H.F. Durrant-Whyte. Achieving Integrity in an INS/GPS Navigation Loop for Autonomous Land Vehicle Applications. *In Proceedings of IEEE International Conference on Robotics and Automation*, 4:3437 –3442 vol.4, May 1998.

[96] R.E. Walpole, R.H. Myers, S.L. Myers, and K. Ye. *Probability and statistics for engineers and scientists*, volume 6. Prentice Hall Upper Saddle River, NJ:, 1998.

[97] O.A. Basir. "Markovian model for predicting the impact of observation conditions on the reliability of sensory systems, innovations in theory, practice and applications". *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, 3:1646–51, 1998.

[98] H. Lee. Reliability indexed sensor fusion and its application to vehicle velocity estimation. *Journal of Dynamic Systems, Measurement, and Control*, 128:236–243, Jun. 2006.

[99] L. Gui, T. Val, and A. Wei. Improving Localization Accuracy Using Selective 3-Anchor DV-hop Algorithm. In *Vehicular Technology Conference, 2011. VTC 2011-Fall. IEEE 74th*, Sept. 2011.

[100] S. Fujii, A. Fujita, T. Umedu, H. Yamaguchi, T. Higashino, S. Kaneda, and M. Takai. Cooperative Vehicle Positioning via V2V Communications and Onboard Sensors. In *Vehicular Technology Conference, 2011. VTC 2011-Fall. IEEE 74th*, Sept. 2011.

[101] T. Hastie, R. Tibshirani, and J.H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Verlag, 2009.

[102] J.-S. Lee and I.-S. Oh. Binary Classification Trees for Multi-class Classification Problems. In *Proceedings of the Seventh International Conference on Document*

Analysis and Recognition - Volume 2, ICDAR '03, pages 770–, Washington, DC, USA, 2003. IEEE Computer Society.

[103] S.C. Han, J.H. Kwon, and C. Jekeli. Accurate absolute gps positioning through satellite clock error estimation. *Journal of Geodesy*, 75(1):33–43, 2001.

[104] R.M. Alkan and E. Arslan. GPS Standard Positioning Service Performance after Selective Availability Turned Off. *International Symposium on GIS, Istanbul-Turkey*, Sep 2002.

[105] J.E. Jackson and J. Wiley. *A User's Guide to Principal Components*. Wiley Online Library, 1991.

[106] J.A. Hartigan. *Clustering Algorithms*. John Wiley & Sons, Inc., 1975.

[107] G.W. Irwin, K. Warwick, and K.J. Hunt. *Neural network applications in control*, volume 53. Iet, 1995.

[108] C. Takenga, T. Peng, and K. Kyamakya. Post-processing of fingerprint localization using kalman filter and map-matching techniques. In *Advanced Communication Technology, The 9th International Conference on*, volume 3, pages 2029 –2034, feb. 2007.

[109] I. Chih-Lin, L.J. Greenstein, and R.D. Gitlin. A Microcell/Macrocell Cellular Architecture for Low- and High-Mobility Wireless Users. In *Global Telecommunications Conference, 1991. GLOBECOM '91. 'Countdown to the New Millennium. Featuring a Mini-Theme on: Personal Communications Services*, pages 1006 –1011 vol.2, dec 1991.

[110] E.H. Mamdani and S. Assilian. An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller. *International Journal of Man-Machine Studies*, 7(1):1–13, 1975.

[111] L.A. Zadeh. Fuzzy Sets*. *Information and control*, 8(3):338–353, 1965.

[112] F.O. Karray and C.W. De Silva. *Soft computing and intelligent systems design: theory, tools, and applications*. Addison-Wesley, 2004.

[113] A.K.M.M. Hossain and W.-S. Soh. Cramer-rao bound analysis of localization using signal strength difference as location fingerprint. In *INFOCOM, 2010 Proceedings IEEE*, pages 1 –9, march 2010.

[114] J. Machaj, P. Brida, and B. Tatarova. Impact of the number of access points in indoor fingerprinting localization. In *Radioelektronika (RADIOELEKTRONIKA), 2010 20th International Conference*, pages 1 –4, april 2010.

[115] R.A. Rae. *Augmenting Vehicle Localization with Visual Context*. PhD thesis, University of Waterloo, 2009.

[116] G. Dissanayake, S. Sukkarieh, E. Nebot, and H. Durrant-Whyte. The Aiding of a Low-Cost Strapdown Inertial Measurement Unit using Vehicle Model Constraints for Land Vehicle Applications. *IEEE Transactions on Robotics and Automation*, 17(5):731 –747, Oct. 2001.

[117] H.C. Tijms. *A First Course in Stochastic Models*. John Wiley and Sons, 2003.

[118] A.P. Dempster. Upper and lower probability inferences based on a sample from a finite univariate population. *Biometrika*, 54(3-4):515–528, 1967.

[119] G. Shafer. *A mathematical theory of evidence*, volume 1. Princeton university press Princeton, 1976.

[120] P. Smets and R. Kennes. The transferable belief model. *Artificial intelligence*, 66(2):191–234, 1994.

[121] J. Sudano. Pignistic probability transforms for mixes of low-and high-probability events. 2001.

[122] J.W. Guan and D.A. Bell. Approximate reasoning and evidence theory. *Information sciences*, 96(3):207–235, 1997.

[123] H. Zhu. *Data Fusion Using Neural-Fuzzy Embedded Evidential Reasoning*. PhD thesis, University of Waterloo, 2004.

[124] L.A. Zadeh. A simple view of the dempster-shafer theory of evidence and its implication for the rule of combination. *AI magazine*, 7(2):85, 1986.

[125] R.R. Yager. On the dempster-shafer framework and new combination rules. *Information sciences*, 41(2):93–137, 1987.

[126] K. Sentz, S. Ferson, and Sandia National Laboratories. *Combination of evidence in Dempster-Shafer theory*. Citeseer, 2002.

[127] P.M. Pardalos and H.E. Romeijn. *Handbook of global optimization*, volume 2. Springer, 2002.

[128] O.A. Basir and H.C. Shen. Goal-Driven Task Assignment and Sensor Control in Multi-Sensor Systems: A Multicriteria Approach. In *Robotics and Automation, 1993. Proceedings., 1993 IEEE International Conference on*, pages 559 –566 vol.2, may 1993.

[129] M. Sami, D. Sciuto, C. Silvano, and V. Zaccaria. An Instruction-Level Energy Model for Embedded VLIW Architectures. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 21(9):998 – 1010, sep 2002.

[130] S. Fowler. Study on Power Saving Based on Radio Frame in LTE Wireless Communication System using DRX. In *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*, pages 1062 –1066, Dec. 2011.

[131] J. Paek, J. Kim, and R. Govindan. Energy-Efficient Rate-Adaptive GPS-Based Positioning for Smartphones. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*, MobiSys '10, pages 299–314, New York, NY, USA, 2010. ACM.

[132] J. Yao, A.T. Balaei, M. Hassan, N. Alam, and A.G. Dempster. Improving Cooperative Positioning for Vehicular Networks. *Vehicular Technology, IEEE Transactions on*, 60(6):2810 –2823, july 2011.

[133] N. Alam, A.T. Balaei, and A. G. Dempster. Relative Positioning Enhancement in VANETs: A Tight Integration Approach. *Intelligent Transportation Systems, IEEE Transactions on*, PP(99):1 –9, 2012.