

# **Environmental Context Detection for Adaptive Navigation using GNSS Measurements from a Smartphone**

HAN GAO and PAUL D. GROVES

University College London, London, UK

*ABSTRACT: The signals that are available for navigation depend on the environment. To operate reliably in a wide range of different environments, a navigation system is required to adopt different techniques based on the environmental contexts. In this paper, an environmental context detection framework is proposed, building the foundation of a context adaptive navigation system. Different land environments are categorised into indoor, urban and open-sky environments based on how Global Navigation Satellite System (GNSS) positioning performs in these environments. Indoor and outdoor environments are first detected based on the availability and strength of GNSS signals using a hidden Markov model. Then the further classification of outdoor environments into urban and open-sky is investigated. Pseudorange residuals are extracted from raw GNSS measurements in a smartphone and used for classification in a fuzzy inference system alongside the signal strength data. Practical test results under different kinds of environments demonstrate an overall 88.2% detection accuracy is achieved.*

**KEY WORDS:** Environments, Context Detection, Hidden Markov Model, Fuzzy Inference System.

## INTRODUCTION

The operation of navigation and positioning systems is inherently dependent on the environmental context because it determines the types of signals available [1]. For example, GNSS reception is good in open-sky environments, but poor indoors and in deep urban areas. Wi-Fi signals are not available in rural areas, in the air or at sea. In an underwater environment, most radio signals do not propagate at all. Processing techniques can also depend on the environments. Terrain referenced navigation typically determines terrain height using radar or laser scanning in the air, sonar or echo sounding at sea and a barometer on land [1]. In an open-sky environment, non-line-of-sight (NLOS) reception of GNSS signals or multipath interference may be detected using consistency checking techniques based on sequential elimination [2]. In dense urban areas, more sophisticated algorithms are required for GNSS positioning in the presence of severe multipath interference and NLOS reception [3].

In order to provide greater accurate and reliable positioning services in more challenging environments, environmental context becomes important for three reasons. Firstly, there is a move towards navigation systems that can operate in a wider range of different contexts. For many applications, the context can change, particularly for smartphones, which move between indoor and outdoor environments. Similarly, a micro air vehicle (MAV) may be required to operate from above, amongst buildings, or even indoors. Secondly, over the past fifteen years, many navigation and positioning techniques have been developed [4], such as Wi-Fi positioning [5][6], multiple-constellation GNSS [7] and GNSS shadow matching [8]. However, no single current technique is capable of providing reliable and accurate positioning in all contexts. Finally, as the navigation systems become more complex, there is a need to re-use hardware and software modules across multiple applications to take advantage of computational power [9]. Therefore, in order to operate reliably across different environments,

a navigation system is required to be able to detect its operating context and adapt the techniques it uses accordingly [1][10].

A universal navigation system might be expected to provide position within 3 meters in any context with a very high reliability [10]. Picking up an inappropriate technique will increase the danger of obtaining wrong positioning information while using too many sensors for positioning simultaneously will consume too much power. With the environment detected, the optimum set of techniques can be selected based on the context, giving better accuracy and reliability. For example, a smartphone is able to get a positioning precision within 5 meters in an open-sky environment, but degrades to tens of meters in dense urban. But using 3D mapping GNSS aided on a smartphone, the precision can be improved to within 7m [3]. For an indoor environment, Wi-Fi positioning or pedestrian dead reckoning (PDR) is an appropriate option, while in an environment with lots of features, visual positioning can work well.

Moreover, environmental context can also provide some useful information for a navigation application. For instance, a vehicle detected indoors may imply it is located in a car park. The smartphone itself may choose to turn on or off sensors automatically depending on application requirements and operating environments [11]. For example, Wi-Fi signals are likely to be too weak or unavailable in an open environment so repeated scanning of Wi-Fi networking would be wasteful.

The increasing availability of embedded sensors and built-in radio modules makes it possible to obtain the environmental context with a smartphone. In previous studies, researchers have focused environment detection on indoor and outdoor classification. In IODetector proposed by Zhou [12], indoor/outdoor context is classified into “indoor”, “semi-outdoor” and “outdoor” and is determined by using a combination of cellular signals, light sensors, magnetometers and proximity sensors. However, the usage of hard thresholds for each

sensor feature in the approach made it hard to be applied across different devices. In addition, cellular signals were also used for indoor and outdoor detection on their own [13]. Using the same sensors, Radu et al. [14] considered employing co-training, one of the semi-supervised learning methods for detection and provided a 90% accuracy even in unfamiliar environments.

Besides the sensors mentioned above, other sensor features, including strength of Bluetooth signals [15] and sound frequency received from the microphone [16], were also utilised for indoor and outdoor detection.

At the same time, GNSS signals have also been used for indoor/outdoor detection. Researchers began to distinguish indoors and outdoors on GNSS receivers in [17][18][19]. In 2013, according to the preliminary experiments by University College London (UCL) [1], the feasibility of environmental context detection using GNSS and Wi-Fi signals from a smartphone was examined. The results showed that GNSS  $C/N_0$  measurement can be used to distinguish indoor from outdoor environments, while Wi-Fi measurements have been shown to be unreliable for indoor and outdoor classification when used alone. In [20], Parviainen et al. presented an environment recognition implementation for smartphones. They extracted features from GPS, Wi-Fi and Bluetooth modules and applied supervised machine learning algorithms to recognise indoor and outdoor scenarios.

Although there has been substantial research into determining indoor and outdoor environments, a more detailed categorization is still in its infancy. A context framework must be designed that is fit for its purpose. For a navigation system, a good categorization of environmental context is expected to provide an indicator of the availability of signals and other features that may be used for determining position. A simple indoor and outdoor classification is far from sufficient for a navigation application. For example, a conventional GNSS technique works well in an open-sky environment, but degrades seriously in deep urban

areas. In a shallow indoor environment, some GNSS and cellular signals are available but they are not when deep inside a building. Except indoor and outdoor environments, there are also intermediate areas where a mixture of indoor and outdoor positioning signals are available. Thus, the transition between indoor and outdoor environments is smooth, not sharp.

In May 2016, Google announced that raw GNSS measurements can be accessed from the Android “Nougat” operating system by the developers [21], which opens up the possibilities of making use of pseudorange, Doppler and carrier-phase measurements and extending GNSS-based environmental context determination methods to more detailed classes.

This paper aims to determine the environmental context for navigation using the GNSS module on a smartphone. Some preliminary results on basic indoor/outdoor context determination were presented in [22]. A categorization of environmental context based on positioning signals and a suitable detection framework is first proposed. Environmental contexts are categorized into indoor, intermediate, urban and open-sky environment based on GNSS positioning accuracy and detected using the characteristics of the received GNSS signals. The GNSS data collection and the extraction of suitable features for context determination from the GNSS measurement data are then described. Next, the method for indoor-outdoor environmental context detection using a hidden Markov model is described, followed by the further classification of outdoor environments into urban and open-sky using a fuzzy inference method. Practical test results of this hybrid context detection framework are then presented and discussed. Finally, the conclusions and plans for future work are summarised.

## ENVIRONMENTAL CONTEXT CATEGORIZATION FOR NAVIGATION

A navigation system may not fully benefit from a context framework that is designed for more general purposes. The context categorization framework for navigation and positioning must be designed especially in order to be fit for navigation purposes. A good environment categorization for navigation is expected to provide an indication of the positioning techniques applicable for determining positions.

Generally, the environmental context may be divided into several different broad classes: on land, on water, underwater, air and space [1]. As the smartphone is used as the sensing device in this study, it is not applicable to be used for positioning purposes underwater, in the air or in space. Therefore, here the range of environmental contexts is limited to scenarios on land because a common mobile user spends most of their time in daily life on land.

The environment categorization is proposed in Table 1, with the characteristics of GNSS signals and positioning accuracy described.

Table 1—Categorization of environments based on GNSS reception

Category		Characteristics	Accuracy
Indoor	Deep indoor	No GNSS reception	N/A
	Shallow indoor	Some GNSS reception	Tens of meters
Intermediate		Poor GNSS reception	~ 30m
Outdoor	Urban	Some disruption to GNSS	> 20m
	Open-sky	Good GNSS reception	< 3m

For land navigation, locating whether the user is indoor or outdoor is a basic but prerequisite task because indoor and outdoor positioning depend on inherently different techniques. For

example, in an outdoor environment, GNSS or its enhanced techniques performs well while Wi-Fi positioning or Bluetooth positioning are better options when staying inside a building. Note that in reality, the boundaries between indoor and outdoor environments can be ambiguous, rendering some scenarios hard to classify as either indoor or outdoor [22]. Thus, the intermediate type, where a client is adjacent to a building or in a semi-open environment, is included in the framework. Typical examples of intermediate environments are shown in Figure 1, whose top side is covered by the building and at least one surrounding side of the area is open. Moreover, the “intermediate” environment can serve as a bridge between indoor and outdoor categories to smooth the transition between the two. In an intermediate environment, indoor positioning techniques (e.g. Wi-Fi and Bluetooth) can still work well, while direct line-of-sight (LOS) GNSS reception can be limited.

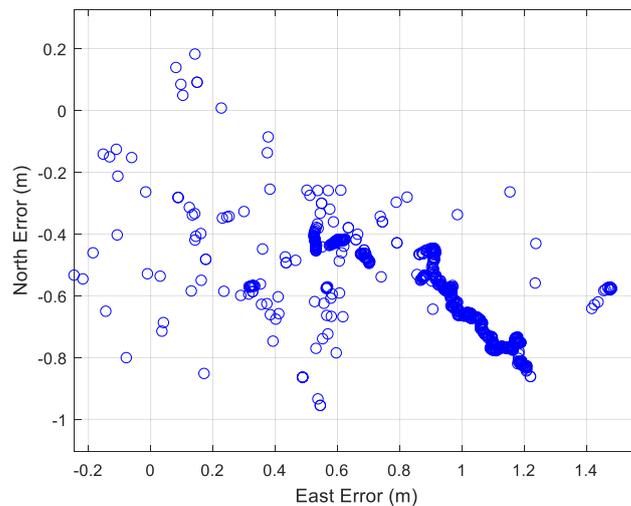


*Figure 1—Examples of intermediate environments*

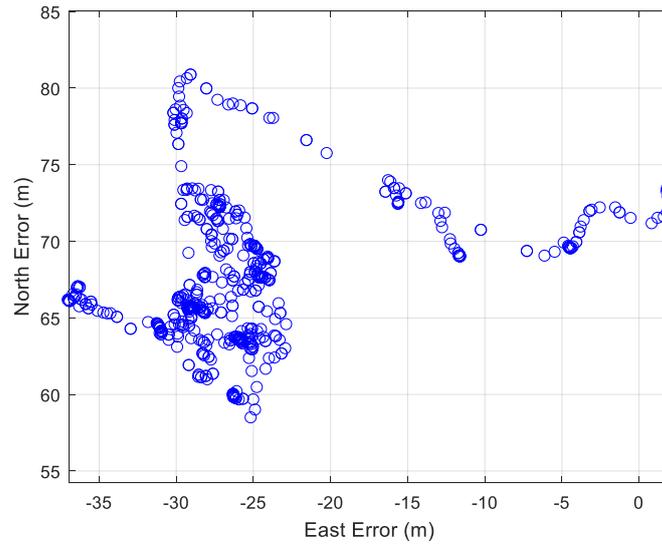
Outdoor environments are further categorized into open-sky and urban categories based on the characteristics of GNSS reception. As GNSS uses line-of-sight ranges between the navigation satellites and receivers to derive position solutions, its signals are subject to severe degradation in the presence of reflection and multipath [23]. In an open-sky environment, a conventional GNSS positioning technique is able to provide a positioning accuracy within 3m on a smartphone. Figure 2 shows an example in Regent’s Park, an open-sky area in London.

However, in urban areas, some line-of-sight signals would be blocked by tall buildings or walls, and some received signals might be from the reflecting surfaces. In such scenarios, the localization accuracy degrades dramatically. This is demonstrated in Figure 3, where in a dense urban area in central London, the horizontal errors can be as high as 80 meters. So such deteriorated solution should not be used for navigation directly in applications and should be enhanced or altered by other techniques.

For an indoor environment, many navigation techniques (e.g. Wi-Fi positioning, PDR and map matching) have better performance than GNSS and they are not greatly affected by whether the environment is deep or shallow indoors. Therefore, detailed indoor classification will not significantly bring any benefit a navigation system, so it will not be discussed in this paper.



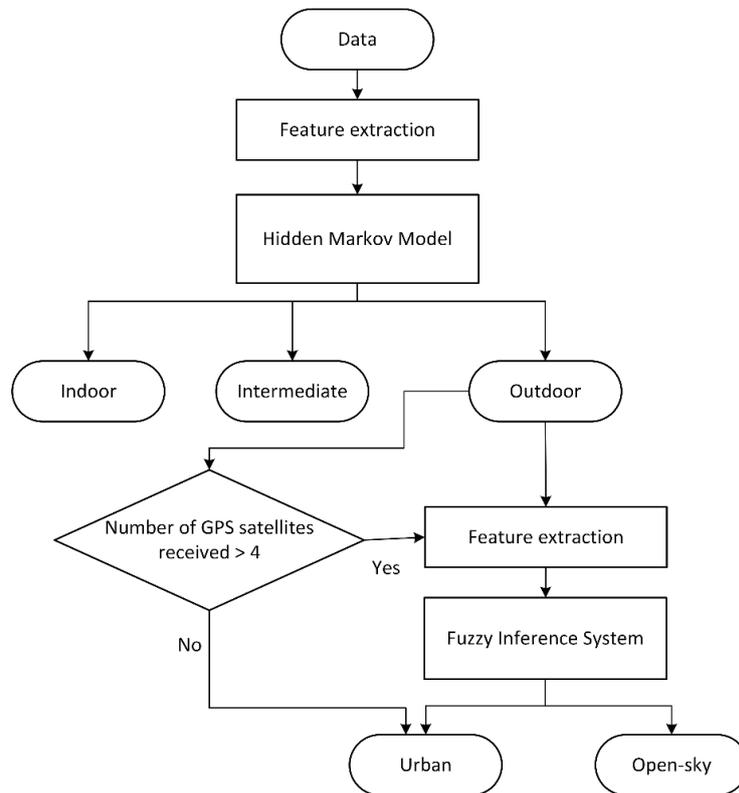
*Figure 2—Example: GNSS positioning errors in Regent's Park, London (the location shown in Figure 8)*



*Figure 3—Example: GNSS positioning errors in the City of London, London (the location shown as P2 in Figure 7)*

Different smartphone sensors whose outputs vary with features of the environment can be potentially used as detectors and each sensor used for environment detection has its advantages and drawbacks respectively. A cellular module detects cellular signal strengths from a cellular network, but at the same time the signals strongly depend on the proximity of cellular base stations in the network. A Wi-Fi module can receive signals broadcast from access points. However, tests [1][10] found that the assumption, the number of access points are larger and strength of signals are stronger indoors, does not always stand. And it was not sufficient to distinguish indoor environments from outdoor environments directly using Wi-Fi signals. A GNSS module, with GPS (Global Positioning System) and GLONASS (GLObal NAVigation Satellite System) chips in most current smartphones, is chosen as the main detector for this research, because the availability and accuracy of satellite signals tend to be less affected by factors other than the environment type. More importantly, the globally distributed properties of GPS and GLONASS ensure that we can infer environments from the availability and strength of GNSS signals anywhere on Earth. Note that Galileo and BeiDou can also be used

once the constellations are completed and smartphone GNSS chips start processing their signals. The main drawback of GNSS is its high-power consumption compared to other smartphone sensors. As the research advances, other sensors can be added into the context determination framework to improve the environment detection using GNSS module.



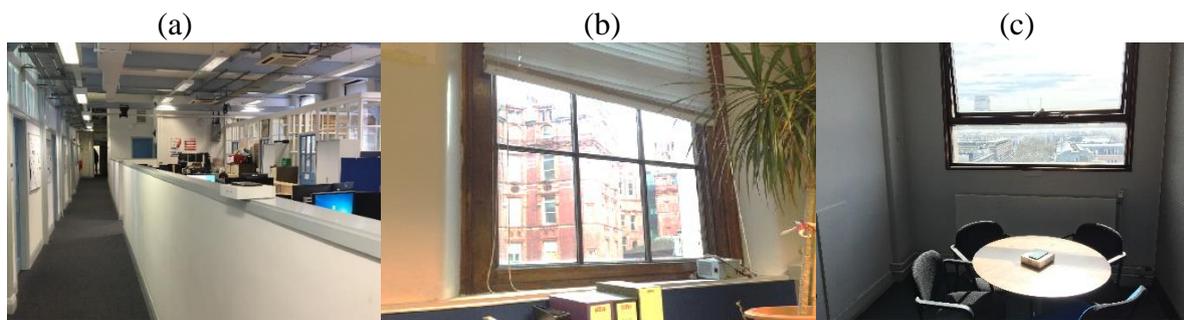
*Figure 4—Workflow of the environmental context detection algorithm*

Based on the proposed categorization, environmental contexts will be detected in two phases. As shown in Figure 4, the data will first be used to classify the environment as indoor, intermediate or outdoor. When an outdoor environment is detected, the open-sky and urban categories will be further distinguished. The features used in each phase will be introduced in detail in this paper.

## ENVIRONMENTAL DATA COLLECTION

GNSS measurements, comprising GPS and GLONASS data, were collected at 1 Hz using a Google Pixel smartphone running an Android data logging application. Time tags, the  $C/N_0$  measurements, satellite azimuths and elevations, the conventional GNSS position solutions and raw GPS measurements when available can all be logged in files for processing.

The dataset covers different kinds of scenarios in indoor, intermediate, urban and open-sky environments. The indoor data was collected at different indoor sites, covering deep indoor, inner room, office with window and by the window scenarios. Some of the sites are illustrated in Figure 5. The whole indoor dataset was split into two parts: one part of sites for training the model and one part for testing.



*Figure 5—Selected indoor data collection sites*

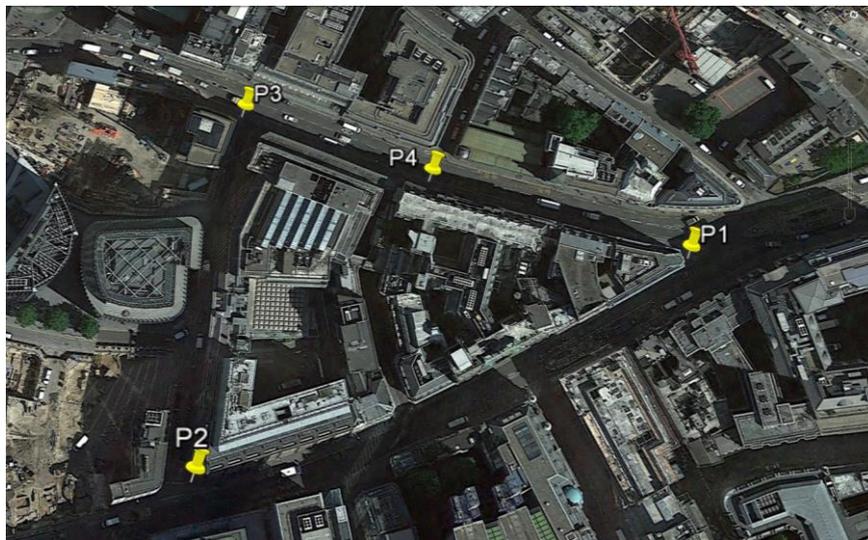
The data for the intermediate category was collected at left, middle and right side of the portico of UCL's Wilkins building as shown in Figure 6. Outdoor data collection was performed using the same device, including four sites in urban area and four sites in open-sky environments. Figure 7 to Figure 9 illustrate these sites.

For the intermediate and outdoor sites, the data was logged statically for about 10 mins and two rounds of data collection were performed. The time between two rounds of data collection was longer than one hour, so they can be considered to be independent of each other due to the satellite motions between the two rounds. The first round of data is used for training and tuning

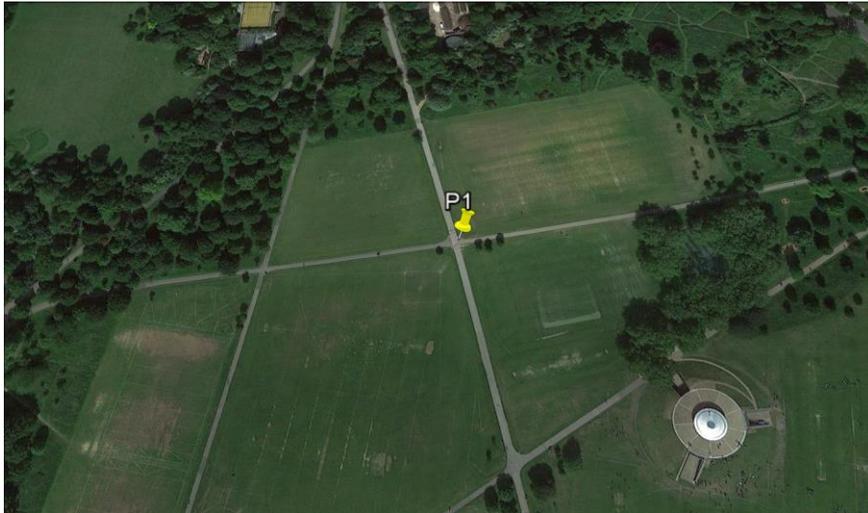
the parameters of the context detection algorithms; while the second round of data is used for testing the model. As true reference positions of outdoor sites are required for detailed outdoor environment classification, they were established using models on Google Earth to identify landmarks and tape measure to measure the relative position of the user from those identified landmarks.



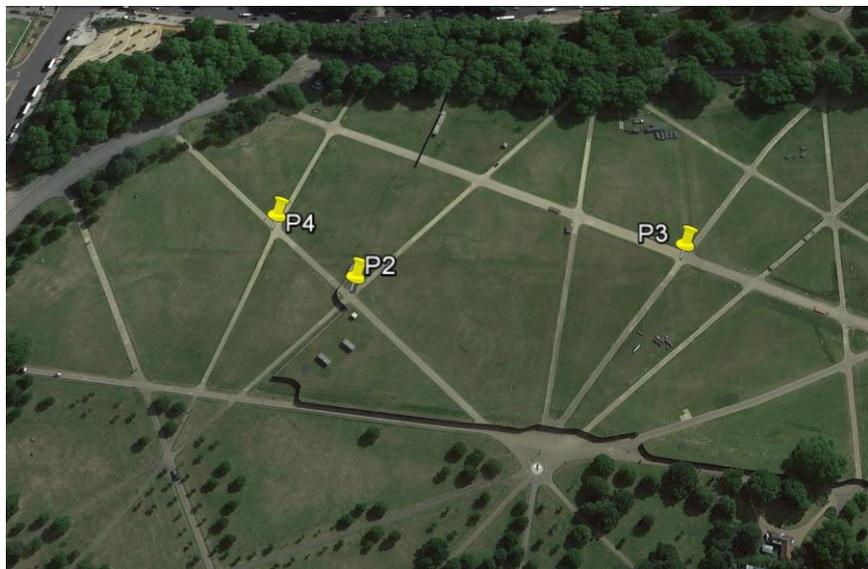
*Figure 6—Intermediate environment data collection sites on the portico of UCL's Wilkins building*



*Figure 7—Urban data collection sites in the City of London (GoogleTM earth)*



*Figure 8—Open-sky data collection site in Regent’s Park in London (GoogleTM earth)*



*Figure 9—Open-sky data collection sites in Hyde Park in London (GoogleTM earth)*

## **FEATURE EXTRACTION FROM THE GNSS DATA**

The features extracted from the smartphone GNSS measurements are described in this section and will be used as inputs to the environmental context classifiers as shown in Figure 4. In an indoor environment, where the GNSS signals are obscured by ceilings and walls, most GNSS signals are attenuated by the structure of the building or received by NLOS paths, rendering

them weaker or unavailable indoors. Thus features based on the availability and strength of GNSS signals are extracted and will be implemented for indoor and outdoor detection.

In urban areas, the presence of NLOS and multipath inference can affect the carrier-power to noise-density ratio ( $C/N_0$ ) and ranging measurement errors [4][24]. Therefore, when GPS raw measurements from more than four satellites are available in outdoor environments, a feature based on the pseudorange residuals will be computed and used for open-sky and urban detection with a combination of the  $C/N_0$  related feature.

### ***Features Based on the Availability and Strength of Signals***

A set of GNSS measurements was collected by smartphone over the transition from an outdoor to an indoor environment. The person holding the smartphone walked from an outdoor to an indoor environment at about the 30th second. Figure 10 and Figure 11 demonstrate the differences in availability and strength of GNSS signals, respectively, in the indoor and outdoor environments. In Figure 10, the number of satellites received decreased gradually after moving indoors, as more satellite signals were blocked by the building.  $C/N_0$ , expressed in decibel-Hertz (dB-Hz), is a good indicator of signal strength in the absence of significant interference and adopted as a standard output of GNSS receivers. Figure 11 shows the  $C/N_0$  outputs from three satellites. A drop of about 5 dB-Hz was observed when the person was nearing the building, following by a sharp decrease when they entered. It was also noted that most of the satellite signals indoors were weaker than 20 dB-Hz and PRN 83 lost track after about 90s.

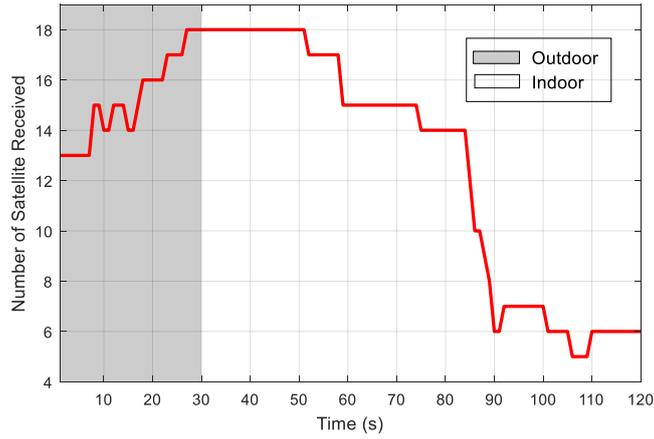


Figure 10—Number of satellites received during outdoor-indoor transition

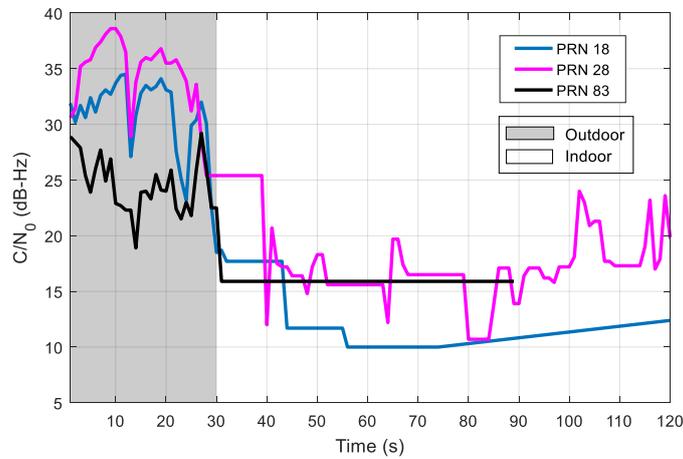
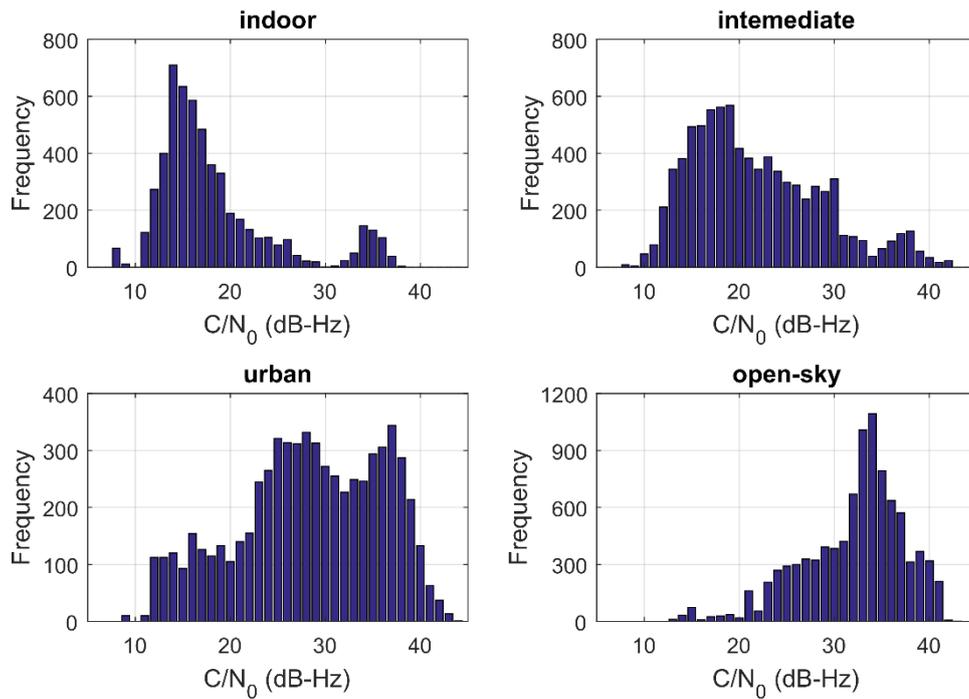


Figure 11—Selected  $C/N_0$  values during outdoor-indoor transition

Figure 12 presents frequency histograms showing the distributions of  $C/N_0$  measurements over 500s from different environmental categories for illustration. A number of trends may be identified from the histograms. As expected, the average received  $C/N_0$  is lower in indoor environments than in urban and open-sky environments, and the value of the intermediate environment is between indoors and outdoors. The double peak in the indoor histogram shows some LOS signals and strong reflected signals are available indoors. In outdoor environments, a signal with a higher  $C/N_0$  is more likely to be LOS than NLOS [8]. In urban areas, more NLOS signals are received due to the reflection from the surface of buildings, therefore the average  $C/N_0$  is normally lower in urban than open-sky areas. By comparing the GNSS  $C/N_0$

distributions, it can also be seen that the proportions of signals weaker than 25 dB-Hz vary between different environment types. Most of the signals received in an indoor environment are weaker than 25 dB-Hz while increasing proportions of signals stronger than 25 dB-Hz are observed from intermediate to urban and open sky environments.



*Figure 12— $C/N_0$  measurement distributions under different environments (indoor data was collected at the site shown in Figure 5(b); intermediate data was collected at P1 in Figure 6; urban data was collected at P1 in Figure 7; open-sky data was collected at P1 in Figure 8.)*

To estimate the availability and strength of GNSS signals, the number of satellites received and the total measured  $C/N_0$ , summed across all the satellites received at each epoch, were considered. Note that as the average number of satellites received indoor is normally less than those received outdoor, the summed  $C/N_0$  is considered instead of the average value. These two metrics are shown in Figure 13 and Figure 14, based on the same set of data shown in Figure 12. It can be observed that the number of satellites received in the intermediate

environment was similar to that in the open-sky environment, while the number of satellites received in the urban and indoor environments were also similar to each other. In Figure 14, although open-sky and indoor environments can be clearly distinguished from others based on  $C/N_0$  measurements, it is hard to tell the differences between intermediate and urban environments based on the same measurements. In summary, these two metrics are clearly unreliable and cannot be used for indoor and outdoor recognition.

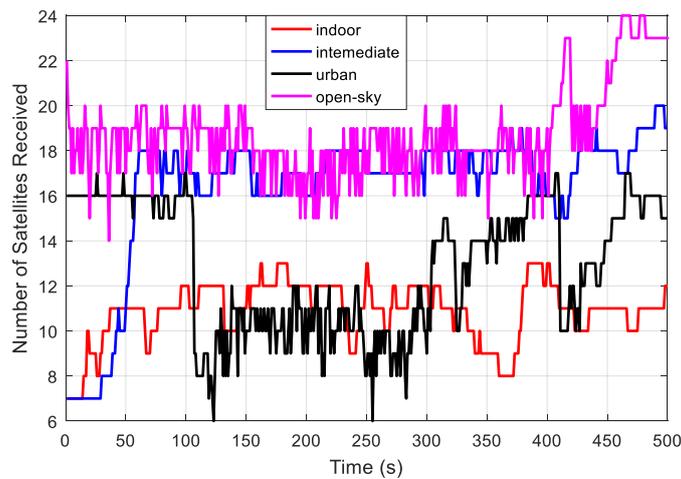


Figure 13-- Number of satellites received

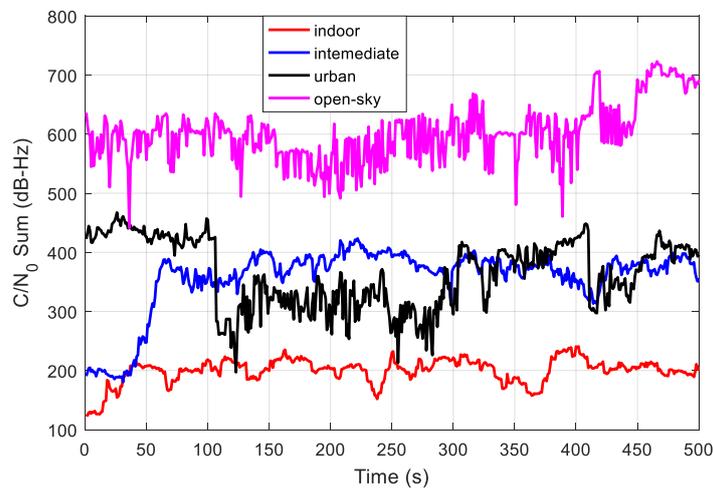


Figure 14—Total  $C/N_0$ , summed across all satellites received

As a larger percentage of “weak” signals (less than 25 dB-Hz) are received indoors than outdoors, to enlarge the differences in the classification features between environments, these

signals are deducted from the observations. Thus, two new features,  $numCNR_{25}$  and  $sumCNR_{25}$ , are proposed, which are defined by

$$numCNR_{25} = \sum_i H(CNR_i), \quad (1)$$

$$sumCNR_{25} = \sum_i CNR_i \cdot H(CNR_i), \quad (2)$$

where  $CNR_i$  indicates the  $C/N_0$  value of the  $i$ -th satellite received at the current epoch and the function  $H(\cdot)$  is defined as:

$$H(x) = \begin{cases} 1, & \text{if } x \geq 25 \text{ dB-Hz} \\ 0, & \text{otherwise} \end{cases}. \quad (3)$$

Comparing the features plotted in Figure 15 and Figure 16 with the ones in Figure 13 and Figure 14, indoor, intermediate and outdoor (urban + open-sky) environments are not overlapped anymore and can be more clearly distinguished, which shows the effectiveness of the proposed features. To further verify the reliability of the features, they will be further tested using a dataset collected at different indoor and outdoor sites in this paper. Based on the whole dataset we collected, it is worth noting that  $sumCNR_{25}$  is typically less than 100 dB-Hz indoors and greater than 200 dB-Hz outdoors. For the observations between 100 and 200 dB-Hz, their specific environment types need to be distinguished using more information, such as measurements from other sensors or the temporal relationship between measurements.

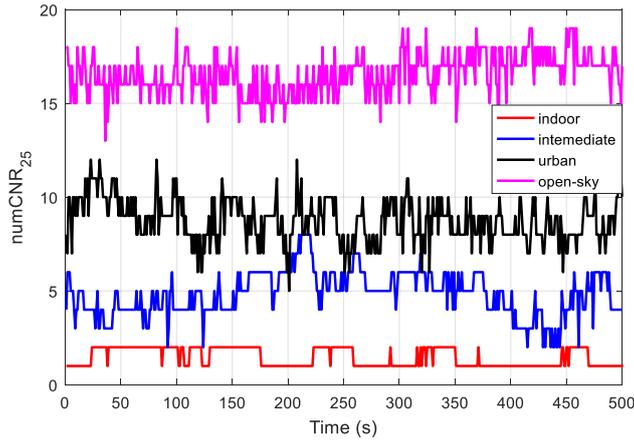


Figure 15--Number of satellites with signals above 25 dB-Hz

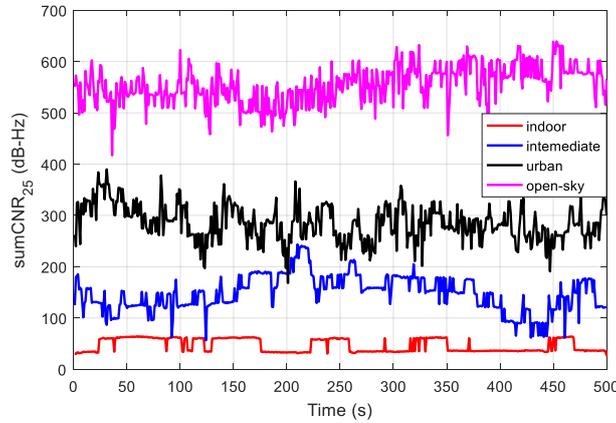


Figure 16--Total  $C/N_0$ , summed across the signals above 25 dB-Hz

### Features Based on Pseudorange Residuals

The receiver actually measures the pseudorange  $\tilde{\rho}_a^s$ , which is also perturbed by the satellite and receiver clock errors,  $\delta t_c^s$  and  $\delta t_c^a$ , and is equal to the difference between the arrival  $sa$  and transmission time  $st$  multiplied by the speed of light  $c$  [4]. Thus,

$$\begin{aligned}\tilde{\rho}_a^s &= r_{as} + (\delta t_c^a - \delta t_c^s)c \\ &= (\tilde{t}_{sa,a}^s - \tilde{t}_{st,a}^s)c\end{aligned}\quad (4)$$

where  $r_{as}$  is the true range between the satellite  $s$  and the antenna  $a$ ;  $\tilde{t}_{st,a}^s$  is the time of signal transmission and  $\tilde{t}_{sa,a}^s$  is the time of signal arrival; the subscript  $c$  means clock. Note that the

receiver actually measures the time delay applied to a replica of the satellite ranging code that is synchronized with the incoming signal.

The pseudorange can also be estimated from position and clock offset solution [4], given by

$$\begin{aligned}\hat{\rho}_a^s &= \hat{r}_{as} + \delta\hat{\rho}_c^a(\hat{t}_{sa,a}^s) \\ &= \left| \mathbf{C}_e^I(\hat{t}_{st,a}^s) \hat{\mathbf{r}}_{es}^e(\hat{t}_{st,a}^s) - \hat{\mathbf{r}}_{ea}^e(\hat{t}_{sa,a}^s) \right| + \delta\hat{\rho}_c^a(\hat{t}_{sa,a}^s)\end{aligned}\quad (5)$$

where  $\hat{r}_{as}$  is the estimated range between antenna  $a$  and satellite  $s$ , which can be calculated from the estimated satellite positions  $\hat{\mathbf{r}}_{es}^e$ , obtained from the ephemeris, and the estimated antenna position solution  $\hat{\mathbf{r}}_{ea}^e$ .  $\mathbf{C}_e^I$  is the transformation matrix from an ECEF (Earth-centered earth-fixed) frame to an ECI (Earth-centered inertial) frame, synchronized at the time of signal arrival.  $\delta\hat{\rho}_c^a$  is the estimates of the receiver clock offset, and  $\hat{t}_{st,a}^s$  and  $\hat{t}_{sa,a}^s$  are the estimated times of signal transmission and arrival.

The difference between measured and estimated pseudorange is called the pseudorange residual. NLOS signals cause errors in the position solution, which result in large residuals for the “clean” LOS measurements. In theory, the larger of the sum of squared pseudorange residuals, the more likely they contain NLOS signals. Therefore, the feature  $zPRR$  used for classification is expressed as the sum of squared residuals of the pseudoranges divided by the degrees of freedom:

$$zPRR = \left( \sum_{i=1}^N \left| \tilde{\rho}_a^s - \hat{\rho}_a^s \right|_i^2 \right) / (N - 4) \quad (6)$$

where  $N$  is the number of satellites received at the epoch and  $i$  denotes the  $i$ -th satellite signal. It is worth mentioning that since the computation of pseudorange residuals are based on the calculation of receiver’s position, they can only be calculated when at least five satellites received. Otherwise, there are not enough satellite signals for positioning or not enough

redundancy for calculating pseudorange residuals. Only GPS residuals were used here because the Qualcomm GNSS receiver chip in Pixel smartphone does not support GLONASS raw measurement outputs.

## **INDOOR AND OUTDOOR DETECTION**

Features based on the availability and strength of GNSS signals have been introduced. In this section, a hidden Markov model that infers the current environment types (indoor, intermediate and outdoor) from those features is described. At the end of this section, test results of the indoor-outdoor detection algorithm are presented.

### ***Hidden Markov Model (HMM)***

A hidden Markov model is a temporal pattern recognition algorithm, which assumes a Markov process [25] with the states (indoor, intermediate or outdoor environment in this study), so it is capable of modelling the movements of a device from one environment to another according to observations. Within an HMM, the probability that the system is in each of three states are estimated, so that the navigation system knows the certainty of the decision. In general, a HMM comprises five elements as follows:

1) The state space  $\mathbf{S}$  that consists of three hidden states: indoor, intermediate and outdoor, which are denoted as  $S_1$ ,  $S_2$  and  $S_3$  respectively. At each epoch  $k$ , hidden states satisfy the condition

$$\sum_{i=1}^3 P(X_k = S_i) = 1 \quad (7)$$

where  $X_k$  refers to the environmental context at that epoch.

2) The set of observations at each epoch  $k$ ,  $\mathbf{z}_k = \{z_{1,k}, z_{2,k}, \dots, z_{\ell,k}, \dots, z_{m,k}\}$ , where  $z_{\ell,k}$  is the  $\ell$ -th observation at epoch  $k$  and  $m$  is the number of observations. In this study,  $z_{1,k}$  refers to  $numCNR_{25}$  while  $z_{2,k}$  is  $sumCNR_{25}$ .

3) The matrix of state transition probabilities  $\mathbf{A} = \{A_{ij}\}$ . Each element of the state transition probabilities matrix,  $A_{ij}$ , defines the probability that a state  $S_i$  at the immediately prior epoch transits to another state  $S_j$  at the current epoch.

4) The matrix of emission probabilities  $\mathbf{B} = \{B_i(k)\}$  that defines the conditional distributions  $P(\mathbf{z}_k | S_i)$  of the observations from a specific state.

5) An initial state probability distribution  $\mathbf{\Pi} = \{\pi_i\}$  that defines the probability of state  $S_i$  being at the first epoch.

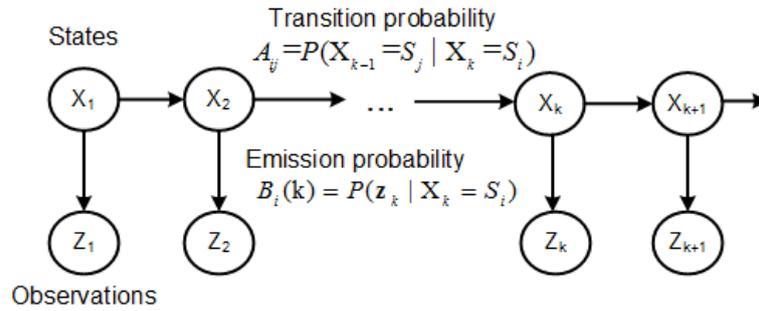


Figure 17—Structure of a first-order HMM

In this paper, we use the first-order HMM, which assumes the current environmental context is only affected by the immediate previous context. Figure 17 is an illustration of a first-order hidden Markov model. Given the sequence of the observations, the probabilities of each context at each epoch can be inferred using the Viterbi algorithm [25][26]. The probabilities of the model are determined as follows.

- Transition Probability

Since the sample interval here is 1s, when a user was previously indoors, the current state is highly likely to be indoor and might be intermediate, but is not likely to be outdoor. This is

because the user rarely moves directly from indoors to a fully outdoor GNSS reception environment, noting that GNSS signals exhibit transitional effects immediately outside buildings as shown in Figure 11. However, when the user is at the intermediate state, he/she can move directly to either of the other states. Based on these assumptions and with reference to the parameters applied in IODetector [12], the values of transition probability were list in Table 2. Note that the values are selected in order to obtain an experimental balance between responsiveness to change and vulnerability to noise. There is no intention here to model realistic transition probabilities as this would result in the context determination algorithms taking a long time to respond to changes.

Table 2—Transition probabilities of HMM

$k+1 \backslash k$	Indoor	Intermediate	Outdoor
Indoor	2/3	1/3	0
Intermediate	1/3	1/3	1/3
Outdoor	0	1/3	2/3

- Initial Probability

As there is no prior information about the initial state, we have to make a judgement based on the available initial observations. Clearly, the indoor and outdoor contexts occur much more frequently than the intermediate context. However, if there is insufficient information to correctly determine the context, it is better to select the intermediate context than to incorrectly select the indoor or outdoor context. The initial probabilities were therefore set as follows:

$$\begin{aligned}
 P(X_1 = S_1) &= P(X_1 = S_3) = 0.4 \\
 P(X_1 = S_2) &= 0.2
 \end{aligned}
 \tag{8}$$

- Emission Probability

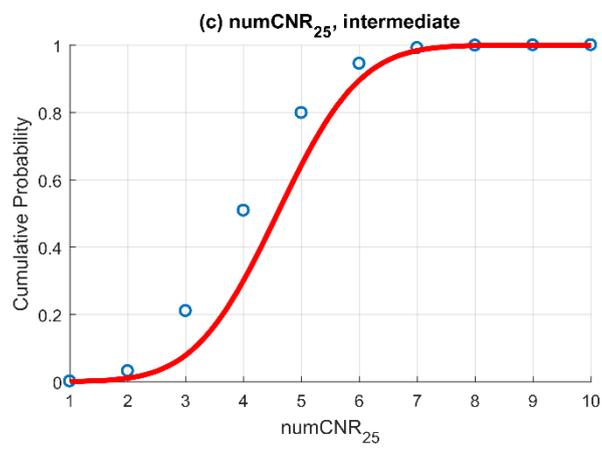
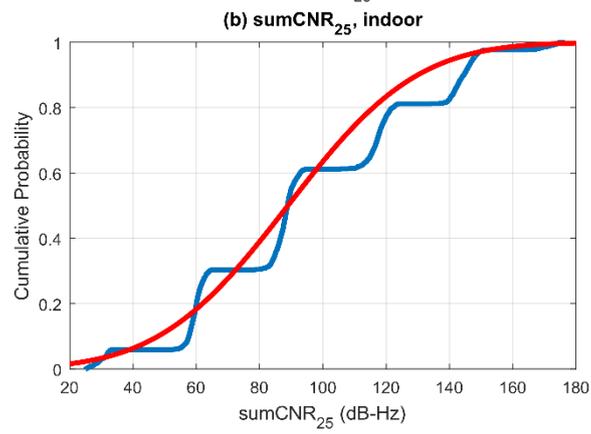
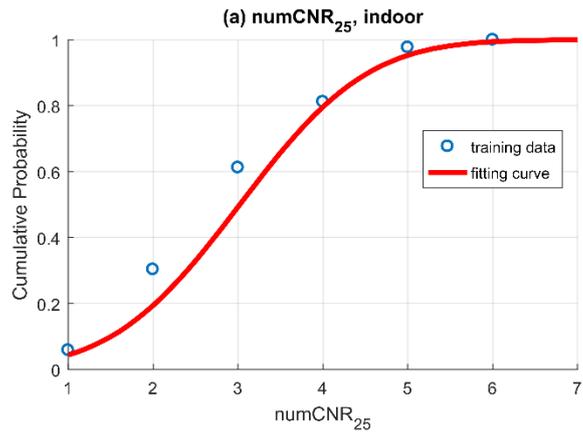
The emission probabilities describe the probability distribution of the observations for each of the three states. The emission probabilities of each hidden states (indoor, intermediate and outdoor) are obtained by fitting the observations (features of each environment) according to the training dataset described in the previous section.

The observations are then modelled as Gaussian distributions, whose means and variances are fitted to the datasets collected at different indoor and outdoor sites. Note that as both  $numCNR_{25}$  and  $sumCNR_{25}$  distributions for the outdoor data are bimodal, using a single Gaussian distribution is obviously unrealistic, so the emission probabilities are modelled by two mixture Gaussian distributions with equal weights.

The fitting results are depicted from Figure 18(a) to Figure 18(f) correspondingly. Table 3 shows the emission probabilities of each environment to each feature, where the Gaussian distributions are denoted by  $N(\mu, \sigma^2)$  with mean  $\mu$  and variance  $\sigma^2$ .

Table 3—Emission probability

	$numCNR_{25}$	$sumCNR_{25}$ (dB-Hz)
<b>Indoor</b>	$N(3.02, 1.4)$	$N(88.95, 1025.37)$
<b>Intermediate</b>	$N(4.58, 1.26)$	$N(142.55, 625)$
<b>Outdoor</b>	$N(7.77, 3.25)$	$N(242.08, 2697.4)$
	$N(17.33, 4.58)$	$N(607.35, 5218.4)$



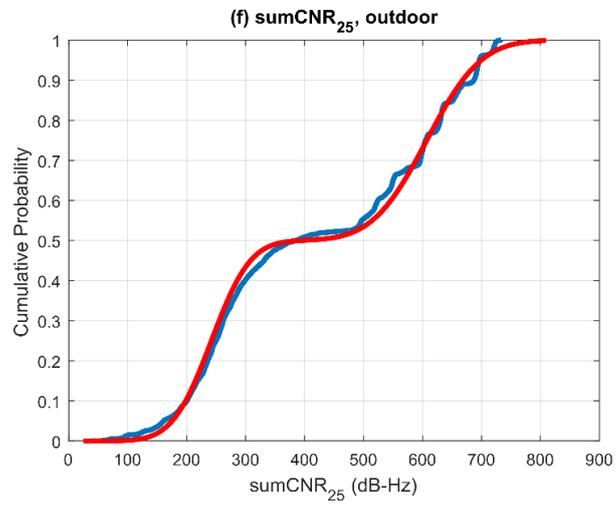
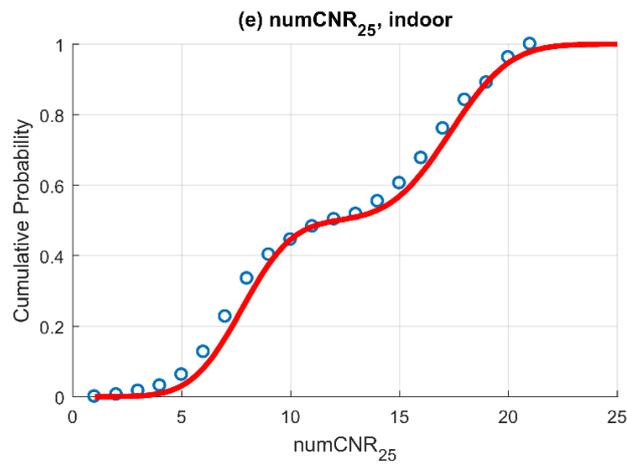
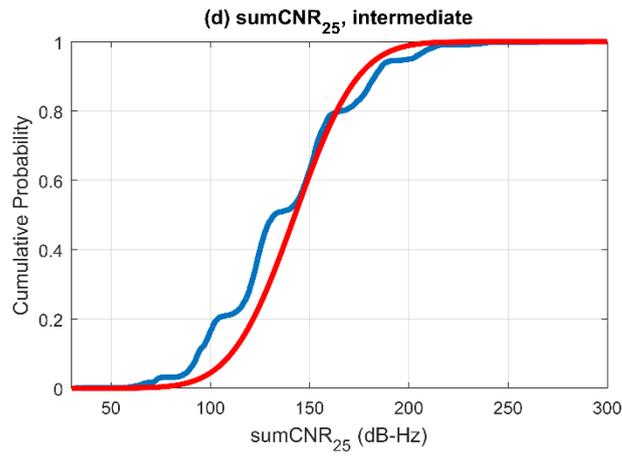
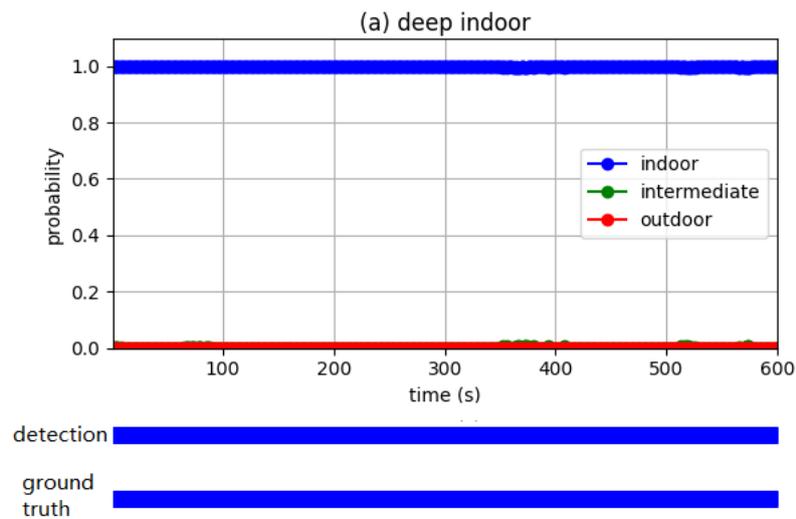
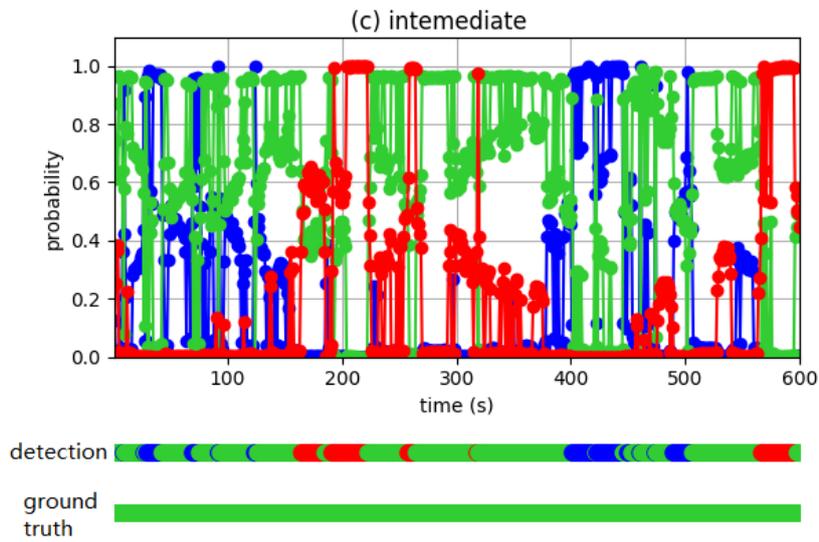
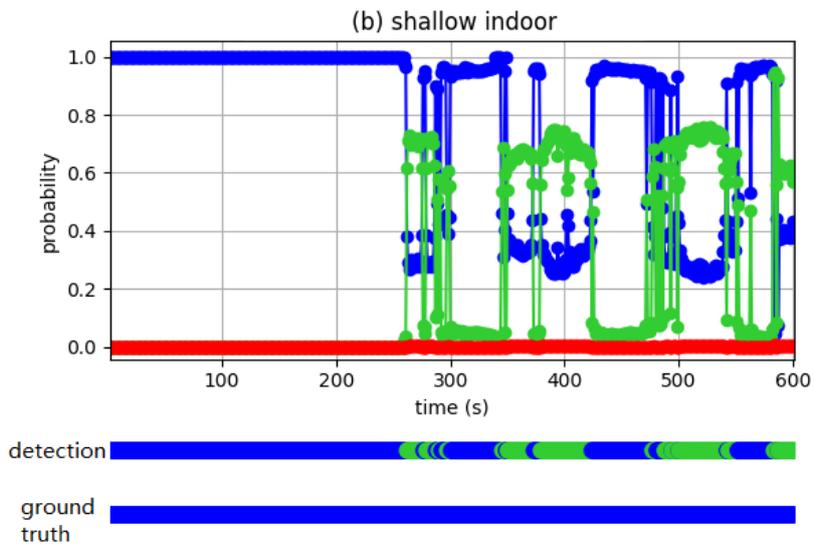


Figure 18—Fitting results of emission probabilities

**Results**

To test the indoor/outdoor detection ability under different GNSS reception conditions, the proposed detection method was examined by both static and kinematic experiments. Five different locations were chosen from the test database to carry out static tests – deep indoor (indoor), shallow indoor (indoor), intermediate, urban (outdoor), open-sky (outdoor). The respective classification results for these environments are depicted in Figure 19.





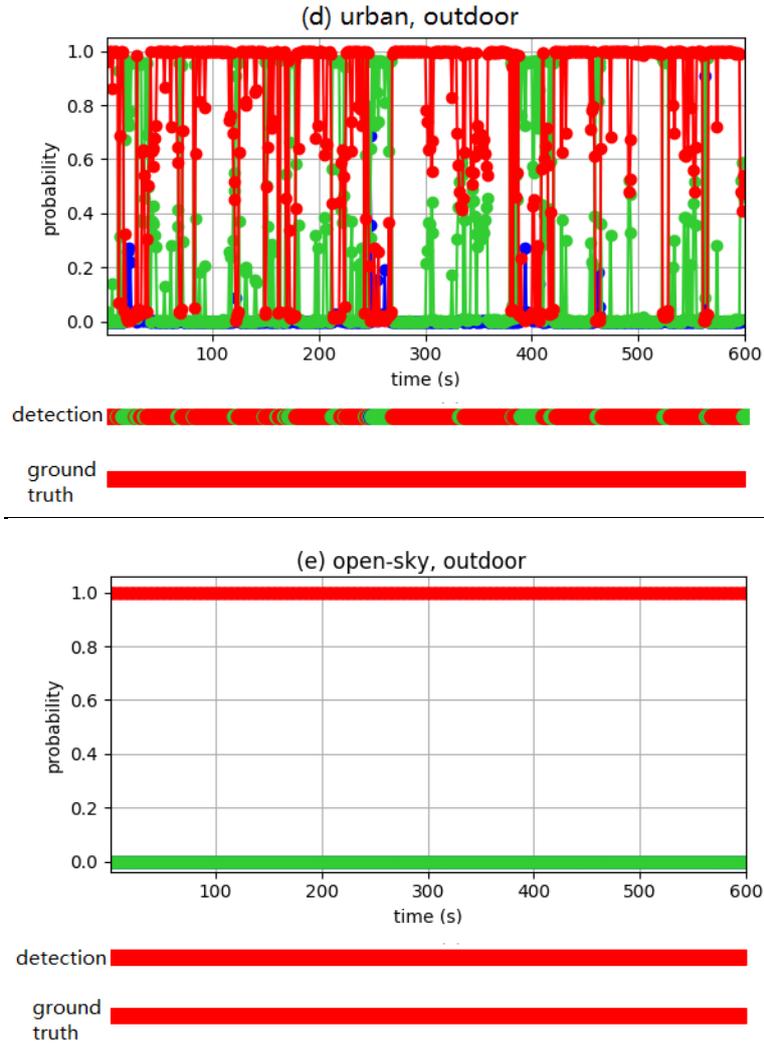
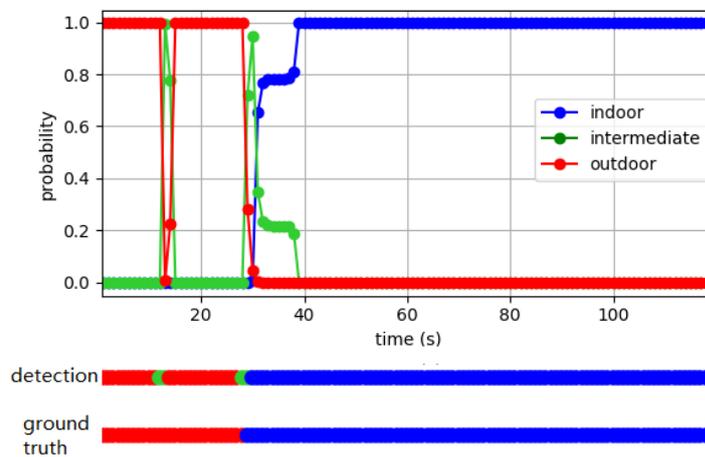


Figure 19—Static experimental results of the indoor-outdoor detection algorithm (deep and shallow indoor data was collected at the sites shown in Figure 5(a) and 5(b) respectively; intermediate data was collected at P3 in Figure 6; urban data was collected at P2 in Figure 7; open-sky data was collected at P3 in Figure 9.)

In the case of the open-sky and deep indoor environments, the detection results are very accurate as all samples of these scenarios are successfully detected with almost 100% probability. The shallow indoor scenario is a little challenging for the detector as more LOS signals and some strong reflected signals can be received through the window. It can be observed from Figure 19(b) that most samples are classified to indoor correctly but with some intermediate detections occasionally appearing among them. Meanwhile, from the probabilistic

output, it can be seen that the detection results are much less certain than the deep urban and open-sky scenarios. A similar behaviour is observed phenomenon happens for the data collected in an urban area, which can be explained by the fact that some signals are blocked by the tall buildings around and NLOS signals are also received. In the case of the intermediate environment, more signals are blocked by the roof and side walls, but some NLOS signals can still be received from the side without a wall. This makes recognition of the intermediate environment more challenging. The decision certainty is thus lower than the other scenarios and some measurements are classified as indoor or outdoor.

A kinematic test was carried out using the same outdoor-indoor transition data shown in Figure 10 and Figure 11. Test results with the ground truth are shown in Figure 20. For most of the time, the detection results are consistent with ground truth. In particular, the gradual increase of indoor probability after 30s shows the proposed method is able to recognise an indoor-outdoor transition with the help of transition probabilities within HMM.



*Figure 20—Kinematic test results*

## **OPEN-SKY AND URBAN ENVIRONMENT DETECTION**

In an open-sky environment, with no major obstacles between the receiver and the satellites, there are enough direct LOS signals for a good positioning solution. However, in urban environments, where the sky view is obscured by the surrounding buildings, only a limited number of satellites are directly visible. Moreover, a smartphone GNSS antenna uses linear polarization, making it especially susceptible to multipath interference and making it more difficult to detect NLOS reception from the signal strength [8]. This can affect the pseudorange and  $C/N_0$  measurements, causing deterioration in positioning accuracy.

As there is no clear boundary between an urban and an open-sky environment, it is more useful for the navigation system to provide a continuous measure of the density of the urban environment. An urban index (UI) is therefore introduced. A method is then needed to compute the UI as a continuous function of the input features. Qualitative relationships between the input features and UI have been identified, which can be expressed using fuzzy logic in a fuzzy inference system (FIS).

### ***Fuzzy Inference System***

Most data classification methods follow Boolean logic, a sample either belongs to a class or it does not. However, for the urban and open-sky environment classification problem, the boundary between these two contexts is not clearly defined, which limits the applicability of a hard data classification method.

Fuzzy logic, proposed in [27], uses fuzzy set theory to describe the degree of belongingness of a sample to a class. Fuzzy logic is characterized by membership functions and fuzzy rules. Membership functions define how an input or output value is mapped to the degree of membership of the relevant classes and fuzzy rules are a set of *if-then* statements that describe

how the FIS should make a decision from the input memberships. Thus, the rules enable the degree of output membership to be determined from the input memberships.

Figure 21 shows the architecture of the proposed fuzzy inference system for urban and open-sky environment classification. As described before, features of  $sumCNR_{25}$  and  $zPRR$ , extracted from smartphone GNSS measurements, are used as indicators of satellite signals for environment prediction. The output of the fuzzy inference system is a numeric UI between zero and one, computed from a linear combination of output memberships. The UI value describes the degree of likelihood that a context belongs to an urban environment and is applied to classify outdoor contexts. A higher rating value of UI indicates a higher likelihood of urban environment.

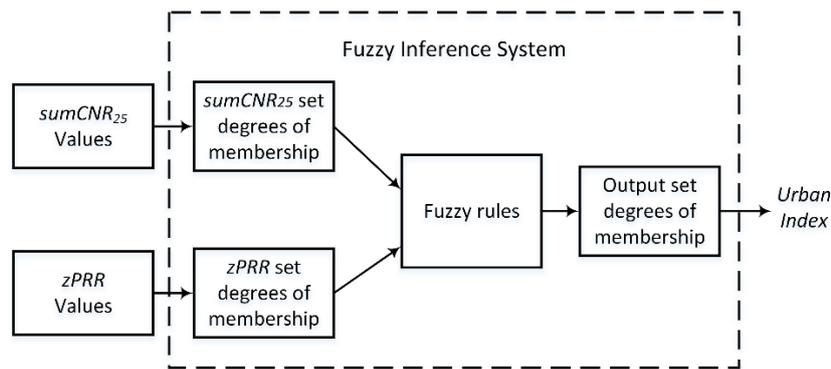


Figure 21—The architecture of the fuzzy inference system

Once the input and output variables of FIS are determined, the next step is to determine the membership functions. For the input membership functions, the triangle and Gaussian membership functions are used to assign the input variables into low, medium and high classes. For the output membership functions, five triangles with ranging from zero to one and overlaps between sets, were used to handle every combination of input fuzzy sets and provide gradual outputs from open to dense urban environments. The membership functions are shown in Figure 22. It should be noted that the parameters and architectures for the membership

functions were tuned and optimized based on the outdoor training dataset and the proposed fuzzy inference system is tested using the testing dataset as described in next section.

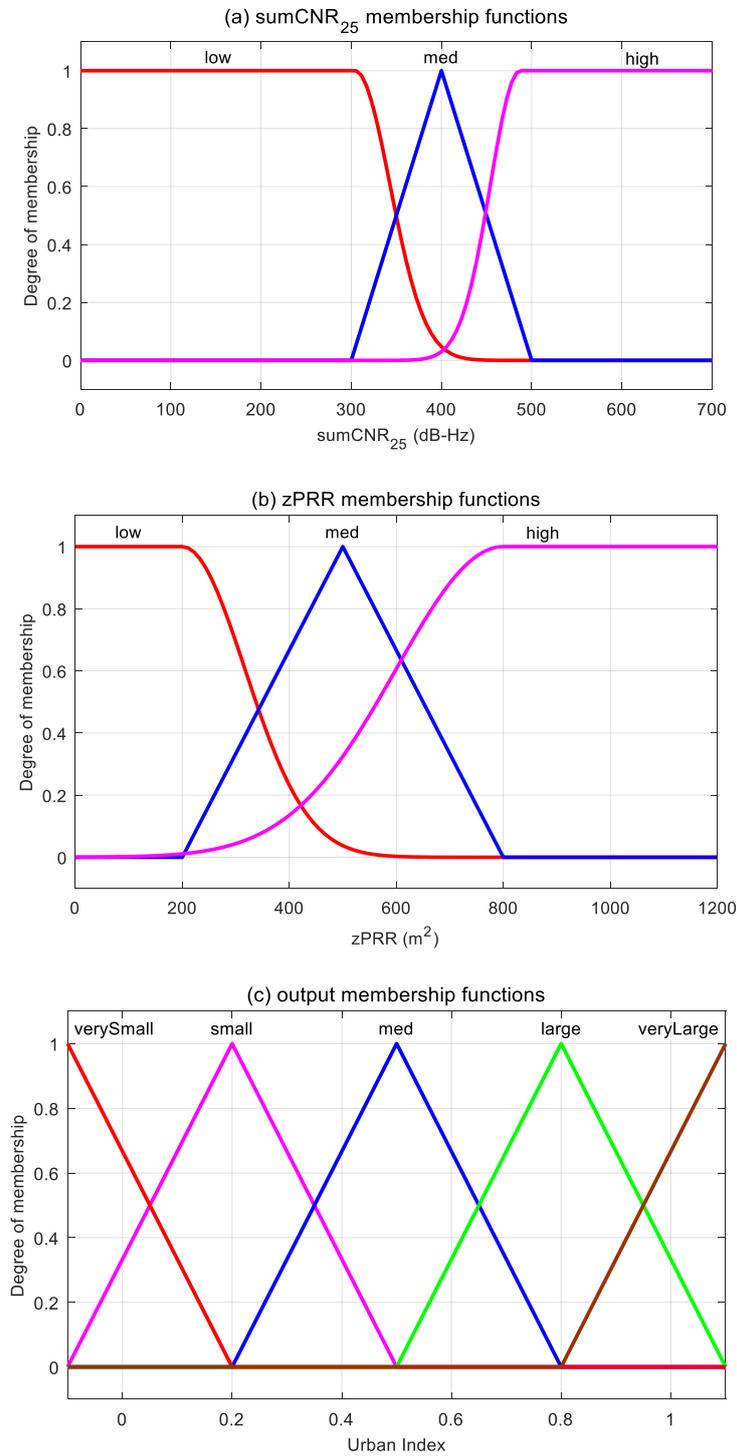


Figure 22—Membership functions used in fuzzy inference system

At the same time, to describe the relationship between the inputs and the output, a set of *if-then* rules were established as shown in Table 4. The rules are developed based on the basic knowledge of signal qualities in different environments. For example, if signals are strong with small residuals, the environment must have excellent GNSS reception, so it can be presumed to be an open-sky environment.

Table 4—Fuzzy rules used in fuzzy inference system

---

<b>R<sub>1</sub></b> : if $sumCNR_{25}$ is HIGH and $zPRR$ is HIGH then UI is MED
<b>R<sub>2</sub></b> : if $sumCNR_{25}$ is HIGH and $zPRR$ is MED then UI is SMALL
<b>R<sub>3</sub></b> : if $sumCNR_{25}$ is HIGH and $zPRR$ is LOW then UI is VERY SMALL
<b>R<sub>4</sub></b> : if $sumCNR_{25}$ is MED and $zPRR$ is HIGH then UI is LARGE
<b>R<sub>5</sub></b> : if $sumCNR_{25}$ is MED and $zPRR$ is MED then UI is MED
<b>R<sub>6</sub></b> : if $sumCNR_{25}$ is MED and $zPRR$ is LOW then UI is SMALL
<b>R<sub>7</sub></b> : if $sumCNR_{25}$ is LOW and $zPRR$ is HIGH then UI is VERY LARGE
<b>R<sub>8</sub></b> : if $sumCNR_{25}$ is LOW and $zPRR$ is MED then UI is LARGE
<b>R<sub>9</sub></b> : if $sumCNR_{25}$ is LOW and $zPRR$ is LOW then UI is MED

---

Once membership functions and fuzzy rules are defined, an inference procedure is applied to derive the output fuzzy set. In this research, the most commonly used one, Mamdani-type fuzzy inference system [28] is used. To better illustrate how a fuzzy inference system gets UI value from inputs, an implement of a sample with 450 dB-Hz  $sumCNR_{25}$  and 400 m<sup>2</sup>  $zPRR$  is shown in Figure 23.

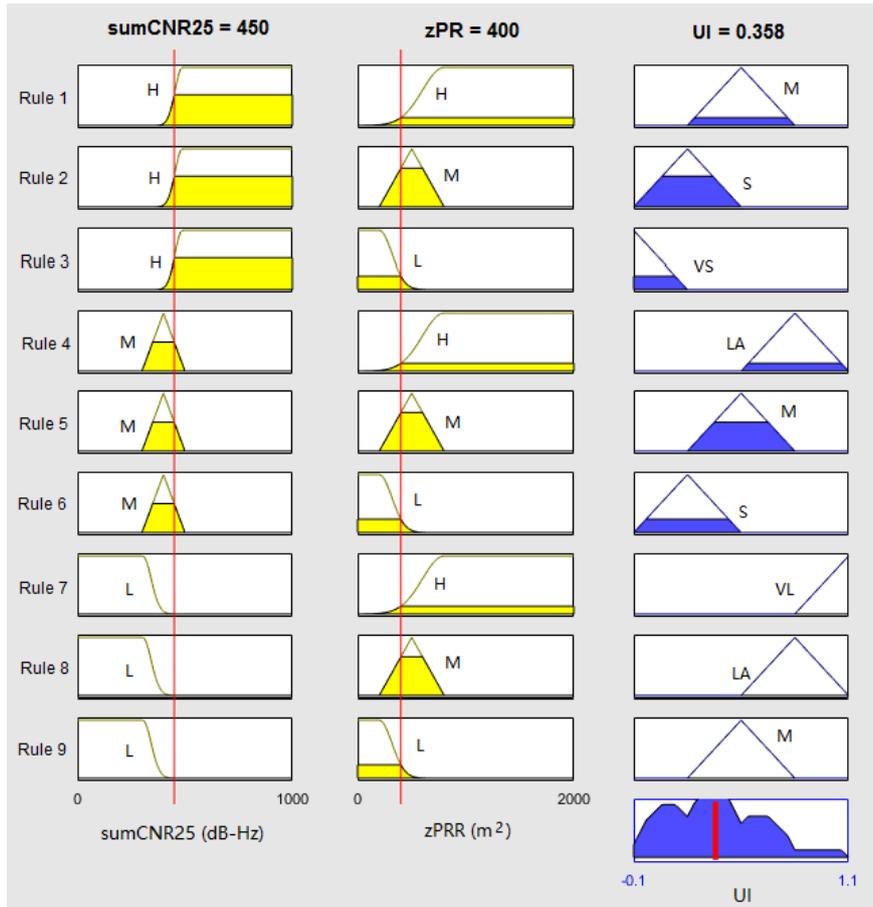


Figure 23—Example of a Mamdani-type fuzzy inference system ( $H$  = high,  $L$  = low,  $M$  = medium,  $S$  = small,  $LA$  = large,  $VL$  = very large,  $VS$  = very small)

Shown in Figure 24 are the outputs of the tuned fuzzy inference system versus the horizontal position errors based on the training data. The results demonstrate that the outdoor environmental contexts can generally be distinguished from each other. Based on these results, a threshold value of UI as 0.45, shown by the dashed line in the figure, was set with training samples with a UI smaller than threshold classified as an open-sky environment while samples with a UI larger than 0.45 are classified as an urban environment. From the figure, it is also interesting to mention that the positioning solutions with UI values smaller than the threshold value will always have horizontal position errors within 4.5 meters.

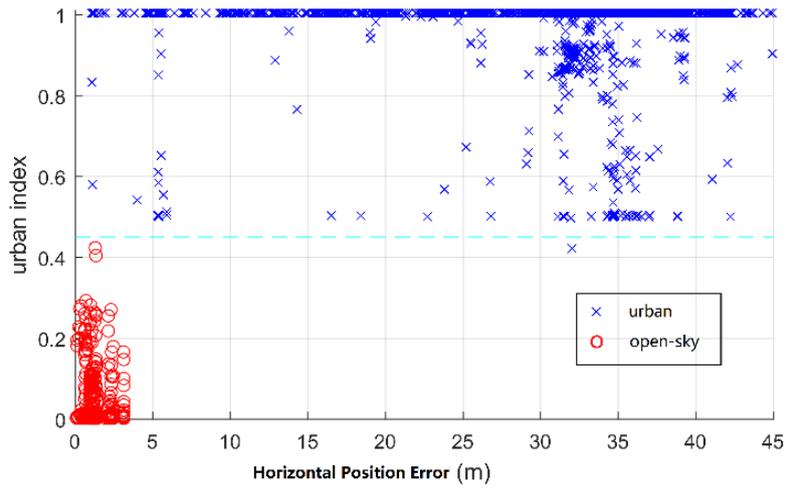


Figure 24—Data classification of training data using fuzzy inference system

### Results

To further verify the fuzzy inference system, the test data was processed by that system. The corresponding UI value versus horizontal position error is presented in Figure 25. The results show that all urban samples and most open-sky samples are correctly classified by the proposed system while about 0.6% (17 out of 2709) of open-sky environment data are misclassified. As a result, the reliability of the proposed fuzzy inference system has been demonstrated. Note that when applying in the actual application, depending on the requirements, the navigation system could be supplied with the urban indexes instead of binary classification results.

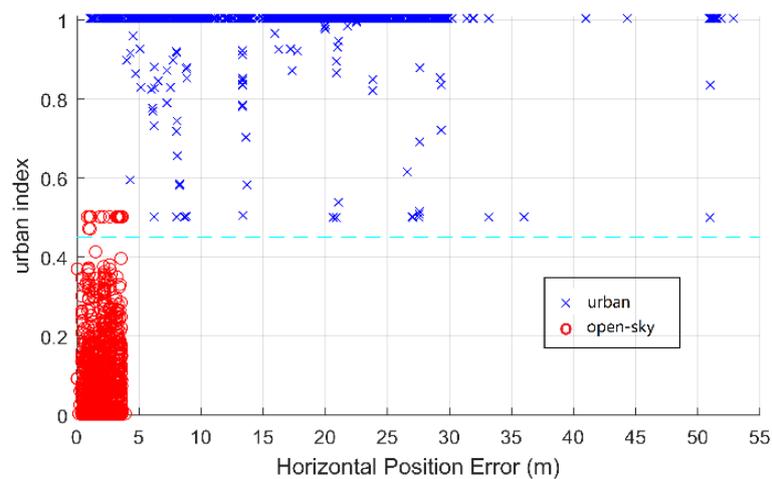


Figure 25—Test results of outdoor environment data classification

## HYBRID ENVIRONMENT DETECTION

The performance of environmental context detection following the proposed framework in Figure 4 is presented in the form of a confusion matrix, as shown in Table 5. A confusion matrix is a classification result table with each row representing the true class and each column representing the class output by the algorithm. The results show that the system achieves an overall 88.2% accuracy with this data, demonstrating that this approach can distinguish most of the test environments correctly. Note, however, that a thorough performance assessment would require testing over a much wider range of environments with a commensurably large training dataset. It can also be observed that some classes are more difficult to be detected than others. Most indoor and open-sky environment data can be classified correctly. Since the emission probabilities of the intermediate category are partly overlapped with indoor and outdoor ones, it makes the intermediate type the hardest one to detect due to similar signal properties with shallow indoor and dense urban scenarios.

Table 5—Confusion matrix of hybrid environment detection

<b>Actual</b>	<b>Predicted</b>			
	<b>Indoor</b>	<b>Intermediate</b>	<b>Urban</b>	<b>Open-sky</b>
<b>Indoor</b>	<b>2070</b>	113	0	0
<b>Intermediate</b>	307	<b>1305</b>	217	0
<b>Urban</b>	3	389	<b>1716</b>	0
<b>Open-sky</b>	0	0	17	<b>2692</b>

## CONCLUSIONS AND FURTHER WORK

This paper demonstrates the determination of the environmental context for navigation purposes using the GNSS module on a smartphone. Based on the detected environment, the optimum navigation and positioning techniques can be selected accordingly.

Environmental contexts are categorised as indoor, intermediate, urban or open-sky based on different characteristics of GNSS reception. In an open-sky environment, the horizontal GNSS position error is within 5m using a smartphone. In an urban area, some satellite signals may be blocked and reflected by the buildings around, so the positioning accuracy using conventional GNSS positioning algorithms degrades dramatically. In an indoor environment, GNSS may incur tens of meters' error or even no position solution at all. An intermediate environment type is also included in the framework, which can also serve as a connection between the indoor and outdoor categories. According to the proposed categorization, environmental context is first distinguished between indoor and outdoor scenarios. Two features based on the availability and strength of GNSS signals are extracted from the smartphone outputs. Then a detection scheme using a hidden Markov model is applied for classification. The outdoor context is then further categorised into urban and open-sky environments. To distinguish the differences in GNSS signal qualities, an additional feature based on the pseudorange residuals are calculated from the measurements. Since the boundary between urban and open-sky areas is not clearly defined, a fuzzy inference system is implemented. In the model, membership functions and fuzzy logic rules are set, with an urban index as the output. Finally, the proposed environmental context detection method is tested as a whole, achieving an overall 88.2% accuracy. Most indoor and open-sky environment data can be classified correctly while the intermediate type is the hardest one to detect due to similar signal properties with shallow indoor and dense urban scenarios.

In future work, we plan to add more smartphone sensors into the context detection framework to enhance the reliability of environment detection. More work also need to be done to determine a suitable boundary between urban and open areas. This will be integrated with urban positioning techniques, such as shadow matching and 3D-mapping-aided GNSS method [3], for better position solutions in urban areas.

## ACKNOWLEDGEMENTS

This work is funded by the UCL Engineering Faculty Scholarship Scheme and the Chinese Scholarship Council.

## REFERENCES

- [1] Groves, P.D., Martin, H., Voutsis, K., Walter, D., and Wang, L., “Context Detection, Categorization and Connectivity for Advanced Adaptive Integrated Navigation,” *Proceedings of the 26th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2013)*, Nashville, TN, September 2013, pp. 1039–1056. Also available from <http://discovery.ucl.ac.uk/>.
- [2] Groves, P.D., and Jiang, Z., “Height Aiding, C/N<sub>0</sub> Weighting and Consistency Checking for GNSS NLOS and Multipath Mitigation in Urban Areas,” *Journal of Navigation*, Vol. 66, No. 05, 2013, pp. 653–669. Also available from <http://discovery.ucl.ac.uk/>.
- [3] Adjrad, M., and Groves, P.D., “Intelligent Urban Positioning using Shadow Matching and GNSS Ranging Aided by 3D Mapping,” *Proceedings of the 29th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS*

- 2016), Portland, Oregon, September 2016, pp. 534–553. Also available from <http://discovery.ucl.ac.uk/>.
- [4] Groves, P.D., *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*, Second Edition, Boston London: Artech House, 2013.
- [5] Ching, W., Teh, R.J., Li, B., and Rizos, C., “Uniwide WiFi Based Positioning System,” *IEEE International Symposium on Technology and Society (ISTAS)*, June 2010, pp. 180-189.
- [6] Bell, S., Jung, W.R., and Krishnakumar, V., “WiFi-Based Enhanced Positioning Systems: Accuracy Through Mapping, Calibration, and Classification,” *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*, ACM, November 2010, pp. 3-9.
- [7] Betz, J.W., *Engineering Satellite-Based Navigation and Timing: Global Navigation Satellite Systems, Signals, and Receivers*, John Wiley & Sons, 2015.
- [8] Wang, L., Groves, P.D., and Ziebart, M., “Smartphone Shadow Matching for Better Cross-Street GNSS Positioning in Urban Environments,” *Journal of Navigation*, Vol. 68, No. 03, 2015, pp. 411–433. Also available from <http://discovery.ucl.ac.uk/>.
- [9] Groves, P.D., “The Complexity Problem in Future Multisensor Navigation and Positioning Systems: A Modular Solution,” *Journal of Navigation*, Vol. 67, No. 02, 2014, pp. 311-326. Also available from <http://discovery.ucl.ac.uk/>.
- [10] Groves, P.D., Wang, L., Walter, D., Martin, H., Voutsis, K., and Jiang, Z., “The Four Key Challenges of Advanced Multisensor Navigation and Positioning,” *IEEE/ION PLANS 2014*, Monterey, California, May 2014, pp. 773-792. Also available from <http://discovery.ucl.ac.uk/>.

- [11] Capurso, N., Song, T., Cheng, W., Yu, J., and Cheng, X., “An Android-based Mechanism for Energy Efficient Localization depending on Indoor/Outdoor Context,” *IEEE Internet of Things Journal*, 2016.
- [12] Zhou, P., Zheng, Y., Li, Z., Li, M., and Shen, G., “IODetector: A Generic Service for Indoor Outdoor Detection,” *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, ACM, November 2012, pp. 113-126.
- [13] Wang, W., Chang, Q., Li, Q., Shi, Z., and Chen, W., “Indoor-Outdoor Detection Using a Smart Phone Sensor,” *Sensors*, Vol. 16, No. 10, 2016, pp. 1563.
- [14] Radu, V., Katsikouli, P., Sarkar, R., and Marina, M.K., “A Semi-Supervised Learning Approach for Robust Indoor-Outdoor Detection with Smartphones,” *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, ACM, November 2014, pp. 280-294.
- [15] Zou, H., Jiang, H., Luo, Y., Zhu, J., Lu, X., and Xie, L., “BlueDetect: An iBeacon-Enabled Scheme for Accurate and Energy-Efficient Indoor-Outdoor Detection and Seamless Location-Based Service,” *Sensors*, Vol. 16, No. 2, 2016, pp. 268.
- [16] Sung, R., Jung, S.H., and Han, D., “Sound Based Indoor and Outdoor Environment Detection for Seamless Positioning Handover,” *ICT Express*, Vol. 1, No. 3, 2015, pp. 106-109.
- [17] Lin, T., O’Driscoll, C., and Lachapelle, G., “Development of a Context-Aware Vector-Based High-Sensitivity GNSS Software Receiver,” *Proceedings of the 2011 International Technical Meeting of The Institute of Navigation (ITM ION)*, San Diego, CA, January 2011, pp. 1043–1055.
- [18] Shafiee, M., O’Keefe, K., and Lachapelle, G., “Context-Aware Adaptive Extended Kalman Filtering using Wi-Fi Signals for GPS Navigation,” *Proceedings of the 24th*

*International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2011)*, Portland, Oregon, September 2011, pp. 1305–1318.

- [19] Shivaramaiah, N.C., and Dempster, A.G., “Cognitive GNSS Receiver Design: Concepts and Challenges,” *Proceedings of the 24th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2011)*, Portland, Oregon, September 2011, pp. 2782–2789.
- [20] Parviainen, J., Bojja, J., Collin, J., Leppänen, J., and Eronen, A., “Adaptive Activity and Environment Recognition for Mobile Phones,” *Sensors*, Vol. 14, No. 11, 2014, pp. 20753-20778.
- [21] Android, “Google Location Services API,”  
<https://developers.google.com/android/reference/com/google/android/gms/location/package-summary>.
- [22] Gao, H., and Groves, P.D., “Context Determination for Adaptive Navigation using Multiple Sensors on a Smartphone,” *Proceedings of the 29th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2016)*, Portland, Oregon, September 2016, pp. 742–756. Also available from  
<http://discovery.ucl.ac.uk/>.
- [23] Kaplan, E., and Hegarty, C., *Understanding GPS: Principles and Applications*, Second Edition, Artech House, 2005.
- [24] Groves, P.D., Jiang, Z., Rudi, M., and Strode, P., “A Portfolio Approach to NLOS and Multipath Mitigation in Dense Urban Areas,” *Proceedings of the 26th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS*

2013), Nashville, TN, September 2013, pp. 3231–3247. Also available from

<http://discovery.ucl.ac.uk/>.

- [25] Bishop, C.M., *Pattern Recognition and Machine Learning*, Vol. 1, New York: Springer, 2006.
- [26] Viterbi, A., “Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm,” *IEEE Transactions on Information Theory*, Vol. 13, No. 2, 1967, pp. 260-269.
- [27] Zadeh, L.A., "Fuzzy Sets," *Information and Control*, Vol. 8, No. 3, 1965, pp. 338-353.
- [28] Mamdani, E.H., and Assilian, S., "An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller," *International Journal of Man-Machine Studies*, Vol. 7, No. 1, 1975, pp. 1-13.