Context Determination for Adaptive Navigation using Multiple Sensors on a Smartphone

Han Gao, Paul D Groves
University College London

BIOGRAPHY

Han Gao is a PhD student at University College London (UCL) in the Engineering Faculty's Space Geodesy and Navigation Laboratory (SGNL). He received a Bachelor's degree in Aerospace Engineering from Shanghai Jiao Tong University (SJTU) in 2014. He is interested in multi-sensor contextual navigation and positioning techniques. (han.gao.14@ucl.ac.uk)

Dr. Paul Groves is a Senior Lecturer (associate professor) at UCL, where he leads a program of research into robust positioning and navigation within SGNL. He joined in 2009, after 12 years of navigation systems research at DERA and QinetiQ. He is interested in all aspects of navigation and positioning, including improving GNSS performance under challenging reception conditions, advanced multisensor integrated navigation, and novel positioning techniques. He is an author of more than 80 technical publications, including the book *Principles of GNSS, Inertial and Multi-Sensor Integrated Navigation Systems*, now in its second edition. He holds a bachelor's degree and doctorate in physics. (p.groves@ucl.ac.uk)

ABSTRACT

Navigation and positioning is inherently dependent on the context, which comprises both the operating environment and the behaviour of the host vehicle or user. No single technique is capable of providing reliable and accurate positioning in all contexts. In order to operate reliably across different contexts, a multi-sensor navigation system is required to detect its operating context and reconfigure the techniques accordingly. This paper aims to determine the behavioural and environmental contexts together, building the foundation of a context-adaptive navigation system.

Both behavioural and environmental context detection results are presented. A hierarchical behavioural recognition scheme is proposed, within which the broad classes of human activities and vehicle motions are detected using measurements from accelerometers, gyroscopes, magnetometers and the barometer on a smartphone by decision trees (DT) and Relevance Vector Machines (RVM). The detection results are further improved by behavioural connectivity. Environmental contexts (e.g., indoor and outdoor) are detected from GNSS measurements using a hidden Markov model.

The paper also investigates context association in order to further improve the reliability of context determination. Practical test results demonstrate improvements of environment detection in context determination.

1. INTRODUCTION

Navigation and positioning systems are inherently dependent on the context, which comprises both the operating environment and the behaviour of the host vehicle or user [1]. Environments and behaviours reveal different aspects of context implicitly. Environmental context is concerned with the spatial information (e.g., indoor or outdoor) of the navigation system while behavioural context is more related to the mobility information (e.g., stationary, walking or running). For many applications, the context can change, particularly for smartphones, which move between indoor and outdoor environments and can be stationary, on a pedestrian, or in a vehicle. To meet the growing demand for greater accuracy and reliability in a wider range of challenging contexts, many navigation and positioning techniques have been developed or improved, such as Wi-Fi positioning [2][3], multiple-constellation global navigation satellite system (GNSS) [4], GNSS shadow matching [5][6] and pedestrian dead reckoning (PDR) using step detection [4][7]. However, no single current technique is capable of providing reliable and accurate positioning in all contexts. Therefore, in order to operate reliably across different contexts, a multi-sensor navigation system is required to be able to detect its operating context and reconfigure the technique it uses accordingly, which is referred to as context-adaptive navigation [1][8].

Figure 1 illustrates a possible architecture for a multi-sensor context adaptive navigation or positioning system. In an autonomous context adaptive navigation system, behavioural and environmental context categories are identified independently and associated using available sensors in the mobile device. Based on the contexts detected, different sensors may be selected and their measurements may be processed in different ways within each subsystem. Consequently, the integration module can adapt itself and export positioning results by selecting proper subsystems and varying the tuning of the algorithms. Then the positioning results can be used for location based services and improvement of context detection.

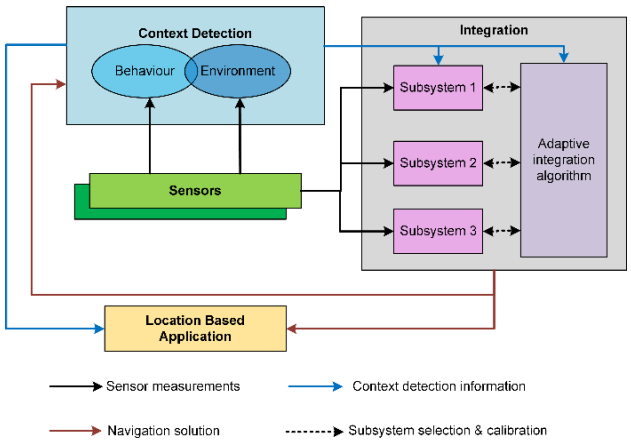


Figure 1: An example of a context adaptive navigation system

Environmental context is the central component of a contextual navigation or positioning system because it determines the types of signals available. For example, GNSS reception is good in open environments, but poor indoors and in deep urban areas. Wi-Fi signals are not available in rural area, in the air or at sea. In an underwater environment, most radio signals are not available at all. In addition, processing techniques 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 environment, non-line-of-sight (NLOS) reception of GNSS signals or multipath interference may be detected using consistency checking techniques based on sequential elimination [9]. In dense urban areas, more sophisticated algorithms are required for NLOS/multipath detection.

Behavioural context is also important because it can contribute additional information to the navigation solution. It will help mobile devices to understand what the user is doing under particular circumstances [10]. A stationary pedestrian or a land vehicle indicates a fixed location and will not need to update its velocity and position. Land vehicles normally remain on the ground, effectively removing one dimension from the position solution. Similarly, boats, ships and underwater vehicles can all be on land, but only exhibit some specific types of behaviours. Within a GNSS receiver, the behaviour can be used to set the bandwidths of the tracking loop and coherent integration intervals, and to predict the temporal characteristics of multipath [11].

Previous work on contextual navigation and positioning has focused on individual positioning techniques. For instance, there has been substantial research into determining the motion types for indoor positioning applications [12][13]. Researchers have also investigated context-adaptive GNSS to adjust the processing strategies and parameters of GNSS receivers [11]. Moreover, despite contextual awareness having been applied for different tasks within a mobile device [14][15][16], most of the related services are provided for non-navigation purpose. Context frameworks designed in general may not be suitable for context adaptive navigation. A context framework for navigation and positioning must be designed especially in order to be fit for navigation purpose; otherwise, it serves no purpose.

Recently, a number of researchers have investigated different approaches towards a context adaptive navigation system. The initial attempt was made in [17], in which a Location-Motion-Context (LoMoCo) solution was proposed using Bayes reasoning, to determine the users' context information from the locations and motion states. In 2013, the concept and framework of 'context adaptive navigation' was first introduced systematically in [1] by UCL, with the preliminary behavioural and environmental context detection results following. Following the initial proof of concept, an adaptive activity and environment recognition algorithm for context model parameters was proposed in [16]. In [18], it was shown that a hidden Markov model (HMM) can be used to integrate the location based motion states for inferring the mobility context of pedestrians. In [19], a framework for inferring typical on-campus contexts was developed, which took locations, timespans and the user's mobility contexts into consideration and achieved an 88.8% success rate using a Naïve Bayes classifier.

Existing research have demonstrated the relevant context detection techniques and built the foundations of a context adaptive navigation system. To further carry on the research, many new lines of research are required to be pursued, including:

- Definition of a set of context categories for navigation and positioning purposes;
- Development of behavioural and environmental context detection algorithms using a wide range of sensors;
- Development of a robust context determination algorithm;
- The practical demonstration of a basic multi-sensor context adaptive integrated navigation system.

This paper aims to determine behavioural and environmental context as whole with multiple sensors on the smartphone. Section 2 presents a study of behaviour recognition. The recognition framework and the performance of supervised machine learning classification algorithms are described, along with the temporal connectivity of behaviours. Section 3 proposes a new method for indoor-outdoor environmental context detection using GNSS measurements under a hidden Markov model. Section 4 considers behavioural and environmental context as whole and investigates context association to improve the reliability of context determination. Conclusions are presented in Section 5.

2. BEHAVIOUR RECOGNITION

2.1. Framework

The behavioural context may be divided into several broad classes: human activity, land vehicle, water vehicle, aircraft and spacecraft [1]. Each class contains detailed subdivisions. To provide robust and accurate classification of behaviours, a hierarchical detection frame is proposed to proceed from a coarse-grained recognition towards fine-grained subtasks. The top level of classifier is designed to distinguish between the broad classes while the bottom level of classifiers is responsible for recognising the subclasses within each broad class. Compared with a single classifier dealing with all behavioural scenarios, this hierarchical framework has two benefits. Different features and machine learning algorithms can be used in different classifiers for better recognition performance. Moreover, a flexible scheme is offered as new classes (e.g. water vehicle, aircraft) or subclasses can be added within the framework, as illustrated in Figure 2**.** The classifiers and classes in solid boxes indicate those are already implemented within the current context determination framework, while those in dash boxes indicate potential future extensions.

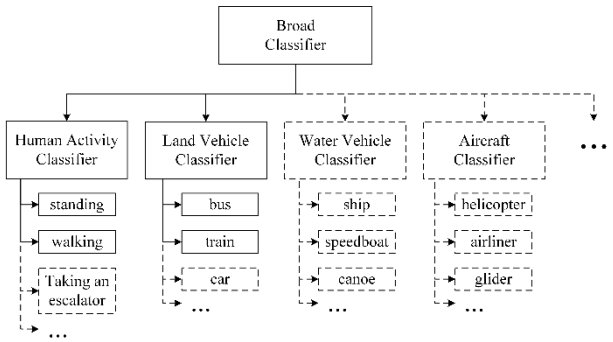


Figure 2: Extensive framework of behaviour detection

To illustrate the effectiveness of the framework in the initial stage of the study, human activity and land vehicle classes are included within the scope of the research. The current detection system consists of three classifiers: a human-vehicle classifier, a human activity classifier and a vehicle motion classifier, which are organized into a hierarchy as in Figure 3. A human-vehicle classifier is designed at the start of the system to distinguish between motorised vehicle motions and non-motorised activities. When motorised transport is recognised, the detection system proceeds to the vehicle motion classifier for classification of different vehicle motions. Otherwise, it proceeds to the human activity classifier. A set of detailed categories currently included within each classifier are introduced in Table 1. Note that distinct from human activities, motorised land vehicles, propelled by internal combustion engines or electric motors, sometimes combinations of the two, can be identified by the vibrations from the frequency spectrum of the accelerometers. Engine vibration applies mainly to internal combustion engines, whereas road-induced vibration applies to all land vehicles. The vehicles covered in this study by now include diesel trains, diesel buses and underground trains. All underground trains are electric for safety reasons. The hybrid vehicles were not included in the current study. The mode of stationary vehicles with the engine on is included within the category because it can play a significant role in context association to minimise impossible behavioural context transitions, such as from a moving vehicle to another moving vehicle directly, or one human activities connected to a moving vehicle without intermediate categories.

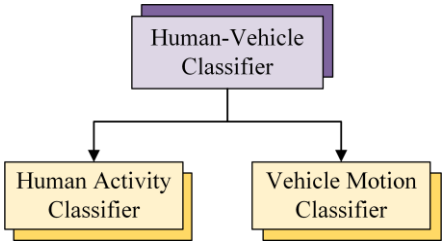


Figure 3: Overview of behaviour recognition system

Human activity types	Vehicle motion types
Stationary; Walking; Running; Ascending stairs; Descending stairs.	Stationary vehicles with the engine on; Moving diesel trains; Moving diesel buses; Moving underground trains.

Table 1: Detailed types of behaviours

2.2. Methodology

Behavioural context recognition systems consist of four main phases: sensing, preprocessing, feature extraction and either training or classification. They are described as follows:

1) Sensing: In this step, sensors collect the raw data at specific sampling rates.
2) Preprocessing: Subsequently, the raw data can be processed in various ways, such as cleansing and filtering. Then, a windowing scheme is applied to segment the data into successive pieces for further calculation.
3) Feature extraction: Various features, representing the main characteristics of behaviours, are extracted from the segmented data as the inputs of classifiers.
4) Training and Classification: In the training stage, the recognition classifiers for classification are constructed and the parameters of the model are learned from training sets. In the classification stage, the trained classifiers are used to recognise different behaviours. A detailed description of the machine learning algorithms used in this study is presented in Section 2.2.4.

2.2.1. Sensing

As previous research [20][21] has already proved, among the sensors in a smartphone, measurements from the

inertial sensors are capable of taking the leading roles in motion recognition. The accelerometer and gyroscope signals are able to track kinematic motions indirectly by measuring the specific force and angular rate. Motion can also be inferred from some sensors that measure magnetic features. Magnetometers sense the Earth's magnetic field, enabling changes in heading to be detected. A barometer, also called a barometric altimeter, measures the ambient air pressure, from which the heights can be estimated and the changes in height can be derived [4]. Therefore, in this study, accelerometers, gyroscopes, magnetometers and a barometer, found in most smartphones, are used for behavioural recognition.

2.2.2. Pre-processing

Prior to feature extraction, the raw sensor data are divided into small segments using sliding windows. The selection of an appropriate window length is important, and different durations can be set for it. At a sampling frequency of 100 Hz, a 500 sample window is suitable based on previous studies [22][23]. It was shown that a window length of four seconds was an effective and sufficient value for behaviour recognition, neither too short to capture enough features, nor too long to avoid mixing multiple contexts in a single window. A 4s sliding window with a 50% overlap is used for training and testing to avoid missing information between successive windows. Note that to get quicker responses, a 75% overlap is adopted in Section 2.4 and 4.2, thus context can be determined every second.

For accelerometers, gyroscopes and magnetometers, they are made in three dimensions, referred to as the *x*-axis, *y*-axis and *z*-axis. However, the recognition performance may be affected by orientation changes if the model is trained only for a specific orientation [20][24]. In order to minimise such effects, the magnitudes of the sensors are calculated from the outputs of three axes, *x*, *y* and *z*, thus

$$magnitude = \sqrt{x^2 + y^2 + z^2} \quad . \tag{1}$$

However, the existence of a sequence with a non-zero mean can hide important information in the frequency domain, so the means of the magnitudes are removed from each segment prior to computing the frequency-domain features.

2.2.3. Feature Extraction

Once the data pre-processing is completed, features need to be extracted from the segmented data to be used for training and classification. A good set of feature measurements can often provide accurate and comprehensive descriptions of patterns from which the differences between context categories are easily discerned. In this study, both time-domain and frequency-domain features are extracted for behavioural recognition.

Time-domain features describe temporal variations of motions during the epoch. The time-domain features selected include range, variance, skewness and kurtosis extracted from all sensors. The effectivenesses of these features for behaviour classification have been shown in different studies [20][25][26]. Zero-crossing rate (ZCR) is also extracted from the preprocessed accelerometer signals, which is used to differentiate different periods of human activity changing with the time. They are expressed as follows and summarized in Table 2:

$$range = max\{x\} - min\{x\} \tag{2}$$

$$\sigma = \sqrt{E\{(x-\mu)^2\}} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(x_n - \overline{x})^2} \tag{3}$$

$$skewness = \frac{E\{(x-\mu)^3\}}{\sigma^3} = \frac{1}{N\sigma^3}\sum_{n=1}^{N}(x_n - \overline{x})^3 \tag{4}$$

$$kurtosis = \frac{E\{(x-\mu)^4\}}{\sigma^4} = \frac{1}{N\sigma^4}\sum_{n=1}^{N}(x_n - \overline{x})^4 \tag{5}$$

$$ZCR = \frac{1}{N-1}\sum_{n=1}^{N-1}\mathbb{I}\{x_n x_{n+1} < 0\} \tag{6}$$

where N is the length of the sample window, μ is the mean, x_n represents the n-th epoch of data in the window and the indicator function $\mathbb{I}(.)$ is 1 if its argument is true and 0 otherwise.

	Range (4)	**Variance (4)**	**Skewness (4)**	**Kurtosis (4)**	**ZCR (1)**	**Number of features**
Human-Vehicle	√	√	√	√		16
Human Classifier	√	√	√	√	√	17
Vehicle Classifier	√	√	√	√		16

Table 2: Time-domain features for each classifier. Range, variance, skewness and kurtosis are extracted from accelerometers, gyroscopes, magnetometers and the barometer respectively while ZCR is only extracted from the accelerometer signals.

Frequency-domain features describe the periodic characteristics of motion during the sample window. In frequency-domain analysis, peaks are centered on different frequency values for different behaviours after a fast Fourier transform (FFT). For this reason, features in the frequency spectrum can reveal significant information on motion periods and vibrations. In the human-vehicle classifier and human activity classifier, the frequency of the largest peak and related spectrum peak magnitude of accelerometers and gyroscopes, are extracted to capture the differences between human and vehicles, and the main temporal periodicity of different human activities. Specifically, according to [1][8], the vehicles always exhibit one or more peaks between 20 Hz and 40 Hz due to vibration and little peaks below 10 Hz when the vehicle is

not moving. Thus all frequency domain features of vehicle classifier are estimated in the following sub-bands instead of the whole spectrum: 0-10 Hz, 10-20 Hz, 20-30 Hz, 30-40 Hz and 40-50 Hz.

The Power Spectral Density (PSD) of signals shows the strength of the energy distributed in the frequency spectrum, thus the PSD of accelerometers is adopted in the vehicle motion classifier to distinguish different vehicle motions with diverse vibrations. For finite time series x_n sampled at discrete time Δt for a total measurement period $T = N\Delta t$, the PSD is defined by

$$S_{xx}(\omega) = \frac{(\Delta t)^2}{T}\left|\sum_{n=1}^{N} x_n e^{-i\omega n}\right|^2 . \quad (7)$$

A summary of the frequency domain features for each classifier is presented in Table 3.

	Peak Magnitude	Peak Position	PSD	Number of features
Human-Vehicle	2	2		4
Human Classifier	2	2		4
Vehicle Classifier	10		5	15

Table 3: Number of frequency-domain features for each classifier. In the human-vehicle and human activity classifier, the features are the frequencies of the largest peak and related spectrum peak magnitudes, extracted from accelerometers and gyroscopes, respectively. In the vehicle classifier, the largest peak magnitudes and the PSDs of the accelerometers and gyroscopes are estimated in the sub-bands.

2.2.4. Supervised Machine Learning

Supervised classification methods learn a model of relationships between the target vectors and the corresponding input vectors consisting of training samples and then utilize this model to predict target values for the test data [28]. Note that in the algorithms described in this section, assume that there are L possible behavioural categories $\mathbf{C}=\{C_k \mid k=1,2,\cdots, L\}$. Given a training dataset $\mathbf{X}=\{X_{i,j} \mid i=1,2, \cdots, N; j=1,2, \cdots, M\}$, each sample $\mathbf{X}_i=\{X_{i,1}, X_{i,2}, \cdots, X_{i,M}\}$ is assigned to a target value $y_i \in \mathbf{C}$. M is the number of features and N is the number of the samples in the dataset.

Due to the limited space, only the algorithms used in the framework, decision tree and relevance vector machine, are introduced in detail. The reason why they are selected will be explained in Section 2.3.2.

- **Decision Tree (DT)**

A decision tree is a method that performs a recursive binary partitioning of the feature space to reach a decision. Given training samples and the corresponding class labels, the dataset is split into branch-like segments such that samples with the same labels are grouped together. The root is the starting point of the tree while the nodes without outgoing lines are the terminals. The samples are classified while navigating from the root down to the terminals. Along the path, the internal nodes split the data into two or more segments according to decision criteria based on features until all samples at a node belong to the same class.

Let the training dataset at node m be represented by Q. For each split at the node, one feature value $X_{i,j}$ and a threshold θ are required to partition the data into two subsets:

$$\begin{aligned} Q_{left}(\theta) &= \left\{(\mathbf{X}, y) \in Q \mid X_{i,j} \leqslant \theta\right\} \\ Q_{right}(\theta) &= \left\{(\mathbf{X}, y) \in Q \mid X_{i,j} > \theta\right\} \end{aligned} \quad (8)$$

The impurity at node m, a measure of the homogeneity of the labels, is computed using an impurity function $H(.)$,

$$G(Q,\theta) = \frac{n_{left}}{N_m} H(Q_{left}(\theta)) + \frac{n_{right}}{N_m} H(Q_{right}(\theta)) \quad (9)$$

where N_m is the number of the samples at node m, and n_{left} and n_{right} indicate the number of samples splitted into the left and right branch, respectively. To choose the thresholds that best splits the samples at each step, the parameters that minimise the impurity are selected.

$$\theta^* = arg\min_{\theta} G(Q,\theta) \quad (10)$$

This is repeated for subsets $Q_{left}(\theta^*)$ and $Q_{right}(\theta^*)$ until the maximum size of tree is reached, or N_m=1. Note that the choice of an impurity function depends on the task being solved. In this study, information entropy [27] is used for its computational simplicity:

$$H(Q) = -\sum_k p_{mk} log_2 p_{mk} \quad (11)$$

where p_{mk} is the proportion of samples Q belonging to category C_k at node m.

Amongst the supervised machine learning methods, the decision tree has various advantages. The model is simple and clearly explained by Boolean logic. Also, a large amount of data can be trained and tested within a reasonable time. However, the training process of decision tree methods cannot guarantee to return the globally optimal tree. Moreover, the method can create a biased tree if some classes dominate the training data.

- **Relevance Vector Machine (RVM)**

Fundamentally, RVM is a binary classifier ($y \in \{0,1\}$) under a Bayesian probabilistic framework [28]. The relationship

of the input vector and their real-valued predictions $t(\mathbf{X}_i)$ are modelled by a linearly weighted function

$$t(\mathbf{X}_i;\mathbf{w}) = \sum_{i=1}^{N} w_i \phi(\mathbf{X}_i) = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{X}_i) \quad (12)$$

where $\mathbf{w}$ denotes the weights of samples and $\phi(\mathbf{X}_i)$ is a nonlinear basis function. The input data samples $\mathbf{X}_i$ are classified according to the sign of $t(\mathbf{X}_i)$. To infer the function $t(\mathbf{X}_i)$, we need to define the basis function and to estimate the weights as well. In here, the radial basis kernel function is used, so that:

$$\Phi_{ij} = \phi^T(\mathbf{X}_i)\phi(\mathbf{X}_j) = \exp\left(-\frac{\left\|\mathbf{X}_i - \mathbf{X}_j\right\|^2}{2\sigma^2}\right). \quad (13)$$

A Bayesian probabilistic framework infers a distribution over the weights. According to Bayes rule, the posterior probability of $\mathbf{w}$ is

$$p(\mathbf{w}\,|\,\mathbf{y},\boldsymbol{\alpha}) = \frac{p(\mathbf{y}\,|\,\mathbf{w},\boldsymbol{\alpha})\,p(\mathbf{w}\,|\,\boldsymbol{\alpha})}{p(\mathbf{y}\,|\,\boldsymbol{\alpha})} \quad (14)$$

where $\mathbf{y} = (y_1, \cdots, y_N)^{\mathrm{T}}$, α_i represents the precision of the corresponding parameter w_i, and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_M)^{\mathrm{T}}$. $p(\mathbf{y}|\mathbf{w},\boldsymbol{\alpha})$ is the likelihood of the target values given the training dataset. The conditional prior probability distribution $p(\mathbf{w}|\boldsymbol{\alpha})$ in Equation (14) is modelled by a Gaussian function where the parameters w_i are weighted by parameters α_i.

$$p(\mathbf{w}\,|\,\boldsymbol{\alpha}) = \prod_{i=1}^{M} \frac{\sqrt{\alpha_i}}{\sqrt{2\pi}} \exp\left(-\frac{\alpha_i w_i^2}{2}\right) \quad (15)$$

Because $y \in \{0,1\}$ is a binary variable, the likelihood function can be described by a Bernoulli distribution:

$$p(\mathbf{y}\,|\,\mathbf{w},\boldsymbol{\alpha}) = \prod_{i=1}^{N} [\sigma(t(\mathbf{X}_i;\mathbf{w}))]^{y_i} [1-\sigma(t(\mathbf{X}_i;\mathbf{w}))]^{1-y_i} \quad (16)$$

where $\sigma(y)=1/(1+e^{-y})$ is the logistic sigmoid link function. Equation (14) with the probability densities given by (15) and (16) cannot be solved analytically. Therefore, a numerical method, the Laplacian approximation, is proposed to find the maximum a posterior (MAP) weights $\mathbf{w}^*$ based on the training dataset,

$$\begin{aligned} &\ln\{p(\mathbf{y}\,|\,\mathbf{w})\,p(\mathbf{w}\,|\,\boldsymbol{\alpha})\} \\ &= \sum_{i=1}^{N} [y_i \ln t_i + (1-y_i)\ln(1-t_i)] - \frac{1}{2}\mathbf{w}^T \mathbf{A}\mathbf{w} \end{aligned} \quad (17)$$

where $\mathbf{A} = diag(\alpha_i)$. By computing the maximum value of (17) with respect to $\boldsymbol{\alpha}$ and $\mathbf{y}$, the mean $\mathbf{w}^*$ and covariance $\boldsymbol{\Delta}$ of the Laplacian approximation are obtained:

$$\begin{aligned} \mathbf{w}^* &= \mathbf{B}\boldsymbol{\Delta}\boldsymbol{\Phi}^T\mathbf{y} \\ \boldsymbol{\Delta} &= (\boldsymbol{\Phi}^T\mathbf{B}\boldsymbol{\Phi} + \mathbf{A})^{-1} \end{aligned} \quad (18)$$

where $\mathbf{B} = diag(\beta_1, \beta_2, \ldots, \beta_N)$ is a diagonal matrix with $\beta_i = \sigma(y_i)[1-\sigma(y_i)]$. After obtaining $\mathbf{w}^*$, the parameters α are iteratively updated using

$$\alpha_i^{new} = \frac{1-\alpha_i^{old}\Delta_{ii}}{\mu_i^2} \quad (19)$$

where μ_i is the i-th posterior mean weight $\mathbf{w}_i^*$ and Δ_{ii} is the i-th diagonal element of the covariance. The procedure is repeated until it converges to a fixed value or the maximum number of iterations is reached.

In order to tackle multiclass situations using the RVM method, two possible strategies could be used [28]. The first one is the 'one-against-all' strategy. L binary classifiers will be created for an L-class classification and each classifier is trained to separate one class from the others. The second strategy is 'one-versus-one'. There are $L(L-1)/2$ binary classifiers created to separate every two classes. In this study, the first method is adopted as it is more computational efficient.

2.3. Testing

2.3.1. Datasets

Behavioural data was collected from several individuals and different vehicles using a Samsung Galaxy S4 smartphone. This comprised both human and land vehicle behaviours. About 30 minutes of data was collected for each behaviour. The behavioural motions were recorded using the 3-axis accelerometers, 3-axis gyroscopes, 3-axis magnetometers and barometer of the smartphone. In the data collection, a higher sampling rate provides more samples in each window but more processing is needed. By balancing the amount of data required per window and the power consumption, the accelerometers, gyroscopes and magnetometers were sampled at 100 Hz while the barometer was set at its maximum sampling rate, 6.25 Hz.

For the human activity dataset, eight participants, including both females and males of age range 23 to 35, were enrolled to collect daily human activities, comprising stationary (including standing, sitting still and placed on a table), walking, running, climbing stairs and descending stairs. During each data collection, the smartphone was placed in the front pockets of the trousers and no instructions were given about its orientation. All participants were asked to perform each activity as flexibly as usual without any restrictions.

For the vehicle motion datasets, data were collected separately on buses, underground trains and diesel trains. Data were collected in both dynamic and stationary (with the engine on) scenarios. During the collection, the smartphone was put on a seat within the vehicle, where noise conditions were typical.

2.3.2. Results

- **Comparisons with Different Algorithms**

To determine the most suitable algorithm for each classifier, a wide range of common supervised machine learning algorithms were compared. In addition to the DT and RVM described in Section 2.2.4, an artificial neural network (ANN), Bayesian network (BN), Naïve Bayes (NB) algorithm and support vector machine (SVM) were assessed. The ANN, BN, NB and SVM algorithms are described in [28] and their capabilities for sensing behavioural contexts are discussed in [29].

To carry out the evaluations for the comparison, a 6-fold cross-validation strategy was applied to train and test each of the three classifiers in the framework individually. Using this method, the database is randomly divided it into 6 equally sized folders. Each time, 5 folders are used as training sets while the remaining one is used as a test set. This procedure is repeated 6 times to ensure that all the samples are used equally in testing, while maintaining independence of training and testing data for model learning.

After each folder is tested, the algorithms are evaluated based on statistical metrics. Two commonly used measures are precision and recall: precision P is the number of results correctly attributed to the class divided by the total number attributed to that class, recall R is the number of results correctly attributed to the class divided by the number that truly belong to that class. In this research, the overall accuracy of the classification results is evaluated using F_1 score, the harmonic mean of precision and recall, defined in Equation (22).

$$P = \frac{T_P}{T_P + F_P} \tag{20}$$

$$R = \frac{T_P}{T_P + F_N} \tag{21}$$

$$F_1 = 2 \cdot \frac{P \cdot R}{P + R} \tag{22}$$

In the equations, T_p indicates the number of true positives or correctly classified results, F_N is the number of false negatives and F_p is the number of false positives.

Algorithm	Human-Vehicle	Human Activities	Vehicle Motions
ANN	97.4	96.4	88.2
BN	90.2	94.4	85.6
DT	98.9	91.4	87.6
NB	89.4	91.7	80.3
RVM	96.4	97.6	91.0
SVM	97.9	98.3	92.0

Table 4: F_1 score of different supervised machine learning algorithms for each classifier (%)

The performance of the different supervised machine learning techniques in three classification tasks is presented in Table 4. Note that each result listed in the table is the best one achieved using that algorithm by tuning the parameters. For the human-vehicle classifier, the decision tree shows better performances than the others, achieving an F_1 score of nearly 99%. As the decision tree is also simple structured and computational efficient for both training and testing, it is therefore selected for the human-vehicle classifier. The classification results of the human and vehicle classifiers suggest that RVM and SVM are both excellent candidates, with SVM performing slightly better than RVM. However, the outputs of RVM are probabilities, while those of the SVM are Boolean, so the RVM provides an indication of the uncertainty of the classification decision, which is useful for context-adaptive navigation. Therefore, the RVM is chosen for both the human activity and vehicle motion classifiers.

- **Performance of Behavioural Classifier**

To evaluate the performance of the overall recognition system with three classifiers working together, the whole dataset was divided into two parts: 200 samples of each category in the dataset were randomly selected as test samples; the others (about 700 samples for each category) were used as training samples. The behaviour recognition results of our approach are shown in the confusion matrix, presented 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 classification algorithms.

The results show that the system achieves an overall F_1 score of 95.1%, demonstrating that this approach can distinguish most of the behaviours. It can be observed from Table 5 that the misclassification rate between human activities and vehicle motions is less than 1% due to the hierarchical classification scheme. However, some categories are more difficult to detect. For example, many moving bus samples are misclassified as other vehicle motions due to the presence of similar road-induced and engine vibrations.

Actual	Predicted								
	S	**W**	**R**	**A**	**D**	**V**	**U**	**T**	**B**
S	**192**	0	0	0	0	8	0	0	0
W	0	**194**	0	3	3	0	0	0	0
R	0	0	**200**	0	0	0	0	0	0
A	0	3	0	**194**	3	0	0	0	0
D	0	3	0	5	**192**	0	0	0	0
V	7	0	0	0	0	**186**	4	2	1
U	0	0	0	0	0	5	**183**	7	5
T	0	0	0	0	0	5	2	**190**	3
B	0	0	0	0	0	10	2	8	**180**

Table 5: Confusion matrix

Note: S=stationary, W=walking, R=running, A=ascending stairs, D=descending stairs, V=stationary vehicles with the engine on, U=moving underground trains, T=moving trains, B=moving buses.

2.4. Behavioural Connectivity

One way of reducing incorrect behaviour determination is to consider the likelihood of behaviour connectivity. Connectivity describes the temporal relationship between the current behaviour category and the previous ones. If a direct transition between two categories can occur, they are connected; otherwise, they are not [1]. For example, stationary vehicle and pedestrian behaviour can be connected directly, whereas moving vehicle behaviour is not because a vehicle must normally stop to enable a person to get in or out.

Behavioural connectivity is represented in a probabilistic way. Comparing with Boolean results, there are two advantages. First, a Boolean implementation may occasionally result in the decisions being stuck on incorrect context categories following a faulty selection. This can occur when the correct context category is not directly connected to the incorrectly selected category and the other categories are poor matches to the measurement data. But probabilities are more flexible to increase the directly connected category and minimise the unlikely one. Second, a probabilistic scheme permits the transitions between context categories that are rare, but not impossible.

To illustrate the temporal relationships, the likelihoods of connections between behaviours are listed in Table 6, where the permitted direct connections are set to 0.9 and the unlikely connections are set to 0.1.

Prev / Now	H	V	U	T	B
H	0.9	0.9	0.1	0.1	0.1
V	0.9	0.9	0.9	0.9	0.9
U	0.1	0.9	0.9	0.1	0.1
T	0.1	0.9	0.1	0.9	0.1
B	0.1	0.9	0.1	0.1	0.9

Table 6: Behavioural connection matrix (**C**)
Note: H = human activities, including stationary, walking, running, ascending and descending stairs; V=stationary vehicles with the engine on; U=moving underground trains; T=moving diesel trains; B=moving buses.

As the behaviours between two concessive epochs are not independent, a straight smoothing method is first applied. As in Equation (23), the smoothed estimates are obtained by combining the normalised outputs from the classification algorithms at epoch *k* and the estimates at epoch *k*-1 using filter gain α.

$$\hat{\mathbf{x}}_k^- = \alpha \cdot \mathbf{z}_k + (1-\alpha) \cdot \hat{\mathbf{x}}_{k-1} \tag{23}$$

where $\hat{\mathbf{x}}_k^-$ and $\hat{\mathbf{x}}_{k-1}$ are, respectively, the estimates of behaviours at epoch *k* before connectivity updating and estimates at epoch *k*-1 and $\mathbf{z}_k$ is the detected probability of behaviours at epoch *k* across the detection algorithms. α=0.5 is used here, which indicates the measurements at epoch *k* and the estimates at epoch *k*-1 are weighted equally. Then the relationships between estimates are constructed based on a linear assumption by a transfer matrix $\mathbf{\Omega}_k$, as shown in Equation (24).

$$\hat{\mathbf{x}}_k^- = \mathbf{\Omega}_k \cdot \hat{\mathbf{x}}_{k-1} \tag{24}$$

The transfer matrix is a quantitative representation to describe the response of estimate at epoch *k* to the previous one. However, connectivity implies that some transitions are more likely than others, thus the transfer matrix should be re-estimated using the connectivity constraints, as shown in Equation (25).

$$\hat{\mathbf{x}}_k^+ = (\mathbf{\Omega}_k \circ \mathbf{C}) \cdot \hat{\mathbf{x}}_{k-1} \tag{25}$$

In Equation (25), notation ° denotes matrix element-wise multiplication, satisfying $(\mathbf{\Omega} \circ \boldsymbol{C})_{i,j} = \mathbf{\Omega}_{i,j}\mathbf{C}_{i,j}$. Note that in most practical cases, the dimensions of vector $\hat{\mathbf{x}}_k^-$ and $\hat{\mathbf{x}}_{k-1}$ are larger than one, thus Equation (24) becomes an underdetermined equation. To obtain the transfer matrix, the minimum (Euclidean) norm of the transfer matrix constraint is imposed as it is able to control the propagation to the perturbations in the estimates [30][31]. To calculate the matrix, a Moore-Penrose pseudoinverse [32] of vector $\hat{\mathbf{x}}_{k-1}$ is applied:

$$\mathbf{\Omega}_k = \hat{\mathbf{x}}_k^- \cdot (\hat{\mathbf{x}}_{k-1})^\dagger \tag{26}$$

In Equation (26), superscript † is the operator of pseudoinverse (right inverse in this case), which satisfies

$$\hat{\mathbf{x}}_{k-1} \cdot (\hat{\mathbf{x}}_{k-1})^\dagger = \mathbf{I}. \tag{27}$$

The final step is to re-scale the likelihood of each category to obtain a probability using

$$\hat{x}_{k,i} = \frac{\hat{x}_{k,i}^+}{\sum_I \hat{x}_{k,i}^+} \tag{28}$$

where $\hat{x}_{k,i}$ is the probability of behaviour *i* at epoch *k*.

To test the performance of the proposed connectivity method, a piece of continuous underground train data was collected on a London underground train (District line) for about 5 minutes, with the vehicle moving and stopping at the stations. It was processed and classified using the same method described in Section 2.2.

A comparison of context recognition results with and without connectivity is shown in Figure 4. Note that most of the misclassified samples are corrected to the right ones, showing that the connectivity constraint is able to reduce the number of incorrect context selections and improve the performance of behavioural detection. Comparing with the reference line, it can also be seen that there were one to two-second response delays after the behaviour changed.

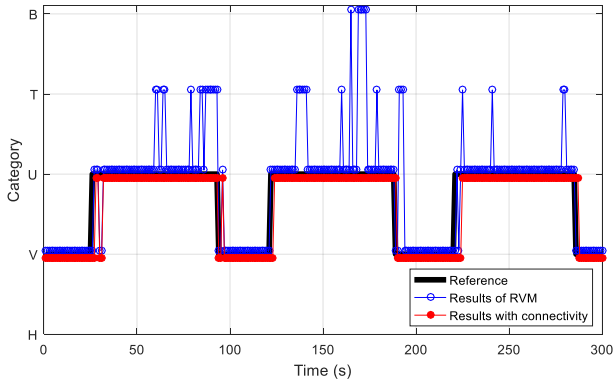


Figure 4: Results of behaviour connectivity

3. ENVIRONMENT DETECTION

Generally, the environmental context may be divided into several different broad classes: on land, on water, underwater, air and space [1]. A good environment categorisation for navigation can provide an indication of the positioning techniques applicable for a context adaptive navigation system. As the smartphone is used as the context detection device in this study, it is not applicable to be used for positioning purposes underwater, in the air or in space. In this paper, 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.

As the related research on environment detection is still in its infancy, locating whether the user is indoor or outdoor is considered as a prerequisite task of a navigation system because indoor and outdoor positioning depend on inherently different techniques. An effective indoor and outdoor detection algorithm can provide essential information for context adaptive navigation applications. For instance, if a stationary car with the engine on is detected outdoors, this may imply it is waiting at traffic lights while an indoor context may reveal it is located in a car park.

Different smartphone sensors whose outputs vary with 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][8] show that it is not sufficient for indoor and outdoor detection based on the numbers of access points and the strengths of signals. A GNSS module (including GPS and GLONASS) 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. 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 detection of indoor and outdoor environments with GNSS signals.

Figure 5: The portico of UCL's Wilkins building, an example of the intermediate category

In reality, the boundaries between indoor and outdoor environment can be ambiguous, rendering some scenarios hard to classify as either one. The portico of UCL's Wilkins building in Figure 5 is a typical example. This is covered by the roof of the building, but there is only one wall and the other three sides of this area are open. For a practical detection system, an uncertain decision is better than a wrong classification. Because an uncertain environment decision can be used in other ways (e.g. environment connectivity, environment and behaviour association) to improve the classification, but a wrong classification cannot. Similarly, it is better to inform a context-adaptive navigation system that the environment is uncertain than to provide it with an incorrect context. Therefore, to have a smooth transition between indoor and outdoor categories and reduce the likelihood of wrong classification, a new environment category of "intermediate" is introduced to serve as a bridge between the indoor and outdoor categories.

3.1. Overview of GNSS Signals

GNSS measurements were collected at 1 Hz from both GPS and GLONASS signals received by the smartphone. The data was collected at different locations of various indoor and outdoor environments, such as deep indoor, urban, outer indoor and open sky. About 200s of static data was collected at each site. Figure 6 presents histograms showing the normalised distributions of signal-to-noise ratio (SNR) measurements from four types of environment.

A number of trends may be identified from the histograms. A signal with a higher SNR is more likely to be LOS (Line-of-Sight) than NLOS (Non-Line-of-Sight). As expected, the average received SNR is lower in indoor environments than in deep urban and open sky environments, which is useful for environmental context detection. By comparing the GNSS SNR distributions, it can also be seen that the proportions of signals weaker than 25 dB-Hz vary between different environment types. Almost all the signals received in deep indoor environments are weaker than 25 dB-Hz while increasing proportions of signals stronger than 25 dB-Hz are observed for outer indoor, deep urban and open sky.

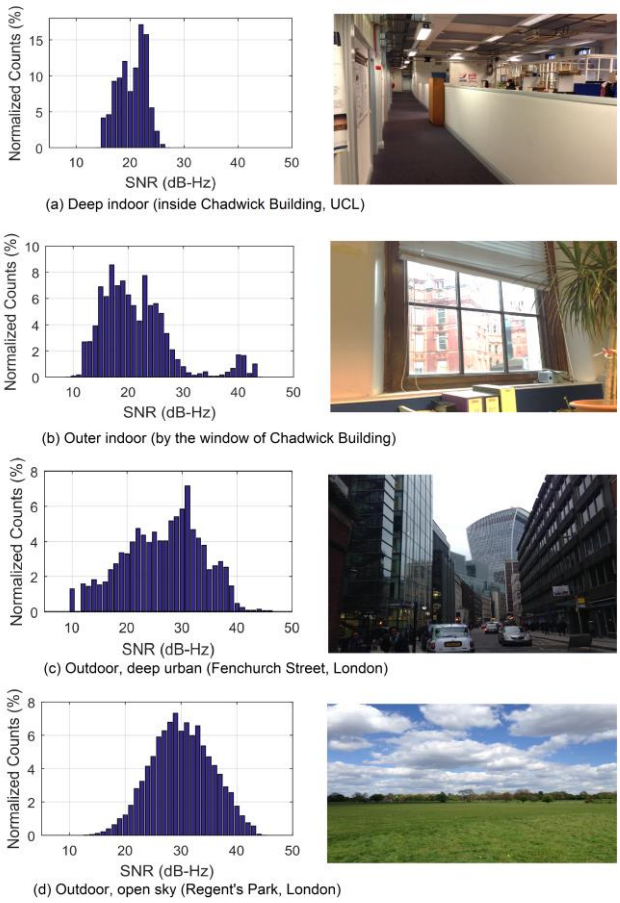


Figure 6: SNR measurement distributions under different environments

3.2. GNSS Based Features

The number of satellites received and the total measured SNR, summed across all the satellites received at each epoch, were also considered as features for the environmental classification algorithm. As normally the average number of satellites received indoor is no more than those received outdoor, the summed SNR is considered instead of the average value. These features are shown in Figure 7, based on the same set of data shown in Figure 6. Note that it takes time for a GNSS receiver to acquire satellite signals at the start of a test period, so there is an increase in satellite numbers during the first few seconds. It can be observed that open sky and deep indoor environments can be clearly distinguished from the others.

However, it is hard to distinguish "outer indoor by the window" and "deep urban" from each other based on these observations. Approximately the same number of satellites was received and there were only slight differences in the total SNR measurements. Therefore, neither feature is a reliable metric for indoor and outdoor classification.

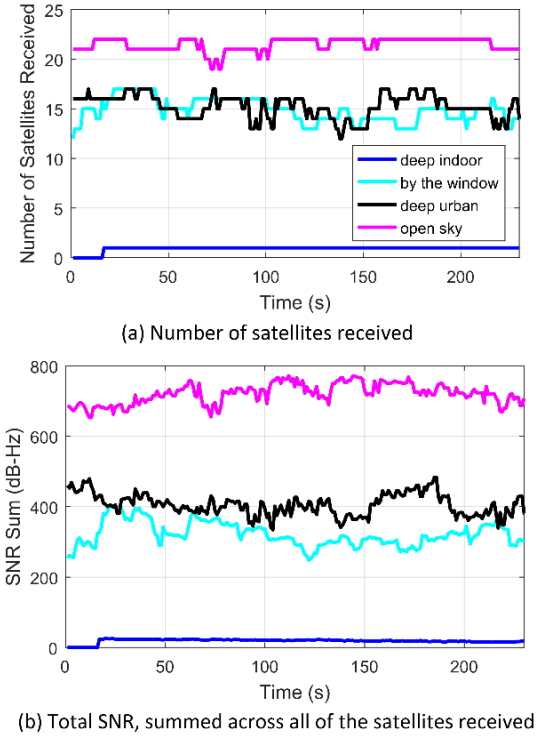


Figure 7: Features derived from satellite signals

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, $numSNR_{25}$ and $sumSNR_{25}$, are proposed, which are defined by

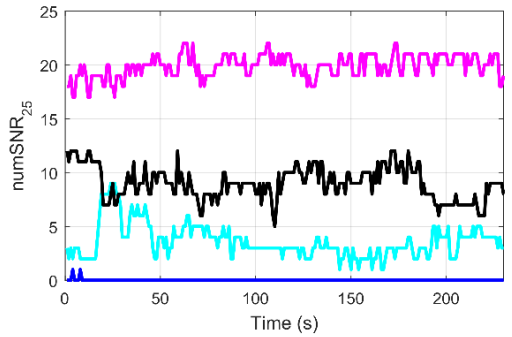
$$numSNR_{25} = \sum_i H(SNR_i), \tag{29}$$

$$sumSNR_{25} = \sum_i SNR_i \cdot H(SNR_i), \tag{30}$$

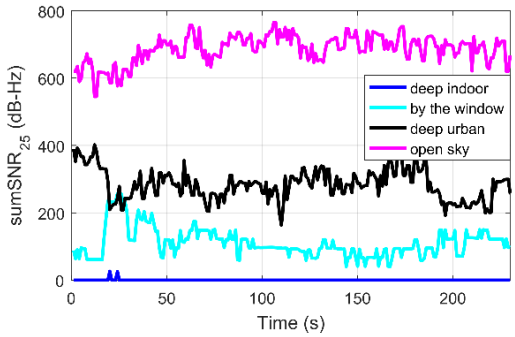
where SNR_i indicates the SNR value of the i-th satellite received at the current epoch and the function $H(.)$ is defined as:

$$H(x) = \begin{cases} 1, & \text{if } x \geqslant 25 \text{ dB-Hz} \\ 0, & \text{otherwise} \end{cases}. \tag{31}$$

Comparing the features plotted in Figure 8 with the ones in Figure 7, more satellites with signals above 25 dB-Hz are received in the deep urban environment than indoor by the window, leading to larger differences in the summed SNR features. Meanwhile, the deep indoor and open sky environments can still be clearly distinguished based on the observations. To verify the effectiveness of the features, they will be further tested using a dataset collected at different indoor and outdoor sites in Section 3.4. It is worth noting that $sumSNR_{25}$ 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 environment types need to be distinguished using more information, such as measurements from other sensors or the temporal relationship.



(a) Number of satellites with signals above 25 dB-Hz



(b) Total SNR, summed across the signals above 25 dB-Hz

Figure 8: Features based on signals above 25 dB-Hz

3.3. Hidden Markov Model (HMM)

The features $numSNR_{25}$ and $sumSNR_{25}$ can be computed sequentially from the outputs of a GNSS receiver module. A hidden Markov model is used in this study to determine the environmental context by integrating the observations over time. The HMM assumes a Markov process with the states that cannot be visible directly [28] (indoor, intermediate or outdoor environment in this study), so that it is capable of modelling the inherent dynamic temporal relationships of environments. In general, a HMM includes five elements as follows:

1) The state space **S** that consists of N hidden states $\mathbf{S}=\{S_1, S_2,\ldots, S_N\}$. In this research, there are only 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

$$\sum_{i=1}^{N} P(X_k = S_i) = 1 \qquad (32)$$

where X_k refers to the environmental context at that epoch.

2) The set of observations at each epoch k, $Z_k = \{z_{1,k}, z_{2,k},\ldots, z_{\ell,k},\ldots, z_{m,k}\}$, where $z_{\ell,k}$ is the ℓ-th observation at epoch k and m is the number of observations. In this study, $z_{1,k}$ refers to $numSNR_{25}$ while $z_{2,k}$ is $sumSNR_{25}$.

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 the state transits from a value S_i at the immediately prior epoch to another value S_j at the current epoch.

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

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

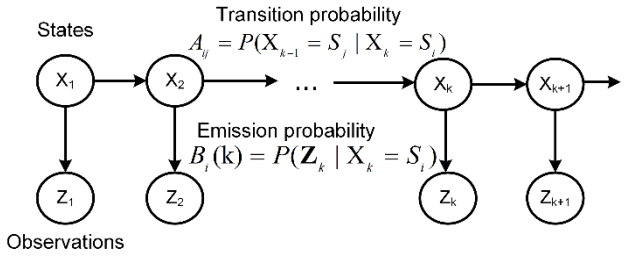


Figure 9: Structure of a first-order HMM

In this paper, we use the first-order HMM, which assumes the current environmental context is only affected by the immediate previous context. Figure 9**:** Structure of a first-order HMM is an illustration of a first-order hidden Markov model. Given the sequence of the observations, the most likely sequence of hidden states can be inferred using the Viterbi algorithm [28][33]. The probabilities of the model are determined as follows.

Transition probability. When a user was previously indoor, the current state is highly likely to be indoor and might be intermediate, but is not likely to be outdoor. Because the user rarely moves directly from indoor to fully outdoor. It is similar when a user is at door. However, when the user is at the intermediate state, he/she can move directly to either of the other states. Based on the above assumptions and tuning the parameters, the transition probabilities are listed in Table 7.

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

Table 7: Transition probabilities of HMM

Initial probability. As there is no prior information about the initial state, we have to make a judgement based on the available initial observations. On one hand, if the observations may enable either an indoor or outdoor environment to be determined with high confidence, the algorithm shall allow the initial state to follow the determination based on the initial observations. Thus, the initial probability of either indoor or outdoor shall not be set to zero. On the other hand, if the observations at the initial epoch cannot disclose the state confidently, the initial state is preferred to be uncertain. In this case, the initial state is presumed to be most likely the intermediate state, so that it offers an equal probability of transiting to

the indoor or outdoor states at the next epoch. Therefore, the initial probabilities are set as follows:

$$P(X_1 = S_1) = P(X_1 = S_3) = 0.25$$
$$P(X_1 = S_2) = 0.5 \quad (33)$$

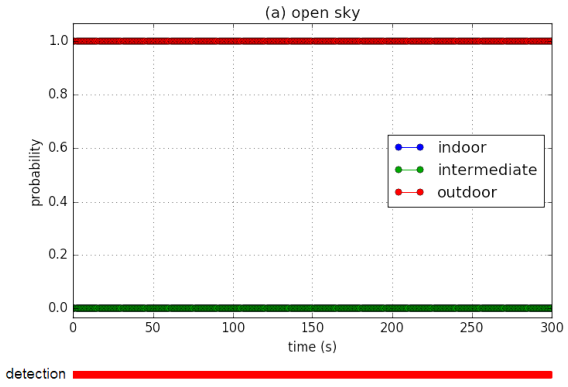
Emission probability. The emission probability describes the measurement likelihood of making an observation in different states. Our observations are modelled by Gaussian process, whose means and variances are fitted to the dataset collected at different indoor and outdoor sites, which covers different kinds of environment scenarios. Note that the Normal distribution in Equation (34) and (35) is denoted by N(μ, σ^2) with mean μ and variance σ^2.

$$P(z_{1,k} \mid X_k = S_1) \sim N(4,1.6)$$
$$P(z_{1,k} \mid X_k = S_2) \sim N(7.5,1.36) \quad (34)$$
$$P(z_{1,k} \mid X_k = S_3) \sim N(9,4)$$

$$P(z_{2,k} \mid X_k = S_1) \sim N(50,2500)$$
$$P(z_{2,k} \mid X_k = S_2) \sim N(150,625) \quad (35)$$
$$P(z_{2,k} \mid X_k = S_3) \sim N(350,5000)$$

3.4. Test Results

Four different kinds of environment types were chosen to test the detection ability of the proposed detection method under different GNSS reception conditions. The data for open sky (outdoor), deep urban (outdoor), outer indoor and deep indoor environments are as depicted in Figure 6.
Figure 10 presents the detection results of the static experiments in different environments. In the case of open sky and deep indoor, the detection results are very accurate as all samples of these scenarios are successfully detected with almost 100% probability. Deep urban is a little challenging for the detector as more signals are blocked or reflected by the tall buildings around. It can be observed from the figure that most samples are classified to outdoor correctly but with some intermediate states occasionally appearing among them. A similar thing happens for the outer indoor environment by a window. As some direct signals can still be received by the window, the measurements between 20s and 30s are erroneously classified as an outdoor environment.



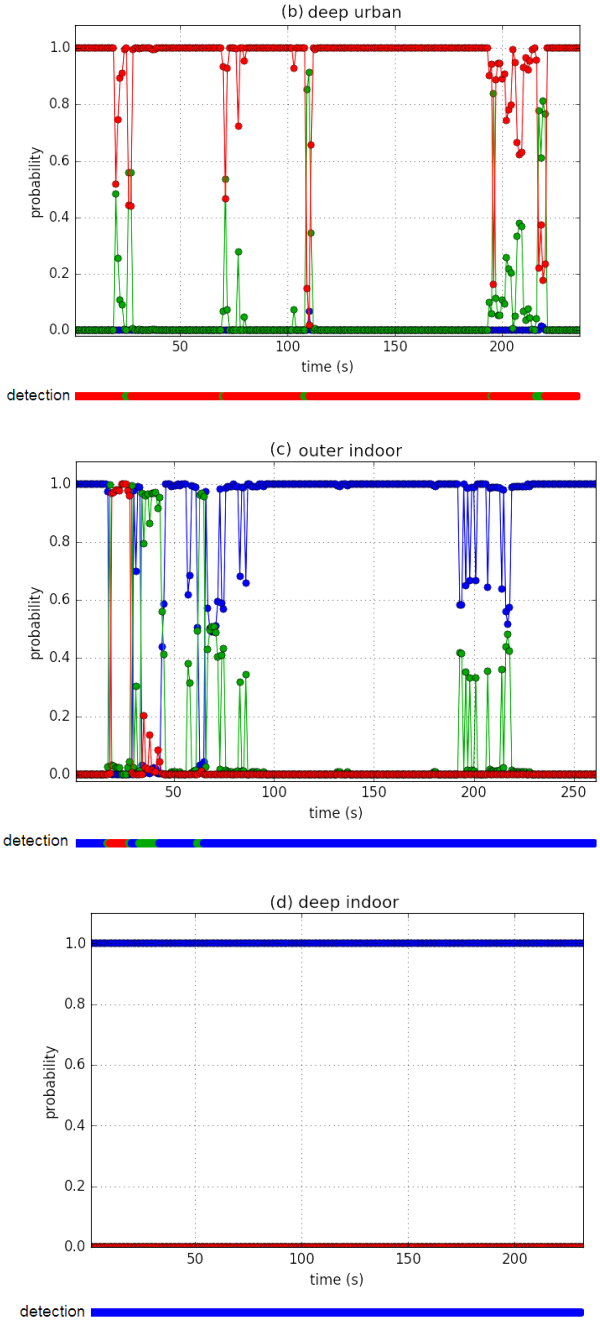


Figure 10: Static experiment results of the environment detection algorithm

4. ASSOCIATION OF ENVIRONMENTAL AND BEHAVIOURAL CONTEXT

Although behavioural and environmental context are treated separately in the proposed context detection framework, they are not completely independent in reality. Certain activities are associated with certain environments [1]. For example, land vehicles normally remain on the road instead of in the sky; boats or ships can be on land, but only exhibit some specific types of behaviours. As more contexts are included in the framework, this

information can be used to eliminate combinations of environment and behaviour that are not associated in practice, which reduce the chances of the context determination system selecting an incorrect context.

4.1. Mathematical Model

The update process of context association is expressed mathematically as follows. The likelihoods of the environmental and behavioural context categories are reweighted to take into consideration of association information. It is assumed that only associated combinations of vehicle motions and human activities are considered. The reweighted likelihoods of environment i and behaviour j, $\Lambda_{i,e}$ and $\Lambda_{j,b}$, respectively, are given by

$$\begin{aligned} \Lambda_{i,e} &= p_{old,i,e}\left(\sum_{J}(a_{i,j}\, p_{old,j,b})\right) \\ \Lambda_{j,b} &= p_{old,j,b}\left(\sum_{I}(a_{i,j}\, p_{old,i,e})\right) \end{aligned} \tag{36}$$

where $p_{old,i,e}$ and $p_{old,j,b}$ are the probability of environment i and behaviour j before association. $a_{i,j}$ is the association likelihood of environment category i and behaviour category j, given in Table 8. Note that the parameters in the table are based on daily experience and their reliability may be further enhanced by location-dependent association.

Envir. \ Behav.	H	V	U	T	B
Indoor	1.0	1.0	0.9	0.6	0.9
Intermediate	1.0	0.8	0.6	0.4	0.6
Outdoor	1.0	1.0	0.8	1.0	1.0

Table 8: Association factors (α)

Note: H=human activities, V=stationary vehicles with the engine on, U=moving underground trains, T=moving diesel trains, B=moving buses.

Finally, the probabilities of environments and behaviours are updated, respectively, using

$$\begin{aligned} p_{new,i,e} &= \frac{\Lambda_{i,e}}{\sum_{I}\Lambda_{i,e}} \\ p_{new,j,b} &= \frac{\Lambda_{j,b}}{\sum_{J}\Lambda_{j,b}} \end{aligned} \tag{37}$$

4.2. Experimental Results

To assess feasibility of the proposed method, a practical test was conducted on a bus. The bus was driven along South Colonnade Street in the Canary Wharf district of London and stopped at the bus station under the bridge for about 10 seconds, as shown in Figure 11. This route was designed to incorporate both indoor and outdoor environments, as well as moving and stationary vehicle motions.

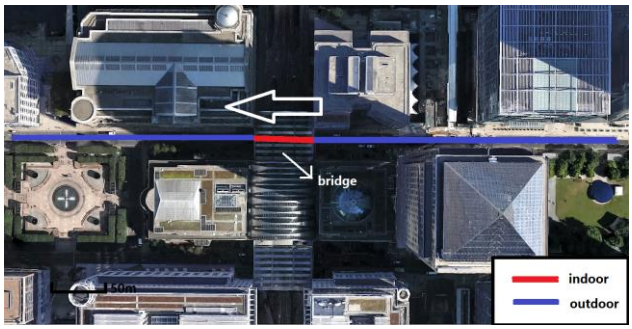


Figure 11: An aerial view of the experiment route (satellite image from Google Earth)

The results of behaviour detection are presented in Figure 12. Significant improvements can be observed that most misclassified samples are corrected by connectivity. As there is no degradation or further improvement of performances after association, only the results with connectivity and association are plotted in red in the figure. The corresponding environment detection probabilities before and after association are illustrated in Figure 13. Using association, some slight improvements of the results can be observed when the detected probabilities of the correct environment are not high. This shows that the association mechanism, to some extent, helps enhance the robustness of context determination.

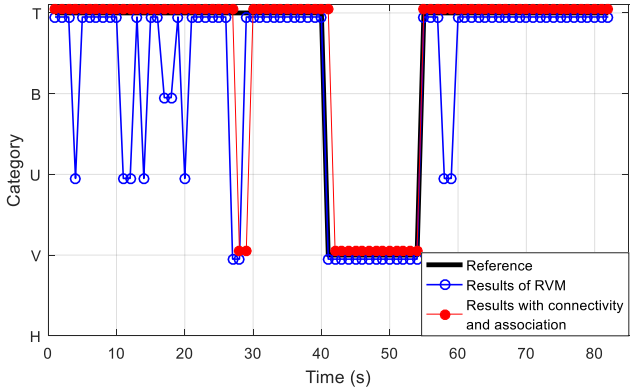
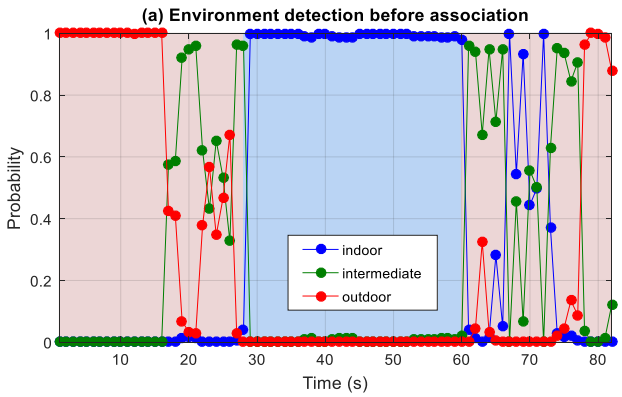


Figure 12: Performance of behavioural context detection



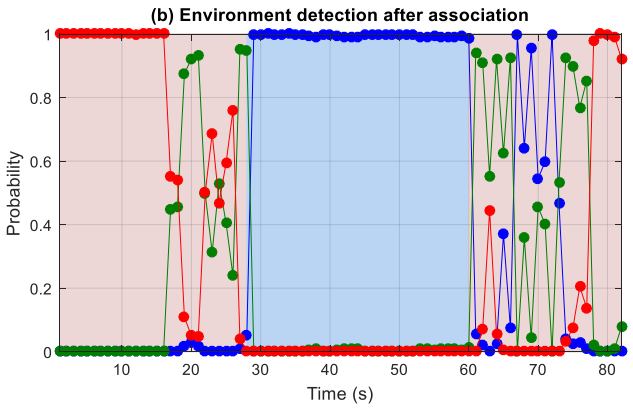


Figure 13: Comparisons of performances of environment detection before and after context association

5. CONCLUSION

This paper demonstrates the determination of both the operating environment and the behaviour of a host vehicle or human user using multiple sensors on a smartphone, building the foundation of a context-adaptive navigation system.

Detection of behavioural context using accelerometers, gyroscopes, magnetometers and the barometer has been presented. A hierarchical recognition scheme has been demonstrated, within which the behaviours are classified from broad classes to detailed types by decision trees and relevance vector machines respectively. The results have shown that the system achieves an overall F_1 score of 95.1%, with some vehicle motions easier to distinguish than others. It has also been shown that the performance can be further improved by considering behavioural connectivity.

Environmental context detection has focused on indoor and outdoor classification. The properties of GNSS signals under indoor and outdoor environments have been analysed and two features are extracted for classification. Then a detection scheme under a hidden Markov model has been applied to indoor and outdoor classification. It has been shown that this method can be used to distinguish indoor and outdoor environments under different GNSS reception conditions.

In addition, context association provides a way of linking environments to behaviours. Results of practical results have demonstrated that this can slightly improve the environment detection within the context determination process.

ACKNOWLEDGEMENT

This work is funded by the UCL Engineering Faculty Scholarship Scheme and the Chinese Scholarship Council. The authors would like to thank Dr. Mark Herbster and Dr. Simon Julier of UCL's Department of Computer Science for their useful comments and suggestions.

REFERENCES

[1] Groves, Paul D., et al. "Context detection, categorization and connectivity for advanced adaptive integrated navigation." *ION GNSS+ 2013*, (2013): 1039-1056.

[2] Ching, William, et al. "Uniwide WiFi based positioning system." *IEEE international symposium on technology and Society (ISTAS)*. 2010.

[3] Bell, Scott, Wook Rak Jung, and Vishwa Krishnakumar. "WiFi-based enhanced positioning systems: accuracy through mapping, calibration, and classification." *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*. ACM, 2010.

[4] Groves, Paul D. *Principles of GNSS, inertial, and multisensor integrated navigation systems*. Second Edition, Artech house, 2013.

[5] Groves, Paul D. "Shadow matching: A new GNSS positioning technique for urban canyons." *Journal of Navigation* 64.03 (2011): 417-430.

[6] Wang, Lei, Paul D. Groves, and Marek K. Ziebart. "Smartphone shadow matching for better cross-street GNSS positioning in urban environments." *Journal of Navigation* 68.03 (2015): 411-433.

[7] Pratama, Azkario Rizky, and Risanuri Hidayat. "Smartphone-based pedestrian dead reckoning as an indoor positioning system." *System Engineering and Technology (ICSET), 2012 International Conference on*. IEEE, 2012.

[8] Groves, Paul D., et al. "The four key challenges of advanced multisensor navigation and positioning." *Position, Location and Navigation Symposium-PLANS 2014, 2014 IEEE/ION*. IEEE, 2014.

[9] Groves, Paul D., and Ziyi Jiang. "Height aiding, C/N 0 weighting and consistency checking for GNSS NLOS and multipath mitigation in urban areas." *Journal of Navigation* 66.05 (2013): 653-669.

[10] Chen, Ruizhi, et al. "Inferring Human Activity in Mobile Devices by Computing Multiple Contexts." *Sensors* 15.9 (2015): 21219-21238.

[11] Lin, Tao, Cillian O'Driscoll, and Gérard Lachapelle. "Development of a context-aware vector-based high-sensitivity GNSS software receiver." *Proc. ION ITM* (2011): 1043-1055.

[12] Pei, Ling, et al. "Using LS-SVM based motion recognition for smartphone indoor wireless positioning." *Sensors* 12.5 (2012): 6155-6175.

[13] Elhoushi, Mostafa, et al. "Motion Mode Recognition for Indoor Pedestrian Navigation Using Portable Devices." *Instrumentation and Measurement, IEEE Transactions on* 65.1 (2016): 208-221.

[14] Bricon-Souf, Nathalie, and Conrad R. Newman. "Context awareness in health care: A review." *International journal of medical informatics* 76.1 (2007): 2-12.

[15] Schwinger, W., et al. "Context-awareness in mobile tourism guides–A comprehensive survey." *Rapport Technique. Johannes Kepler University Linz*(2005).

[16] Parviainen, Jussi, et al. "Adaptive activity and environment recognition for mobile phones." *Sensors* 14.11 (2014): 20753-20778.

[17] Pei, Ling, et al. "Human behavior cognition using smartphone sensors." *Sensors* 13.2 (2013): 1402-1424.

[18] Liu, Jingbin, et al. "Reciprocal Estimation of Pedestrian Location and Motion State toward a Smartphone Geo-Context Computing Solution." *Micromachines* 6.6 (2015): 699-717.

[19] Chen, Ruizhi, et al. "Inferring Human Activity in Mobile Devices by Computing Multiple Contexts." *Sensors* 15.9 (2015): 21219-21238.

[20] Shoaib, Muhammad, et al. "Fusion of smartphone motion sensors for physical activity recognition." *Sensors* 14.6 (2014): 10146-10176.

[21] del Rosario, Michael B., Stephen J. Redmond, and Nigel H. Lovell. "Tracking the evolution of smartphone sensing for monitoring human movement." *Sensors* 15.8 (2015): 18901-18933.

[22] Bao, Ling, and Stephen S. Intille. "Activity recognition from user-annotated acceleration data." *Pervasive computing*. Springer Berlin Heidelberg, 2004. 1-17.

[23] Yang, Jhun-Ying, Jeen-Shing Wang, and Yen-Ping Chen. "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers." *Pattern recognition letters* 29.16 (2008): 2213-2220.

[24] Sun, Lin, et al. "Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations." *Ubiquitous intelligence and computing*. Springer Berlin Heidelberg, 2010. 548-562.

[25] Saeedi, Sara, Adel Moussa, and Naser El-Sheimy. "Context-aware personal navigation using embedded sensor fusion in smartphones." *Sensors* 14.4 (2014): 5742-5767.

[26] He, Weihua, et al. "Recognition of human activities with wearable sensors." *EURASIP Journal on Advances in Signal Processing* 2012.1 (2012): 1-13.

[27] Duda, Richard O., Peter E. Hart, and David G. Stork. *Pattern classification*. John Wiley & Sons, 2012.

[28] Bishop, Christopher M., *Pattern recognition and machine learning*. Company New York, 2006.

[29] Guinness, Robert E. "Beyond where to how: A machine learning approach for sensing mobility contexts using smartphone sensors." *Sensors* 15.5 (2015): 9962-9985.

[30] Hansen, Per Christian. "Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems." *Numerical algorithms* 6.1 (1994): 1-35.

[31] Rump, Siegfried M. "Verified bounds for least squares problems and underdetermined linear systems." *SIAM Journal on Matrix Analysis and Applications* 33.1 (2012): 130-148.

[32] Ben-Israel, Adi, and Thomas NE Greville. *Generalized inverses: theory and applications*. Vol. 15. Springer Science & Business Media, 2003.

[33] Viterbi, Andrew J. "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm." *Information Theory, IEEE Transactions on* 13.2 (1967): 260-269.