HABITS: A Bayesian Filter Approach to Indoor Tracking and Location

Eoghan Furey, Kevin Curran, Paul Mc Kevitt

Faculty of Computing & Engineering,
University of Ulster, Derry, Northern Ireland
Email: furey-e1@email.ulster.ac.uk

Abstract. Knowledge of the location of people and things has always been a valuable commodity. The explosion of new devices and techniques has brought people and item tracking out of the experimental stage and into the wider world. Using Wi-Fi signals is an attractive and reasonably affordable option to deal with the currently unsolved problem of widespread tracking in an indoor environment. HABITS (History Aware Based Indoor Tracking System) models human movement patterns by applying a discrete Bayesian filter to predict the areas that will, or will not, be visited in the future. We outline here the operation of the HABITS Real-Time Location System (RTLS) and discuss the implementation in relation to indoor Wi-Fi tracking with a large wireless network. Testing of HABITS shows that it gives comparable levels of accuracy to those achieved by doubling the number of access points. These probabilistic predictions may be used as an additional input into building automation systems for intelligent control of heating and lighting. It is twice as accurate as existing systems in dealing with signal black spots and it can predict the final destination of a person within the test environment almost 80% of the time.

1. Introduction

To perform accurate indoor tracking of people and things, some sort of frame of reference is normally required and a number of waypoints need to be established. In satellite positioning the satellites themselves provide these waypoints, their position relative to each other and relative to the ground is known, therefore then location of an unknown device can be calculated relative to these. The same principle is used in the majority of positioning systems. Problems arise when these reference points are either too few in number, or those that are there do not have a clear line of sight to the object being tracked. One of the reasons for the success of satellite positioning is that generally there is nothing of substance between the satellites (waypoints) and the device to be tracked. Consider the example of a sailor who was trying to use a lighthouse as a waypoint for navigation but the lighthouse cannot be seen due to an obstacle such as a mountain being in the way. In that case the lighthouse is unusable as a reference point. With indoor positioning a similar

 ISSN:2041-6407

problem exists. If the exact distance, angle or signal strength to a point cannot be established then localisation will be difficult or impossible. Occlusion, attenuation, reflection and refraction are the cause of many errors in Real-Time Location Systems (RTLS). Often these problems make it extremely difficult or impossible to accurately establish location. For these reasons, intelligent techniques and ‘tricks’ are utilized in an attempt to improve performance of the RTLS. In some cases these have greatly improved accuracy and yield (the ability to get position fixes in all environments) but the improvements are accompanied by drawbacks in terms of time and cost. While no solution works perfectly in all environments, within reason, almost anything can be tracked to any desired resolution if enough resources are available. These resources can be quantified in terms of financial cost and vastly reduce the scalability of the RTLS. Innovative methods are required to improve accuracy levels and to allow positioning to be achieved for a reasonable cost in terms of time and infrastructure. Ekahau RTLS is an example of a commercial off the shelf (Cots) piece of software which was developed by a Finnish company for WiFi localisation in an indoor environment. For Ekahau to work, an existing 802.11 wireless network must be in place in the test area. Ekahau contains a number of components. The Ekahau Positioning Engine (EPE acts as a server controlling all location updates. It needs to be located on a server which has access to the existing WLAN.

If the knowledge of the location of a person or thing is a valuable commodity, then knowledge of the future location is doubly so. Information relating to where a person will be at some time in the future enables actions to be taken. The application of such knowledge could be far reaching. Fields were this is applicable range from logistical and social to environmental and entertainment. 100% accuracy in prediction is unattainable, however, scientific methods have enabled educated guesses to be made with a high degree of accuracy given certain conditions. As humans we make future location predictions all the time. Consider the following scenario. If a parent knows that their child returns from school every day at approximately the same time, then they can make a prediction that tomorrow the child will arrive home at around the same time, all else being equal. Knowing that this is likely to happen, the parent can take action armed with this information. They can cook dinner, starting it at an earlier time, so that it is ready at approximately the same time as the child returns from school. The parent has no proof that the child will be home at that time, but has enough evidence to make a prediction that this will be the case. Having made the prediction, action is taken by cooking the dinner. If the prediction is correct, then the parent is rewarded by time saved. While this is not the optimal situation, the consequences of an incorrect prediction are not severe. A prediction in this case, if correct had a positive outcome and if incorrect had no or only insignificant consequences.

Predictions of future movement indoors is even more sparsely researched. The Smart Doorplate project and the Augsburg location trials [1] are among the leading attempts to tackle this challenging area. Other research in this area have made use of different Machine Learning techniques. Neural Networks (NN) [2] and Hidden Markov Models (HMM) [3] have been applied with some success to compute where people will go next. In the field of robotics, artificial intelligence techniques are applied to make predictions of next location, generally to try and improve upon the accuracy of existing estimates. Widely used techniques such as the Kalman and Particle filters are probabilistic approaches to taking educated guesses of the future given relevant information. This paper seeks to outline a system called HABITS (History Aware Based Indoor Tracking System) which aims at overcoming weaknesses in existing RTLS by using the human approach of making educated guesses about future location. The hypothesis of this proposal is that knowledge of a person's historical movement habits enables future location predictions to be made in the short, medium and long term. The research questions that are foremost are whether the tracking capabilities of existing RTLS can be improved automatically by knowledge of historical movement and by the application of a combination of artificial intelligence approaches. The rest of the paper is structured as follows. Section 2 provides an overview of intelligent approaches to predicting future location. Section 3 provides an overview of the HABITS system approach to modelling prediction. Section 4 provides an example of tracking users while section 5 concludes.

2. Methods for Predicting Location

To accurately position an object in addition to the techniques and technologies previously discussed, intelligent prediction is required. These methods enable the accuracy levels of the estimates to be increased. When a human makes estimation about where a person or object will be located in the future, they automatically perform the complete calculation. To enable computers to replicate these calculations and to allow them to work with a number of different objects requires a number of artificial intelligence techniques. Technologies such as GPS and mobile phones have made this a hot area of research in recent years. This has been driven by commercial location based software, from Satellite Navigation devices to targeted advertising on mobile phones. Application to indoor environments is a largely under researched area.

Outdoors, using GPS traces to try and learn next location has been attempted by Han [4] for someone on foot and in [5] for vehicles on a road. More recently data gathered from mobile phone records has been mined to try and find patterns of movement which could be used to try and make next location predictions [6]. Indoors, this is a largely under researched area, however a number of ‘smart environments’ have been set up such as the work [1]. Here specific sensors on doors were utilised to provide movement patterns. A Hidden Markov Model (HMM) and a Neutral Network (NN) were applied to the data and successful predictions were made. Ashbrook and Starner [7], used a markov chain model and K-means clustering algorithm to attempt to predict future movement. They clustered GPS data to find significant locations and then built a first and second order markov models using location as state to try and predict future movement. It is possible to create n^{th} order Markov model where probability of the next state is dependent not only on the current state but on the previous n-1 states. For some examples, considering the 2^{nd} order can yield more accurate results as in the case of probability of transition from A → B is 70% but the probability of transitioning from B → A → B is 81%. This could be explained by a situation where A was a Shop and B was Home. If the shop was on the main road from Home then the probability of going from A to B (Shop to Home) is 70%. However, if the journey started at home and went to the shop, return to home could be more probable (perhaps getting something for dinner?). This demonstrates a situation when higher order models are useful and give extra information. It raises the question of which order of model is suitable for prediction. Ashbrook and Starner [7] conclude that this depends on the quantity of data available. Other factors affecting their probabilities were due to the large distances travelled and the fact that their tests took place outdoors. They also found that changes in routine would take a long time to show up in their model and they suggested a possible method of weighting certain updates, but warned that this could lead to model that was somewhat skewed. Han [4] attempts to build upon the work of [7] by using a Self Organising Map (SOM) as a means of learning without pre-knowledge. To use a supervised learning method to learn patterns of movements, pre-knowledge of the person is required, however a SOM can overcome this. An SOM is an "unsupervised learning neural network" which can preserve the topology of a map as it creates it. Sang uses an SOM to convert sequences of raw GPS data into meaningful patterns which are in turn applied to a markov chain approach. They used the output from the SOM to learn a first order markov model and to try and make predictions of next location from it. Their data was gathered based on a university campus. While their method looks promising, their results are very sparse and their conclusion of 'acceptable' prediction accuracy is of little value.

2.1 Indoor Predictive Tracking

In indoor localisation, the area of movement prediction is sparsely researched. This is due to the fact that any sort of indoor localisation is a relatively recent phenomenon, however a number of research studies have been conducted in this area. One of the first research projects that considered future movement was Microsoft Research's RADAR project [8]. This was the first significant attempts to track indoors using 802.1 Wi-Fi signals. Due to the severe problem of signal attenuation it was difficult to get an accurate fix on position using Received Signal Strength (RSS) measurements alone. Position was occasionally reported in locations that were not possible or at least highly unlikely. An effort to overcome these problems is described in Bahl and Padmanabhan [9] paper. They concluded that the next location position should be close to the last reported one. Their Viterbi-like tracking algorithm deals with a situation of when two physically separate locations are close together in signal space (due to aliasing). The shortest path is depicted in bold. The likely trajectory is calculated based on the previous unambiguous location and a guess of somewhere in between the two is given. Between vertices i and j there is an edge d_{ij} whose weight is calculated based on the Euclidian distance between the locations i and j. This approach has been shown to significantly reduce the accuracy error in locating a user who is walking. They tested the Viterbi-like approach against an NNSS (Nearest Neighbour in Signal Space) and an NNSS-AVG (where the three nearest neighbours in signal space were averaged to estimate location) and it was found to significantly outperform the others. Median distance error for NNSS (3.59 m) and NNSS-AVG (3.32 m) are 51% and 40% worse, respectively compared with Viterbi (2.37 m)" [9].

Anticipating or predicting a future situation has been attempted through the use of a number of learning techniques. Hidden Markov Models (HMM) are a popular technique which has been successfully applied in numerous different fields. The application of HMM to speech recognition has been examined by Rabiner [10]. In speech recognition predicting the next possible words can greatly increase accuracy. Rabiner examined HMM from their simplest form (discrete markov chain) to more sophisticated approaches such as continuous density models and those of variable duration. These techniques have been in widespread use for many years in speech recognition software. Computational biology is another field that has seen widespread application of predictive machine learning. Medical diagnosis, treatment and approaches to drug design all require techniques that can predict sequences. The use of HMM for gene prediction in sequences of DNA has been reviewed by Birney [11]. A new method for predicting the secondary structure of RNA using HMM was proposed by Yoon and Vaidynathan [12]. They demonstrated very accurate, secondary structure prediction using their proposed model with a low computational cost.

Gellert and Vintan [3] have analysed the data available from the Augsburg dataset to try and predict next room using HMMs. In their results the HMM outperformed the other techniques applied, namely NN and markov predictor. They obtained an average accuracy of 84.81% using 4 state confidence automata based HMM. This result was based on movement prediction from every room except, notably, the test subjects own room. As presumably a significant portion of journeys originated at their own room, this would not give a complete picture of the accuracy of the HMM. A study by Petzold et al., [13] converted algorithms normally used as "branch prediction techniques for current high performance microprocessors" to handle next context prediction of a person. These were applied to previously gathered behaviour patterns. The predictors were stimulated with patterns of behaviour of people walking indoors as the workload. They found that while these predictors worked well they were not consistent in their ability to handle complex patterns or in the training and retraining speeds. In another study by Petzold et al., [1] he evaluated these 'predictors by movement sequence' with real people in an office building [Augsburg trials]. Values of 59% and 98% accuracy in next location prediction were achieved without pre-training and with pre-training respectively. This extremely high accuracy achieved with pre-training can be accounted for by the fact that journeys from the test subjects' own offices were again not considered. A study by Mozer [14] proposed an Adaptive Control of Home Environments (ACHE) project. ACHE attempts to predict the next actions taken by the inhabitants by observing their actions taken, monitoring the environment and attempting to learn to anticipate their needs. The next action is predicted by means of a feed forward NN. They used these predictions to try and control energy use in a prototype house.

3. HABITS Modelling

 ISSN:2041-6407

Past movement habits have been shown to be repeated by humans, usually to do necessary tasks or just to take what is felt to be the path of least resistance. These habits are often linked to particular tasks that need to be done regularly. Movement habits are the same as other types of habits in that they tend to be regularly repeated. While each of us has a number of habits or patterns that appear to be unique to us, much more probable is that we share habits with others. In our approach, HABITS (History Aware Based Indoor Tracking System) is use to enhance an existing tracking system from Ekahau. The technology of the underlying tracking system or the positioning methods used are not relevant. HABITS is designed to be generic with application to many potential domains. The three main components of HABITS are a connected graph, a discrete Bayesian filer and a set of logic rules. The focus of HABITS is to combine these three methods in a novel way, allowing for predictions of human movement habits. These predictions overcome the latency of updates from currently available systems and enable them to make predictions of likely future movement. The underlying principle of our approach involves representing the movement areas as a graph which in turn is represented by a number of matrices; incidence, distance and transition. These constraints show where it is possible for a user to go and where not, the distance between points of interest (for our purposes) and eventually represent the probability of going from one area to another. Methods of modelling the travel environment exist and of these, a graph structure closely represents the possible paths. The nodes in the graph can be positioned to represent areas of interest, decision points or places where the user stops. In between these locations are the paths that may be travelled. The paths are edges and those locations of interest are the nodes/vertices of a connected graph. The graph structure clearly represents the connections between nodes and therefore areas in the real building. It shows which locations are connected either directly or indirectly.

When studying a building plan or road map this information is normally clear to see however, in a new location, different methods need to be used to identify these areas of interest. Areas where a user stops for some reason may be thought of as '*base nodes*'. Stopping for reasons such as sleep, eating, call of nature or work are some of the main reasons why humans would habitually stop at the same location. While for many people these may be in the same room or adjacent rooms, in the developed world, relatively large houses exist and these functions often occur in a number of different rooms with travel paths between. Examples of these rooms could be bedroom, kitchen, bathroom and living room. Movement between these rooms is often only possible by one or two different routes. The layout of a typical house (in the developed world) may be represented as a connected graph. In Figure 1 the green nodes represent stopping locations and the blue nodes represent decision points. A connected graph or topological map of these nodes is shown in Figure 2. Learning the locations of these points can be performed automatically in a number of ways, all of which require an underlying tracking system to be installed.

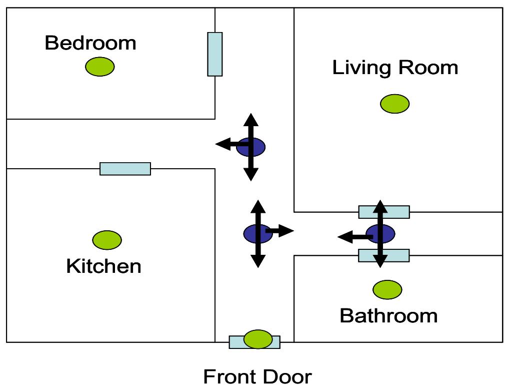

Figure 1: Node positions in house

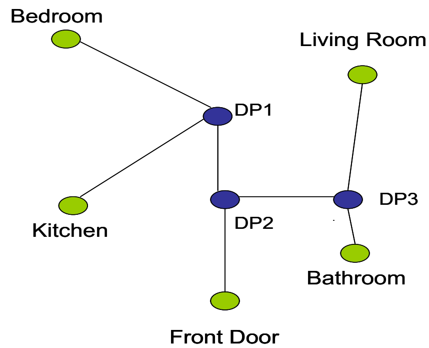

Figure 2: Connected graph with node connections

Learning these significant locations can be carried out automatically by computers. One methods of achieving this is to plot the locations where there was a significant delay between movements. These would indicate the areas where a person was stationary. Even within the same room these points are not all likely to be in the exact same location. To extract wait nodes from a large number of estimates, clustering techniques are used to group the updates together, revealing the main stopping locations. When the nodes have been discovered and coded with numbers for names (Figure 3) they may be represented as an n x n adjacency matrix where n is the number of nodes and the matrix details specific information about the graph. Figure 4 shows the adjacency matrix corresponding to the connected graph which in turn corresponds to the node positions in the sample house (Figure 2). If a connection exists between the nodes then in the matrix location ij which represents the connection from i to j place a 1, if no connection exists then place a zero. This enables the paths between nodes to be represented mathematically and the matrix can easily be processed by a computer program.

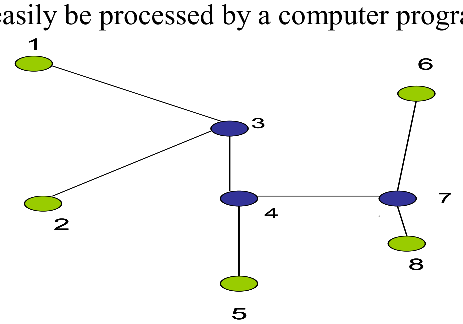

Figure 3: Node names replaced with numbers

	1	2	3	4	5	6	7	8
1	0	0	1	0	0	0	0	0
2	0	0	1	0	0	0	0	0
3	1	1	0	1	0	0	0	0
4	0	0	1	0	1	0	1	0
5	0	0	0	1	0	0	0	0
6	0	0	0	0	0	0	1	0
7	0	0	0	1	0	1	0	1
8	0	0	0	0	0	0	1	0

Figure 4: Adjacency matrix for nodes in sample house

When the node locations have been discovered and the distance between two nodes is known, travel time between nodes may also be calculated automatically by the underlying tracking system Ekahau. Average walking or travelling speed for each user is estimated by using: speed = distance /time. Knowledge of the relative travel times between nodes is then used to generate a distance matrix with distances between each node being calculated based on average user speed. The distance matrix values are in the same positions in the matrix as the 1's are in the adjacency matrix. A transition matrix showing the probabilities of travelling from one node to the other is built up by monitoring the person's travel through the nodes. Again, a number of methods exist to solve this, but a straight forward method is to use the sequence of all nodes traversed for a day, a week or all travel time (depending on the application). String identification tools can be used giving the sequences of nodes and from this mathematical functions can generate a transition matrix. As before, the size of the matrix corresponds to n x n and at each location (node) a count is kept of the movement through it and where it goes to next. In the sample house scenario, consider movement from the kitchen, through a decision point to either the bedroom or another decision point. Hypothetically, it could be found that the probability of going from the kitchen to the bedroom was 12/50. This would equate to a situation where out of 50 times leaving the kitchen, 12 of these journeys were to the bedroom. 12/50 would give a probability of 0.2 of travelling to the bedroom meaning that 0.8 or 38 journeys went the other way to the next decision point. This is how transition matrices are created and knowledge of them gives a first order Markov chain.

At any time along the chain, only the current location gives the probability of going to the next location. A simple markov chain like this gives some idea of the next node but alone it would not be enough to model real human movement habits. Raising the order of the model to consider the previous two nodes would help in some locations but [5] proved this needs to be done with a large dataset which takes a considerable time to generate. Maintaining a separate transition matrix for each day and/or each time period would improve the accuracy slightly but the system would not be expandable to a large area due to becoming overly complex. To predict the most likely next location with a useful degree of accuracy requires more than just a simple one state markov chain. The movement habits of people are dependent on a variety of factors and to improve the accuracy of any model requires that more of these factors are considered. A discrete Bayesian filter had been shown to work well for data fusion [15]. The underlying Ekahau tracking system gives the initial location, $bel(x_{t-1})$. The transition matrix provides the belief, $bel(x_t)$ when combined with the information in the Perceptual Model and the System Dynamics. This outputs the probability of moving to the next node when given just the previous one and no other information.

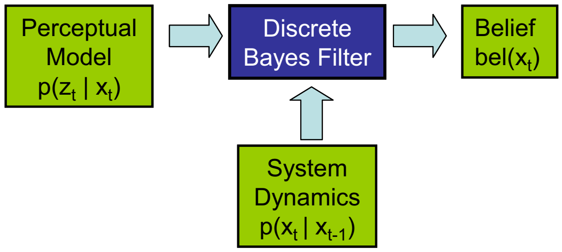

Figure 5: Bayes Filter Components

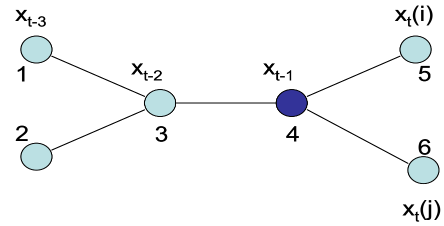

Figure 6: HABITS graphical operation scenario

HABITS uses more information than just that provided by the first order Markov chain. As a Bayesian filter only works for instances that hold to the markov assumption (meaning only a single order model), substantial information is being left out about commonly travelled paths or sequences of nodes. Froehlich and Krumm [5] found that the more nodes they had information about (previously travelled), the higher the chances of predicting their final location. If an order (3 for example) Markov model was used, then for some paths, the predicted location probability would be much higher, however it would also take into account shorter journeys and could have sequences like 2-4-2 which would include changing direction completely. Taking into account higher order models results in overcomplicated calculations. The notion of '*preferred paths' (PP),* however allows for the same information to be gathered without keeping track of every path.

As part of the definition of a habit, it states that they are routines of behaviour that are repeated regularly. An approach to viewing habits could be that they take places between distinct locations, but it does not mean that those locations are necessarily adjacent locations. The paths may go through a number of intermediate nodes and a common journey could be kitchen to toilet in the example in Figure 2. This would involve travelling through 4 different nodes but may be repeated a number of times a day. If a pattern occurred more often than a set number of times then it could be considered habitual. Habitual journeys of this sort we call 'preferred paths' and they can be mined from the string of all nodes visited. There could also be a temporal link between taking these preferred paths and a certain time period. This information can be used to adjust the output of HABITS prediction. It can also help with the identification of final destination which is another aim of HABITS. A preferred path is also stored as a vector and may be temporally linked to a specific time period if required. Some would be more frequently travelled at particular times than others. When on a preferred path, the information is used to increase the accuracy of the future location estimate. If we assume it is known that Node 1 was the node visited at time, t_{-3}. This would now give a sequence of nodes 1-3-4 leading up to the decision point. The preferred path vector for that particular sequence would be the probability of going to node 5 or 6 from that point. We assume that preferred paths only consider movement to new nodes and do not consider backward movement. We now have a vector showing:

$$PP\_Node4 = \begin{bmatrix} 0.66 \\ 0.33 \end{bmatrix}$$

This tells us that when the sequence of nodes visited was 1-3-4, the likelihood of being on the preferred path 1-3-4-5 is 0.66 and the likelihood of being on the preferred path 1-3-4-6 is 0.33. The method used to combine these two probabilities multiplies them together and adds the results to the initial belief from the Bayesian filter.

Table 1: Addition of probabilities from Bayes filter and preferred paths

Next node	Initial belief	Preferred path belief	Product of two beliefs	New belief
5	**0.45**	**0.66**	**0.297**	**0.747**
6	**0.35**	**0.33**	**0.116**	**0.466**

The new belief gives a much higher probability of going to node 5 next than of going to node 6. A last influencing factor to be considered in some instances is a rule that takes into account when people change their habits depending on who they are with. In largely populated environments certain people's movements have an influence on others. In, for example, going for lunch it may be that a particular person is a common factor in most locations. This is discovered by checking to see if people travel routes matched up temporally and if so, was one dominant over the other? When this is the case, a rule is applied in the same manner as the preferred paths, influencing the prediction. HABITS combines a number of different elements to produce future location predictions. The inputs to the Bayesian filter include the Motion Model showing where it is possible to go in the next step, the Sensor model giving the accuracy of the updates from the underlying tracking system, the learnt Historical belief and the location updates from the base system. When the filter has all the necessary information to give a prediction, it is run through a set of rules to improve the accuracy of its estimates. The HABITS approach described in this paper is designed to be able to operate on any type of tracking system to enable it to track between its updates and to give future predictions.

HABITS does not attempt to improve the underlying Wi-Fi positioning system but is used in conjunction with it to improve overall performance. While HABITS uses the same radio signals and equipment as other systems, it enables positioning and continuous real time tracking with increased accuracy, and in areas that were not previously possible. However, HABITS will only work in certain environments where people follow particular habitual movement patterns. Examples include work environments such as factories or hospitals.

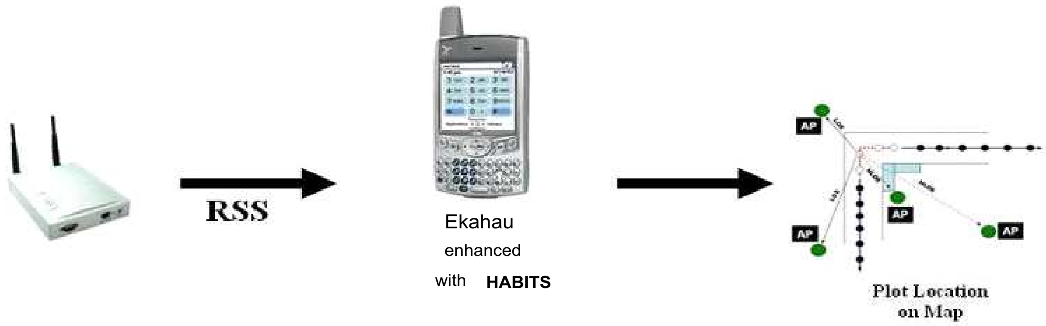

Figure 7: Context of HABITS

Figure 7 shows the context in which HABITS can be used. When a mobile device is tracked by the Ekahau RTLS and the HABITS algorithm is applied, it can still be tracked when it is no longer within line of sight (LOS) of three or more Access Points (AP). This is normally the minimum required for accurate localisation. The highest frequency rate of position updates from the Ekahau RTLS has been found to be 5 s [16, 17, 18]. These updates are often up to 15 seconds apart. Each update is sent to HABITS along with the learnt historical movement data and from this an intelligent prediction of the next likely location is given. Short term predictions effectively fill in the blanks in between updates from the Ekahau system. HABITS does not try and improve on the RSS positioning methods currently in use, but instead uses knowledge of the movement habits of users as a means of adding intelligence to existing tracking systems. This knowledge is then used to overcome signal black spots where existing systems fail (Figure 8) and to predict where the tracked user will travel to next. At time, t_1 (Figure 8) the Ekahau RTLS can give a position estimate that is close to the true position. At time t_2 both the standard Ekahau RTLS and the HABITS system also give a accurate estimate. However, at time t_3, the Ekahau system is no longer accurate due the user travelling through a signal black spot. This is where HABITS can dramatically improve standard location tracking systems and provide accurate updates of where the user is located.

4. Evaluating Long term predictions with HABITS

Long term predictions are related to the likelihood that a particular node will be visited during a particular time period. This could be later the same day or later in the week. For example, HABITS tells us with 85% confidence that during the lunchtime period that users 1, 2 and 5 will all leave their base nodes and will exit the building through the front door.

Table 2: Frequency of Preferred paths during Time periods for user 1

	PP	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Mon	Morning	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	3
	Lunch	0	0	0	0	4	3	0	0	0	0	0	0	4	3	1	1	2	3	2	0	0	0
	Evening	0	0	0	0	2	2	0	0	0	1	0	0	2	2	1	1	1	1	3	2	3	2
Tue	Morning	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	3
	Lunch	0	0	0	0	2	2	0	0	0	0	1	1	2	2	0	0	2	2	0	0	0	1
	Evening	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	2	2	4	2
Wed	Morning	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	Lunch	0	1	0	0	3	3	0	0	0	0	0	0	3	3	0	0	3	3	1	1	1	1
	Evening	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	3	2
Thur	Morning	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	Lunch	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0	1	1	0	0	0	1	0
	Evening	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	2	0
Fri	Morning	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	2
	Lunch	0	0	0	0	1	3	0	1	0	1	0	0	1	3	0	0	2	1	0	0	1	0
	Evening	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0

To calculate these movements, the repeatability of a PP within a time period (TP) is considered. Table 2 shows the frequency of each PP in each TP for User 1. The frequency tables for all users are automatically extracted from the learning data by HABITS. This works on the principle that if a journey has occurred every Tuesday morning for three weeks, then there is a high probability that it will occur the next Tuesday, all else being equal. There is no guarantee that this will occur but evidence from the tests show that it is highly probable. Users 3 and 4 both travel to node 3 (canteen in test area) on 13 out of the 15 days used for testing. These patterns allow HABITS to predict who will go where, when, for commonly repeated journeys with a useable degree of accuracy. The green boxes in Table 2 show preferred paths that have occurred twice during the same TP on the same day and the yellow boxes show those that have occurred on 3 or more occasions. Applying this data to the test data for User 1 yields the results shown in Figure 9. These show that when a PP has only been observed twice, the successful predictions occur 64% of the time meaning that 36% of the time the predictions are incorrect. However, when a

PP has been observed three or more times within a time period during a particular day, then the predictions are correct 78% of the time.

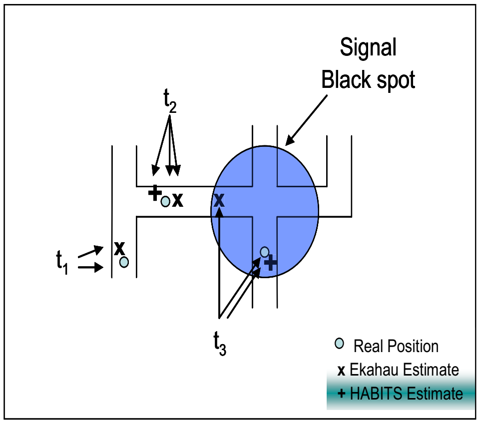

Figure 8: HABITS overcomes need for extra APs

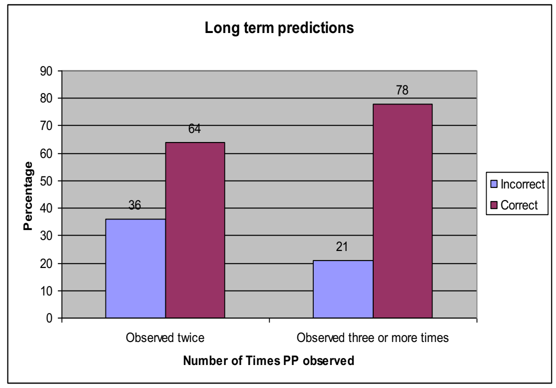

Figure 9: Long term predictions from User 1

Table 3 lists the overall average predictions for all of the test subjects. These are compiled by running HABITS on the test data available. Overall User 2, an RA is the most predictable. The short and medium term predictions for all subjects are similar however, the long term predictions are much lower for User 3 (Student) and User 5 (Academic).

Table 3: Predictions for all Test Subjects

Test Subject	**Job**	**Predictability**		
		Short	**Medium**	**Long (3 or more)**
User 1	**Research Student**	**82%**	**83%**	**78%**
User 2	**Research Associate**	**85%**	**88%**	**82%**
User 3	**Research Student**	**81%**	**77%**	**55%**
User 4	**Research Associate**	**87%**	**82%**	**76%**
User 5	**Academic**	**80%**	**78%**	**62%**
Average		**83%**	**81%**	**70%**

The users Base Node (desk) is the key to making predictions with HABITS. Of the total number of journeys made during the test period, 42% had the base node as the destination and 47% had the base node as the starting point. This means that 89% of all journeys undertaken by our test subjects involved travel to or from their base node. All of the test subjects showed very high (> 89 %) predictability when travelling to their own work station. When travelling from the base station, the final destination was more difficult to predict. However, HABITS still predicted the correct destination over 60% of the time for all users. User 4, the RA, was still predictable in over 90% of their journeys from their base station. Other journeys in the building had a much lower predictability. Some small patterns were apparent such as going to the toilet after the canteen, but overall these journeys proved to be beyond the predictability of HABITS. The average predictability of final destination of any of the test subjects was almost 80%. This means, in our test week, for four out of every five journeys taken, HABITS correctly predicted the final destination. It must be noted that these results are for journeys of greater than two nodes. The testing of HABITS revealed a number of interesting facts. HABITS is suitable in environments were people follow particular movement patterns. The two RAs (User 2 and User 4) proved to have much more predictable habits than the other three test subjects. It was concluded that this was because they were paid to sit in the same spot each day and had set times for breaks. User 5 (Academic) and Users 1 and 3 (Ph.D. students) did follow repeating movement patterns but these did not follow a rigid timetable. The conclusion from this was that the Academic had a changeable meeting schedule, whereas the student made particular journeys when he/she felt like it.

5. Conclusion & Future Work

Widely used techniques such as the Kalman and Particle filters are probabilistic approaches to taking educated guesses of the future given relevant information. This paper outlines a system HABITS which aims at overcoming weaknesses in existing Real-Time Location Systems (RTLS) by using the human approach of making educated guesses about future location. The hypothesis of this proposal is that knowledge of a person's historical movement habits allows for future location predictions to be made in the short, medium and long term. The research questions that were foremost are whether the tracking capabilities of existing Real Time Locating Systems can be improved automatically by knowledge of historical movement and by the application of a combination of artificial intelligence approaches. We also considered whether this approach can allow for intelligent prediction of future locations. We conclude that HABITS improves on the standard Ekahau RTLS in term of accuracy (overcoming black spots), latency (giving position fixes when Ekahau cannot), cost (less APs are required than are recommended by Ekahau) and prediction (short, medium and longer term predictions are available from HABITS). These are features that no other indoor tracking system currently provides.

References

1. Petzold, J., Bagci, F., Trumler, W. et al.: Next Location Prediction within a Smart Office Building. 1st International Workshop on Exploiting Context Histories in Smart Environments (ECHISE'05) at the 3rd International Conference on Pervasive Computing, Munich, Germany, May, (2005)
2. Vintan, L., Gellert, A., Petzold, J. et al.: Person Movement Prediction using Neural Networks. (2004)

3. Gellert, A., & Vintan, L.: Person Movement Prediction using Hidden Markov Models. Studies in Informatics and control, 15 (2006) 17
4. Han, S. J.: Learning and Predicting User's Movements for Intelligent Location-Based Services. (2004)
5. Froehlich, J., & Krumm, J.: Route Prediction from Trip Observations. SAE SP, 2193 (2008) 53
6. González, M. C., Hidalgo, C. A., Barabási, A. L.: Understanding Individual Human Mobility Patterns. Nature, 453 (2008) 779-782
7. Ashbrook, D., & Starner, T.: Learning Significant Locations and Predicting User Movement with GPS. Wearable Computers, 2002.(ISWC 2002).Proceedings.Sixth International Symposium on, (2002) 101-108
8. Bahl, P., & Padmanabhan, V. N.: RADAR: An in-Building RF-Based User Location and Tracking System. INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, 2 (2000) 775-784 vol.2
9. Bahl, P., Padmanabhan, V. N., Balachandran, A.: Enhancements to the RADAR User Location and Tracking System. Microsoft Research, (2000)
10. Rabiner, L. R.: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE, 77 (1989) 257-286
11. Birney, E.: Hidden Markov Models in Biological Sequence Analysis. IBM Journal of Research and Development, 45 (2001) 449-454
12. Yoon, B. J., & Vaidyanathan, P.: RNA Secondary Structure Prediction using Context-Sensitive Hidden Markov Models. (2004) S2/7/INV-S2/7/1-4
13. Petzold, J., Bagci, F., Trumler, W. et al.: Hybrid Predictors for Next Location Prediction. LECTURE NOTES IN COMPUTER SCIENCE, 4159 (2006) 125
14. Mozer, M. C.: An Intelligent Environment must be Adaptive. Intelligent Systems and their Applications, IEEE, 14 (1999) 11-13
15. Fox, D., Hightower, J., Kauz, H. (2003) Bayesian Techniques for Location Estimation.
16. Furey, E., Curran, K., Mc Kevitt, P. (2008) HABITS: A History Aware Based Wi-Fi Indoor Tracking System. PGNET 2008 - The 9th Annual Postgraduate Symposium: The Convergence of Telecommunications, Networking and Broadcasting 2008, Liverpool, John Moores University, UK, 23rd-24th June 2008, pp: 224-229
17. Furey, E., Curran, K. and Mc Kevitt, P. Learning Indoor Movement Habits for Predictive Control, International Journal of Space-Based and Situated Computing, Vol. 1, No. 2, pp:, July-August 2011, ISSN: 2044-4894, InderScience Publishing
18. Curran, K., Furey, E., Lunney, T., Santos, J., Woods, D. and Mc Caughey, A. An Evaluation of Indoor Location Determination Technologies, Journal of Location Based Services, Vol. 5, No. 2, pp:, June 2011, ISSN: 1748-9725, Taylor & Francis