ORIGINAL RESEARCH

Probabilistic indoor human movement modeling to aid first responders

Eoghan Furey · Kevin Curran · Paul Mc Kevitt

Abstract The arrival of new devices and techniques has brought tracking out of the investigational 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. Here we present a system 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 primary research question that is foremost is whether the tracking capabilities of existing real time locating systems can be improved automatically by knowledge of previous movement especially in the short term in the case of emergency first responders by the application of a combination of artificial intelligence approaches. 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 term predictions are available from HABITS). These are features that no other indoor tracking system currently provides and could prove crucial in future emergency first responder incidents.

Keywords Real time locating systems · Location determination · Emergency services · Wireless networks · Bayesian networks

E. Furey · K. Curran (✉) · P. M. Kevitt
Faculty of Computing and Engineering,
University of Ulster, Derry, Northern Ireland, UK
e-mail: kj.curran@ulster.ac.uk

1 Introduction

First responders have to rapidly assess any new emergency situation. Questions that need answered quickly include; Who is in the building? How do we get to them? What evacuation plan are they following? The answers to these questions can be readily available thanks to the application of integrated electronics and artificial intelligence (AI) techniques (Furey et al. 2011). What is needed to achieve such a system is some wearable device which will help pinpoint the current location of each occupant of the building. It turns out that to achieve tracking 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 way point for navigation but the lighthouse could not 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 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 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. The recent proliferation of mobile communications devices is allowing tracking and analysis of large groups of people to be carried out. Where available, this data is being mined searching for patterns and trends which facilitate prediction (Petzold et al. 2005; Vintan et al. 2004; Gellert and Vintan 2006). This paper seeks to outline a 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 allows for 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 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 consider whether this approach can allow for intelligent prediction of future locations. A practical application of our system would be for use by first-responders inside buildings.

2 Predicting location using intelligent techniques

To accurately position an object in addition to the techniques and technologies previously discussed, a certain amount of intelligence is required. These methods allow the accuracy levels of the estimates to be increased. When a human makes estimation about where an object will be located in the future, they automatically carry out the complete calculation in their brains. To enable computers to repeat these feats and to allow them to work with a number of different objects requires a number of artificial intelligence techniques. In an indoor environment this is a largely under researched area. Being able to predict with some degree of accuracy, the future location of a person or thing has many applications. This seemingly magical quality is one of the 'holy grails' of localisation. The prediction does not need to be 100 % certain, and indeed certainty is almost impossible to establish. Never the less,

the next location can be predicted with a certain degree of accuracy as has been demonstrated in previous research projects. Outdoors, using GPS traces to try and learn next location has been attempted by Han (2004) for someone on foot and in Froehlich and Krumm (2008) 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 (González et al. 2008). Indoors, this is a largely under researched area, however a number of 'smart environments' have been set up such as the work Petzold et al. (2005). 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. Since around 2006, due to its commercial value, most of the leading research in next location prediction is not being conducted by academic institutions and therefore everything that has been learnt may not be in the public domain. However, this review covers all that has been published but research may exist that has not been published, for instance, that which is carried out by defence contracts or commercial entities. Ashbrook and Starner (2002), 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 nth 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 2nd order can yield more accurate results as in the case of probability of transition from A $\rightarrow$ B is 70 % but the probability of transitioning from B $\rightarrow$ A $\rightarrow$ 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 (2002) 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 (2004) attempts to build upon the work of (Ashbrook and Starner 2002) 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.

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 carried out in this area. One of the first research projects that considered future movement was Microsoft Research's RADAR project (Bahl and Padmanabhan 2000). 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 et al. (2000). 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 nearest neighbour in signal space (NNSS) 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) (Bahl et al. 2000). 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 (1989). 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 (2001). A new method for predicting the secondary structure of RNA using HMM was proposed by Yoon and Vaidyanathan (2004). They demonstrated very accurate, secondary structure prediction using their proposed model with a low computational cost. A study by Petzold et al. (2006) 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. A study by Mozer (1999) 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 Modelling an indoor tracking system

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. The acronym history aware based indoor tracking system (HABITS), describes the approach of this research. The past movement history of tracked humans is used to enhance an existing tracking system. The technology of the underlying tracking system or the positioning methods used are not important. The HABITS approach can be designed to be generic. The three main components of HABITS are a connected graph, a discrete Bayesian filer and a set of logic rules. One of the contributions in this work involves combining 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. The graph allows us to represent this information in mathematical terms and as numbers which may then be processed by a machine/computer. Methods of modelling the travel environment exist and of these, a graph structure closely represents the travelable paths. The nodes in the graph can be positioned to represent areas of interest, decision points or places where the user stops. Only areas that are in use and where access is unrestricted are represented. In between these locations are the paths that may be travelled between. The paths may be considered as edges and those locations of interest could be 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. As one-way pedestrian systems are uncommon indoors all the graphs used for this study are undirected. 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 rooms with travel paths between. Examples of rooms could be bedroom, bathroom and living room. Movement between 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 Fig. 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 Fig. 2. Learning the locations of these points can be done automatically in a number of ways, all of which require an underlying tracking system to be installed.

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.
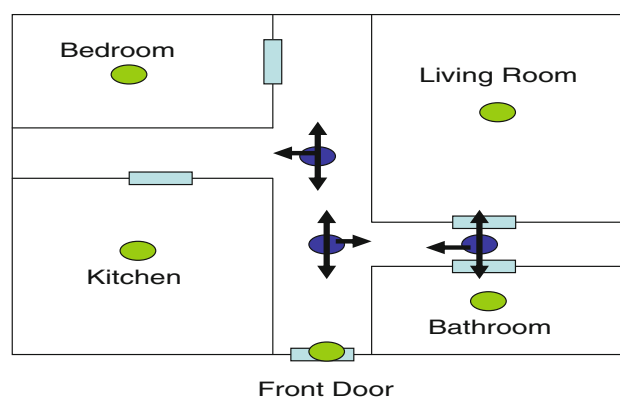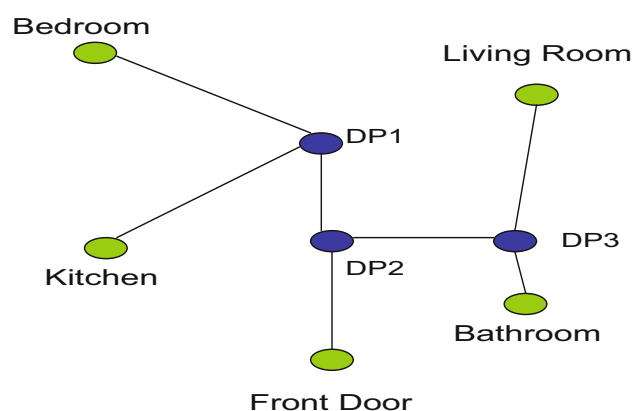

Fig. 1 Node positions in house

Fig. 2 Connected graph with node connections

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 (Fig. 3) they may be represented as an $n \times n$ 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 (Fig. 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.

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. 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
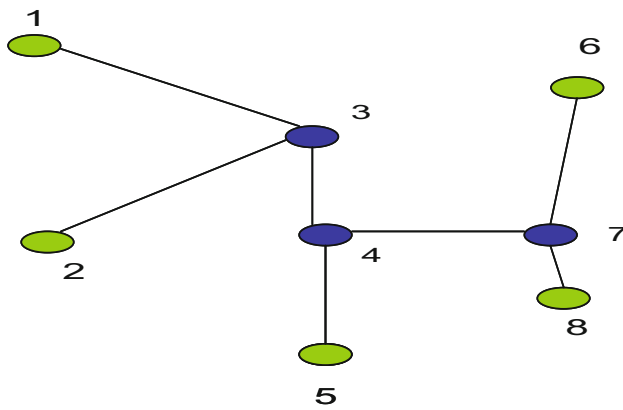
Fig. 3 Node names replaced with numbers

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

Fig. 4 Adjacency matrix for nodes in sample house

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 persons travel through the nodes. Again, a number of methods exist to do this, but a straight forward one is to use the sequence of all nodes traversed through 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 × 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 Froehlich and Krumm (2008) proved this needs to be done with a large dataset which takes a long 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 to be considered. A discrete Bayesian filter had been shown to work well for data fusion (Fox et al. 2003). The underlying 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. 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), a great deal of information is being left out about commonly travelled paths or sequences of nodes. Froehlich and Krumm (2008) 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 makes the calculations overcomplicated. 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 a necessarily adjacent locations. The paths may go through a number of intermediate nodes and a common journey could be kitchen to toilet. This would involve travelling through four 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. Preferred paths are also stored as a vectors 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).

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 were with. In largely populated environments certain people's movements have an influence on other peoples. If, 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 allow it to track between its updates and to give future predictions.

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 (Furey et al. 2008). These updates are often up to 15 s 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 (Fig. 5) and to predict where the tracked user will travel to next. At time, t_1 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 good 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 Human movement indoors with habits

Ekahau RTLS is a commercial off the shelf (Cots) piece of software which was developed by a Finnish company for Wi-Fi 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. Once the server is in place, the Ekahau Site Survey (ESS) model must be created. A number of steps are required for this. First, a map (JPEG) of each floor is uploaded to the ESS application. Figure 5 shows a signal strength map for the ISRC. The dark green indicates areas of good signal strength where tracking capability should be good. These

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

Fig. 5 First floor plan—signal strength map
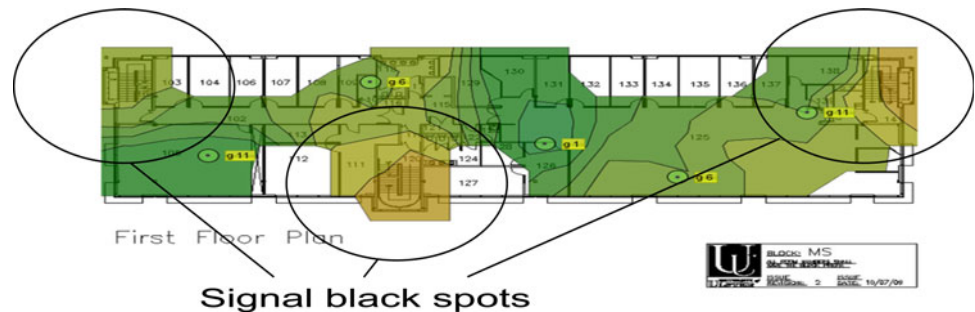
heat maps also show areas signal strength is weak. These are target areas for HABITS.

In order to collect historical movement data a topological map of the test area is created. A topological map is one which consists of a number of nodes representing places of interest which are connected by edges representing paths where a user may travel. These areas are covered by zones in Ekahau which allow for reporting of when a person carrying a mobile Wi-Fi device enters or leaves them. The zones represent areas that are passed through frequently on the ground and first floors in the ISRC. Each of these zones can be considered to be a node in a connected graph. The positioning of these zones is a manual process based on expert knowledge of where a user is likely to stop and areas where they would pass through often. Also used are locations where a user has a number of options of where next to travel. The locations of these zones relative to one another can now be represented as an adjacency matrix and hence a connected graph. To do this each node in the graph representing a zone is given a unique ID between 1 and n, were n is the number of zones. By querying the EPE the list of zones is retrieved and each zone is allocated a unique ID. For the two floors in the MS building there are 19 zones in total. Using the zone map, an adjacency matrix of size n × n is manually created. The corresponding zone-node list allows all zone data from the EPE to be manipulated as if each zone was the node in the connected graph. Figure 5 shows a connected graph representation of the two floors in the ISRC. The edges between nodes show paths that may be travelled and represent the movements of Wi-Fi tracked people in the building. The numbers on the nodes are those used by the zone to node conversion table. Knowledge of where the user is, whether they are in motion or not and the exact time are essential for HABITS to function. This is the only live information that HABITS processes.

Once the live user information is received, HABITS checks the matrices to see what constraints on movement exist. Combining these constraints with the data from Ekahau allows an initial probability prediction to be made from the transition matrix. Information about node types and preferred paths is now added to HABITS to further improve the accuracy of the predictions. Predictions from HABITS depend on the time scale required and may be short term, (a few seconds), medium term, (a few minutes–end of current journey) or longer term, (later that day or week). The last stage of the operation of HABITS involves taking a particular action based on the predictions provided if the probability confidence is high enough or plotting these future locations on a map. The following is a sample scenario taken for real movement data within the ISRC. The user has left his desk. HABITS is attempting to predict where he will go to next. The pseudo code and associated explain the scenario and the steps HABITS takes in order to make a prediction follow.

1. If tag = Eoghan
2. **Node = 5 and previous node = 4**
3. Node 5 NOT = wait node
4. Action = calc next node,
5. On preferred path–NO
6. **Next node = Either 2, 3, 6, 7, 8 %, all have non-zero probability**
7. Check time period = Lunch
8. **If time = Lunch THEN next node is 6 or 3 %. Probability > 80 %**
9. Check other users in area
10. **If with John THEN next node = 6 %. John doesn't go to the canteen**
11. **If with Mary THEN next node = 3 %. Mary usually goes to the canteen**
12. If alone then next node = 6 (40 %) OR 3 (40 %) % wait for more info
13. Use speed and distance to calculate position at time t
14. **Calculate and show positions at t + 1, t + 2···t + n**

An update is now received from the Ekahau RTLS. This update is not in a zone so using the nearest neighbour search the closest node is found to be node 16. Given that the last zone entered update was from the wait node 18, HABITS now possesses three bits of new information.

x_t = node 16, x_{t-1} = node 18 and user Eoghan is 'in-motion'.

Table 2 Probabilities from the transition matrix at node 16

Node no	15	13	12	11	14	17	19	5
Bel(x_t)	0.27	0.05	0.05	0.01	0.05	0.05	0.01	0.01

From this new information HABITS can offer only a general prediction of the next possible node that will be visited. Eight different options are possible at this stage. Node 16 is not a 'wait node' therefore HABITS does not consider this to be the end of the current journey. The options at node 16 are listed in Table 2.

While node 15 is given as having a much higher probability that the other possible next nodes, 0.27 is too low to make any kind of reasonable prediction so HABITS must wait for more information from Ekahau. If another location update is received between nodes 15 and 4. HABITS knows from the previous bel(x_t) that node 15 was one of the last possible nodes, therefore, even though node 15 is slightly closer to the update than node 4 (from the k nearest neighbours search, 15 closest, followed by 4), it chooses node 4 as the next node from which to make a calculation. At this point due to a lack of other options and the fact that HABITS has information about the previous four nodes 18–16–15–4, a fairly confident prediction can be made that the next node to be visited is will be node 5.

When a third update is received which is very close to node 5, given the previous nodes sequence, the distance from the third update to node 5 and the previous prediction, HABITS carries out the next prediction based on node 5 as its current node. As node 5 is also a transition node, possible predictions from the transition matrix give five possible next nodes as Table 3 shows. At this stage no prediction can be made as node next node has a clear higher probability. Probabilities change when the preferred paths for the current time period are checked. Here, as the time period is equal to 'lunch' and the previous node sequence is quite extensive at this stage, two nodes, 3 and 6

Table 3 Probabilities from the transition matrix at node 5

Node Number	3	2	6	7	8
Bel(x_t)	0.17	0.02	0.31	0.15	0.02

now possess a significantly higher probability that the other possible nodes. The preferred path statistics are shown in Table 4.

HABITS can now say with over 80 % confidence that the next node to be visited will be either node 3 or node 6. To further increase the accuracy of the predictions, HABITS checks to see if there are any other users in the area. Figure 6 depicts a position update being received from John. Examining the preferred path of John reveals that John does not go to node 3. As the paths of Eoghan and John are regularly together, combining John's 'preferred path' probability with Eoghan's gives a new prediction that Eoghan will go to node 6 with 80 % confidence.

5 Overcoming signal black spots

With the predictions available from HABITS, the location of a tracked person can be intelligently guessed when Ekahau loses it in a signal black spot. Table 5 shows the improvements due to HABITS at particular black spots. The three black spots tested below are the three stairwells in the research centre. When using Ekahau alone the accuracy of a position estimates is very low in these areas. However, on journeys through the building, HABITS was able to more than half the error in estimate of location within these black spots. While a subject is travelling through a black spot, HABITS estimates were within the figures listed 95 % of the time. It was concluded that the 5 % of the time the estimate were above the stated accuracy were due to stopping and turning around or just stopping midway through the journey.

Table 6 describes the difference between Ekahau when set up with a low number of APs (the amount required for data communication), the same setup with the addition of HABITS and Ekahau alone but with many more APs installed. The extra APs were installed to test the improvements to Ekahau when many extra APs where available. For the purposes of these tests, an extra 5 APs were temporally installed and a number of calibration and test surveys were conducted. These did improve the accuracy of the Ekahau RTLS to almost 2 m which is comparable to the overall average accuracy achievable when HABITS is in operation. The yield of the Ekahau
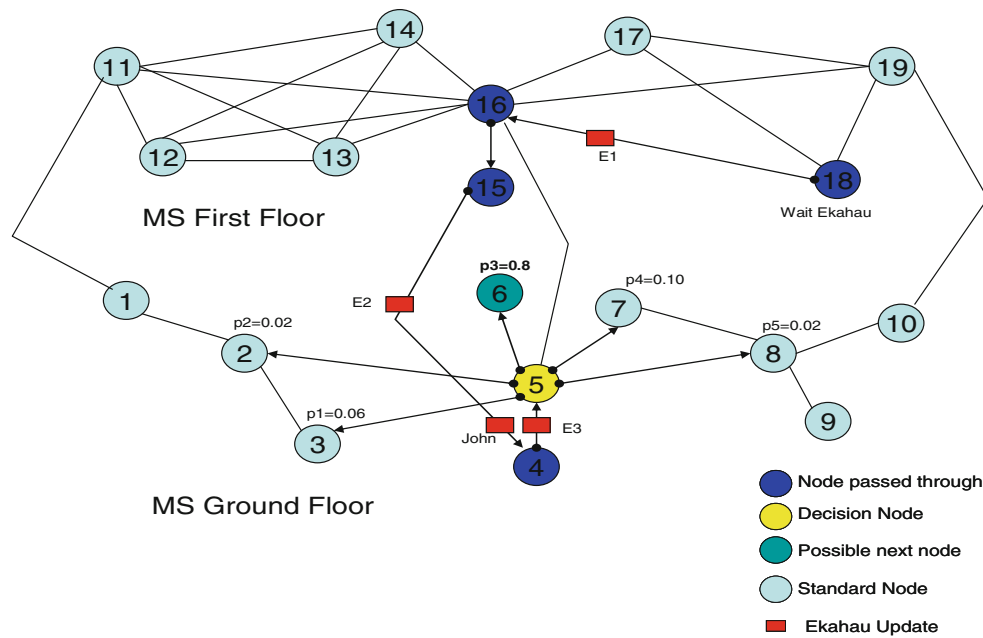
Table 4 Preferred Path options at node 5

1st node	2nd node	3rd node	4th node	5th node	6th node	No. of times repeated	Time period	Probability
18	16	15	4	5	3	2	Lunch	0.25
18	16	15	4	5	6	6	Lunch	0.75

Fig. 6 Checking other users rule, John does not go to the canteen so node 6 becomes most probable

Table 5 Accuracy of HABITS in signal black spots

	Accuracy (95 %)	
	Ekahau (m)	HABITS (m)
Black spot 1 (left stairwell)	5.0	2.2
Black spot 2 (centre stairwell)	4.6	2.1
Black spot 3 (right stairwell)	6.1	3.0

RTLS is also increased by HABITS to a level that is comparable with doubling the amount of APs.

These results show that when implementing a system such as Ekahau without redesigning the AP layout, the average accuracy achieved of 4 m is well below the level of 1 m which is claimed by Ekahau. Adding substantially more APs (five per floor) did improve the average accuracy to approximately 2 m in our test area. However, this improvement came with a significant extra cost in terms of: installation, calibration time and deployment cost as each AP is around €100. Application of HABITS showed a marked improvement compared to just using the Ekahau system by itself.

Overall HABITS improves on the standard Ekahau RTLS in terms of: Accuracy, Yield, Latency, Cost and Predictive ability. Accuracy is improved as HABITS is able to overcome signal black spots and give higher location accuracy than Ekahau alone. The Yield is improved as HABITS enables positioning to be carried out in locations where Ekahau previously failed or gave low levels of accuracy. Signal black spots are overcome by the operation of HABITS. Latency is improved as HABITS predictions allow for continuous location updates. Savings are made in terms of Cost as when employing HABITS, fewer extra APs are required to give improved accuracy levels. HABITS improves overall positioning accuracy to a level that is normally achievable only by investment in extra infrastructure. Finally, HABITS provides a predictive ability which is not available in any existing RTLS. These predictions are available in the short term and medium and long term behaviour suggestions are also available.

A number of future opportunities exist to enhance HABITS. If a large scale deployment of HABITS was to be created then issues relating to scalability and algorithm complexity and efficiency would need to be addressed. Methods dealing with the changing of subject's habits and

Table 6 Results of testing HABITS

	Accuracy (m) approx	Yield (%)	Latency (s)	Cost
Ekahau (APs configured for data communication)	4	84	5–15	Ekahau RTLS
Ekahau plus HABITS	2	97	1	Ekahau RTLS
Ekahau with 5 extra APs per floor	2	100	5–15	Ekahau RTLS plus €100 per AP

how the underlying graph and matrices are updated and modified over a long period of time could also be examined. Considerations of the social and ethical implications of movement data collection are obviously important to the success of the HABITS approach. Options available could include either forcing people (as happens in the army or in areas where security is paramount) or encourage them by offering them some form of reward or benefit in return for their location data (as is currently used by many social networks). Questions' regarding the right and wrong of people tracking is another subject in itself, however, HABITS is envisaged initially as being used in areas where the benefits outweigh the drawbacks.

A potential useful area of application for HABITS could be in control systems, specifically those that are dependent on the movement of people. Bolick (2010) reports that Lighting and Heating, Ventilation and Air Conditioning (HVAC) account for approximately 60 % of a buildings energy costs. HABITS gives short (<15 s), medium (15 s– a few minutes) and long (a few hours or days) term predictions on the general movement habits of people in a work environment. Knowledge of where people will travel within a building and when, also gives information regarding where they are not likely to go! This knowledge could be used as input to an intelligent control system for heating and lighting in a large building. In the short term, if a system knew what room or area a person would travel to next, then the lights could already be on or in some standby mode to facilitate quick power up. This way they would not have to stay on standby continuously. Areas which were infrequently travelled could be put into low energy mode or switched off completely, thereby saving energy consumption costs.

With heating systems a similar but longer term approach could be applied. If the automatically controlled heating system knew that at a certain time of day, e.g., lunch, many people stood in the canteen or corridor then the heat could be adjusted up or down depending on the outside temperature and number of people. Conversely, if the system knew areas were people rarely travelled, then the heating could be turned off and would not be wasted while the area was vacant. While various sensors can currently control this, they only work when activated, i.e., when someone walks past them. HABITS could control the system in advance and could learn when the movement patterns changed. Existing sensors (motion) on doors could either be used in conjunction with HABITS or could be replaced by HABITS.

If a long term study was carried out or was simulated then the number of kilowatt hours saved could be calculated and this should prove to be substantial. The system could also be linked into controlling air conditioning systems in areas with hot climates. HABITS is suitable for such an application as it gives predictions of varying degree of accuracy and would not be suitable for life critical applications as there is a large element of probability involved. However, in building automation control systems a certain degree of inaccuracy would be acceptable if the overall energy savings were significant.

6 Conclusion

This paper has outlined the HABITS approach to the tracking and modeling of people's movement patterns in an indoor environment. HABITS aims to overcome weaknesses in existing RTLS by using the human approach of making educated guesses about future location. 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 term predictions are available from HABITS). These are features that no other indoor tracking system currently provides and could provide crucial in future emergency first responder incidents.

Knowledge of who is where within a building could potentially provide emergency first responders with information that could save lives. Quickly finding casualties and rapid but safe evacuation of a building are primary aims for emergency services and HABITS offers a new approach that could significantly reduce the time required to carry out these tasks. Integrating HABITS into the alarm systems could provide emergency services with the locations of people within a building even before those people knew there was an emergency. This could allow for greater control of the situation and could help avoid panic.

References

Ashbrook D, Starner T (2002) Learning significant locations and predicting user movement with GPS. In: Proceedings of sixth international symposium on wearable computers (ISWC 2002), pp 101–108

Bahl P, Padmanabhan VN (2000) RADAR: an in-building RF-based user location and tracking system. In: INFOCOM 2000. Nineteenth annual joint conference of the IEEE computer and communications societies, vol 2, pp 775–784

Bahl P, Padmanabhan VN, Balachandran A (2000) Enhancements to the RADAR user location and tracking system. Microsoft research technical report: MSR-TR-00-12

Birney E (2001) Hidden markov models in biological sequence analysis. IBM J Res Dev 45:449–454

Bolick J (2010) A wireless solution for energy control in existing buildings. http://www.automatedbuildings.com/news/apr10/arcles/adura/100329095808adura.htm

Fox D, Hightower J, Kauz H et al (2003) Bayesian techniques for location estimation. In: Proceedings of the 2003 workshop on location-aware computing

Froehlich J, Krumm J (2008) Route prediction from trip observations. SAE SP 2193:53

Furey E, Curran K, Mc Kevitt P (2008) HABITS: a history aware based Wi-Fi indoor tracking system. In: PGNET 2008, The 9th annual postgraduate symposium: the convergence of elecommunications, networking and broadcasting 2008

Furey E, Curran K, Mc Kevitt P (2011) A Bayesian filter approach to modelling human movement patterns for first responders within indoor locations. In: CIDM-2011 second international workshop on computational intelligence for disaster management, Fukuoka, Japan, 30 Nov–2 Dec 2011

Gellert A, Vintan L (2006) Person movement prediction using Hidden Markov models. Stud inform control 15:17

González MC, Hidalgo CA, Barabási AL (2008) Understanding individual human mobility patterns. Nature 453(7196):779–782

Han SJ (2004) Learning and predicting user's movements for intelligent location-based services. MSc. Dissertation, Yonsei University

Mozer MC (1999) An intelligent environment must be adaptive. In: Intelligent systems and their applications, IEEE, pp 11–13

Petzold J, Bagci F, Trumler W et al (2005) Next location prediction within a smart office building. In: First international workshop on exploiting context histories in smart environments (ECHISE'05) at the third international conference on pervasive computing, Munich, Germany, May 2005

Petzold J, Bagci F, Trumler W et al (2006) Hybrid predictors for next location prediction. Lecture notes in computer science, p 125

Rabiner LR (1989) A tutorial on Hidden Markov models and selected applications in speech recognition. In: Proceedings of the IEEE, pp 257–286

Vintan L, Gellert A, Petzold J et al (2004) Person movement prediction using neural networks. In: First workshop on modeling and retrieval of context, Ulm, Germany

Yoon BJ, Vaidyanathan P (2004) RNA secondary structure prediction using context-sensitive hidden markov models. In: Proceeding of IEEE international workshop on biomedical circuits an systems, Singapore