

LEAPS : A Location Estimation and Action Prediction System in a Wireless LAN Environment

Qiang Yang¹, Yiqiang Chen², Jie Yin¹, and Xiaoyong Chai¹

¹ Department of Computer Science
Hong Kong University of Science and Technology
Clearwater Bay, Kowloon Hong Kong, China
{qyang, yinjie, carnamel}@cs.ust.hk

² Shanghai Division, Institute of Computing technology
Chinese Academy of Sciences
Building 13 Block 116, Lane 572, Bi Bo Road, Pu Dong District
Shanghai, China, 201203
yqchen@ict.ac.cn

Abstract. Location estimation and user behavior recognition are research issues that go hand in hand. In the past, these two issues have been investigated separately. In this paper, we present an integrated framework called LEAPS (location estimation and action prediction), jointly developed by Hong Kong University of Science and Technology, and the Institute of Computing, Shanghai, of the Chinese Academy of Sciences that combines two areas of interest, namely, location estimation and plan recognition, in a coherent whole. Under this framework, we have been carrying out several investigations, including action and plan recognition from low-level signals and location estimation by intelligently selecting access points (AP). Our two-layered model, including a sensor-level model and an action and goal prediction model, allows for future extensions in more advanced features and services.

1 Introduction

In recent years, the research area of indoor location estimation and user behavior prediction have attracted intense attention. Much work has been done in the computer network and pervasive computing areas on using the signal strength values from the access points (AP) to determine locations using various geometric and probabilistic knowledge. Similarly, different statistical models have been proposed in artificial intelligence and data mining area for recognizing and predicting a user's behavior and plans. However, no work has attempted to combine location estimation with high-level behavior recognition, and to use machine learning and probabilistic reasoning to help with low-level location estimation. In this paper, we survey the work in location estimation and high-level plan recognition, and present our own integrated framework for accomplishing both tasks. Through our integrated system known as LEAPS (Location Estimation and Action Prediction System), we present our research results in three different tasks, including using machine learning methods for access point selection and using location estimation for high-level goal recognition.

Our two-layered model for LEAPS is shown in Figure 1, which includes a sensor model, and an action and goal recognition model. This framework allows several different tasks to be accomplished, ranging from low level to high level. It also allows for more advanced extensions in the future. Below, we discuss the different layers in turn.

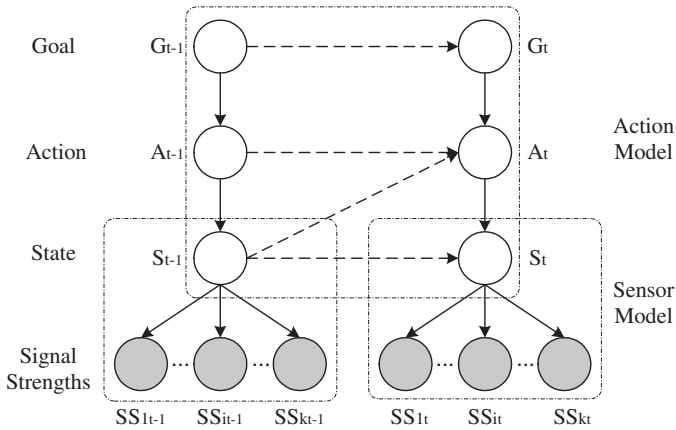


Fig. 1. Two level model for the LEAPS framework

2 Wireless Environment for LEAPS

Our experimental test bed is set up in the faculty office area of Computer Science Department in the Academic Building of Hong Kong University of Science and Technology. The building is equipped with IEEE 802.11b wireless Ethernet network in the 2.4 GHz frequency bandwidth. The layout of the floor is shown in Figure 2. Experiments were carried out in the five hallways (HW1~HW5) and two printing rooms as labeled in the figure. The five hallways are labeled as HW1 to HW5. There are also two rooms in the area.

There are a total of 25 base stations that are detectable in the environment, of which three base stations that are distributed within the area are marked with concrete circles in the figure. Among the other 22 bases stations, some are located on the same floor outside of the area while the others are located on the different floors. The IEEE 802.11b standard works over the radio frequencies in the 2.4 GHz band. However, accurate location estimation using measurements of signal strength is a longstanding complex and difficult task due to the *noisy characteristics* of signal propagation. Subject to reflection, refraction, diffraction and absorption by structures and even human bodies, signal propagation suffers from severe multi-path fading effects in an indoor environment. As a result, a transmitted signal can reach the receiver through different paths, each having its own amplitude and phase. Figure 3 gives a typical example of the normalized histogram of the signal strength received from a base station at a fixed location.

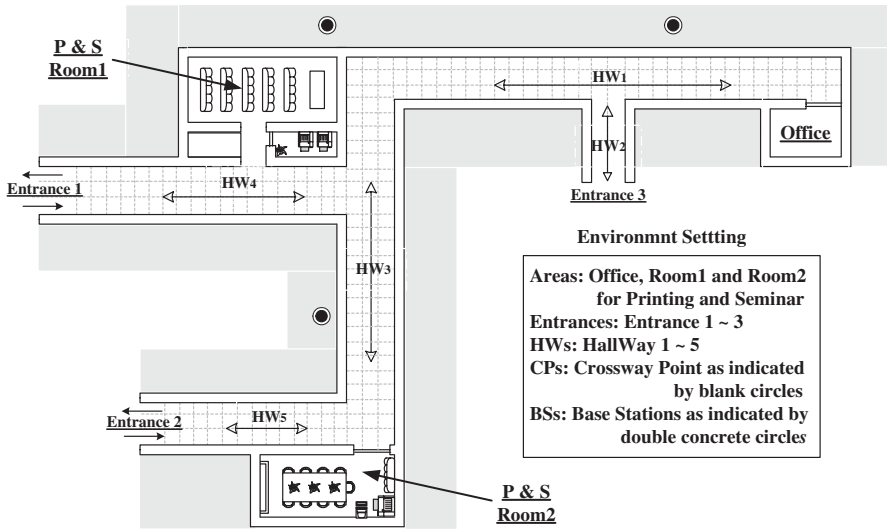


Fig. 2. The layout of the HKUST office area for the Wireless LAN experiments

3 Indoor Location Estimation at the Sensor Level

Deterministic Techniques In general, the location estimation research can be classified into two categories: deterministic techniques and probabilistic techniques. Deterministic techniques [2, 8] use deterministic inference methods to estimate a user's location. The RADAR system by Microsoft Research [2] proposes nearest neighbor heuristics and triangulation methods to infer a user's location. It maintains a radio map which tabulates the signal strength received from different access points at selected locations. Each signal-strength measurement is then compared against the radio map and the coordinates of the best matches are averaged to give the location estimate. The accuracy of RADAR is about three meters with fifty percent probability. The LANDMARC system [8] exploits the idea of reference points to alleviate the effects caused by the fluctuation of RFID signal strength. The accuracy is roughly one to three meters. However, the placement of reference tags should be carefully designed since it has a significant effect on the performance of the system. Moreover, the RFID readers are so expensive that it is infeasible for localization in a large area.

Probabilistic Framework Another branch of research is the probabilistic techniques [14, 13, 11, 6] construct a conditional probability distribution over locations in the environment of interest. In [6], Ladd et al. uses probabilistic inference methods for localization. They first use Bayesian inference to compute the conditional probability over locations, based on received signal-strength measurements from nine access points in the environment. Then a postprocessing step, which utilizes the spatial constraints of a user's movement trajectories, is used to refine the location estimate and reject the results with significant change in the location space. Depending on the postprocessing step is used or not, the accuracy of this method is 83% and 77% within 1.5 meter.

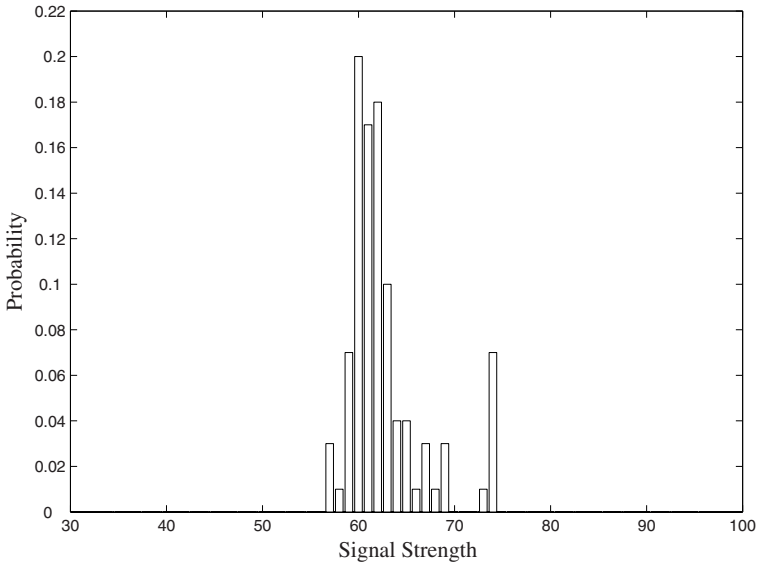


Fig. 3. An example of signal strength distribution

The LEAPS Framework to Location Estimation Our LEAPS framework for location estimation is divided into two phases. The first phase is done offline, where the main purpose is to perform intelligent AP selection. We divide this phase into the following steps:

1. First, a feature selection algorithm is applied to find a subset S of AP's that can give the best performance. This subset will then be used as the basis for subsequent computation.
2. A subsequent clustering analysis is then applied to the set S and data collected in the offline phase, in order to partition the grid space into clusters. Each cluster will then provide a subsequent location model.
3. Finally, a decision tree model is constructed for each cluster, based on the AP's given in S . For each cluster only a subset of AP's from S is selected, which further reduces the number of AP's needed for location estimation within each cluster.

The second phase is an online phase, in which a new trace of signal strengths is taken as input and the current location is estimated. This phase is done in two steps:

1. First, the signal strength values from the selected AP's from the set S is used to determine the cluster the current client is most likely located within.
2. Then, the decision tree from the identified cluster is used to determine, at a finer level, which grid the client belongs to. This step will use a subset of the AP's given in S , which further reduces the number of AP's used in a computation. In addition, the AP's that are used only involve arithmetic comparison, which is one of the cheapest computations as computational energy is concerned.

As one of our experimental results, we show in Figure 4 the relation between different AP selection methods and accuracy. We can see that in order to achieve the accuracy of 90%, the number of access points required by each location estimation are 12, 18 and 24, according to information gain (*InfoGain*), MLE used in previous work of [14] (*MaxMean*) and AP selection in the opposite order of information gain ordering (*MinMean*), respectively. It can be concluded that the *InfoGain* criterion uses the fewest access points to achieve the same level of accuracy, which achieves much better result as compared to previous methods.

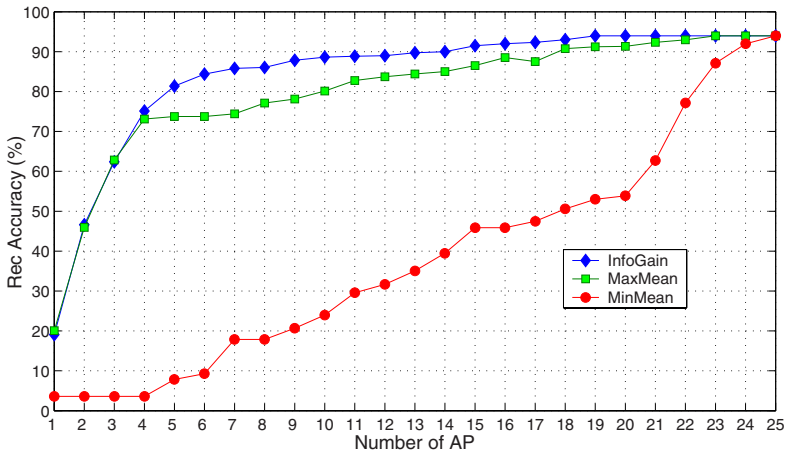


Fig. 4. AP selection and its impact on accuracy

4 User Behavior Recognition in LEAPS

Being able to predict a user's location is one thing, but the eventual purpose of location estimation is to infer users' high-level behavior patterns from low-level sensory data, and provide useful services. Such a task is what we call *location-based plan recognition*. Being able to accomplish this task is critical to many applications. For people suffering from various cognitive limitations in hospitals and care facilities, the technique can discover when a person's behavior is out of the norm and provide help in a timely manner [9]. For shoppers in a busy business environment such as a shopping mall, services and products can be offered not only according to people's current location, but also according to their intended goals and actions.

In our current work, we have taken a first step in inferring high-level user goals from low-level mobile data in an indoor environment, where a wireless LAN is available. In this section, we summarize our statistical model for goal recognition, where a full report can be found in [12], where a two-level dynamic Bayesian network model (DBN) is applied. This model integrates DBN with and a fast-inferencing model based

on n-grams for goal inference. We show that this architecture allows us to incorporate domain knowledge, whereby we achieve a nice tradeoff between model accuracy and inferencing efficiency.

Previous Work on Plan Recognition In the artificial intelligence area, recognizing complex high-level behaviors has traditionally been the focus of plan recognition [5, 7]. A Bayesian network was used for plan recognition in story understanding [4]. In [3], a corpus-based N-gram model was introduced to predict the goal from a given sequence of command actions in the UNIX domain. In addition, other advanced stochastic models for recognizing high-level behaviors were proposed such as Dynamic Bayesian Networks (DBN) [1] and Probabilistic State Dependent Grammars [10]. However, most of the work in plan recognition has been restricted to the high level for inference, and the challenge of dealing with low-level sensor models has not been addressed. Only in recent years, attempts have been made to integrate high-level behavior models with low-level sensor models. The work of [9] presents an approach by applying a Bayesian model to predict a user's transportation mode based on location readings from GPS devices in an urban environment.

The Dynamic Bayesian Network Model To represent different degrees of uncertainty in a time series, researchers have proposed various probabilistic models. Among them, the dynamic Bayesian network model has been shown to be well suited for goal recognition tasks. It includes two time slices numbered t and $t - 1$ respectively. The shaded nodes SSi_t represent the strength variables of signals received from multiple base stations, which are directly observable. All the other variables - the physical location S_t of the user, the action A_t the user is taking and the goal G_t the user is pursuing - are hidden, with the values to be inferred from the raw signals.

To balance between recognition accuracy and computational complexity, we propose a novel two-level model. The complete DBN model is separated into two different models. The lower part of the DBN starting from the observation layer to the action layer corresponds to the lower level of the architecture. An N-gram model is applied above the action model to infer goals.

In this framework, a low-level DBN model is responsible for computing the most plausible sequence of actions A_1, A_2, \dots, A_t from the observations o_1, o_2, \dots, o_t obtained up to time t . The task is carried out by the method of smoothing, which estimates the hidden states of the past given all the evidence up to the current time point t , $P(A_\tau | o_1, o_2, \dots, o_t)$, where $\tau \leq t$. To test the validity of the model, we have collected 570 traces for 19 goals of a professor to be modeled in the office area. Figure 5 shows the recognition process of one trace belonging to the goal "Seminar-in-Room1", with respect to three other goals among the 19 goals.

5 Conclusions and Future Work

In this paper, we presented an integrated approach towards addressing the problem of inferring high-level goals from low-level noisy signals, as well as estimating a users' current locations, in a complex indoor environment using an RF-based wireless LAN.

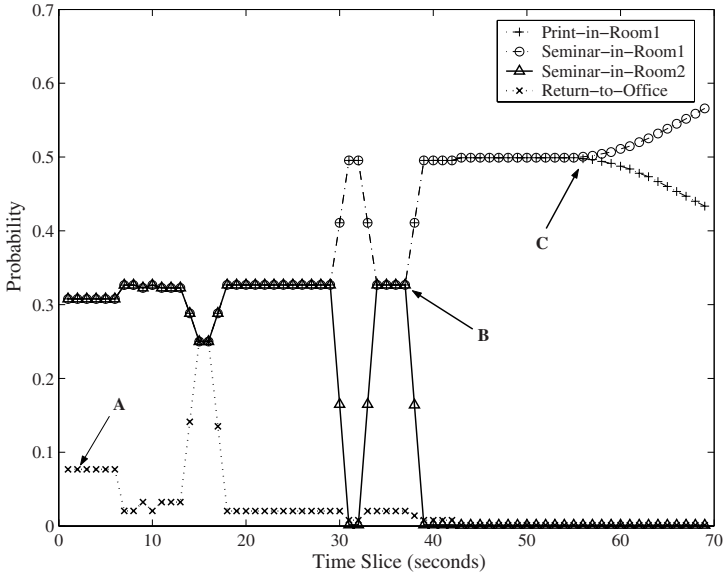


Fig. 5. Behavior recognition for four goals in our experiment

For location estimation, we applied machine learning techniques to select an optimal subset of access points that allow the best balance in accuracy and power saving. For the task of action prediction and plan recognition, we model the problem based on the framework of dynamic Bayesian network, using a two-level n-gram based model. The experiments demonstrate that the two-level approach is more efficient than the whole DBN solution while the accuracy is comparable. The three-layered model, including the sensor model, the action model and the goal model, provides an integrated framework for location estimation and high-level goal recognition.

Acknowledgement

The authors are supported by a grant from Hong Kong RGC and the Hong Kong University of Science and Technology. This research is conducted at the Hong Kong University of Science and Technology.

References

- [1] D. Albrecht, I. Zukerman, and A. Nicholson. Bayesian models for keyhole plan recognition in an adventure game. *User Modelling and User-adapted Interaction*, 8(1–2):5–47, 1998.
- [2] P. Bahl and V. N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *Proceedings of IEEE INFOCOM2000*, pages 775–784, 2000.
- [3] N. Blaylock and J. Allen. Corpus-based statistical goal recognition. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*, pages 1303–1308, Mexico, August 2003.

- [4] E. Charniak and R. Goldman. A bayesian model of plan recognition. *Artificial Intelligence Journal*, 64:53–79, 1993.
- [5] H. Kautz and J. F. Allen. Generalized plan recognition. In *Proceedings of AAAI1986*, pages 32–38, 1986.
- [6] A. Ladd, K. Bekris, G. Marceau, A. Rudys, L. Kavraki, and D. Wallach. Robotics-based location sensing using wireless ethernet. In *Proceedings of MOBICOM2002*, Atlanta, Georgia, USA, September 2002.
- [7] N. Lesh and O. Etzioni. A sound and fast goal recognizer. In *Proceedings of IJCAI95*, pages 1704–1710, Montreal, Canada, 1995.
- [8] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil. Landmarc: Indoor location sensing using active rfid. In *IEEE International Conference in Pervasive Computing and Communications 2003 (PerCom 2003)*, Dallas, TX, USA, March 2003.
- [9] D. J. Patterson, L. Liao, L. Fox, and H. Kautz. Inferring high-level behavior from low-level sensors. In *Proceedings of UBIComp2003*, Seattle, WA, 2003.
- [10] D. V. Pynadath and M. P. Wellman. Probabilistic state-dependent grammars for plan recognition. In *Proceedings of the Sixteenth Conference on UAI*, pages 507–514, San Francisco, CA, 2000.
- [11] T. Roos, P. Myllymaki, H. Tirri, P. Misikangas, and J. Sievanen. A probabilistic approach to WLAN user location estimation. *International Journal of Wireless Information Networks*, 9(3):155–164, July 2002.
- [12] Jie Yin, Xiaoyong Chai, and Qiang Yang. High-level goal recognition in a wireless lan. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04)*, San Jose, CA USA.
- [13] M. Youssef and A. Agrawala. Handling samples correlation in the horus system. In *IEEE InfoCom 2003*, Hong Kong, March 2004.
- [14] M. Youssef, A. Agrawala, and U. Shankar. WLAN location determination via clustering and probability distributions. In *Proceedings of IEEE PerCom2003*, March 2003.