

Inferring Long-term User Properties based on Users' Location History

Yutaka Matsuo and Naoaki Okazaki and Kiyoshi Izumi
and Yoshiyuki Nakamura and Takuichi Nishimura and Kôiti Hasida

National Institute of Advanced Industrial Science and Technology
y.matsuo@aist.go.jp, {kiyoshi, nmura, taku}@ni.aist.go.jp, hasida.k@aist.go.jp

Hideyuki Nakashima

Future University, Hakodate
h.nakashima@fun.ac.jp

Abstract

Recent development of location technologies enables us to obtain the location history of users. This paper proposes a new method to infer users' long-term properties from their respective location histories. Counting the instances of sensor detection for every user, we can obtain a *sensor-user matrix*. After generating features from the matrix, a machine learning approach is taken to automatically classify users into different categories for each user property. Inspired by information retrieval research, the problem to infer user properties is reduced to a text categorization problem. We compare weightings of several features and also propose sensor weighting. Our algorithms are evaluated using experimental location data in an office environment.

1 Introduction

Context-aware computing are gaining increasing interest in the AI and ubiquitous computing communities. To date, numerous approaches have been taken to recognize and model a user's external context, for example one's location, physical environment, and social environment, to provide context-dependent information. Though "context" is a slippery notion [Dourish, 2004], it is promising if we can recognize and adapt to aspects of users such as their activities, general interests, and current information needs [Jameson and Kruger, 2005]. Such user models are useful for personalized information services in ubiquitous computing.

Recently, location information has become widely available both in commercial systems and research systems. Devices such as Wi-Fi, Bluetooth, low-cost radio-frequency tags and associated RFID readers, and ultrasound devices all provide location-based information support in various situations and environments. One early and famous project was Active Badge [Want *et al.*, 1992]. Since that work, numerous studies of users' activity recognition and location-aware applications have been developed using location and other sensory information in the context of ubiquitous and mobile systems [Wilson, 2003; Hightower *et al.*, 2005; Ashbrook and Starner, 2003; Liao *et al.*, 2005; Lester *et al.*, 2005; Wilson and Philipose, 2005]. In these studies, user models are sometimes implicitly assumed. For example, being

in a laboratory might imply working behavior for laboratory members. While for different types of users such as guests for a campus tour, being in a laboratory implies sightseeing behavior. Therefore, user modeling and behavior detection are mutually complementary: if we have a more precise user model, we can guess more precisely the user behavior, and vice versa. Automatically obtaining a user model will bootstrap activity recognition in a ubiquitous environment to enable context-aware information services.

Toward user modeling for ubiquitous computing, several studies have been done in recent years. Heckmann proposes the concept of *ubiquitous user modeling* [Heckmann, 2005]. He proposes a general user model and context ontology GUMO and a user model and context markup language *UserML* that lay the foundation for inter-operability using Semantic Web technology. Carmichael *et al.* proposes a user-modeling representation to model people, places, and things for ubiquitous computing, which supports different spatial and temporal granularity [Carmichael *et al.*, 2005].

This paper describes an algorithm to infer a user's long-term properties such as gender, age, profession, and interests from location information. The system automatically learns patterns between users' locations and user properties. Consequently, the system can infer properties: it can automatically produce a user model of a new user coming to the environment. We show that some properties are likely to be inferred and others are difficult to infer. The algorithm is useful in various ubiquitous computing environments to provide user modeling for personalized information services.

We address users' long-term properties, especially among many user-modeling dimensions. Kobsa lists frequently found services of user-modeling, some of which utilize users' long-term characteristics such as knowledge, preference, and abilities [Kobsa, 2001]. Jameson discusses how different types of information about a user, ranging from current context information to the user's long-term properties, can contribute simultaneously to user adaptive mechanisms [Jameson, 2001]. In the ontology GUMO, long-term user model dimensions are categorized as demographics such as age group and gender, personality and characteristics, profession and proficiency, or interests such as music or sports. Some are basic and therefore domain independent, whereas others are domain dependent.

Our algorithm is so simple that it is applicable to many

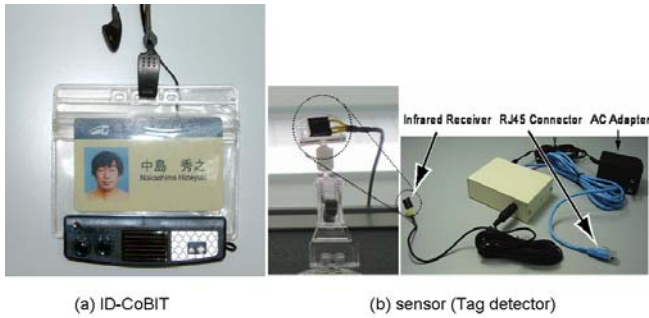


Figure 1: ID-CoBIT and sensors

types of location data. Inspired by research on information retrieval, we regard the problem of inferring users' properties as text categorization problems. Support vector machine (SVM) is applied to the problem with various feature weighting methods compared in the paper. Our study is evaluated on empirical data obtained through one-week experiments at our research institute. We collected location data of more than 200 users (staffs and guests). The *ID-CoBIT* system consisting of namecard-type infrared transmitters and sensors is installed in the environment to recognize users' location.

The paper is organized as follows. The next section introduces the *ID-CoBIT* system and describes location data. The proposed algorithm for user modeling from location information is explained in Section 3. Analyses of the results are made in Section 4. We also propose measurements of sensor importance in Section 5. Related works and discussion are described in Section 6. Finally, we conclude the paper.

2 Location information

In our research, location information is obtained using the *CoBIT* system. This section briefly overviews the *ID-CoBIT* system and describes experiments to collect location data.

2.1 ID-CoBIT

The Compact Battery-less Information Terminal (*CoBIT*) is a compact information terminal that can operate without batteries because it utilizes energy from the information carrier [Nishimura *et al.*, 2004]. The *ID-CoBIT* is a terminal integrating *CoBIT* with an infrared (IR) ID tag and a liquid crystal (LC) shutter. Figure 1(a) depicts an *ID-CoBIT*, which is useful as a namecard holder. The ID detector for *ID-CoBIT* is a single detector type IR sensor, as shown in Fig. 1(b). The sending cycle of a tag is about 3 s. The effective distance is 3–5 m. Detailed specifications are available in [Nakamura *et al.*, 2003].

The *ID-CoBIT* system provides location-based information support in the environment such as exhibitions, museums, and academic conferences [Nishimura *et al.*, 2004]. Users can download information depending on their location and orientation, mainly via voice information. The entire architecture is shown in Fig. 2. Although the *ID-CoBIT* system has multiple communication channels, in this study we use only the IR LED on an *ID-CoBIT* and IR sensors in the environment. We specifically address obtaining locations of users without disturbing usual daily behavior.

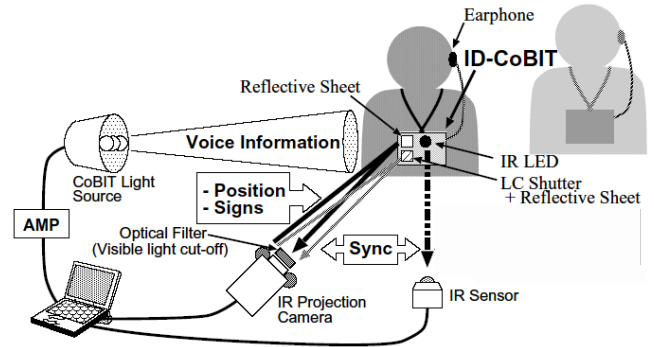


Figure 2: *ID-CoBIT* system.

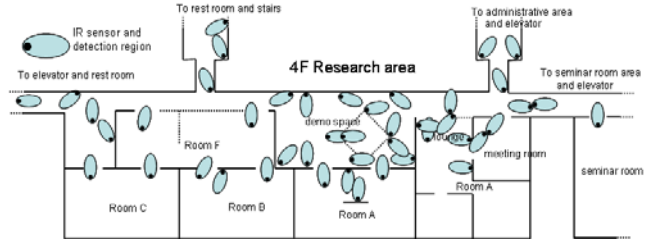


Figure 3: Sensor allocation map for a part of the fourth floor.

2.2 Experiments

The *ID-CoBIT* system is installed in an office environment of our research institute. Location data were collected at AIST Tokyo Waterfront Research Center, from February 16, 2004 (Mon) through February 20 (Fri). In all, 94 sensors are installed at the first floor and the fourth floor. On the first floor, we have entrances, reception areas, lobbies, and lounges. The fourth floor is our main office, consisting of a research area and an administrative area. The sensor allocation map on the fourth floor is shown partly in Fig. 3, which is about 1000 square meters, or about a third of the entire covered area. Every working staff member on these floors was delivered an *ID-CoBIT*, which they continued to wear during the period.

We also delivered *ID-CoBIT*s to and obtained location information of 170 guests who visited the institute temporarily during the period.

After the experiments, we analyzed the location data. The detection instances of all sensors were 24317 times: 20273 times of staff, 4044 times of guests. The number was almost constant each day. On average, a staff member was detected 431.3 times; a guest was detected 23.8 times. Because the location information and user properties of staffs are quantitatively and qualitatively better than those of guests, we use the staff information in this paper.

For obtaining users' long-term properties, we manually surveyed user attributes for all staff such as age, work frequency, room, and whether they smoke or not. The user properties used in our study are shown in Table 1. We chose demographic properties such as AGE and POSITION, domain-dependent properties such as TEAM and WORK-FREQUENCY, and user-specific properties such as COFFEE and SMOKING.

We elaborate these properties considering usefulness in our domain and also versatility in other domains: First, AGE and

Table 1: User properties.

user property	range
AGE	under24, 24-29, 30-34, 35-39, over40
POSITION	sc*, full-time researcher, part-time researcher :
	technical-staff, temporary-staff
TEAM	research-group-A, -B, -C, -D, secretaries, administrators
WORK-FREQUENCY**	high, middle, low
COFFEE***	high, middle, low
SMOKING	yes, no
ROOM ⁺	A, B, C, D, E, F
COMMUTING ⁺⁺	station-A, station-B

* SC stands for steering committee. ** Because of the free time system of this work environment, working time and commuting frequency depend on the person. *** How often one drinks coffee. ⁺ Working room at one's desk. ⁺⁺ Two train stations on two lines are accessible from this Institute.

POSITION are important properties especially in Japan; in the Japanese culture, age and position make large differences in communication such as using respect language and behavior. As it is often inappropriate and impolite to ask a user about the age and position directly, it is useful for the system to infer such properties. TEAM and WORK-FREQUENCY can be seen as users' interests in the research domain. Because team organization is flexible in our institute, they reflect well the researcher's interest. COFFEE and SMOKING are useful for guests. If the system can recognize that a guest likes coffee or smoking, it can suggest appropriate restaurants or cafes in break time. Because we are often asked "do you like coffee or tea?" or "do you smoke or not?" (in Japan), it indicates the usefulness of the properties in our daily lives. Lastly, ROOM and COMMUTING are for navigation. Knowing the properties, the system can infer in which room a researcher might be (even if he does not wear the COBIT), or recognize whether he/she goes home or not.

3 Inference of User Properties

In this section, we propose our algorithm to infer user properties based on their respective location histories. We first describe how to reduce the prediction problem of user properties into a text categorization problem. Then, the feature design for machine learning is explained.

3.1 Reduction to a Text Categorization Problem

When a sensor detects a user, the *SensorID* and *UserID* are obtained each time a sensor detects a user. Counting the number of detections, we can build a matrix that represents how many times each sensor detects each user. We call it a *sensor-user matrix*. Denoting the number of users as n and the number of sensors as m , the sensor-user matrix is an $n \times m$ matrix W . We denote W_{ij} as the element of W , i.e., the number of detections of user u_j by sensor s_i . The illustration of sensor detection to a sensor-user matrix is shown in Fig. 4.

Next, we consider user properties. For example, a user property of whether the user drinks coffee or not (the COF-

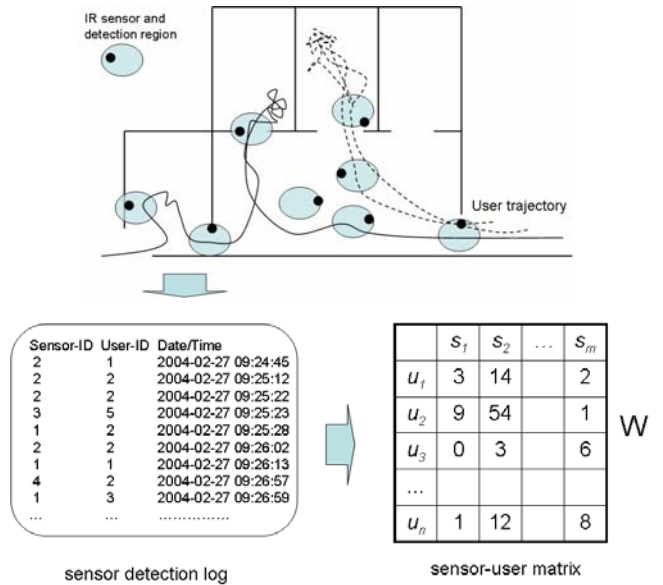


Figure 4: Illustration of sensor detection and the sensor-user matrix.

FEE property) can be represented as $\{yes, no\}$ or $\{1,0\}$. Assuming that three users have the values *yes*, *no*, and *yes* for the property, we have the following table as a training set.

	s_1	s_2	s_3	s_4	COFFEE
u_1	1	2	2	4	1
u_2	1	0	2	0	0
u_3	3	2	0	0	1

Then, when a new user u_4 comes and the detection frequencies are observed, a prediction problem arises. Is the COFFEE property of the user 1 or 0?

	s_1	s_2	s_3	s_4	COFFEE
u_4	2	2	0	0	?

From the training set, classification can be learned using machine-learning techniques. If we take nearest neighbor method, the most similar one to u_4 seems to be u_3 (though it depends on the similarity measure). Therefore, the method outputs 1.

This approach is justified using the following example: Let us consider a situation in which sensor s_2 is installed in front of a coffee server. Then, frequent detection by s_2 means that the user comes frequently to the coffee server, which might imply that the user likes coffee. We cannot know in advance which sensors are important for classification; they might be those in front of a coffee server, the ones in front of a vending machine, or those that are completely unexpected. In any case, classification is learned through sensor detection data and the performance is evaluated by k -cross validation or the leave-one-out method, where each part of the training data is used repeatedly as both initial training data and test data.

The obtained problem closely resembles a text categorization problem. A document is often represented by a word vector (or a *bag of words*) in which each word in the document is weighted by some word weighting; all structure and linear ordering of words in the document are ignored. The term-document matrix (or a document-by-word matrix [Manning

and Schütze, 2002]) resembles our sensor-user matrix W in that we have n documents and m words in which each user corresponds to a document and each sensor corresponds to a word. In a text categorization task, categories are annotated to each document, which can be considered as user properties in our problem. The classification is learned and used to infer the category based on the word vectors. Therefore, the user-modeling problem from location information is reduced to a (multi-label) text categorization problem under the proper assumptions and simplifications.

Text categorization is typically attained using several classification techniques. We employ support vector machine (SVM) as a learner, which creates a hyperplane that separates the data into two classes with the maximum-margin [Vapnik, 1995]. The SVMs offer two important advantages for text categorization: term selection is often unnecessary because SVMs tend to be fairly robust to overfitting. In addition, there is a theoretically motivated, “default” choice of parameter setting [Joachims, 1998]. These benefits are also provided by our user-modeling problem.

3.2 Feature Design

In the context of text categorization, *tf-idf* is frequently used as feature weighting, which encodes the intuition that (i) the more often a word occurs in a document, the more it is representative of its content, and (ii) the more documents in the word occurs in, the less discriminating it is. In our studies, it is rephrased as follows: (i) the more often a sensor detects a user, the more it is representative of the user’s characteristics, and (ii) the more users a sensor detects, the less discriminating it is.

The *tf-idf* weighting function tailored to our case is defined as $tfidf(s_i, u_j) = freq(s_i, u_j) \times idf(s_i)$ where $freq(s_i, u_j)$ is the number of detections of users u_j by sensor s_i . The $idf(s_j)$ is defined as $idf(s_i) = \log(n/uf(s_i))$ where $uf(s_i)$ is the number of users that sensor s_i detects (corresponding to document frequency). A sensor that detects many users has high $uf(s_i)$ value, and therefore a low $idf(s_i)$ value. In an extreme case, a sensor detecting all n users has a zero idf value as $\log(n/n) = 0$.

Aside from *tf-idf* weighting, several ways of feature weighting are possible. We compare typical weighting methods that are often used and compared in information retrieval. The following is a list of feature weighting methods that we use:

- Frequency (number of detections): $w_{ij} = freq(s_i, u_j)$
- Binary: $w_{ij} = \begin{cases} 1 & \text{if } freq(s_i, u_j) > f_{thre} \\ 0 & \text{otherwise} \end{cases}$ where f_{thre} is a threshold. In this paper, we determine $f_{thre} = 1$ through preliminary experiments.
- IDF: $w_{ij} = \begin{cases} idf(s_i) & \text{if } freq(s_i, u_j) > f_{thre} \\ 0 & \text{otherwise} \end{cases}$
- TFIDF: $w_{ij} = tfidf(s_i, u_j)$

For the weights to fall in the $[0,1]$ interval and for the vectors to be of equal length, the weight can be normalized by cosine normalization, given as $w_{ij}^{normalized} = w_{ij} / \sqrt{\sum_{i=1}^m (w_{ij})^2}$.

Table 2: Classification performance depending on various feature weighting

	F-value(%)	Recall(%)	Precision(%)
FREQ	44.45	73.56	37.75
BINARY	43.92	65.83	41.62
TFIDF	44.28	71.62	37.38
IDF	44.37	68.45	45.33
N-FREQ	54.46	68.83	49.23
N-BINARY	40.73	63.80	40.97
N-IDF	41.23	61.02	41.46
N-TFIDF	53.00	65.50	47.88

Therefore, we have 4×2 feature weighting methods, which are compared in the next section. We call those: FREQ, BINARY, IDF, TFIDF, N-FREQ, N-BINARY, N-IDF, and N-TFIDF. Although normalized *tf-idf* (N-TFIDF) is known to perform well for text categorization, different results are revealed in our user-modeling problem.

4 Evaluation

For each user property shown in Table 1, a categorization problem is generated. More exactly, because SVM is fundamentally applicable to the two-classes problem, a problem is generated for each value of the property. For example, the AGE property can take five values: five classification problems are generated. We make positive and negative classes for each value, say `under24`, i.e., those who are under 24 and those who are not. Thus the obtained classifier will classify people into those who are under 24 and those who are not. The SVM is used to learn the categorization and the performance is evaluated by leave-one-out. We employ a radius basis function (RBF) kernel, which performs well in our preliminary experiments.¹

Average performances on all categorization problems are shown in Table 2. For example, if we use FREQ as a feature weighting, the recall is 73.56%, meaning that we can detect 73.56% of persons with a property having a certain value. As a baseline, we investigate the performance of the straightforward classifier that always outputs positive; F-value is 38.2% and precision is 23.93%. Thus our method is much better than the baseline. As for feature weighting, N-FREQ has the highest F-value, and N-TFIDF is the second best. Normalization seems to function effectively for either feature weighting: it might alleviate the difference of detection frequency among users that was caused by the difference on the working time or individual device/usage characteristics. Depending on feature weighting, the performance varies as much as 10 points, thereby emphasizing the importance of feature weighting for user modeling.

In text categorization, normalized *tf-idf* works well and normalized frequency does not compete [Joachims, 1998]. In our case, N-TFIDF performs well, but N-FREQ performs the best. Thus the result is not completely identical to those in the text categorization literature. The reason can be considered as follows: compared to documents that have many functional

¹Other kernels, such as linear and polynomial kernels, produce similar results overall; the results worsen by a few points.

words and popular words with less information, location data suffers less from such a problem. Therefore, a naive approach using a normalized frequency might work well.

Generally, recall is about 70% and precision is less than 50%. However, we have much better results for a particular set of user properties. For **SMOKING**, **ROOM** and **COMMUTING**, the F-values are as high as 64.13%, 67.00%, and 61.86% respectively with about 60-90% recall and 50-80% precision. The **under24** and **over40** values of **AGE**, most values of **TEAM**, and high value of **COFFEE** are all more than 10 points greater than the baseline. Some of them has more than F-values of 80% with 70-100% recall and 50-80% precision.

In summary, some user properties, such as **TEAM** and **ROOM**, can be predicted effectively using solely location information. To some degree, **AGE**, **COFFEE**, and **SMOKING** are also predictable. **POSITION** and **WORK-FREQUENCY** are difficult to predict.

We investigated feature weights from the learned models and found that some sensors are unexpectedly important; if we take **COFFEE** property for example, the ones around a coffee server are of the fifth and ninth importance among all sensors. The most important one was that in front of a small table where people gather for break. Some corridors are also recognized as important. Surprisingly a sensor exactly in front of the coffee server was slightly negatively weighted; it may be because around the sensor there is a copy machine and a door, thus the detection has little information for the property. These results show the limitation of our presumption for user behaviors and effectiveness of our approach.

5 Sensor Weighting

In actual use-cases, it is not always possible to prepare training data consisting of users' location histories and user properties. Then, the question arises: is there a way to find out whether a sensor is useful for future user modeling in advance without training data? In the real world situation, we often change sensor locations depending on actual user behaviors. Therefore it is useful if we can know the importance of sensors for future user modeling so that we can properly choose sensors to fix locations. This section describes an approach to measure the usefulness of sensors using only location histories. It is similar to *keyword extraction* for indexing documents for future retrieval.

5.1 Importance of Sensors

A sensor that does not detect users at all is almost useless, at least for user modeling purposes. Therefore, one definition to measure the usefulness, or the importance, of a sensor is simply the total number of detections: its frequency of detection. Alternatively, sensors that detect many different users might be important.

The importance of a sensor is understood as follows: The user-modeling performance becomes better than the other sets of sensors if we have a set of more "important" sensors. In the context of information retrieval, several studies have examined finding good indexing terms for document categorization. Better indexing terms improve categorization performance [Mladenic *et al.*, 2004].

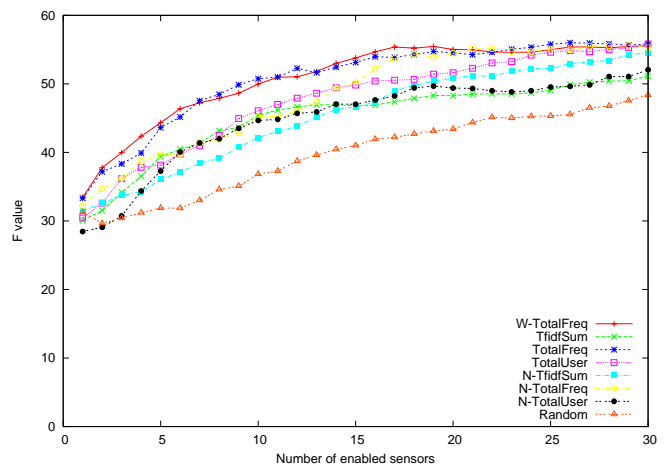


Figure 5: Number of enabled sensors versus F-value.

We compare several importance measures for a sensor derived from text categorization studies. The importance of sensor s_i is defined as follows: (i) Overall frequency (TOTALFREQ): $w(s_i) = \sum_{j=1}^n freq(s_i, u_j)$ (ii) Total detected users (TOTALUSER): $w(s_i) = uf(s_i)$ (iii) Total *Tf-idf* sum (TFIDFSUM): $w(s_i) = \sum_{j=1}^n tfidf(s_i, u_j)$ These functions can be normalized and taking a summation over every user, denoted as N-TOTALFREQ, N-TOTALUSER, N-TFIDFSUM.

In addition, we use another importance measure, called weighted frequency (W-TOTALFREQ), following the intuition that a sensor that detects users who are detected by fewer sensors might be more important, as (iv) Weighted frequency (W-TOTALFREQ): $w(s_i) = \sum_{j=1}^n freq(s_i, u_j) \times \log(m/sf(u_j))$, where $sf(u_j)$ represents the number of sensors that detect u_j . This can be regarded as a *tf-idf* measure on the transposed sensor-user matrix. In summary, we have seven sensor importance measures.

5.2 Comparison and Results

Assume that we tentatively disable all sensors and enable them one by one in decreasing order of sensor importance. Eventually all sensors are enabled and the results will coincide in any case. However, if a sensor weighting method is superior to the others, the performance will improve faster. If sensor weighting is poor, the performance grows no faster than random selection of sensors does. This approach to evaluate feature selection is found in text categorization research [Joachims, 1998; Mladenic *et al.*, 2004].

Figure 5 shows how the categorization performance changes over the number of enabled sensors. We used N-FREQ for feature frequency, and we used user properties that are shown to be predictable as shown in Section 4. In the figure, the undermost line (RANDOM) is plotted by selecting sensors randomly: it is shown as a baseline. The best performance is obtained by W-TOTALFREQ and TOTALFREQ. Other methods such as TOTALUSER, TFIDFSUM, and normalized series (N-TOTALFREQ, N-TOTALUSER, and N-TFIDFSUM) are better than RANDOM, but not as much as the best two.

Sensor weighting is beneficial in several situations: we can

identify which sensors might contribute to user modeling before learning the user properties. We can move sensors with low importance to elsewhere. These configurations of sensor allocation yield better performance during future user modeling.

6 Related Works and Discussion

Hightower distinguishes symbolic location systems and physical positioning technologies [Hightower and Borriello, 2001]. Our algorithm is applicable to symbolic location data. Therefore, some preprocessing is necessary for physical positioning data. For that purpose, studies to cluster position data into significant locations [Ashbrook and Starner, 2003] are useful. Anonymous sensors are applicable to our approach if they are used with ID-sensors, as proposed in [Schulz *et al.*, 2003].

We discard timestamps of sensor detection. We are aware that this is a crucial abstraction. Nevertheless, we persisted in our approach for two reasons: First, there are numerous alternatives for converting timestamped sensory data into features. Tailored heuristic rules might improve the results, but we want to retain simplicity in our algorithm to protect its general applicability to many location data and many domains. Second, our algorithm is mainly inspired by works in information retrieval. We discard the ordering of sensor detections so that correspondence of the data structures is maximized. Considering that we would have increasing amount of location data in the real world, simplicity and scalability of information retrieval methods are of great use. For example, in Japan people use RFID cards when taking trains and shopping; almost every cell phone has GPS and broad band communication. In this environment, a vast amount of user location data is potentially available, which needs simple and scalable processing. However, we do not disregard the usability of timestamps; actually they have much information and can be used to improve our results. Our contribution in this paper is to show a bridge between techniques in information retrieval and ubiquitous computing.

Our algorithm can infer user properties if given location data history. A promising application domain of our algorithm is event spaces [Nishimura *et al.*, 2004] such as conferences and business showcases, and large-scale shopping malls. Frequently, data mining of sales data is conducted using user demographic properties, which can be inferred by location data.

7 Conclusion

This paper proposes a new method to infer long-term user properties from a user's location history. Only the detection frequency is used. Machine learning techniques are applied to learn the pattern. Some user properties are well predictable. We also propose sensor weighting, which enables better allocation of sensors for future uses of modeling.

The algorithms in this paper are inspired by information retrieval. Because of the (proposed) structural similarity between sensor-user matrix and term-document matrix, we consider that many information retrieval techniques are applicable to sensory data. User modeling in ubiquitous computing

research will contribute greatly in AI studies for modeling and recognizing human behaviors.

References

- [Ashbrook and Starner, 2003] Daniel Ashbrook and Thad Starner. Using GPS to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing*, 7(5), 2003.
- [Camichael *et al.*, 2005] D. J. Camichael, J. Kay, and B. Kummerfeld. Consistent modelling of users, devices and sensors in a ubiquitous computing environment. *User Modeling and User-Adapted Interaction*, 15(3-4):197–234, 2005.
- [Dourish, 2004] P. Dourish. What we talk about when we talk about context. *Personal and Ubiquitous Computing*, 8(1), 2004.
- [Heckmann, 2005] D. Heckmann. *Ubiquitous Use Modeling*. Ph.d thesis, University of Saarland, 2005.
- [Hightower and Borriello, 2001] Jeffrey Hightower and Gaetano Borriello. Location systems for ubiquitous computing. *IEEE Computer*, 34(8):57–66, August 2001.
- [Hightower *et al.*, 2005] Jeffrey Hightower, Sunny Consolvo, Anthony LaMarca, Ian Smith, and Jeff Hughes. Learning and recognizing the places we go. In *Proc. UbiComp2005*, 2005.
- [Jameson and Kruger, 2005] A. Jameson and A. Kruger. Preface to the special issue on user modeling in ubiquitous computing. *User Modeling and User-Adapted Interaction*, (3–4), 2005.
- [Jameson, 2001] A. Jameson. Modeling both the context and the user. *Personal Technologies*, 5, 2001.
- [Joachims, 1998] T. Joachims. Text categorization with support vector machines. In *Proc. ECML'98*, pages 137–142, 1998.
- [Kobsa, 2001] A. Kobsa. Generic user modeling systems. *User Modeling and User-Adapted Interaction*, 11:49–63, 2001.
- [Lester *et al.*, 2005] J. Lester, T. Choudhury, N. Kern, G. Borriello, and B. Hannaford. A hybrid discriminative/generative approach for modeling human activities. In *Proc. IJCAI-05*, 2005.
- [Liao *et al.*, 2005] L. Liao, D. Fox, and H. Kautz. Location-based activity recognition using relational markov networks. In *Proc. IJCAI-05*, 2005.
- [Manning and Schütze, 2002] C. D. Manning and H. Schütze. *Foundations of statistical natural language processing*. The MIT Press, London, 2002.
- [Mladenic *et al.*, 2004] D. Mladenic, J. Brank, M. Grobelnik, and N. Milic-Frayling. Feature selection using linear classifier weights: interaction with classification models. In *Proc. SIGIR2004*, pages 234–241, 2004.
- [Nakamura *et al.*, 2003] Yoshiyuki Nakamura, Takuichi Nishimura, Hideo Itoh, and Hideyuki Nakashima. Id-cobit: A battery-less information terminal with data upload capability. In *Proc. of IECON2003*, 2003.

- [Nishimura *et al.*, 2004] T. Nishimura, H. Itoh, Y. Nakamura, Y. Yamamoto, and H. Nakashima. A compact battery-less information terminal for real world interaction. In *Proc. Pervasive 2004*, pages 124–139, 2004.
- [Schulz *et al.*, 2003] Dirk Schulz, Dieter Fox, and Jeffrey Hightower. People tracking with anonymous and ID-sensors using Rao-Blackwellised particle filters. In *Proc. IJCAI-03*, pages 921–928, 2003.
- [Vapnik, 1995] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag, 1995.
- [Want *et al.*, 1992] Roy Want, Andy Hopper, Veronica Falcao, and Jon Gibbons. The active badge location system. *ACM Transactions on Information Systems*, 10(1):91–102, January 1992.
- [Wilson and Philipose, 2005] D. Wilson and M. Philipose. Credible and inexpensive rating of routine human activity. In *IJCAI-05*, 2005.
- [Wilson, 2003] Daniel H Wilson. The narrator : A daily activity summarizer using simple sensors in an instrumented environment. In *Proc. UbiComp 2003*, 2003.