This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier’s archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
Exploiting indoor location and mobile information for context-awareness service

Hyun-Yong Noh, Jin-Hyung Lee, Sae-Won Oh, Keum-Sung Hwang, Sung-Bae Cho*

Department of Computer Science, Yonsei University, 134 Shinchon-dong, Seodaemun-ku, Seoul 120-749, South Korea

A R T I C L E   I N F O

Article history:
Received 28 December 2009
Received in revised form 6 February 2011
Accepted 20 February 2011
Available online 12 March 2011

Keywords:
Bayesian network
Decision tree
Context awareness
Indoor localization
Mobile device

A B S T R A C T

Personal mobile devices such as cellular phones, smart phones and PMPs have advanced incredibly in the past decade. The mobile technologies make research on the life log and user-context awareness feasible. In other words, sensors in mobile devices can collect the variety of user's information, and various works have been conducted using that information. Most of works used a user's location information as the most useful clue to recognize the user context. However, the location information in the conventional works usually depends on a GPS receiver that has limited function, because it cannot localize a person in a building and thus lowers the performance of the user-context awareness. This paper develops a system to solve such problems and to infer a user's hidden information more accurately using Bayesian network and indoor-location information. Also, this paper presents a new technique for localization in a building using a decision tree and signals for the Wireless LAN because the decision tree has many advantages which outweigh other localization techniques.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Mobile devices in the early generation played a role in calling only, providing limited services and functions. Those equipments have advanced significantly in the past decade, and now provide various functions including not only calling, but also camera, MP3 player, and the connection to Internet. In addition, various sensors in mobile devices enable to collect more information concerning our daily life. That information collected in mobile devices triggered in-depth investigation on user-context awareness (Abowd, Ebling, Hunt, Lei, & Gellersen, 2002; Dey, 2001; Dey & Abowd, 1999; Gemmell, Bell, & Lueder, 2006; Gemmell, Williams, Wood, Lueder, & Bell, 2004; Ljungstand, 2001; Loke, 2006; Raento, Oulavirta, Petit, & Toivonen, 2005; Nokia LifeBlog, http://www.nokia.com/lifeblog).

Many researchers on context awareness especially focused on the inference of user's hidden information which cannot be obtained explicitly. They usually used users' location information as one of the major clues, which is collected from a GPS receiver. However, GPS receivers cannot work in buildings, so indoor-location information is unavailable in existing works on context awareness. The lack of indoor-location information prevents the accurate inference of hidden information because most people usually perform a certain activity in a building. In addition, a building has many areas which are closely linked to a user's context.

Consider a building which has various areas such as cloth stores, a theater, restaurants, a bookstore, a workplace and so on. GPS devices cannot recognize the exact area in a building, which is closely linked to his context. For example, if the user...
stays in a workplace, we can infer that he/she works. In contrast, we can imagine that the user is having a lunch if he/she stays in a restaurant. However, GPS devices cannot provide those valuable evidences.

In order to overcome this drawback of traditional context awareness using GPS, this paper proposes a simple, yet powerful, technique for indoor localization specific to context awareness. In addition, this paper deals with an entire system to infer user’s context by adopting the Bayesian network and proposed indoor localization technique together.

2. Related works

There are many works on the life log and context awareness. In this section, we describe them briefly by dividing them into two groups: information collection and management, and inference of hidden information (see Table 1).

2.1. Information collection and management

The smartphone recently emerged as one of the promising mobile devices because of its versatility. It takes a role not only in a cellular phone, but also in a small computer. The ContextPhone provides a software platform which was developed for the smartphone to manage user context more efficiently. The software platform especially helps develop user-context-related applications on mobile phones more easily.

MyLifeBits is a Microsoft research project on the management of everyday information. Contemporary people live in avalanches of information such as written documents, photos, videos, and real-time information which is collected in various personal mobile systems. Inspired by the desire for organizing and manipulating personal data more efficiently, MyLifeBits proposed an idea to manage information efficiently by providing quick search, annotation clustering and so on.

There is another work which yokes SenseCam and MyLifeBits together. The work manipulates everyday information by using various sensors. This work uses SenseCam which consists of a camera and sensors to collect information. In addition, the information collected by the SenseCam is applied to MyLifeBits to manage user’s daily life efficiently.

Nokia offered an idea to deal with life logs which are collected by a smartphone. As mentioned before, mobile devices can collect information concerning various pieces of everyday information. However, that information is too affluent to manipulate effectively. Nokia provided a means to collect and to manipulate everyday information in a mobile device effectively.

2.2. Inference of hidden information

Beyond the simple organization and manipulation of everyday information, other works focused on the inference by using user’s unknown information. Whereabouts Diary uses information such as time, location and web-retrieved information to infer the type of a place (Castelli, Mamei, & Rosi, 2007). In other words, this work presented ideas to define meaningful semantic labels of the places (e.g., home, bar, workplace and school) by using GPS information, web-service, the behavioral pattern of users and Bayesian networks.

AniDiary is another research on the inference of user’s activities in his/her daily life (Cho, Kim, Hwang, & Song, 2007). In addition, based on inferred information, AniDiary automatically generate a user’s diary using cartoons and animations. There is another work which provided methodology to generate a story using cartoons. The research uses Petri-net which is one of the mathematical modeling languages and mobile context for generating a story (Lee & Cho, 2011).

Table 1
Summary of related works.

<table>
<thead>
<tr>
<th>Type</th>
<th>Research</th>
<th>Author (year)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information collection and</td>
<td>ContextPhone</td>
<td>Raento et al. (2005)</td>
<td>Provided a software platform which helps develop applications to manage user context effectively</td>
</tr>
<tr>
<td>management</td>
<td>MyLifeBits</td>
<td>Gemmell et al. (2006)</td>
<td>Proposed an application and ideas to manage various pieces of information including traditional records and real-time data in mobile devices</td>
</tr>
<tr>
<td>Passive capture and ensuing</td>
<td>Ggemell et al. (2004)</td>
<td>Offered an idea to manage everyday information by using SenseCam and MyLifeBits</td>
<td></td>
</tr>
<tr>
<td>issues for a personal lifetime store</td>
<td>Nokia Lifeblog</td>
<td>Nokia</td>
<td>Is a blog to mange life logs efficiently</td>
</tr>
<tr>
<td>Inference of hidden</td>
<td>Whereabouts diary</td>
<td>Castelli, Mamei, and Rosi (2007)</td>
<td>Provided an idea to infer the type of a place based on user’s information such as a behavioral pattern, GPS information and web-retrieved information</td>
</tr>
<tr>
<td>information</td>
<td>AniDiary</td>
<td>Cho, Kim, Hwang, and Song (2007)</td>
<td>Presented an idea to infer users’ hidden information based on their information collected in a mobile device. Additionally, it provides a service of the automatically-generated diary by using that information</td>
</tr>
<tr>
<td>Petri-net to generate a story in cartoons</td>
<td>Lee and Cho (2011)</td>
<td>Dealt with a method to generate a story based on Petri-net and mobile contexts</td>
<td></td>
</tr>
<tr>
<td>Modular Bayesian Networks</td>
<td>Hwang and Cho (2006)</td>
<td>Provided a new model of the Bayesian network, called modular Bayesian networks, to infer semantic information and memory landmarks</td>
<td></td>
</tr>
</tbody>
</table>
Hwang proposed an idea to infer memory landmarks (Hwang & Cho, 2006). The research deals with a method to infer semantic information and memory landmarks, which cannot be obtained from a mobile device directly, by using a modular Bayesian networks.

3. The proposed system

In this section, we describe the logical structure of the proposed system (see Fig. 1). The system works through five phases: collection, data manipulation, localization, inference and visualization.

3.1. Collection phase

Data stored in a mobile device can be categorized into one of two types: automatically-generated information and information stored directly by a user. Mobile devices generate data to manage its system or to provide a user with additional information. For example, call logs, such as call duration records and the frequency of call in a day, are stored not by input from a user, but by the mobile device itself. Mobile devices generate and manage various pieces of automatically-generated information (e.g., number of phone numbers, number of SMS sent in a month, and so on).

On the other hand, information can be stored into a mobile device by direct input from a user. For example, a user can store his/her personal information such as age, sex, marital status, family information, job and so on. In the collection phase, the system collects all kinds of information stored in a mobile device for the inference, localization and visualization.

3.2. Data-manipulation phase

This phase plays a role in converting raw data, which is collected in the first phase, into available formatted data and then classifying the data (see Table 2). First, in this phase, the system transforms raw data into the formatted data that can be used in the system. Consider GPS data that is usually stored in the format of NMEA (National Marine Electronics Association). The system needs to convert time data in NMEA format, standard time, into local time. In addition, some pieces of raw data are useless for the system (e.g., delimiter, checksum data and satellite data). Therefore, the system needs additional processes such as the removal of useless part of data and the transformation of raw data to the available data.

Additionally, this phase involves classifying raw data into one of six types depending on its usage. After this phase is finished, the system can appropriately apply them into the operations such as inference by Bayesian network and localization.

![Fig. 1. Phases and data flow in the system.](image-url)
3.3. Localization phase

People usually tend to visit a place to do certain activity which is closely linked to the place. For example, students go to a lecture hall to take a lesson, or visit a gym to take exercise. Hence, location information plays a significant role in inferring user’s information such as activities, emotions, and physical state. However, as mentioned before, a localization using GPS is insufficient for the inference, so we focus on the indoor localization.

There are several ways to detect a user’s location in a building by using signals for wireless communication (Bahl & Padmanabham, 2000; Chen, Lee, & Lee, 2008; Krumm & Horvitz, 2004). Most of the conventional techniques are based on naïve implementation of rule-based system. However, we use the decision tree to detect a place where a user stays because of the following reasons:

1. Existing indoor localization techniques focus on information concerning the point or movement of an object. However, information concerning the role or kind of the area (e.g., library, restaurant, or home) is more valuable for context awareness rather than object-movement-related information.
2. Decision tree can cover a large area if signals for WLAN (Wireless LAN) are available.
3. Decision tree is flexible in terms of the expansion or modification of the tree. In other words, it can effectively adapt to changes in the position of AP (Access Point) by only adding or changing a tree branch, which contains information relevant to the change, without the modification in other part of decision tree.

The localization phase is divided into two steps: tree building and localization.

3.3.1. Tree building

Before detecting the location of a user, decision tree has to be built. There are several steps to build decision trees:

1. The first step is to define the entire area where the context awareness is performed (e.g., a campus, a town or a building).
2. The area defined in the first step is divided into smaller areas according to its type (e.g., library, restaurant, gym, park, and so on). The divided area has the following four characteristics: (a) the area is the unit of the indoor localization; (b) wireless signals should be available in this area; (c) this area has a single function such as a library, a workplace or a restaurant; (d) the area might be a part of the building or a building itself.
3. Decision tree is built based on SSID (Service Set Identifier) and SS (Signal Strength): SSID is an identifier to differentiate each AP and SS means the magnitude of an electronic signal. SSIDs and SSS are collected in some points of a divided area defined in step 2 (e.g., the center of a room or the edges of a room). Surely the more signal information is collected in a place, so the reliable result can be achieved.
4. The same ID is assigned to a group of signal information in a divided area to distinguish from other areas’ information. Then, signal information of each area is grouped depending on the building to which it belongs.
5. There are some algorithms such as CHAID, CART and C4.5 to create the decision tree (Devroye, Györfi, & Lugosi, 1996; Kass, 1980; Quinlan, 1993). Based on one of these algorithms, each decision tree is made using signal information, so each building has its own decision tree.
3.3.2. Localization

After building decision trees, the system can detect the area where a user stays. Two steps follow to recognize the area: the first step is to find out a building where a user stays by using a GPS-equipped mobile device. By checking the GPS device’s disconnection from satellites for minutes or hours, we can recognize whether a user stays in a building.

After a building is detected, indoor localization is performed by using a decision tree specific to the building. SSID and SS in a place where a user stays are applied into the decision tree. The tree searches for the node whose elements have the most analogous to the SS and SSID, and retrieves the location information. There are five steps for indoor localization (see Fig. 2).

---

**Fig. 2.** Algorithm for the indoor localization using a decision tree.

1) WHILE(Set\_stream $\neq \emptyset$)
2) Set\_part\_stream $\leftarrow$ Copy the first $\delta$ data from Set\_stream
3) Set\_stream $\leftarrow$ Set\_stream $-$ Set\_part\_stream
4) Set\_result $\leftarrow$ $\emptyset$
5) WHILE(Set\_part\_stream $\neq \emptyset$)
6) Signal $\leftarrow$ The first element in Set\_part\_stream
7) Set\_part\_stream $\leftarrow$ Set\_part\_stream $-$ Signal
8) Set\_result $\leftarrow$ Set\_result $\cup$ DECISION\_TREE(Signal)
9) RETURN the most frequent location in Set\_result

---

3.3.2. Localization

After building decision trees, the system can detect the area where a user stays. Two steps follow to recognize the area: the first step is to find out a building where a user stays by using a GPS-equipped mobile device. By checking the GPS device’s disconnection from satellites for minutes or hours, we can recognize whether a user stays in a building.

After a building is detected, indoor localization is performed by using a decision tree specific to the building. SSID and SS in a place where a user stays are applied into the decision tree. The tree searches for the node whose elements have the most analogous to the SS and SSID, and retrieves the location information. There are five steps for indoor localization (see Fig. 2).

---

**Fig. 3.** Structure of the Bayesian network for the inference of users’ hidden information.
SSIDs and SSs, which are collected by WLAN, form a stream denoted by Set\text{\textregistered}stream. This stream needs to be divided into slides for applying SSIDs and SSs to the decision tree. (The stream is a sequence of SSIDs and SSs, and the slide is a part of the sequence determined by a parameter.) Parameter s in the Fig. 2 means the size of slide. If the stream contains 100 elements and parameter s is 4, the stream will be divided into 25 slides containing four pairs of SSID and SS. After the first s pairs in a stream are divided, they are stored in the set Set\text{\textregistered}part\text{\textsubscript{stream}}.

The set, denoted by Set\text{\textregistered}result, storing the results of localization by using each pair of SS and SSID in a Set\text{\textregistered}part\text{\textsubscript{stream}} is initialized.

The first pair of SS and SSID in a Set\text{\textregistered}part\text{\textsubscript{stream}} is stored in Signal that is a space to store one pair of SS and SSID.

SS and SSID in the Signal are applied into the decision tree to recognize the location in a building. Results which involve the name or the type of location from the decision tree are stored into Set\text{\textregistered}result.

After all elements in the Set\text{\textregistered}part\text{\textsubscript{stream}} are applied into the decision tree, the most frequently inferred location is selected in the Set\text{\textregistered}result.

Given the accuracy (10–30 m) and the refresh time of a normal civil GPS device, the proposed method for detecting the building where the user stays can be very inaccurate. Moreover, the identity of the building could probably be detected just using WiFi network signals.

### 3.4. Inference phase

In the inference phase, the system infers user-hidden information by using the Bayesian network (Coppola et al., 2005; Hwang & Cho, 2006). The Bayesian network we propose has the logical structure as shown in the Fig. 3 (Hwang & Cho, 2009). The Bayesian network consists of nodes which are connected by arcs. If it is the initial point of an arc, the node plays a role in a clue to infer information which the final node deals with. Consider nodes “Age” and “jobs”, which represent that a user’s age serves as one of the evidences to infer his/her job.

Each node can be categorized into one of three types: cause, intermediate and consequence: A cause node plays a role in only an evidence for intermediate or consequence nodes. The value of this node is set by using information that is collected in the earlier phase. Cause nodes are described in Table 3. The intermediate node can be used as an evidence to infer unknown information, or contains inferred information. For example, activities described in Table 4 can be used as an evidence of the set\text{\textregistered}result.

Inference node is not used as an evidence for other nodes, but only contains information to infer emotions or state. At the same time, activities are also the consequence of inference based on time, age, or contains inferred information. For example, activities described in Table 4 can be used as an evidence to infer whether a user is busy or not. Hence, Moving speed is a major evidence to infer user’s activities, especially, static activities such as “taking a lesson” and “having a meal”.

<table>
<thead>
<tr>
<th>Evidences</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Broadly influences on user’s emotions, physical state, activities, job and so on</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
</tr>
<tr>
<td>Children information</td>
<td></td>
</tr>
<tr>
<td>Address of workplace</td>
<td></td>
</tr>
<tr>
<td>User’s current activity</td>
<td>Means whether a user is currently calling or sending SMS. This information plays a role in an evidence to recognize whether a user is busy</td>
</tr>
<tr>
<td>(calling and sending a SMS)</td>
<td></td>
</tr>
<tr>
<td>Call duration for a month</td>
<td>Indicates how many time a user spent for calls in a month. Job-related information can be inferred based on this information. A salesman, for example, appears to spend more time to call than students</td>
</tr>
<tr>
<td>Frequency of incoming/outgoing calls in a month or day</td>
<td>Involves how many telephone numbers are stored in a mobile device. Like total time for calling, this information can be used to infer job-related information. To illustrate, a salesman would have more telephone numbers than a student</td>
</tr>
<tr>
<td>Total number of phone numbers</td>
<td>Is the most important information to infer user’s state, emotion, and activities because the activity is closely linked to a location where a user stays</td>
</tr>
<tr>
<td>addresses</td>
<td></td>
</tr>
<tr>
<td>Staying time</td>
<td>Indicates how much time passes after a user enters a building. This evidence helps infer whether a user perform certain activity concerning the area. For example, $P(\text{Taking_a_lesson} = \text{True}&amp;\text{Location} = \text{Lecture Hall}&amp;\text{Staying_time} = \text{1_hour}) &gt; P(\text{Taking_a_lesson} = \text{True}&amp;\text{Location} = \text{Lecture Hall}&amp;\text{Staying_time} = \text{5_minutes})$. The first probability is higher because an hour is usually assigned to a course</td>
</tr>
<tr>
<td>User’s speed</td>
<td>Helps infer whether a user stays in an area or not. Hence, Moving speed is a major evidence to infer user’s activities, especially, static activities such as “taking a lesson” and “having a meal”</td>
</tr>
<tr>
<td>Day</td>
<td>Are very important evidences for inference because the behavioral pattern or even emotions vary depending on day and time</td>
</tr>
<tr>
<td>Time</td>
<td></td>
</tr>
</tbody>
</table>
First, the daily-life-tracking program shows the sequence of events which occurred in a daily life. Fig. 4 shows the actual use of the program by applying student’s data collected by a mobile device. The program consists of six windows. Map window displays a user’s trace in a daily life. Each icon in the map is marked as time passes. The green icon indicates the position of a user at the time shown in current–time window, whereas gray icons represent the trace of the user before that time.

User-state window shows state, location and other personal information. In the indoor-location section, the information involving the indoor place where a user stayed is displayed. In this example, the user stayed in the first floor of the engineering building A and this window shows the picture of the first floor in the engineering building. In the picture-list window, all pictures which were taken by the user are arranged. Finally, the photo window shows a photo if there is a photo which was taken at the time shown in current–time window. Also, this section shows messages which are written by the user at that time.

The other program is the automatically-generated blog. It is made based on information which is stored in a mobile device or is inferred by Bayesian networks. This information includes time, visited location and building, state, message and so on. The automatically-generated blog service creates a blog chronologically based on this information.
4. Experiments

Three experiments are conducted to verify the performance and the usability of the system we develop: The performance of the decision tree for localization, the Bayesian network to infer users’ hidden information and poll to measure the system usability.

Experiments were carried out on the laptop with Intel® PRO/Wireless 3945ABG Network connection and Marvell Yukon 88E8055 PCI-E Gigabit Ethernet Controller. NetStumbler v.0.4.0 is used to collect SSIDs and SSs from APs, and WEKA 3.4.13 serves as making a decision tree (The University of Waikato, “WEKA”, www.cs.waikato.ac.nz/ml/weka). In experiments, four buildings having several areas are used. We assume that a building has several areas such as restaurant, cafeteria, conference room and so one (see Table 6).

Table 6
List of locations which are used in experiments.

<table>
<thead>
<tr>
<th>Building</th>
<th>Floor</th>
<th>Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engineering building A</td>
<td>2F</td>
<td>Laboratory, conference room</td>
</tr>
<tr>
<td>Engineering building B</td>
<td>1F</td>
<td>Restaurant1</td>
</tr>
<tr>
<td></td>
<td>2F</td>
<td>lobby1</td>
</tr>
<tr>
<td></td>
<td>5F</td>
<td>Lecture hall</td>
</tr>
<tr>
<td>Central library</td>
<td>B1</td>
<td>Cafeteria</td>
</tr>
<tr>
<td></td>
<td>6F</td>
<td>Library, Reading room</td>
</tr>
<tr>
<td>Students hall</td>
<td>1F</td>
<td>Lobby 2</td>
</tr>
<tr>
<td></td>
<td>4F</td>
<td>Restaurant2</td>
</tr>
</tbody>
</table>

Fig. 4. Daily-life-tracking service and automatically-generated blog service.

Fig. 5. Performance of the decision tree.
4.1. Performance of the decision tree

This experiment is to show the accuracy of localization in a building using transactions, pairs of SSID and SS. For each area, 100 transactions, which have 100 pairs of SSID and SS, are used to make a decision tree by applying them to WEKA. After making a decision tree, 1000 transactions are applied to the tree to evaluate the accuracy of the decision tree. Fig. 5 illustrates the accuracy of inference based on SSID and SS for each area. The graph shows the accuracy of 91.64%, 97.12% and 99.2% when slide size $s$ is 2, 4 and 8, respectively. The results look promising compared with the conventional techniques. Some naïve implementation of rule-based systems did not give us the accuracy of more than 90%. By adjusting the appropriate slide size, we could get acceptable performance with the decision tree.

4.2. Performance of the Bayesian network

This experiment is carried out to show the accuracy of activity inference when a user stays in a building. In this experiment, we use a part of the Bayesian network in Fig. 3. Based on four evidences, the Bayesian network infers one of four activities of user: (1) Studying activity means that a user is studying in a library or reading room. (2) Having a meal is inferred

Table 7
Performance of Bayesian network based on indoor-location information. $#_{TP}$: Number of true positive transactions, $#_{TN}$: Number of true negative transactions, $#_{FP}$: Number of false positive transactions, $#_{FN}$: Number of false negative transactions.

<table>
<thead>
<tr>
<th>Activities</th>
<th>$#_{TP}$</th>
<th>$#_{TN}$</th>
<th>$#_{FP}$</th>
<th>$#_{FN}$</th>
<th>Hit ratio (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Having a meal</td>
<td>960</td>
<td>5804</td>
<td>100</td>
<td>0</td>
<td>99.16</td>
<td>99.27</td>
<td>98.61</td>
</tr>
<tr>
<td>Taking a lesson</td>
<td>120</td>
<td>6706</td>
<td>38</td>
<td>0</td>
<td>98.54</td>
<td>90.57</td>
<td>100</td>
</tr>
<tr>
<td>Studying</td>
<td>2702</td>
<td>4104</td>
<td>20</td>
<td>38</td>
<td>99.45</td>
<td>75.95</td>
<td>100</td>
</tr>
<tr>
<td>Other activities</td>
<td>2924</td>
<td>3820</td>
<td>0</td>
<td>120</td>
<td>98.25</td>
<td>100</td>
<td>98.06</td>
</tr>
<tr>
<td>Total</td>
<td>6706</td>
<td>20,434</td>
<td>158</td>
<td>158</td>
<td>98.85</td>
<td>97.7</td>
<td>97.7</td>
</tr>
</tbody>
</table>

Fig. 6. Probabilities of each activity in a daily life.
when a user is having a meal in a restaurant. (3) Taking a lesson indicates that a user is taking a lesson in a lecture hall. (4) Other activity is a state other than aforementioned activities.

The experiment is conducted using transactions, each of which has four evidences: day, time, staying time and location. Transactions include all possible combination of evidences, which is \( \{\text{week}, \text{weekend}\} \times \{x\ (\text{hour}) | 1 \leq x \leq 24\} \times \{x \times 10\ (\text{minutes}) | 0 \leq x \leq 12\ \text{of staying time}\} \times \{x|x\ \text{is places in Table 6}\} \). To show various activities in the experiment, we assumed that each activity should be made within 2 h when we collected the data. However, this assumption is not critical and can be deleted. The performance is evaluated by showing true positive (TP), true negative (TN), false positive (FP), false negative (FN), hit ratio \((\frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}})\), precision \((\frac{\text{TP}}{\text{TP} + \text{FP}})\) and recall \((\frac{\text{TP}}{\text{TP} + \text{FN}})\). The table 7 shows the result of the experiment. Even though our task has some limitation for practical use, the hit ratio of 98.85% is extraordinary and we can use the proposed Bayesian network for the four activities.

In addition, the Bayesian networks are evaluated using the actual data in a daily life. Fig. 6 shows probabilities of each activity using the data. The figure consists of two parts: the left box includes information such as time, location and activities that are actually carried out by the user. The right box shows the probabilities of each activity that is inferred by the Bayesian network. As can be seen, most of the activities were recognized correctly.

### 4.3. System usefulness test

In this section, the poll on the application’s usability is described. The survey is conducted to evaluate not only the effectiveness and efficiency of the system, but also the degree of satisfaction of a user. The System Usability Scale (SUS) by John Brooke is used to show the usability of the system (Brooke, 1996). This poll consists of 10 questionnaire items on five-point scale as described in Table 8. Twenty-three subjects from the department of computer science at our university were asked to this survey. The subjects were from early twenty to mid thirty years old. To understand better the system usefulness, we will need more varieties of subjects, e.g. computer literacy, age, work type, etc.

For those who are not familiar with the SUS, we categorize the responses of the survey into two parts: positive response and negative response. We define responses to 4 and 5 scales for 1, 3, 5, 7 and 9 items and responses to 1 and 2 scales for 2, 4, 6, 8 and 10 items as positive. Those responses other than 3 scale and positive responses are defined as negative. Fig. 7 shows the difference between the two types of responses. The total SUS score gets to 71.3, which confirms the relative usefulness of the proposed system.

![Fig. 7. Result of survey based on positive and negative responses.](image-url)
5. Concluding remarks

Location information is a major evidence to recognize user context. However, a GPS receiver cannot work in a building, so it is impossible to recognize user context in a building. Furthermore, existing indoor localization techniques seem inappropriate for context awareness because they cannot recognize the role of the place where a user stays. For these reasons, this paper proposes a new indoor localization technique specific to context awareness.

In addition, we develop a system to infer user's hidden information by using the indoor localization and other everyday information, which is used for daily-life-tracking and automatically-generated blog service to manage the information more effectively.

Recently more useful information became available because versatile smartphones prevailed. Though indoor-location information helps infer user's hidden information, further research needs to focus on the manipulation of other valuable information stored in diverse mobile devices to better recognize user contexts.

Acknowledgment

This research was supported by the Converging Research Center Program through the Converging Research Headquarter for Human, Cognition and Environment funded by the Ministry of Education, Science and Technology (2010K001173).

References


H.Y. Noh is a researcher in the Department of Computer Science at Yonsei University. His research interests include context awareness using Bayesian networks and localization technique.

J.H. Lee received his B.S. degree in the Department of Computer Science at Yonsei University. His research interests include Bayesian network, and context-aware computing.

S.W. Oh is a researcher in the Department of Computer Science at Yonsei University. His research interests include the Bayesian network and context awareness.

K.-S. Hwang is a researcher at the LG Electronics, Inc. and the Softcomputing Laboratory in computer science at Yonsei University. His research interests include Bayesian networks and evolutionary algorithms for context-aware computing and intelligent agents. He received his M.S. and Ph.D. degrees in computer science from Yonsei University.

S.-B. Cho received the B.S. degree in computer science from Yonsei University, Seoul, Korea and the M.S. and Ph.D. degrees in computer science from KAIST (Korea Advanced Institute of Science and Technology), Taegun, Korea. He was an Invited Researcher of Human Information Processing Research Laboratories at ATR (Advanced Telecommunications Research) Institute, Kyoto, Japan from 1993 to 1995, and a Visiting Scholar at University of New South Wales, Canberra, Australia in 1998. He was also a Visiting Professor at University of British Columbia, Vancouver, Canada from 2005 to 2006. Since 1995, he has been a Professor in the Department of Computer Science, Yonsei University. His research interests include neural networks, pattern recognition, intelligent man-machine interfaces, evolutionary computation, and artificial life. Dr. Cho was awarded outstanding paper prizes from the IEEE Korea Section in 1989 and 1992, and another one from the Korea Information Science Society in 1990. He was also the recipient of the Richard E. Merwin prize from the IEEE Computer Society in 1993. He was listed in Who’s Who in Pattern Recognition from the International Association for Pattern Recognition in 1994.