

A Location Inference Algorithm Based-on Smart Phone User Data Modelling

Sang-il Kim*, Wan Jung*, Hwa-sung Kim*

*Department of Electronics and Communications Engineering, Kwangwoon University

rlatkd234@kw.ac.kr, jung-wan@kw.ac.kr, hwkim@kw.ac.kr

Abstract— Human life is rapidly changing around the world in terms of smart phones as the number of users is increasing in a surprising rate and the recent development of remarkable performances. And, there is wide range of smart phone services available based on the location information. However, levels of location based services currently provided are not satisfactory due to the GPS error rate. Though there have been many studies offering measurement technologies based on a variety of network infrastructure or technologies to compensate GPS errors, there still exist disadvantages such as large error rates indoors, energy consumption and need of additional infrastructure. Nowadays, location inference study using available user log data on smart phones, not the measurement technology based on a network, is being made to improve these shortcomings. Smart phone has a variety of user log data. Also, it can obtain variety of information about the surrounding environment around the user using the embedded hardware sensors. These user log and collected sensor data can be modelled as an ontology model then meanings of the information can be apprehended to infer the user location. This paper presents an algorithm to infer user location through an ontology modelling of the collected sensor data and the logged user information such as time, day of the week, location, and kind of applications used by user.

Keywords— location based service, location inference, context awareness, ontology, smart phone

I. INTRODUCTION

Due to the explosive growth of smart phone users of late throughout the world, living pattern has been changed in a very fast pace. Smart phone services are provided regardless of user location since there is no limitation of space in case of smart phones. High performance hardware and broadband mobile communications technologies also contributed to these phenomena. On the other hand, context awareness technology with communication and computing ability can apprehend user situation accurately and provide user with useful information. Many studies are being done to apply the context technology into variety of services such as the medical, educational, disaster relief or shopping, to meet future ubiquitous environment.[1] The context awareness technology enables the smart entities to compute and communicate so that they comprehend and adapt to dynamic environment changes to provide personalized services. However, application services currently provided to users are universal services that are not user-specific.

In order to provide accurate user-specific services, user data such as its accurate location and an understanding of possible

activities at that location are needed. Therefore an accurate user location is required in the context awareness services.[2] GPS sensor used in smart phone is relatively accurate in open space with 5 to 20 meters of accuracy, however, does not provide accurate data in indoors or tunnels, or in busy urban area. Variety of infrastructure to support the indoor positioning technology is in works to compensate the disadvantages of GPS, but it takes tremendous initial investment.[3] This study offers a user location inference method to compensate GPS location data with smart phone sensor data, user application logs, or key word search logs. A user position will be located within the error range of GPS by generating ontology through understanding of relations of the acceleration sensor and MIC sensor of smart phone, and application categories and the search keywords used.

This study includes the following contents: Chapter 2 takes a look at the current positioning technology. Chapter 3 defines data elements needed in location inference and generating ontology. Chapter 4 explains the location inference algorithm and evaluates its performances. And chapter 5 presents a conclusion and future study.

II. INDOOR AND OUTDOOR POSITIONING TECHNIQUES

Current location-based services by GPS are relatively accurate; however, services received indoors are not reliable due to signal interferences. More studies are concentrated to solve the problem of indoor positioning techniques than that of outdoor. Study on indoor positioning technique is done to compensate GPS errors or to obtain the user location information through variety networks. Wireless indoor positioning technology utilizes WLAN (Wireless Local Area Network), RFID (Radio Frequency Identification) or UWB (Ultra Wider Band). All three use RSSI (Received Signal Strength Indication). In case of WLAN, position is determined by comparing GPS data, MAC address of AP and the actual location measurement. Since the indoor positioning involves installation of receiver or AP, there exists environmental limitation arising from a lack of infrastructure.[4]

III. DEFINITION OF DATA ELEMENTS FOR POSITIONING & GENERATION OF ONTOLOGY

In this section, data elements of sensor data and user behaviour data provided by a smart phone for the location inference will be defined for the ontology modelling. And then, ontology will be generated based on the defined data elements of sensor data and user behaviour data.

A. Definition of Data Element for Location Inference

As mentioned earlier in the introduction, location inference technique obtains the user location from various data elements attained on the smart phone which represent user's behaviour such as kind of applications used, search keywords entered by a user and smart phone sensor data. This section will define data elements of location inference to generate the ontology.

There are physical and virtual sensors needed in the location inference. The physical sensors are physical measurement of GPS, Wi-Fi, and acceleration sensors and are provided by the smart phone sensors. The virtual sensor data are the kind of data provided by software applications or services running on smart phone, user data stored in the smart phone; such as user schedule, web log data or application usage log.

Table 1 represents the user behaviour data that are attainable by sensors on the smart phone. Using the user behaviour data can help compensate the user location measured by GPS. For example, if GPS measurement is done in the middle of a hospital, a school and a restaurant when there are a hospital, a school and a restaurant, then it is not possible to infer exact user location. With the user behaviour data defined in Table 1 applied, candidates of user location can be approximately determined. Since the user behaviour data is not enough to pinpoint the location, however, the additional user data such as stored search keywords and applications usage log information in the smart phone can be used to position the user location more accurately. Therefore, we defined the 'Keyword categories' and 'Application Categories' in order to use them for location inference as shown in Table 2.

TABLE 1. USER BEHAVIOR DATA ATTAINABLE BY THE SENSORS ON THE SMART PHONE

Sensors	Role	Provided User Behavior Data
GPS	Coordinate detection	User's Location
Wi-Fi	Wireless internet access	User's Location
Bluetooth	Short-range wireless network device access	User's Location, Search for people around
ACC	Motion acceleration measurements	User's movement
Light	Light Detection	Location of the smart phone
Vector	Direction measurement	Direction of the smart phone detection
Proximity	Proximity object detection	Proximity object detection
Magnetic	Orientation Measure	The direction of movement of the user

Table 2 represents the relationship between the location categories, search keyword categories and categories of used applications. Location categories are defined by categories obtained through 'Naver map API' category data. On the other

hand, key word and application categories are defined based on the location category obtained through 'Naver map API' category data. Key word categories are defined based on businesses class such as restaurants or cafes, because all key words are not location-specific and cannot be processed using natural language technique. Application categories are defined by the categories of 'Google Play store' that each application belongs.

For example, if user search words are either educational, academic or related to education, literature or reference material, the chance that user is at either school or university is very high.

TABLE 2. INTERRELATIONS OF LOCATION, KEY WORD AND APPLICATION CATEGORIES

Location Categories	Keyword Categories	Application Categories
High School, University..	Education / Academic	Education, Books & Reference, News & Magazines
Korean food, Japanese food, Families restaurants, Fast food..	Restaurant	Lifestyle & Local
Coffee shops, bakery..	Cafe	Games, Social content, News and Magazines
ATM, KB Kookmin Bank, Woori Bank..	Bank	Finance
GS, Gas station, SK, Gas stations..	Gas station	Travel and Local Information
Large discount stores, Distribution.. Convenience..	Corporate/Shopping	Business and shopping
Accommodation, Theme Park, Travel, Tourist..	Travel	Weather, Travel and Local information
Movies, Concerts, Theatre, Music..	Entertainment	Music and audio, Comics, Media and Video
Balls of sports, Golf..	Leisure & Sports	Health & Fitness
Hospital	Hospital/Medical	Medical
Traffic, Karaoke, reading room ..	Life/Amenities	Traffic
Health, Fitness, Home, wedding halls, beauty	Life/Health	Health & Fitness

B. Ontology Modelling for Location Inference

Ontology model defining the interrelations between each of the data elements for location inference is proposed to achieve location inference using the behaviour data in Table 1 and categories defined in Table 2. A narrative of who, when, what and how based on the obtained sensor data are necessary to position the location of user. 'Who' can be defined by the user

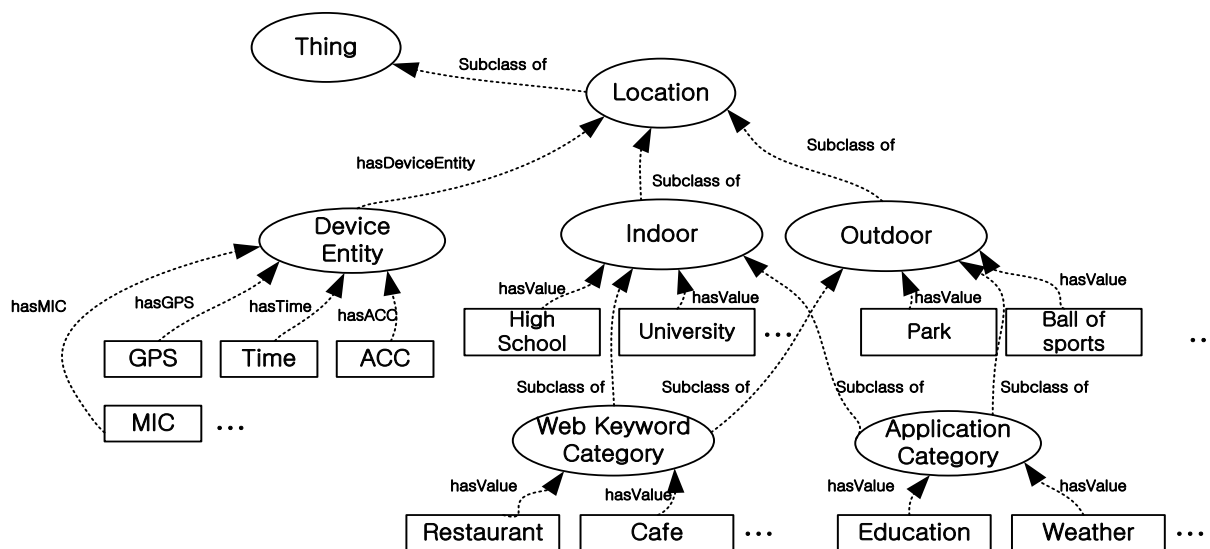


Figure 1. Proposed Location inference ontology

profile data stored in smart phone. ‘When’ can be defined by current time and the time when specific application was executed. ‘What’ can be defined by the category of applications executed by user, web search keywords entered by user for the search and application log. ‘How’ can be defined by relationships between each elements. Examples of each element are: a teenager (who), at 6pm (when), searched educational application (what). Now the location of this user can be inferred as a studying location such as a library or school.

Table 3 represents the relationships between data elements to be used to describe location with the location inference ontology. Location inference ontology modelled based on Table 1, Table 2 and Table 3 is shown in Figure 1. As shown in Figure 1, the following ontology nodes do exist under ‘Location’ node in Location inference ontology.

- **Device Entity** : kind of sensor data obtainable on smart phone
- **Indoor**: kind of locations that belongs to indoor location defined by Location categories in Table 2
- **Outdoor** : kind of locations that belongs to outdoor location defined by Location categories in Table 2
- **Web Keyword Category** : category of location name used for the keyword search on smart phone, and it is one of the Web Keyword Categories defined in Table 2.
- **Application Category** : category of application that was executed by user, and it is one of Application Categories defined in Table 2

‘Device Entity’ node in Figure 1 provides user behaviour data. For example, ‘Time’ literal data expresses ‘when’ defined in Table3. The time when the search was done or the application was executed or the current time may belong to ‘Time’ literal data. And, the other literal data in ‘Device Entity’ node belong to ‘what’ defined in Table 3. On the other

hand, ‘Web Keyword Category’ node and ‘Application Category’ node correspond to ‘when’ defined in Table3 that express the behaviour of user. In the case of ‘who’, which includes a user profile information of a user or surrounding users, it is not considered because of the protection of individual private information in this paper. ‘How’ is expressed as the relationship (property) between ontology nodes.

The relationship between ontology nodes is not hierarchical since they own and share each other’s data. Property is declared in consideration of the relationship. On the other hand, categories below the nodes have a hierarchical relationship with the higher nodes so that they are declared as subclasses. Due to the scalability of ontology, additional device entities are enabled though the simple definition declaration as any situation arises..

TABLE 3. RELATIONSHIPS BETWEEN THE ELEMENTS OF LOCATION INFERENCE ONTOLOGY

Elements of location determination	Meaning of each elements
Who	Data from the user and users in the same area
What	Used applications, web search words, sensor data provided by the smart phones
When	Time the application is used, web searching time, duration of application used
How	Definitions of each elements(Property)

IV. LOCATION INFERENCE ALGORITHM BASED ON THE LOCATION INFERRING ONTOLOGY

A. Location Inference Algorithm

Location inference algorithm is presented in this section with the ontology generated in the previous section. The algorithm works by compensating the user location received by the phone from GPS with the ontology generated in the previous section to solve the problem of the GPS error.

Figure 2 demonstrates the error of user location on a map. 'A' is a university. 'B' is a park. 'C' is a kindergarten. User location is inquired at 'A', but it was not clearly determined whether the user is at 'A', 'B' or 'C'. Now, the inferring algorithm determines the user location with data from the smart phone based on the GPS information. Data used in inferring algorithm are user data stored in the smart phone(GPS log, used applications and web search keywords) and sensor data obtained by the smart phone(GPS, gyroscope and accelerometer, MIC). The following conditions are defined as user locations in the study.

- indoor, outdoor
- busy location, location with not much activity
- noisy location, quiet location
- locations related to the used application category and keyword category



Figure 2. Positioning results

The inferring algorithm is defined within the bases of the above locations. Figure 3 shows the algorithm flow. First filtering in Figure 3 determines data that can be clearly classified as objective data. The range of the first filtering determines the user location data, surrounding locations within the GPS error and whether the location is indoor or outdoor. For example, the first filter will return with a result 'restaurant', if there are restaurant, park and plaza around current user location and the location is classified as indoor according to the strength of GPS signal. The first filtering process works sequentially. Second filtering classifies the

results from the first filtering in consideration of the data provided by the smart phone but not sequentially. The second filtering classifies data based on time, activity and noise. Time specific locations are related to activities during lunch or dinner time. Activity specific locations are determined by whether the user is moving or stationary, and then classify moving direction using a gyroscope sensor. Coffee shops and restaurants are stationary locations and department stores and supermarkets are moving locations. Last step is to determine a location based on the density. Density is inferred by noise around the user. Movie Theatres, coffee shops, office or home can be determined by the noise levels. Also, Density is obtained through the Bluetooth signal around user.

Finally, the results from the second filtering are compensated with used application categories and search keywords to determine the location.

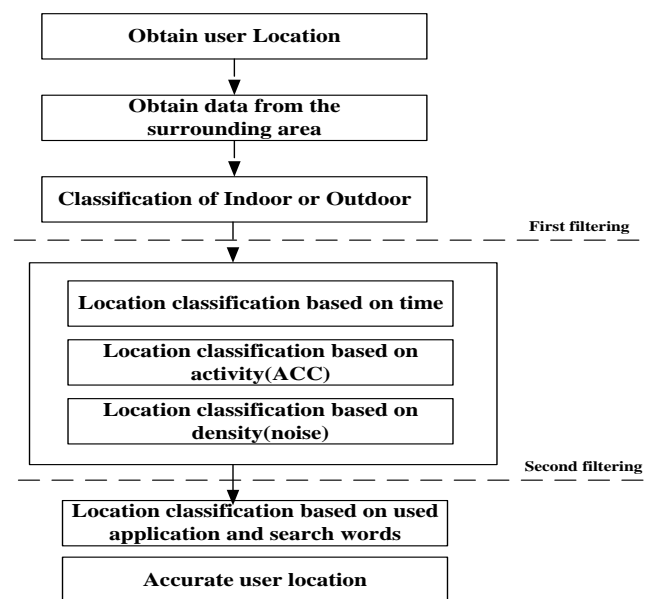


Figure 3. Location inference algorithm

B. Location Inference Algorithm Implementation

In this paper, location inference algorithm is implemented using the previously generated ontology and OWL ontology inferring engine open source called Jena.[5] Figure 4 shows the implemented algorithm. The implemented algorithm distinguishes whether a user is located at indoor or outdoor according to the strength of GPS signal, and searches the corresponding location category. Since then, it returns the location information after inferring on the ontology using the sensor data including ACC, MIC and user data including application, web keyword obtained by smart phone.

Testing the algorithm was performed for the situation as shown in Figure 2. In Figure 2, user is located at 'A' and the ontology inferred locations with the GPS error corrected returns locations 'A', 'B' and 'C'.

Information gathered by the smart phone is that GPS signal is weak, noise is in medium range, density is determined to be high by Bluetooth and the user used education related

application. Now, location can be accurately pin pointed with obtained information. ‘C’ is filtered out by the first filtering since it is outdoor. ‘B’ is eliminated since it has high density. Figure 5 demonstrated the result of location inference using the implemented algorithm.

```

if(DeviceGPS.equals("Low"))
{
    GPSposition = "Indoor";
}
else{
    GPSposition = "Outdoor";
}

System.out.print("Class: " + Loca.getLocalName() + "\n");

for (Iterator Fo = Loca.listSubClasses(true); Fo.hasNext(); num++) {
    OntClass F0c = (OntClass) Fo.next();
    temp_resource[num] = F0c;
    System.out.print(" " + temp_resource[num].getLocalName() + "\n");
}

while (ACCiter.hasNext()) {
    temp_resource_ACCind[z] = (Resource) ACCiter.next().getObject();

    System.out.println(temp_resource_name[number_2].getLocalName() + " hasACC " +
temp_resource_ACCind[z].getLocalName() );
}

while (MICiter.hasNext()) {
    temp_resource_micind[z] = (Resource) MICiter.next().getObject();

    System.out.println(temp_resource_name[number_2].getLocalName() + " hasMIC " +
temp_resource_micind[z].getLocalName() );
}

while (Applicationiter.hasNext()) {
    temp_resource_Applicationind[z] = (Resource) Applicationiter.next().getObject();

    System.out.println(temp_resource_name[number_2].getLocalName() + " hasApplication " +
temp_resource_Applicationind[z].getLocalName() );
}

while (WebKeyworditer.hasNext()) {
    temp_resource_WebKeywordind[z] = (Resource) WebKeyworditer.next().getObject();

    System.out.println(temp_resource_name[number_2].getLocalName() + " hasWebKeyword " +
temp_resource_WebKeywordind[z].getLocalName() );
}
}

```

Figure 4. Implemented algorithm

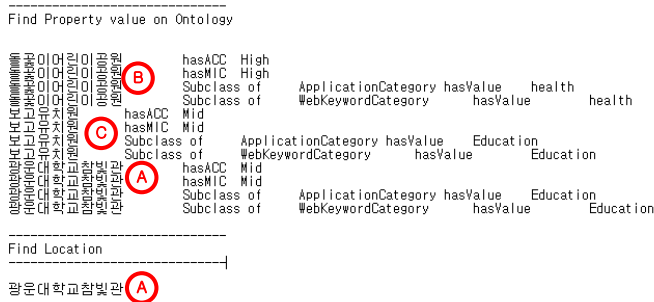


Figure 5. Results from the implemented algorithm

V. CONCLUSIONS

In order to provide the context-aware services for individual users, we propose a method to infer user location by compensating location information provided by GPS with additional information obtained by smart phone sensors, user application usage log and search keywords log. Using the proposed method, a user position will be located within the error range of GPS by means of inference.

Ontology is implemented by modelling the additional information attained from the various sensors of smart phone and user data such as application categories and search keywords used as ontology nodes, and relations between

ontology nodes as property. Accurate user location is inferred from the locations provided by GPS within its error range.

Future study of how to define a variety of factors of location classification would be very helpful for more accurate location inference. Algorithm in this study considers only data elements such as activity, noise and time, however, more diverse data elements are necessary to implement the algorithm. A location inferring algorithm in combination of a learning algorithm through the implementation of time-, day of the week- and location-specific clusters using the user log information of the locations the user had already visited will provide with more reliable algorithm.

ACKNOWLEDGMENT

"This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korean Government(NRF-2011-0025226)

REFERENCES

- [1] D. Siewiorek, A. Smailagic, J. Furukawa, A. Krause, N. Moraveji, K. Reiger, J. Shaffer, F. L. Wong, "SenSay: A Context-Aware Mobile Phone," Proc. of the 7th IEEE International Symposium on Wearable Computers (ISWC'03), p.248, 2003.
- [2] A. Vetek, J. A. Flanagan, A. Colley, T. Keränen, SmartActions: Context-Aware Mobile Phone Shortcuts, Proc. of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: PartI, Aug. 24-28, 2009.
- [3] Y.S. Cho, S.Y. Cho, B.D. Kim, S.H. Lee, J.C. Kim, W.S. Choi, "Technical Trend of Indoor/Outdoor Seamless Positioning," Electronics and Telecommunications Trends, vol 22 no 3, pp. 20-28, June. 2007.
- [4] Hakyong KIM, "Wireless LAN-Based LBS Service," Telecommunications Review, vol 16 no 2, pp. 188-202, April. 2006.
- [5] <http://jena.sourceforge.net/>
- [6] Harry Chen, Tim Finin and Anupam Joshi, "An ontology for context-aware pervasive computing environments", The Knowledge Engineering Review Volume 18 Issue 03, pp 197-207, September 2003



Sang-il Kim is currently enrolled in Ph.D. program at the Dept. of Electronics and communications Engineering, Kwangwoon University, Seoul, Korea. His research interests Semantic web, context reasoning .



Hwa-sung Kim received Ph.D degree at the Dept. of Computer Science, Lehigh University, Bethlehem, PA 18015 USA. He is now a professor at the Dept. of Electronics and communications Engineering, Kwangwoon University, Seoul, Korea. His research interests include mobile network protocol, mobile web computing, embedded software