

# Context-Aware Recommendation Model based on Mobile Application Analysis Platform

Ahyoung Kim<sup>1</sup> · Junwoo Lee<sup>1</sup> · Mucheol Kim<sup>2</sup>

Received: 13 February 2015 / Revised: 19 November 2015 / Accepted: 3 December 2015 /

Published online: 26 December 2015

© Springer Science+Business Media New York 2015

**Abstract** Recently, various types of log data have been collected and used due to the explosive increase of mobile devices. In mobile environment with high portability and mobility, in addition, the user context information is an important factor for recommendation process. This study attempted to analyze usability log data collected from the mobile device through an application analysis platform. We suggested a context-aware recommendation model to recommend mobile applications or contents by recognizing users' context data. The usability data of applications consist of activities which were active during the use of a mobile device. The features of these activities are related with time, location and device information. A model proposed in this study has a flexible structure which can be selectively used depending on user circumstances and performs a usability patterns of the applications based on the collaborative filtering method.

**Keywords** Context-aware · Recommendation system · Collaborative filtering · Log analysis

## 1 Introduction

Lately, context-aware services and systems using diverse mobile devices such as smartphones and tablet computers have been provided. Based on high mobility and portability, mobile environment

---

✉ Mucheol Kim  
mucheol.kim@gmail.com

Ahyoung Kim  
kimay@etri.re.kr

Junwoo Lee  
leejw@etri.re.kr

<sup>1</sup> Electronics and Telecommunications Research Institute, 218, Gajeong-ro, Yuseong-gu, Daejeon, South Korea

<sup>2</sup> Department of Multimedia, Sungkyul University, 53, Sungkyuldaehak-ro, Manan-gu, Anyang, Gyeonggi-do, South Korea

offers various sensors, flexible network connectivity and convenient external resources. With these characteristics of mobile environment, there has been a rising demand for context-aware services. Due to paradigm shifts from ubiquitous environment where flexible network connectivity is what matters to mobile computing environment which provides user-centric personal services, in addition, studies on mobile context-aware have become more important [11]. The analysis technologies with user data have also become more critical as the mobile application market grows. As the number of mobile applications skyrockets, however, it's become more difficult to find user-wanted applications or contents in the mobile market. If an upgraded application with similar contents is released, furthermore, users can easily get bored with previous mobile applications. That's why users keep searching for a new applications with upgraded functions. A recommendation model has been introduced to the application market because of these trends. In this study is described a context-aware recommendation model with mobile application analysis platform. First, for context-aware recommendation, a log data from mobile applications is collected, and a user profile is automatically generated. Then, the profile is automated by collecting log data such as the duration time, the location information and the device information. Second, the recommendation model basically creates a list of candidates for recommendation with analyzing the use frequency, duration and up-to-dateness of the mobile application. The model proposed in this study has a flexible structure designed to generate a list of candidates depending on user circumstances in each period so that a collaborative filtering model aimed to handle with the duration, location and device information is adopted. Lastly, we suggest the mobile applications with the contextual usage data.

This study is structured as follows: In chapter 2, the studies relating to personalization recommendation and context-aware recommendation are introduced. In chapter 3, a big data system-based mobile application analysis platform is explained. In chapter 4, a recommendation model using various contextual data of mobile applications is described. In chapter 5, the results of the test conducted using the proposed model are stated. In chapter 6, a conclusion is given, and future work is discussed.

## 2 Related work

### 2.1 Mobile log analysis

There have been a lot of studies for mobile log analysis. Because a mobile device is highly portable, in particular, data can be collected and analyzed from diverse perspectives.

Xu [24] collected and analyzed user data from diverse perspectives in terms of the mobility and diversity of mobile devices. Based on the ease of use for the mobile device network, packet data were collected and used in usability analysis. However, it is difficult to collect and analyze the user data in a large amount. Girardello [10] proposed a method to recommend a popular application installed by a lot of users by storing the number of installations and deletions as database. This system targeted travelers who are interested in an application which is useful in finding a particular location. Instead of using a general recommendation technique, user location and app's consumption information were only used for recommendation. Yan [25] analyzed how users actually use the app, using the item-based collaborative filtering. Based on analysis on what applications other users around used, an application was recommended. Verkasalo [22] extracted and analyzed the usability of mobile applications. After installing MobiTrack in the mobile device, the usability-related data were automatically collected. Then, they were periodically transferred to a server for synchronization. The collected data were analyzed under five categories only: communication,

usage patterns of the mobile application and navigation through internet, device and multimedia information. Therefore, they weren't able to provide analysis information under specific categories. Kang [12] collected and analyzed diverse usage patterns in the mobile device. Various information (ex: GPS, SMS, battery, etc.) was collected and used for statistical analysis. Then, user status was induced. Cloud-based sensor data were collected and analyzed. Seo [20] collected and analyzed sensor and network information for the purpose of improving user convenience. There has been a rising interest in the studies which are aimed to analyze usage patterns, using the log data [9, 18, 19]. However, it's been very hard to find a study on the analysis and materialization of usage patterns based on cloud after collecting mobile log [4, 6, 8, 16, 26].

## 2.2 Context-aware recommendation system

In mobile environment, there have been a lot of preceding studies on the recommendation of a suitable application or contents depending on user circumstances. Users' context information can be divided into static context information (ex: age, gender, occupation, etc.) and dynamic context information (ex: temperature, weather, location, etc.). As context-aware technology becomes more important with the growth and expansion of mobile industry, in particular, many techniques aimed to figure out user preference using these context data have been proposed in mobile recommendation-related studies [2, 3].

Debnath [7] proposed a method to change adequacy depending on the duration and frequency of the usage after recommending an application suitable to user circumstances and getting feedback on user preference. Lee [17] suggested a methodology to improve recommendation performances by acquiring context information from the user log. They recognize the users' tendency with the time based context information. It is adjusted to collaborative filtering with similar circumstances. Lee [15] did not compare the all user's rating information, but they used above-average users to improve the accuracy of recommendation. After comparing current users' context, high weighted values were applied to similar users for recommendation. Lee [14] attempted to derive user circumstances for dynamic context information (ex: temperature, humidity, time, etc.) as well as static context information by proposing a Fuzzy-Bayesian Network-based music recommendation method. Woerndl [23] suggested a system aimed to recommend a mobile application based on user context through the hybrid recommendation engine under a multidimensional approach.

In addition, there have been various types of studies such as deriving user behavior or circumstance to recommend more customized service or contents in mobile environment or recommending a similar application. However, most of these studies focused on static context information such as user profile, experienced contents and item history [5]. In terms of the configuration of a similar group, in addition, the conventional filtering technique was concentrated. Therefore, there were limitations in effectively composing a similar user group with similar patterns or propensity under diverse circumstances. In a mobile computing environment where various forms of contents are explosively created, lastly, there have been few studies on mobile context-aware recommendation which provides more personalized contents [13, 21].

## 3 A Mobile application analysis platform

The structure of a mobile application analysis platform for context-aware recommendation is shown in Figs. 1 and 2. This platform is operated by the Citrix CloudStack Infrastructure's Big

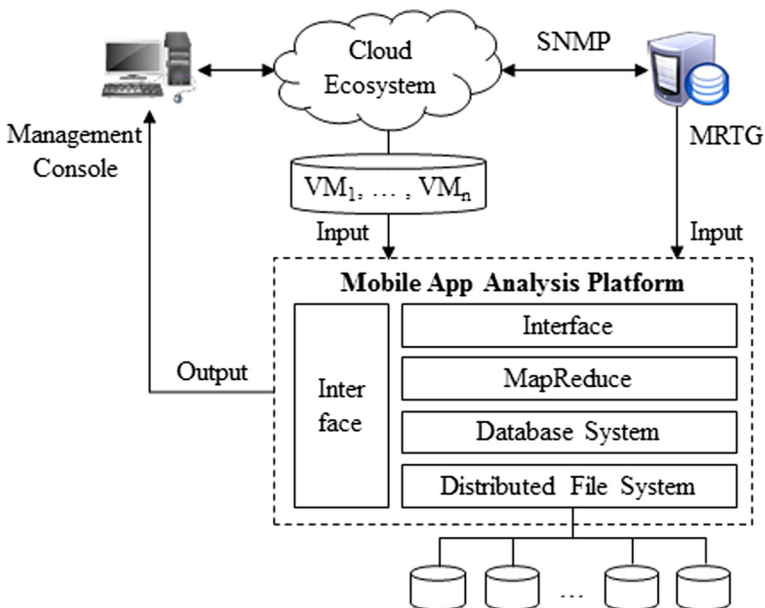
Data System. Hadoop and Hbase to handle a problem of data extensibility. The cloud environment aims to minimize the possibility of system problems which occur because of system extensibility [1].

### 3.1 Client module

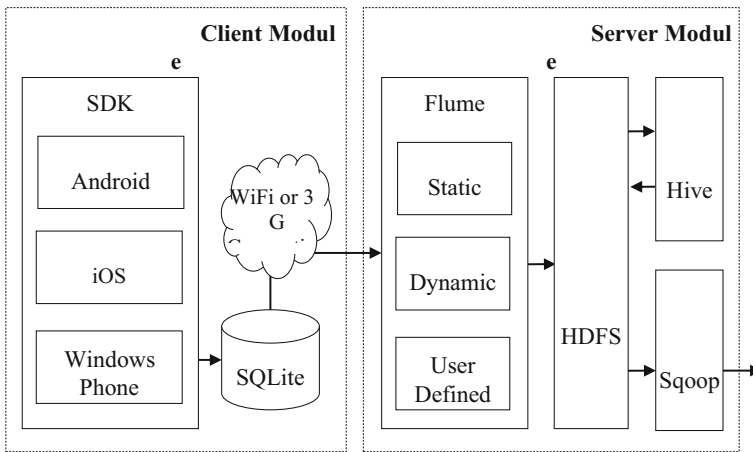
A cloud module collects and transfers sensor data based on mobile users’ activities (Fig. 2). It has SDK library to collect user activities easily through the mobile devices. This module monitors all activities created in the library-included application and collects data which is related the time and location dimensions. Then, the collected data are transmitted to a server module when the mobile device is in WiFi mode, and maximum allowable data size exceeded. In this study, two different SDK versions (Android and iOS) are provided, and user log data are collected in three categories depending on attributes:  $D_s$ ,  $D_d$  and  $D_u$ . First,  $D_s$  includes static data which are collected once such as UUID, title of application, OS version, device model number and resolution. Second,  $D_d$  covers dynamic data (ex: start time, end time, latitude, longitude, etc.) which are collected according to user activities. Third,  $D_u$  includes user-defined data aimed to collect additional sensor data in addition to basic ones. The data collected from the mobile sensors are stored in SQLite and transferred to Flume using WiFi or 3G.

### 3.2 Server module

A server module stores sensor data from the client module, and it analyzes their usability, using the analysis algorithm (Fig. 2). This module consists of Citrix CloudStack Big Data System. The Hadoop-based Big Data system can minimize a problem of system portability which can occur in cloud environment.



**Fig. 1** Structure of the Mobile Application Analysis Platform Module System



**Fig. 2** Configuration of the Mobile Application Analysis Platform Module

A server module gets the sensor data collected by the client module in three different Flume designs depending on attributes. Then, they are processed and analyzed through the big data system. Regardless of data processing environment, this module consists of Tomcat web server, TCP-IP file transfer server and a web service-based data transfer server. The server module receives data from the mobile device and analyzes patterns to check abnormal data transmission. After pattern analysis, the map/reduce module calculates the results derived from the Hadoop or Hbase line by line.

### 4 Context-aware recommendation model

The proposed recommendation model uses a collaborative filtering algorithm considering three aspects: frequency of use, duration of use and up-to-dateness. First, the frequency of the use of the application executed in the mobile device is calculated. The frequency is analyzed with various factors such as time, location and device. The time-based frequency is calculated according to daily, weekly and yearly time. The location-based frequency is estimated depending on latitude and longitude while device-based frequency is calculated depending on model number, manufacturer and resolution. Second, we calculates the duration. Third, the up-to-dateness would be an important factor in usability analysis because a mobile app’s lifecycle is relatively short.

Let  $R$  denote the list of candidates considered for the reliability of the mobile application based on the time, location, and device.

$$R = \{R_{time}, R_{location}, R_{device}\} \tag{1}$$

Where each  $R_{time}$ ,  $R_{location}$ ,  $R_{device}$  are an ordered lists of the mobile applications based on the time, location, and the device type.

$R_{time}$  is recommended by the usage frequency, duration, and recentness of the mobile applications. In addition,  $R_{location}$  and  $R_{device}$  are the means of filtering the candidate applications as an intermediate results from an ordered list selected according with the  $R_{time}$ . These two parameters are acceptable for the model to determine for the refined candidates.

The model determines factors with the user context. For example, if the users want to use the location data to find the mobile applications, the model collects candidates based on the current location with the  $R_{time}$  and  $R_{device}$ .

$$R = \alpha \times R_{time} + \beta \times R_{location} + \gamma \times R_{device} (\alpha + \beta + \gamma = 1) \tag{2}$$

Let  $U$  denote that the user set consists of the usage frequency ( $UFT$ ), duration time ( $UDT$ ), usage time per hour( $URT$ ), usage time per day( $DOW$ ), usage time per month ( $MOY$ ), the operating system version( $OSV$ ), and the location( $L$ ).

$$U = \{UFT, UDT, URT, DOW, MOY, OSV, L\} \tag{3}$$

Let  $UFT$  have an element  $f_{hi}$ , which is the number of executions of the mobile application  $a_j(a_j \in A)$ .  $C_{ij}$  represents execution time for mobile application ( $C_{ij} = 0$  means that there is no relationship between  $f_{hi}$  and  $a_j$ .)

$$\hat{C}_{ij} = \frac{\sum_{f_{hk}} w_{i,k} \cdot C_{k,j}}{\sum_{f_{hk}} w_{i,k}} \tag{4}$$

where  $w_{i,k}$  is the similarity between  $f_{hi}$  and  $f_{hk}$ , is defined as follows.

$$w_{i,k} = \frac{\sum_{a_j \in A} C_{i,j} C_{k,j}}{\sqrt{\sum_{a_j \in A} C_{i,j}^2} \sqrt{\sum_{a_j \in A} C_{k,j}^2}} \tag{5}$$

We could recognize the user preference with this consumption activity( $UFT$ ) because the usage frequency is the indirectly measure which is the propensity to consume the mobile applications. In advance, the proposed approach consider two kinds of types, a weekday or weekend, for the recommendation.

Finally we proposed the location based contextual recommendations of the mobile application. While commuting, most people want to know when they should go to their destination or where the it is located. On the other hand, most people use mobile applications related to their works such as an e-mail service, groupware service, or office suite in office hours.

On the other hand, it is very important to consider not only the usage frequency but also the usage duration because it means the contents popularity. For example, YouTube is often used for a long time owing to their popularity over TV or radio.

The spending activity ( $d_{hi} \in UDT$ ) of mobile app ( $a_j \in A$ ) is denoted as  $S_{ij}$ .

$$w_{i,k} = \frac{\sum_{a_j \in A} S_{i,j} S_{k,j}}{\sqrt{\sum_{a_j \in A} S_{i,j}^2} \sqrt{\sum_{a_j \in A} S_{k,j}^2}} \tag{6}$$

$$\hat{S}_{ij} = \frac{\sum_{d_{hk}} w_{i,k} \cdot S_{k,j}}{\sum_{d_{hk}} w_{i,k}} \tag{7}$$

The lifecycle of a mobile application registered in an market such as Google Play or the Apple AppStore is relatively much shorter than that of the general applications for a PC. Then

we multiply the penalty with the usage rate per month. The consumption activity will be calculated through the following equation. Let  $M$  is the number of month, and  $m_k$  means the durations for consumptions of the application.

$$new \hat{C}_{ij} = \hat{C}_{ij} * \frac{c_{ij}}{\frac{\sum_{m_k \in M} c_{i,j}}{|M|}} \tag{8}$$

We also consider two additional factors, which are OS versions and location based data, for filtering the unnecessary intermediate results. Recently, the hardware and software characteristics of mobile devices tend to be diverse because many companies have introduced new models continuously. Thus, some mobile applications have been generated to showcase these new features. Furthermore location based data is useful to determine the customized candidates.

### 5 Experiments

In this experiment, a background application developed using the Android SDK was installed in 10 mobile devices, and the usability data collected for 3 months were analyzed. The mobile

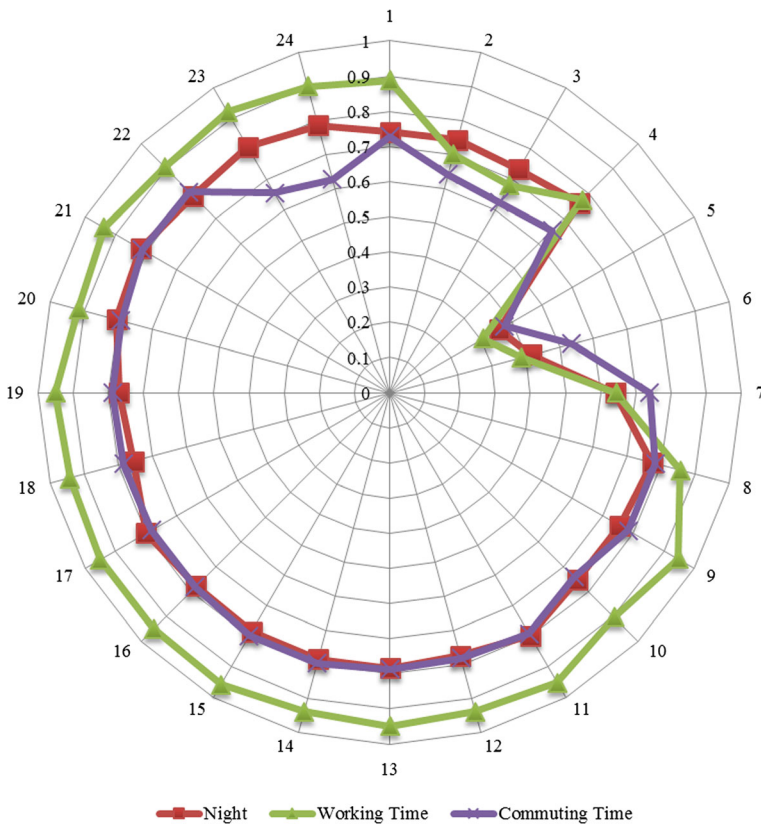


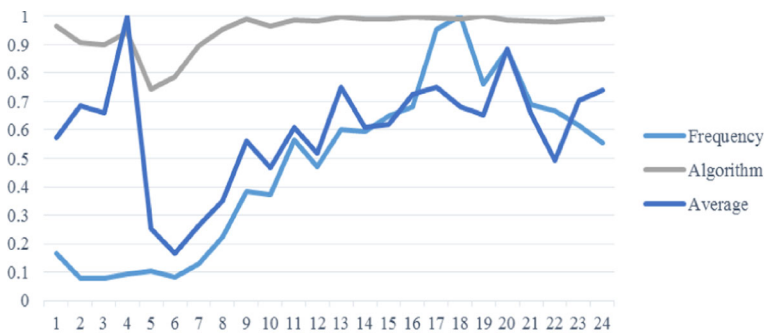
Fig. 3 Relationship with daily activity hours

usability data of 120 applications but Android or middleware application which is unnecessary for usability analysis were collected.

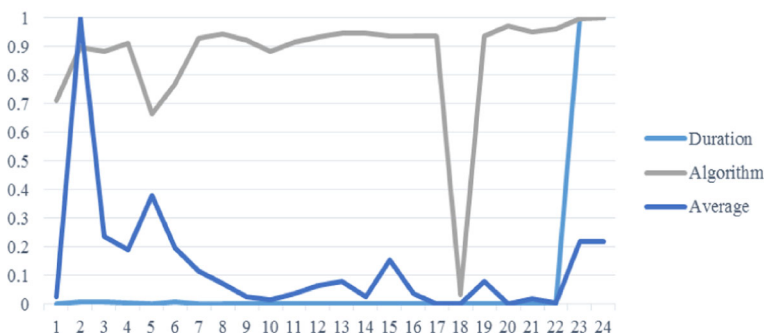
The importance of a mobile application is determined depending on how many people use it. In this study, some attributes were added to normalize differences among usability patterns during daily activity hours in addition to general features. To investigate what time zone is most important against all mobile applications, the relationship with daily activity hours was measured based on the frequency and duration. To analyze the relation during daily activity hours, objective similarities between all applications were normalized and calculated. The normalization process can prevent the mobile application’s importance from being incorrectly analyzed based on the collected total frequency and amount of consumption.

Figure 3 shows that the relationship with daily activity hours based on the frequency of usage. The pattern of commuting time is more similar than the other times such as the night and the dawn. Furthermore the patterns in a working time is similar than the other times in a day. It means that we have to normalize the real data by using this similarity result for shortening the differentiated usability. After the normalization process, we have evaluated the usage pattern with the frequency and the duration. As a result, we could recommend the mobile applications which many people interest in. Figure 4 is introduced severally to explain and demonstrate the feasibility of the recommendation model.

It shows the usage pattern for many applications related with SMS(Short Message Services) and MMS(Multimedia Message Services). The tendency of the usage frequency and average



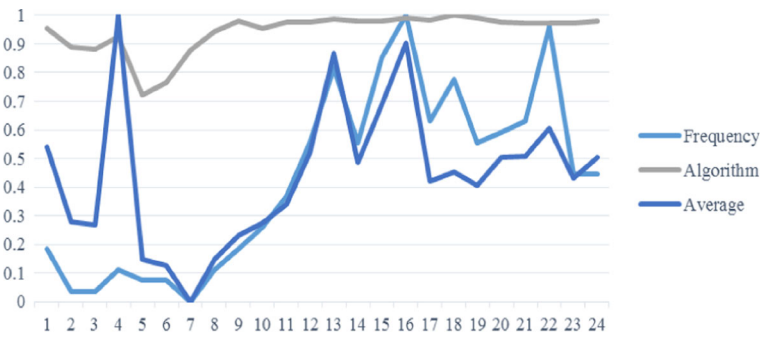
(a) Frequency



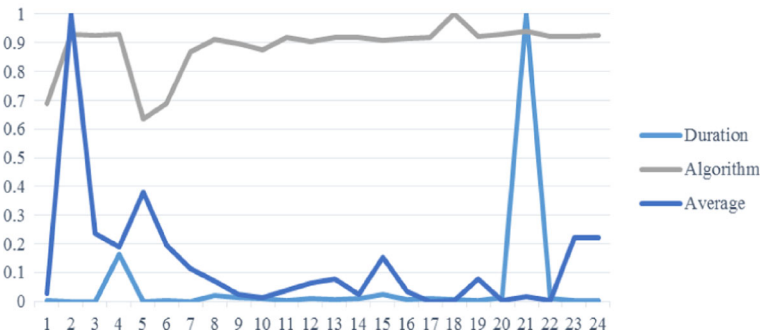
(b) Duration

**Fig. 4** Usage pattern analysis for the SMS applications





(a) Frequency



(b) Duration

**Fig. 5** Usage pattern analysis for the Web Browser applications

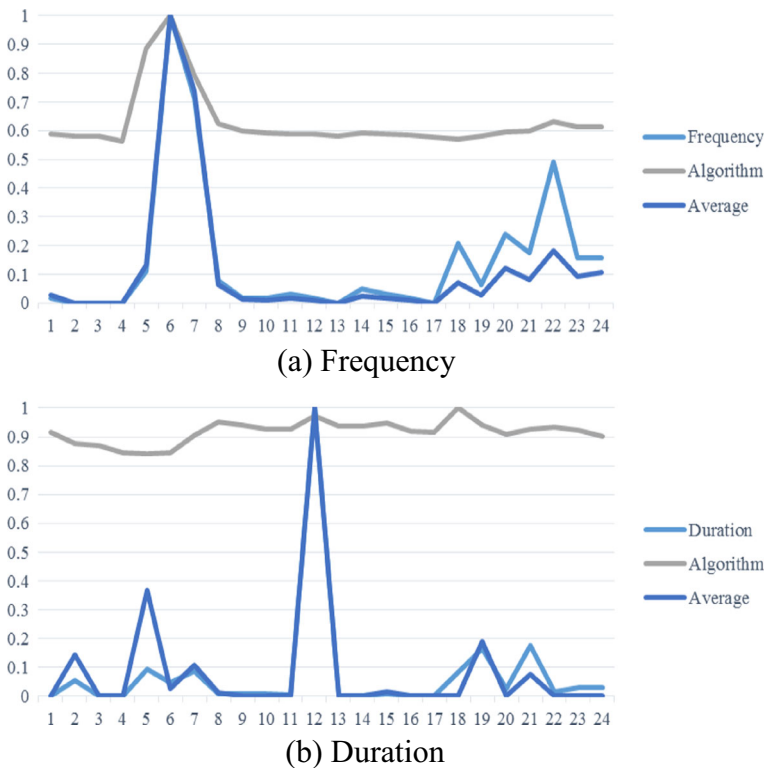
fluctuates heavily according with the time dimension. However, the proposed model shows enhanced result because it normalizes the difference between the hours in a day and recalculates the similarity by using the relative weight for the mobile applications.

Figure 5 shows the usage pattern for the web browser applications. In general, many people are used to consume the time frequently in order to navigate the web site. It shows that the usage duration at night is higher than the other time in a day. Absolutely, we have known that the usage pattern based on the frequency and the duration remains at a high level. It means that web browser is the most popular applications for acquiring information.

Figure 6 shows that many users spend time with a mobile game in a commuting time. For example, the usage frequency displays at a very high level between PM 6 to 8. This means that many users play a mobile game during the commuting after the work. Additionally, the graph of the proposed model is still at a high level. It means the importance of the mobile game is relatively higher than the other applications in a commuting time irrespective of the usage frequency and the usage duration.

## 6 Conclusions

A model proposed in this study overcame the limitations of the conventional recommendation system which would provide simple information without considering a



**Fig. 6** Usage pattern analysis for the Commuter applications

mobile device's features or user circumstances and realized personalized services in mobile environment. In addition, it was designed to provide reliable, intelligent and optimized recommendation services by analyzing cloud-based log data.

This study analyzed mobile users' usability log after configuring a big data-based mobile application analysis platform and suggested a context-aware recommendation model. The mobile application analysis platform consists of a library-type client module and a big data-based server module. The client module collects usability log data from the mobile devices and performs the function needed to send these data to the server. The server module analyzes and processes the data from the client module, using the pre-defined data patterns. The context-aware recommendation model makes a list of candidates using additional context data such as time and place. The preference was measured through usability experience instead of using user profile or user preference in person.

There will be further studies by adding mobile sensing data or analysis algorithm to the cloud-based usability analysis framework and extending them by a category unit, not by a user unit. In addition, there will be continued studies on a faster and more accurate recommendation method through the precise analysis of the data which have increased due to the rapid growth of the number of users, applications and contents.

**Acknowledgments** This work was supported by the Ministry of Science, ICT and Future Planning (MSIP, Korea).

## References

1. Agrawal D, Das S, El Abbadi A (2011) Big data and cloud computing: current state and future opportunities. In Proceedings of the 14th International Conference on Extending Database Technology (pp. 530–533). ACM
2. Böhmer, M., Bauer, G., and Krüger, A. (2010). Exploring the design space of context-aware recommender systems that suggest mobile applications. In 2nd Workshop on Context-Aware Recommender Systems.
3. Böhmer, M., Ganev, L., and Krüger, A. (2013, March). Appfunnel: A framework for usage-centric evaluation of recommender systems that suggest mobile applications. In Proceedings of the 2013 international conference on intelligent user interfaces (pp. 267–276). ACM.
4. Buyya R, Yeo CS, Venugopal S (2008) Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities. In High Performance Computing and Communications, 2008. HPCCom'08. 10th IEEE International Conference on (pp. 5–13). IEEE
5. Davidsson, C., and Moritz, S. (2011, February). Utilizing implicit feedback and context to recommend mobile applications from first use. In Proceedings of the 2011 Workshop on Context-awareness in Retrieval and Recommendation (pp. 19–22). ACM.
6. Dean J, Ghemawat S (2008) MapReduce: simplified data processing on large clusters. *Commun ACM* 51(1): 107–113
7. Debnath S, Ganguly N, Mitra P (2008) Feature weighting in content based recommendation system using social network analysis. In Proc. of the 17th Int. Conf. on World Wide Web, 1041-1042
8. Fontan J, Vazquez T, Gonzalez L, Montero RS, Llorente IM(2008) OpenNEbula: the open source virtual machine manager for cluster computing. Open Source Grid and Cluster Software Conference
9. Foster I, Zhao Y, Raicu I, Lu S (2008) Cloud computing and grid computing 360-degree compared. In Grid Computing Environments Workshop, 2008. GCE'08 (pp. 1–10). IEEE
10. Girardello A, Michahelles F (2010) AppAware: which mobile applications are hot?. In Proceedings of the 12th international conference on Human computer interaction with mobile devices and services. ACM, 431-434
11. Kakousis K, Paspallis N, Papadopoulos GA (2010) A survey of software adaptation in mobile and ubiquitous computing. *Enterprise Inf Syst* 4(4):355–389
12. Kang JM, Hong WK (2011) Usage pattern analysis of smartphones. Asia-Pacific Network Operations and Management Symposium
13. Karatzoglou, A., Baltrunas, L., Church, K., and Böhmer, M. (2012, October). Climbing the app wall: enabling mobile app discovery through context-aware recommendations. In Proceedings of the 21st ACM international conference on Information and knowledge management (pp. 2527–2530). ACM.
14. Lee JS, Lee JC (2006) Music for my mood: a music recommendation system based on context reasoning. *Smart Sens Context Lect Notes Comput Sci* 4272:190–203
15. Lee SI, Lee SY (2011) A recommendation system using context-based collaborative filtering. *J Korean Inst Intell Syst* 21(2):224–229
16. Lee KH, Lee YJ, Choi H, Chung YD, Moon B (2012) Parallel data processing with MapReduce: a survey. *ACM SIGMOD Record* 40(4):11–20
17. Lee DJ, Lee SK, Lee SG (2009) Considering temporal context in music recommendation based on collaborative filtering. *Korea Comput Congress* 36(1(C)):123–128
18. Madden S (2012) From databases to big data. *IEEE Internet Comput* 16(3):0004–0006
19. Padhy RP, Patra MR, Satapathy SC (2011) RDBMS to NoSQL: reviewing some next-generation non-relational database's. *Int J Adv Eng Sci Technol* 11(1):15–30
20. Seo HS, Eum JH, Jeong HS, Yoo TJ, Choi HS, Kim TS (2011) A user behavior analysis application exploiting the sensor information of smartphones. *Korea Comput Congress* 38(2(D)):101–104
21. Shi K, Ali K (2012) GetJar mobile application recommendations with very sparse datasets. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 204–212). ACM
22. Verkasalo H (2010) Analysis of smartphone user behavior. IEEE international conference on mobile business and 2010 ninth global mobility roundtable
23. Woerndl W, Whueller C, Wojtech R (2007) A hybrid recommender system for context-aware recommendations of mobile applications. IEEE 23rd international conference on data engineering workshop, 871-878
24. Xu Q, Erman J, Gerber A, Mao Z, Pang J, Venkataraman S (2011) Identifying diverse usage behaviors of smartphone apps. ACM SIGCOMM conference on internet measurement conference, 329–344
25. Yan B, Chen G (2011) AppJoy: personalized mobile application discovery. In Proceedings of the 9th international conference on Mobile systems, applications, and services ACM 113-126
26. Younge AJ, Von Laszewski G, Wang L, Lopez-Alarcon S, Carithers W (2010) Efficient resource management for cloud computing environments. In Green Computing Conference, 2010 International (pp. 357–364). IEEE



**Ahyoung Kim** is a Researcher in Electronics and Telecommunications Research Institute in Korea. Her research interests include Mobile Business, Pattern Analysis, Context-aware Technology, Internet of Things, and Smart Content Technology. She received the BS, MS degrees from the school of Computer Science and Engineering at Chung-Ang University, Seoul, Korea in 2007 and 2009, respectively.



**Junwoo Lee** is a Principle Engineering Research Staff in Electronics and Telecommunications Research Institute in Korea. His research interests include Network Technology, Internet of Things, and Smart Content Technology. He received the BS, MS degrees from the school of Computer Science and Engineering at Hanyang University, Seoul, Korea in 1996 and 1998, respectively.



**Mucheel Kim** is an Assistant Professor in department of multimedia at Sungkyul University in Korea. He was a Senior Researcher in Korea Institute of Science and Technology Information (KISTI), Daejeon, Korea. He received the BS, MS, PhD degrees from the school of Computer Science and Engineering at Chung-Ang University, Seoul, Korea in 2005, 2007 and 2012, respectively