Appeared in Proceeding of Edutainment 2011, LNCS 6872, 2011.

A Service Platform for Logging and Analyzing Mobile User Behaviors

Po-Ming Chen[1], Cheng-Ho Chen[2], Wen-Hung Liao[1], Tsai-Yen Li[1]

[1]Dept. of Computer Science, National Chengchi University, Taipei, Taiwan
[2]Master's Program in Digital Contents & Technologies, National Chengchi University, Taipei, Taiwan
{radination,chenaho,whliao,li.nccu}@gmail.com

Abstract: Quantitative analysis of user behavior calls for automated collection of large amount of log data over time. In this paper, we present a service platform for logging and analyzing mobile user behaviors. We have implemented a log collection service which records all user operations on the mobile unit. The collected data are transmitted to a remote server based on an energy-aware schedule. By providing a simple web-based interface, researchers can selectively retrieve the log according to the subject of interest. We illustrate the functionalities of the proposed service platform by several sampled logs.

Keywords –mobile devices, user log analysis, mobile user behavior.

1 Introduction

The rapid growth of App platforms coupled with the widespread installation of communication networks have contributed to the proliferation of smart mobile devices all over the globe. In the past few years we have witnessed significant increase in the shipments of smart phones following the introduction of Android and iOS operating systems. The versatile functionalities provided by such devices have fundamentally changed the way people use mobile phones. Voice communication is no longer the dominant function on modern mobile devices. Instead, leisure activities such as gaming, music, photography and wireless services such as web access, location-based service (LBS), and mobile social networks are becoming increasingly popular. New modes of services are being exploited to further take advantage of the new and constantly evolving mobile platforms. It is therefore of great interest to investigate the long-term usage patterns on these devices to study the user behavior and understand their needs to help design a better user experience.

Traditionally, experiments are carefully designed and conducted under controlled environments with specific parameters to examine or identify the influence of certain variables. The operation process is documented either in paper-and-pencil format or using audio-visual recording devices. Because participants are constrained spatially and temporally in a laboratory setting, their behaviors may differ from everyday life.

Moreover, unfamiliarity with the environment as well as the equipments may affect the user performance in specific tasks. Knowing that an observer is watching can sometimes be an annoyance. Apparently, traditional approaches have their limitations when it comes to the study of long term daily user behavior. It is necessary to develop new methodologies to record and interpret user logs on smartphones if we intend to conduct experiments to understand user behavior and assess user preference over a longer period of time.

Many applications on the smart phone do not function independently. They require internet connection to upload and access information, including simple file exchange, social games, or photo sharing. A novel or casual user may take some time to get familiar with the user interface. It is therefore difficult to collect meaningful usage patterns in the laboratory given the time constraint. Additionally, modern smart phones integrate diverse functions such as entertainment, photography, or calendar which may be activated in different contexts. The time or location data, which are considered irrelevant in a lab setting, become an integral part of the user log in a long-term field study.

Many mobile platforms have built-in social network service. With pervasive availability of wireless networks, social activities are no longer constrained by access to computers. Interactions among people now take a new assortment of forms, extending from basic voice communication to text messaging, multimedia sharing and position tracking. These interactions may, in turn, affect the collective behavior of a certain user group. Investigation of the dynamics and characteristics of these mobile social networks also calls for the development of new and powerful log collection and analysis tools.

Prior to the introduction of Android OS, almost all mobile operation systems are proprietary and leave very little room for modification. It is thus difficult to implement a 'monitoring' application to keep track of user operation due to security concerns. The uprise of Android has resolved this issue since Android is an open-source platform that offers great flexibility in modifying and supplementing extra functions for the system. This research takes advantage of the openness of Android OS to develop a log collection service that executes in the background and records detailed user operation in a database. The user log will be transferred to a remote server when internet connection is available. The collected data are stored on a remote database to allow researchers to download specific portion of the log for further analysis using a web-based query interface.

The rest of this paper is organized as follows. In Section 2 we briefly review some related work with special emphasis on long-term log collection and analysis. Section 3 presents the log collection service implemented in this research. It consists of two parts: log collector which runs on the device and log monitor which runs on the remote server. Section 4 concludes this paper with a brief conclusion and outlook on future researches.

2 Related Work

The study of mobile user behavior in real-life settings has gained increasing attention recently. In [1], the authors made a long-term observation of user's personal behavior as well as interactions among members of a group using time and location information recorded by Bluetooth devices. More details regarding the user operation can be obtained by running a background log collection program on the mobile device, as reported in [2]. Such log data reveal individual usage patterns on the smartphone, including text messaging, phone calls, photo shooting, game playing or other leisure activities. The log collection service implemented in our research follows a similar approach, except that we have a broader goal, i.e., to provide a shared platform combining hardware resources, software design and user group management so that researchers from different areas of expertise can work together to explore new research issues.

Online game community also has access to a large amount of log data collected in different game sessions. In [3], the authors built a database containing character's profile such as character ID and location through a client-side user interface. By analyzing the relationship between characters and groups (or guilds), they are able to study the evolution of online communities from a social network perspective. A tool named 'Social Dashboard' was designed to monitor the group dynamics in real-time. Although this research focused on online game community, it inspired us to develop a similar web-based log monitoring tool for researchers to have a quick glimpse of the overall status of the collected data to ensure smooth operation of the experiment.

3 The Proposed Log Collection and Analysis Service

This section presents the log collection and analysis platform implemented in this research. We first describe the initial setup of our experiment, namely, the recruit of volunteers and deployment of machines to have a trial-run to get early feedback. We then proceed to elucidate the details of the log collection service, including the data flow, log table format and service policy.

3.1 User Recruiting and Preliminary Study

Quantitative analysis calls for the collection of sufficient samples. However, we need to set up a trial-run to validate the design of our log collection service before launching a larger scale experiment. Toward this goal, we recruit 15 users to participate in the test phase. To facilitate the research of mobile social networks, we expect certain interactions amongst the users. Therefore, some level of acquaintance is assumed when we recruit the volunteers. The participants are not required to be expert users of smart phones, though. We intend to put together a group with a good mix of novice and veteran users so that we can observe more varieties of user behaviors and investigate possible interactions between different types of users.

The main purpose of conducting a preliminary test is to uncover potential pitfalls of the log collection service through the feedback of early adopters. For examples, it was found that many task management applications on Android whose main function is to monitor and manage resources can interfere with the normal execution of our log collection service. This may result in partial loss of data in some cases, and total failure of the background log collection process in other situations. Consequently, we need to modify the priority of the log collection service and implement an auto-restart policy in case of system breakdown. Additionally, transmission of log data consumes power, which is a valuable asset on mobile devices. During the initial test phase, we can set up different criteria to schedule the delivery of log data and observe the power consumption status to ensure that the logging service will not generate adverse effects on typical mobile phone users.

3.2 Log Collection Service

We utilize the functions provided by the Android Location Manager to record the position information. As for system logs, there is no corresponding API to directly access the low-level statistics. Instead, we make use of the 'logcat' tool included in the Android SDK to obtain raw data and then filter the desired portion according to our needs.

3.2.1 Service dataflow

The log data collected using our service must be stored on the mobile device first. This information will be passed on to a remote server when network connection is available. The overall dataflow is illustrated in Fig 1.

Since all the current participants are college students, they can gain free access to WiFi service while in campus. It is thus guaranteed that the operation logs can be collected without returning the hardware to the laboratory, thereby saving the efforts of researchers. The current approach for data acquisition also causes little interruption and interference to the user so that they can continue to operate the phone in a daily-life setting without constantly worrying about the experiment.

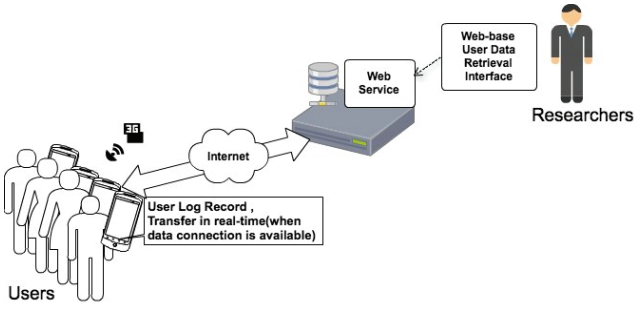

Fig 1. Service dataflow: user logs are saved locally using SQLite and later transmitted to a remote server when 3G/WiFi connection is available. Researcher can query the database through a web-based interface.

3.2.2 Log Table in the Remote Database

In our current implementation, the local SQLite database keeps track of all system logs generated as a result of user operations, as summarized in Table 1. For example, User_ID and Machine_ID are used to identify the person carrying this device. The Name_App field (encoded as Package Name in Table 1) stores the name of the application activated by the user. External data such as URL are kept in ExternalData (encoded as Data). Finally, the coordinates of user location are stored in Latitude and Longitude fields.

Table 1: Log database

User ID	Machine ID	Time	Package Name	Package Activity	Data	Lat.	Long.
U_0029	Machine_027	10/11/30	com.google.andro	.ConversationListAc	content://g	24.980	121.56
U_0029	Machine_027	10/11/30	com.android.mms	.ui.ConversationList	NODATA	24.980	121.56
U_0029	Machine_027	10/11/30	com.android.mms	.ui.ComposeMessag	NODATA	24.980	121.56
U_0029	Machine_027	10/11/30	android	com.android.internal	NODATA	24.980	121.56
U_0029	Machine_027	10/11/30	com.android.htcc	.ContactPhoneMailP	NODATA	24.980	121.56
U_0029	Machine_027	10/11/30	com.android.mms	.ui.ComposeMessag	NODATA	24.980	121.56

3.2.3 Known Limitations of the Log Collection Service

The ability to transfer log data to a remote server in real-time can facilitate instantaneous processing and analysis of user behaviors. However, we also need to be aware of the energy issue since scanning for available access points (AP) and making the connection will consume substantial power. The issue becomes even more critical if precise location is called for because it will require explicit user permission to turn on the GPS service. As reported in [4], GPS module drains a lot of power compared to WiFi or cellular tower ID localization method. To strike a balance between positioning accuracy and power consumption, the default setting is to avoid using GPS unless explicitly requested. In addition, we will adopt a non real-time data transmission policy. Such a policy can effectively limit the frequency of making connection to the remote database and prevent continuous transmission of small, fragmented data. The risk associated with this mechanism is partial loss of information if the user decides to restore the phone to factory settings before the latest logs have been transmitted. When this happens, we can also track the machine based on prior activity statistics.

3.3 Log Monitor and Query Interface

As stated previously, all log data generated as a result of user operation are stored in a central repository so that researchers can retrieve certain section of the data according to their specific research interests. Researchers who are interested in investigating the dynamics of social behaviors will need to define a set of terms containing all social networking activities and retrieve data accordingly. Since the log database is updated incessantly, the proposed platform provides a convenient web-based user interface for

registered researchers to quickly browse and query the content, as elucidated in the following.

3.3.1 Log Charting Service

The amount of data collected using the user logging service is large and continuously expanding. Presenting a large table containing many detailed items is just overwhelming and not very helpful. The log charting service (based on Google Chart) is implemented to assist the researchers to get a quick overview of the log data without much analysis. For example, it is easy to identify the most active users from the user activity chart, as depicted in Fig 2. It is also straightforward to find out which type of application is most often used by a specific user, as illustrated in Fig 3. These statistics can be exported in CVS format for further analysis.

Fig 2. User activity chart (top: last 24-hour period, bottom: the whole experiment period)

Fig 3. Charting the types of application run by a specific user.

3.3.2 Log Query Service

In-depth investigation of specific user behavior requires the collection and analysis of certain subset of the log data filtered by user profile, operation time, or activity type.

The log query service is developed to supply custom-tailored information to researchers. It features a simple web-based interface in which researchers can input some search criteria, such as time interval, activity type, etc., to retrieve certain portions of the log data. When location data is involved, we also present a map (using Google Map API) to indicate the corresponding location, as shown in Fig 4.

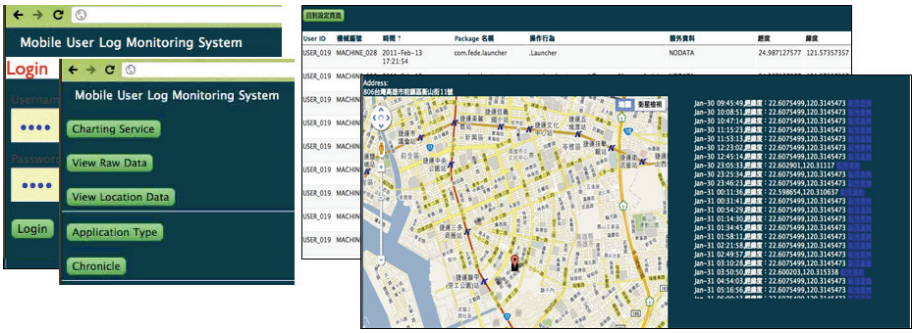

Fig 4. (Left) Log query interface (Right) Returned data

3.3.3 Real-time Log Monitor

User behavior cannot be explained using a single log item. In the past, three different levels of user behavior analysis have been identified, namely, activity, action, and operation [5] An activity usually consists of a sequence of actions and is more related to the motivation behind. An action is a logical unit with a clear goal, and an operation is the most fundamental physical unit. From the researcher's perspective, higher levels of description such as activity or action are of interest. However, a log item can only show us the operations performed. A sequence of logs will constitute higher-level events. Therefore, we need to establish the mapping between an event and its corresponding log sequence. This is achieved by the real-time log monitor service in which the collected logs are listed instantly on a web page in synchronization with user operation, as illustrated in Fig 5.

ACCOUNT	MACHINE_ID	TIME_STAMP	PKG_NAME	STATE	DATA	LATITUDE	LONGITUDE
	356299046685689	1301479982	com.funwish.huayu	.CandidateList	NODATA	24.986974625	121.57350357499999
	356299046685689	1301479985	com.google.android.gm	.ConversationListActivityGmail	NODATA	24.986974625	121.57350357499999
	356299046685689	1301479993	com.google.android.gm	.HtmlConversationActivity	NODATA	24.986974625	121.57350357499999
	356299046685689	1301479999	com.htc.launcher	.Launcher	NODATA	24.986974625	121.57350357499999
	356299046685689	1301480003	mobi.mgeek.TunnyBrowser	.BrowserActivity	NODATA	24.986974625	121.57350357499999
	356299046685689	1301480022	mobi.mgeek.TunnyBrowser	.BrowserActivity	http://www.google.com.tw/	24.986974625	121.57350357499999
	356299046685689	1301480037	com.htc.launcher	.Launcher	NODATA	24.986974625	121.57350357499999
	356299046685689	1301480038	com.htc.android.worldclock	.WorldClockTabControl	NODATA	24.986989457142858	121.57350208571428
	356299046685689	1301480061	com.htc.launcher	.Launcher	NODATA	24.986989457142858	121.57350208571428
	356299046685689	1301480081	com.htc.launcher	.Launcher	NODATA	24.986989457142858	121.57350208571428
	356299046685689	1301480082	com.android.htcdialer	.Dialer	NODATA	24.986989457142858	121.57350208571428

Fig 5. Real-time Log Monitor

4 Conclusion and Future Work

We have developed a service platform for logging and analyzing mobile user behaviors in this research. Our main objective is to complement existing practices such as survey and interview with automated process to facilitate the gathering of objective measures, namely, the user logs to understand and interpret user behavior. The log collection service is executed in the background to cause minimum interference to the user. As a result, long-term daily-life experience can be investigated. The three tools: log charting service, log query service and real-time log monitor work together to supply customized data in an effortlessly manner.

The proposed service platform can be used to gather logs that are relevant to e-learning or education. For example, there exist several e-book reader applications on the Android platform. Using the log collector, we can record and analyze usage patterns to understand reading behaviors and user preferences as well as evaluate the efficacy of the interface design.

The platform is designed to support concurrent experimentations. However, as the number of experiments increases, we need to look out for possible conflicts in the experiment design. Otherwise the log data may be meddled to such an extent that no sensible conclusion can be drawn from the observation. Moreover, we may need to design better scheduling policy for the delivery of log data to preserve power consumption on the mobile device. Finally, we will solicit feedback from researchers who have used our experimental platform to further enhance the quality of the service.

5 References

1. Eagle, N., Pentland, A.: Reality mining: sensing complex social systems. Personal Ubiquitous Comput. 10, 255-268 (2006)
2. Kärkkäinen, T., Vaittinen, T., & Väänänen-Vainio-Mattila, K.: I don't mind being logged, but want to remain in control: a field study of mobile activity and context logging. Proceedings of the 28th international conference on Human factors in computing systems, pp. 163-172. ACM, Atlanta, Georgia, USA (2010)
3. Ducheneaut, N., Yee, N., Nickell, E., Moore, R.J.: The life and death of online gaming communities: a look at guilds in world of warcraft. Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 839-848. ACM, San Jose, California, USA (2007)
4. Constandache, I., Gaonkar, S., Sayler, M., Choudhury, R.R., Cox, L.: EnLoc: Energy-Efficient Localization for Mobile Phones. In: INFOCOM 2009, IEEE, pp. 2716-2720. (Year)
5. Leontjev, A.N.: Activity. Consciousness. Personality. Englewood Cliffs, Prentice Hall (1978)