Smartphone Usage Analysis Based on Actual-Use Survey

Takeshi Kamiyama
NTT DOCOMO, Inc.

Kenji Hisazumi
Kyushu University

Hiroshi Inamura
Future University Hakodate

Teppei Konishi
NTT DOCOMO, Inc.

Ken Ohta
NTT DOCOMO, Inc.

Akira Fukuda
Kyushu University

ABSTRACT

This paper presents a smartphone usage model that will be useful as ground data for works such as proposals on improving services/technologies related to smartphones. It is based on an actual-use survey involving about 700 smartphone users. We conduct web interviews with users to get their demographic data - age, sex, for example, and gather many kinds of usage traces from their device using our logger application. The model describes 1) daily usage pattern (combinations of application usages) and 2) pattern features - demographics, major application usages and so on. Through a cluster analysis of the usage traces, we find there are six typical daily usage patterns and half the users have multiple patterns that differ with the day.

CCS Concepts

•Human-centered computing → Empirical studies in ubiquitous and mobile computing; •Social and professional topics → User characteristics;

Keywords

Smartphone usage; survey; ubiquitous and mobile computing

1. INTRODUCTION

Smartphone users now exceed feature phone users as the smartphone is now the most generally-used mobile terminal in the world[6]. Users can freely choose services/applications that suit their own purposes and interests, and open software development kits are being used by many application developers.

Service development is, however, rather erratic due to the problems and issues raised by each role such as marketing activities by the planning division and research for new technologies by the R&D division. A key difficulty is quantifying the value of a development because there are innumerable directions and possibilities of service/technology evolution. Whichever role we work on, it is important to set a goal and verify its effectiveness using logically defendable grounds. For example, to assess a new service proposal, we need to show who/how many target users there are, sales volume expected and so on. The same is true for technical developments.

In this paper, we present our smartphone usage model. It is based on an analysis of the usage patterns of actual users and their features such as application usages. The model will provide useful ground data for tackling the challenges related to the smartphone. Concretely, our model provides a practical evaluation basis or preconditions when elucidating research issues and validating their effectiveness.

In order to obtain the data needed to conduct the model analysis, we recruited actual smartphone users as the survey subjects. In the surveys, we interviewed each to gather demographic information - age, sex, place of residence and job, and collected usage traces of their device using our logger application for about one month. The trace mainly includes application usages, system configuration and device state such as that of a wireless interface, display, battery and so on. These data can be widely used for market analysis especially to identify target users and specify evaluation parameters such as preconditioned system configuration values for software improvement.

The rest of this paper is organized as follows. In Section 2, we discuss the data needed to advance existing studies in order to define the requirements of our smartphone usage model. We also overview the model. Section 3 describes the survey and summarizes the data obtained from the users. In Section 4, from the discussion in Section 2, we analyze daily smartphone usage patterns based on cluster analysis using categorized application usage. We then show demographic features of each pattern. Finally, we conclude the paper in Section 5.

2. SMARTPHONE USAGE MODELS

In order to elucidate the requirements *(R.1 - R.3)* of our usage model, this section discusses the points to be considered given related works on smartphone usage.

2.1 Related Works

We start with a case study of the commercialization of mobile phones for people past middle age[13]. The authors indicate that traditional commercialization processes tend to take too long in developing an understanding of product concepts and functional/quality requirements and result in products that are far from the customers' needs because

MobiCASE 2016, November 30-December 01, Cambridge, Great Britain

DOI 10.4108/eai.30-11-2016.2267052

the marketing process and research/development process are quite separate. As the case study was conducted in 2001, although almost all mobile phone users were young adults, they successfully created a new market for past middle age as well as improved productivity. The key to their success was that even researchers and developers in the R&D division participated in the marketing activities of planning division, and analyzed the needs of the targets in depth as well as the technical challenges. Recent smartphones have such high service capabilities, it is critical **to analyze/define target users clearly *(R.1)***.

As one recent service trend, many location-based services are being offered in the world[14], such as a location-dependent discount coupon delivery. However, location information is important private information as is name and postal address[12]. Therefore, location service is deactivated by default on most devices. This service must be manual activated by the user. Modern smartphones offer several methods for detecting device location such as GPS, wireless base stations, and so on. Thus service activation must be subdivided on the basis of which methods can be activated. Therefore, releasing a location-based service may not achieve the number of target users expected or service qualities may get worse unexpectedly. Thus it is critical **to consider the reality of device configurations in advance *(R.2)***.

Next, we mention some research works on technical issues related to smartphones. Dong et al. propose an energy-efficient control method for mobile OLED displays; it changes content color[1]. Noro et al. propose a GPU state control method that estimates the drawing termination, to reduce battery consumption[9]. In these works, the authors showed the efficiencies of their proposals through experiments on actual devices. However, actual effectiveness was not verified objectively because the preconditions of their verification were insufficient and restrictive. For example, it seems unlikely content color change will be accepted by the users because some contents, e.g. movies, are color critical. As another example, user satisfaction may depend on the application, device settings, and the device itself. In order to show actual effectiveness, **preconditions (i.e. application usages, device configuration and state) should be defined on appropriate grounds derived from actual smartphone usage *(R.2, R.3)***. The importance of the requirements is mentioned in other previous studies especially on mobile software engineering such as [5, 8, 4].

At the end of this subsection, we mention the existing studies related to user behavior. Flaski et al. studied actual smartphone usage using log data obtained from 255 users' devices in 2008[2]. They reported on the diversities of smartphone usage through their analysis of the data, e.g. display status, network traffic, battery state, application usage and so on. However, the results of the analysis fail to describe recent smartphone usage for the following reasons; 1) the analysis data is too old and the survey was conducted at an early stage of the worldwide spread of smartphones[11]. 2) the results of their analysis couldn't be used as ground data for developing new services or technical challenges, a goal of this paper, because the results were the simple averages of all subjects. Patil et al. analyzed the impact of user variation in the CPU/GPU workload in the actual usage of applications[10]. Although they studied the diversity of the workload using cluster analysis based on the statistics from 33 actual users, their experimental setup was too limited to adequately demonstrate the effectiveness of their work for two reasons. The number of subjects was too few, and the target applications selected were merely the most popular applications. Our studies in this paper suggest what is needed to create truly comprehensive experiments. Li et al. analyzed smartphone usage using log data obtained from millions of Android users[3]. However, they only studied macro-level features from the viewpoint of application usage, and didn't focus on daily application usage patterns as this paper shows later.

2.2 Requirements of our model

Based on the discussion in the previous subsection, we present the requirements of our smartphone usage model. The model should satisfy the following requirements.

R.1 *Demographics*

R.2 *Device Configuration and State*

R.3 *Application Usage*

In order to clearly explain the usage diversity of each user, the model should be able to handle usage patterns, which is the combination of each item above, instead of simply averaging the items. Moreover, even the same user is expected to have multiple usage patterns that depend on the day such as weekdays/weekend.

Therefore, in order to derive smartphone usage patterns in this paper, we classify "daily usage patterns" based on a cluster analysis that uses daily application usage times as the main explanatory variables. Basically, the smartphone usage patterns can be identified through application usage because all smartphones functionalities are implemented as applications. As mentioned above, we define our smartphone usage model so as to provide quantitative explanations of the above items for each pattern (in other word, cluster).

The expectation is that referring to our model will make it easier to investigate and specify the realities of target users in the marketing activities for new products and services. Also with regard to work on technical challenges, we expect the model will help in preparing the ground data such as the precondition of an evaluation.

3. SURVEY OF SMARTPHONE USAGE

In this section, we describe the actual survey conducted to obtain the data for our model study, and show a summary of each survey item.

3.1 Survey methods

In this survey, which started at the end of Oct. 2013, we recruited actual users of two major models of smartphones from all over Japan as subjects of the survey through a research company. Table 1 shows an overview of the survey. Incidentally, we didn't target underage users due to the company's policy, and the survey was conducted after reaching agreement about privacy and data use with all subjects.

First, we interviewed via a website the 694 users to collect their demographic information. Then, we obtained usage logs from 391 users, selected from among the interviewed subjects, using a logger application that they loaded on to their devices. The logger application ran in the background so as not to affect usage style. It continuously recorded the beginning/end time of each application, device configuration and states.

3.2 Statistical information of the survey

3.2.1 *Demographics*

The demographic distribution of the 694 users is shown in Figures 1-4. We can see that the coverage was wide with no bias, so our model is expected to be useful for various kinds of works. No significant difference in usage between Device A and B was found. 95% of the users had only one device as shown in Figure. 5

Table 1: The overview of the survey

Device	A:Galaxy S4	B:Xperia A
Interview period	2weeks from the end of Oct.2013	
Num. of users	221	473
Logging period	1month from the mid. of Nov.2013	
Num. of users	112	279
Log amount(man-day)	3136	7812

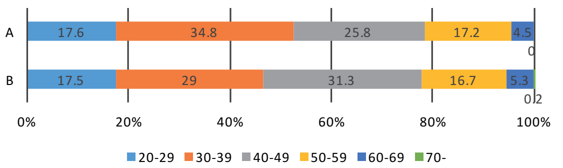

Figure 1: Age distribution of users

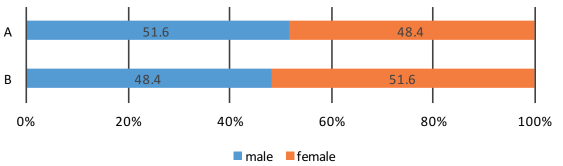

Figure 2: Sex distribution of users

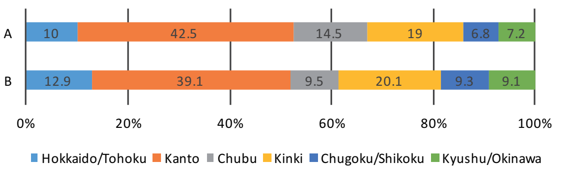

Figure 3: Area distribution of users

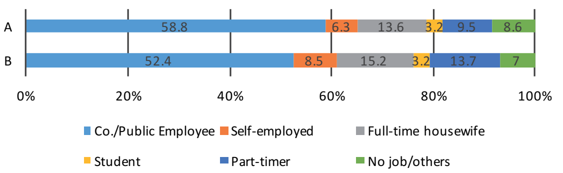

Figure 4: Occupation distribution of users

3.2.2 *Other survey items*

We summarize the survey items other than the demographics below.

Figure 6 shows the percentage of the share of Home applications, which means home screens where shortcut icons of other applications are placed. On both devices, preinstalled home applications were used by more than half the users. However, the post-installed home application ('other' in Figure. 6) occupies about 30% on Device A, but only 9.3% on Device B.

Figure 7 shows usages of the Screen Unlock Types. About 80% of users activate the screen lock in some way, and swipe/touch occupies the greater part.

Periodic Application Update, shown in Figure 8, is the automatic update for the handset vendor's own applications. This configuration is activated by about half the users.

Location Service in Figure 9, which must be 'configured allow' for location-based applications to obtain the current location of users, is permitted by about 60% of users on both devices. Moreover, using Wireless Network shown in Figure 10 is an optional configuration of the Location Service, which enables quick location detection using the information of wireless base stations[7]. using GPS shown in Figure 11 is another optional configuration. It enables location detec-

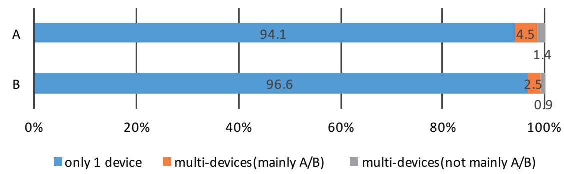

Figure 5: The number of devices owned by each user

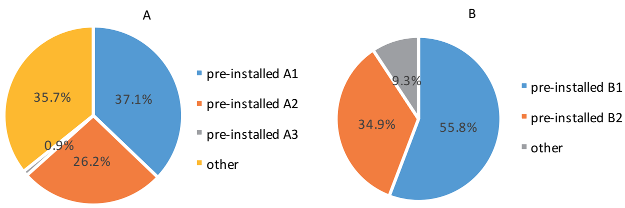

Figure 6: Usages of Home applications

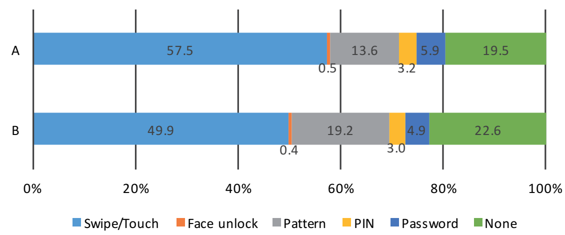

Figure 7: Usage of Screen Unlock Types

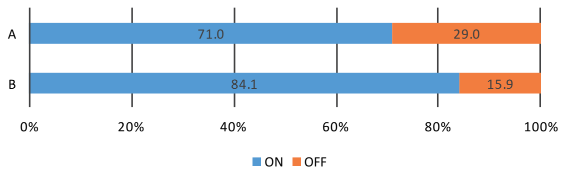

Figure 8: Usage of Periodic Application Update

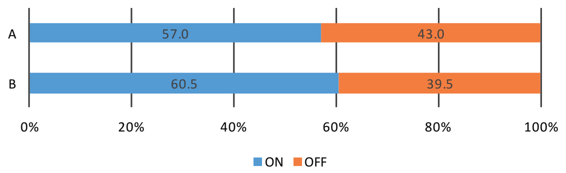

Figure 9: Usage of Location Service conf.

tion using the device's GPS. About 60% of users activate the former, while only about 30% users activate the latter. According to the supplementary interview, 50% of users switch the GPS setting to ON as needed.

As mentioned above, the location-related configurations are initially deactivated and cannot be changed automatically by applications to protect users' privacy. Therefore, although many location-based services are being released, they may fail to reach 30% of the users unless this fact is considered. The services may fail due to insufficient accuracy depending on the GPS setting.

3.2.3 *Application Usage*

In this subsection, we study general features of each application usage using the application usage log. We focus on

- Total usage time of each application
- Average daily usage time of each application

We identified 3339 applications from the data. The "usage" mentioned here means only the active application in the foreground that the user can interact with. Moreover, we also assessed the usage differences depending on the power state, which means battery-powered (discharging) and USB-powered (charging), using the battery state log.

The rest of this paper focuses on device B as it was used by more subjects.

Tables 2 and 3 show the total usage time of each application in the battery-powered state and the USB-powered state. In both states, Browser usage time was the longest followed by Home, Messenger, Mail. All are much longer than the others. Moreover, the data reveals that Messenger has already become a major human-communication tool compared with Mail and Phone call. And also, game applications occupy five of top 20, though the number of users is not so big, at most 41 users.

In addition, the interesting feature for the USB-powered state is, as marked in Table 3, that long-term applications (e.g. Video viewer 1) occupy higher ranks than is true for battery-powered state. A few minor applications such as Slideshow are shown in the table due to a few extremely heavy users.

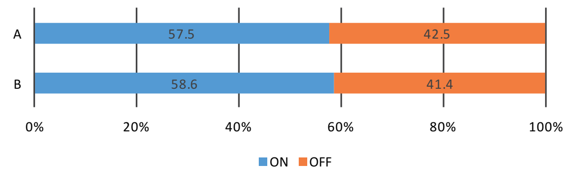

Figure 10: Usage of 'using Wireless Network' optional conf.

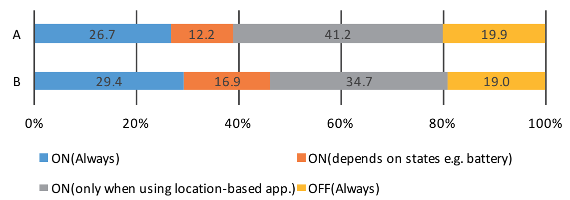

Figure 11: Usage of 'using GPS' optional conf.

Table 2: Total usage time of each application(battery-powered) - Top 20 long usage applications

No.	Application	Unique User	Total usage time hh:mm:ss
1	Preinstalled browser	267	4191:12:43
2	Preinstalled home 1	164	1216:35:23
3	Messenger	202	1159:40:30
4	Mail	278	1133:54:54
5	Preinstall home 2	123	789:46:01
6	Puzzle game 1	43	479:08:16
7	Browser 1	150	470:59:14
8	Web portal	108	453:19:06
9	SNS1	151	402:41:21
10	Phone call	276	354:17:55
11	Puzzle game 2	26	343:02:26
12	Twitter client	61	283:11:13
13	SNS2	15	226:59:36
14	Puzzle game 3	42	218:36:15
15	Puzzle game 4	29	195:43:32
16	Puzzle game 5	38	192:00:07
17	Online video viewer	175	190:50:48
18	Photo album	247	136:39:56
19	Puzzle game 6	3	112:16:05
20	Textboard viewer	14	108:40:41

Next, we study the average daily usage time per person for the top 25 applications in order of the number of users as shown in Table 4. Although it could be said these are typical applications, according to the table, all are, excluding No.10, 17 and 25, applications preinstalled on the device. The three applications are top-hits all over the world, but few other applications are competitive with the preinstalled ones.

Table 3: Total usage time of each application(USB-powered) - Top 20 long usage applications

No.	Application	Unique User	Total usage time hh:mm:ss
1	Preinstalled Browser	248	957:39:52
2	Preinstalled Home 1	162	805:58:07
3	Preinstalled Home 2	116	663:53:53
4	Messenger	193	268:01:01
5	Mail	271	211:53:56
6	Slideshow *	9	208:39:57
7	Browser 1	93	134:34:15
8	Video viewer 1 *	7	115:39:28
9	SNS2	11	104:15:44
10	Web portal	91	96:23:57
11	RPG game	1	95:04:22
12	Media player *	38	93:10:21
13	Exchange trading *	1	91:57:42
14	Screensaver *	1	90:58:01
15	SNS1	105	85:03:34
16	Twitter client 2 *	1	81:49:43
17	Puzzle game 1	41	81:25:59
18	Alarm clock *	172	68:10:10
19	Online video viewer	87	66:03:47
20	Twitter client 1	53	65:03:29

3.2.4 Categorized Application Usage

As mentioned in the previous subsection, we confirmed not only the usage of typical applications used by a lot of users but also that of long-term applications used by a few users through an in-depth analysis of application usages. That is, there could be diversity in the usage patterns as combinations of the applications. However, it is next to impossible to analyze the immense number of the combinations at the application-level; 3339 applications were identified from the usage log. Therefore, we analyze the application usage time at the application category-level to develop a bird's-eye view of the feature of the overall application usages.

We categorized the applications as follows.

1) Categorize according to Google Play categories

2) Make applications that are used by many users for long periods independent categories

3) Manually categorize any remaining uncategorized applications whose usage time is extremely long into the existing categories

In step 1), we adopt 30 categories after subdividing the original Game category into 7 subcategories, according to Google Play as of Jan.,2014.

The reason for step 2) is that major and often-used applications determine the essential characteristics of the existing categories. Step 2) yielded 26 independent categories.

In step 3), 24 uncategorized applications were categorized into the existing categories.

Based on the procedure above, we define 57 categories including the other (uncategorized) as shown in Table 5. The categories under the 11th row are defined in step 2).

Figure.12 shows category-based total usage time. We can see other application usages that couldn't be shown in Table 2, for example, News/Magazines, Shopping and so on.

Table 4: Average daily usage time for each user - Top 25 major applications

No.	Application	Unique User	Total usage time hh:mm:ss	Median
1	Mail	278	00:11:05	00:04:25
2	Phone call	277	00:04:54	00:01:29
3	Google Play	276	00:02:26	00:00:49
4	Camera	273	00:01:57	00:00:37
5	Configurations	273	00:02:07	00:00:26
6	Preinstalled browser	269	00:48:34	00:21:38
7	Photo album	248	00:03:08	00:01:17
8	Web portal 2	245	00:02:05	00:00:39
9	Alarm clock	206	00:02:16	00:00:31
10	Messenger	203	00:18:29	00:09:18
11	Schedule	193	00:02:21	00:00:44
12	Map	191	00:06:46	00:02:15
13	Online video viewer	183	00:16:40	00:05:14
14	Preinstalled home 1	167	00:27:42	00:13:13
15	Calculator	156	00:03:16	00:00:56
16	Browser 1	153	00:22:04	00:07:48
17	SNS1	151	00:12:08	00:06:09
18	Auto. App. Update	148	00:01:54	00:00:22
19	Mail 2	143	00:04:05	00:01:48
20	Video viewer 2	127	00:04:07	00:00:54
21	Web search	124	00:02:21	00:01:05
22	Preinstalled home 2	123	00:28:06	00:12:17
23	Music player 1	119	00:03:33	00:01:02
24	Media player	112	00:23:24	00:03:12
25	Web portal 1	109	00:17:35	00:05:43

Table 5: Application categories

Application Categories		
Lifestyle	Personalization	Productivity
Weather	Photography	Game_Music
Entertainment	Tools	Communication
Social	Comics	Game_Casual
Game_Sports	Media_Video	Books_Refference
Finance	Game_Casino	Travel_Local
Game_Arcade_Action	Shopping	Transportation
Game_Puzzle	Sports	Game_Racing
Libraries_Demo	Business	Education
Health_Fitness	News_Magazines	Medical
Camera	Map	Phone call
Google Play	Mail	Calculator
Online video viewer	Web portal 1	Schedule
System	Browser	Agent 1
Preinstalled home	Configuration	Calendar
Data Backup	Music Player 1	Virus scan
Web portal 2	Clock	Agent 2
SNS1	MMS	Restaurant review
Messenger	Home	The others

4. ANALYSIS OF SMARTPHONE USAGE MODEL

The previous section presented categorized application usage time to study the tendency of application usages as a whole. In this section, we derive the usage patterns, the combination of multiple applications (the categories), based on cluster analysis. We then present the patterns explained with their features as our smartphone usage model.

This time, the target of our modeling is "daily usage" without identifying each user for the following reasons; 1) almost all people follow a similar daily life cycle. 2) even the same person may exhibit in a different life pattern depending on the day such as weekdays/weekend. We evaluate the hy-

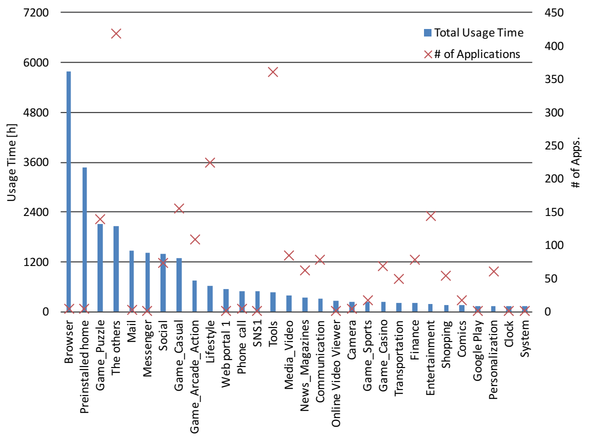

Figure 12: Category-based Application Usages

pothesis of 2) at the end of Section 4.

4.1 Classification of Daily usage patterns

4.1.1 *Explanatory variables for the cluster analysis*

To classify daily usage by cluster analysis, we defined the following explanatory variables for each daily usage. The dataset of variables holds 7784 data points, which equals the number of the target days analyzed.

1) Categorized application usage time (56 variables)

2) Device working time while display is OFF

3) Display brightness

56 Variables associated with 1) were defined based on the application categories in Section 3 excluding "The other (uncategorized)". In addition, values of the variables were normalized and log-transformed due to the skewed distribution. We also defined two more variables for 2) and 3) to offset the lack of each application's own features due to the abstraction from application-level to category-level as demanded by the definition of 1). The concrete reasons are as follows. Actually, the application usage(working) time consists of not only the period in the state of display ON, which users recognize visually as their usage time, but also another period where the application is running in the background, especially in the state of display OFF. We note that the display brightness is a system-level common configuration; each application can change the brightness level in the foreground (while it occupies the display) to suit the contents being displayed. These are very important factors when, for example, we tackle the challenge of optimizing battery usage on devices. However, such application-specific features are lost due to the definition of variables of 1). Therefore, we additionally defined variables 2) and 3) so that model can consider such features of device resource usage.

4.1.2 *Result - clusters*

We present the result of the cluster analysis performed on the dataset shown in 4.1.1. As shown in Table 6, we tried 7 patterns of k-means clustering by changing the number of clusters from 4 to 10.

Table 6: The member composition ratio between clusters

Composition Ratio(%)		The number of clusters(k) 4	5	6	7	8	9	10
Cluster No.	1	32.2	30.1	22.9	21.9	19.3	16.9	16.3
	2	24.2	21.1	18.7	17.9	16.2	15.4	15.0
	3	24.2	17.6	17.3	16.7	15.7	12.8	11.3
	4	19.5	15.9	15.1	14.7	13.6	11.9	11.1
	5		15.2	14.6	14.4	12.3	11.3	10.3
	6			11.4	10.8	10.2	10.3	9.4
	7				3.7	9.0	9.6	8.0
	8					3.7	8.0	7.5
	9						3.7	7.4
	10							3.7

Table 7: Transfer rate of cluster members by changing k from 5 to 6

Transfer Rate(%)		Source Cluster No. (k=5) 1	2	3	4	5
Dest. Cluster No. (k=6)	1	69	5	1	4	1
	2	0	88	0	0	1
	3	0	0	98	0	0
	4	0	0	0	94	0
	5	0	0	0	0	95
	6	31	7	1	1	2

Table 8: Transfer rate of cluster members by changing k from 6 to 7

Transfer Rate(%)		Source Cluster No. (k=6) 1	2	3	4	5	6
Dest. Cluster No. (k=7)	1	95	1	0	0	0	0
	2	0	93	0	0	0	4
	3	0	0	96	0	0	0
	4	0	0	1	96	0	0
	5	0	0	0	0	96	1
	6	0	0	0	0	0	93
	7	4	6	2	2	3	2

Although it is difficult to determine an overall most appropriate number of clusters (k), we focused on the clusters with k=5 and 6 based on the idea that no cluster should be greater larger than or smaller than any of the others.

Next, to determine which cluster set is appropriate, we inspected how many cluster members with k=5 belong to the clusters with k=6. As shown in red in Table 7, there was no big difference other than the members of source cluster 1 were divided into destination clusters 1 and 6.

Furthermore, focusing on member transfer between the source(k=6) and the destination(k=7) in Table 8, a few members of each source cluster constitute a new small destination cluster 7.

As mentioned above, we determined the appropriate number of clusters to be 6, in the other words, there are 6 kinds of smartphone usage models.

4.2 Analysis of each cluster's features

In this subsection, we characterize each smartphone usage pattern. Table 9 shows cross tabulation values, i.e., the daily average values of categorized application usage time, display brightness, battery usage and so on cluster by clus-

Table 9: Features of each cluster

Cluster No.	# of days	Total app us time	Display brightness	Battery Usage(%/h)	Browser	Preinstalled home	Messenger	Mail	Web portal 1	SNS1	Phone call	Clock	Online Video Viewer	Camera	Home
1	1779	135.98	116.72	2.52	52.28	19.34	0.16	9.87	1.68	1.96	2.48	0.69	1.43	1.12	1.16
2	1457	188.20	126.60	3.07	58.16	32.65	23.32	11.55	2.54	6.48	5.81	0.91	2.11	2.88	1.51
3	1346	218.58	135.22	3.56	42.11	24.59	15.55	12.70	6.45	2.12	3.65	1.30	1.25	1.90	0.65
4	1174	241.16	124.74	3.76	51.50	40.46	10.55	16.81	7.64	5.60	3.92	0.93	3.42	3.06	0.34
5	1137	291.35	118.14	4.30	43.74	28.26	13.39	11.31	4.33	5.27	4.03	1.95	3.19	1.52	0.87
6	891	78.17	110.70	1.91	0.31	15.33	3.20	5.50	4.19	1.05	3.88	0.93	0.53	0.79	0.32

Cluster No.	Google Play	System	Config.	Map	Schedule	Web portal 2	Calculator	Virus scan	Music Player 1	Restaurant portal	Agent 1	Agent 2	Calendar	MMS	Data Backup
1	1.09	0.73	0.43	0.66	0.36	0.32	0.53	0.16	0.22	0.13	0.14	0.07	0.02	0.00	0.05
2	0.76	0.87	0.45	0.62	0.29	0.29	0.25	0.34	0.24	0.30	0.05	0.03	0.13	0.00	0.04
3	0.86	1.12	0.37	0.41	0.53	0.16	0.17	0.08	0.21	0.18	0.13	0.07	0.04	0.00	0.02
4	1.27	1.86	0.66	1.38	0.21	0.31	0.20	0.35	0.53	0.16	0.27	0.27	0.04	0.00	0.02
5	1.63	1.67	2.84	0.27	0.21	0.21	0.14	0.13	0.15	0.25	0.47	0.31	0.01	0.00	0.02
6	1.51	0.32	0.25	0.09	0.65	0.56	0.21	0.05	0.31	0.07	0.07	0.08	0.01	0.00	0.01

Cluster No.	Game Puzzle	Game Casual	Social	Game Arcade Action	Lifestyle	Comm.	Game Casino	Tools	News Magazines	Productivity	Media Video	Entertainment	Finance	Game Sports	Transport-ation
1	0.08	0.16	0.37	5.82	5.98	1.73	0.97	6.10	2.82	0.56	1.93	2.23	3.51	0.83	1.73
2	0.04	0.21	0.15	3.90	4.60	1.05	0.53	4.76	4.52	0.90	4.66	2.13	0.42	0.05	2.19
3	64.81	0.13	5.02	7.13	3.08	2.15	1.11	0.90	0.58	0.40	1.17	1.01	0.78	5.24	1.99
4	1.39	0.87	48.95	2.76	6.49	1.10	3.34	3.10	3.82	0.96	4.97	2.24	0.28	0.70	0.83
5	31.93	64.08	18.34	14.07	3.89	7.14	2.97	2.41	2.90	1.04	3.66	1.79	1.09	2.76	2.01
6	0.39	2.74	0.69	0.66	3.49	2.77	2.56	2.48	0.74	0.62	2.04	0.40	2.94	1.23	1.06

Cluster No.	Comics	Shopping	Personaliza-tion	Health Fitness	Books Refference	Photography	Education	Game Music	Business	Travel Local	Weather	Sports	Game Racing	Medical	Libraries Demo
1	0.52	1.13	0.25	0.42	0.38	0.10	0.10	0.27	0.09	0.19	0.09	0.38	0.14	0.00	0.00
2	1.80	0.95	0.80	0.32	0.32	0.20	0.14	0.24	0.43	0.20	0.08	0.03	0.00	0.00	0.00
3	0.31	1.73	2.58	0.39	0.21	0.35	0.18	0.21	0.26	0.10	0.11	0.02	0.00	0.04	0.00
4	0.04	2.75	0.93	1.09	0.66	1.16	0.14	0.17	0.17	0.21	0.19	0.05	0.04	0.00	0.00
5	0.32	0.85	0.02	1.91	0.49	0.76	0.16	0.21	0.15	0.05	0.18	0.00	0.00	0.02	0.00
6	6.14	0.39	2.88	1.02	1.68	0.04	0.47	0.09	0.11	0.18	0.06	0.07	0.02	0.00	0.00

A unit of time : minute

Table 10: Summary of cluster features

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster6
Main features	Tools (little Messenger)	**Browser** **Messenger** SNS1 通話	**Game_Puzzle** Game_Sports Display bright	**Home** **Mail** **Social** Web portal 1 Lifestyle	**Game_Casual** **Game_Arcade/Action** **Communication** heavy battery usage	**Comics** App.Usage time: short
Pattern Name	Mid. use non-Mssenger	Mid. use Messenger day	Puzzle game	Non-voice communication	Heavy use	Light use

ter. That is to say, the table itself shows the raw features of each usage pattern. In addition, the categorized usage times in red/blue are relatively longer/shorter compared with the other cluster.

The light-use cluster, which the total application usage time is the shortest, is cluster 6 for about one hour and 20 min. On the other hand, the heaviest-use cluster is cluster 5 followed by cluster 4. Their total time is about four to five hours a day.

The cluster with the highest display brightness is cluster 3. This cluster follows cluster 4 in terms of total usage time. This is reasonable as cluster 3 includes Puzzle games.

Cluster 5 consumes most battery energy, while cluster 6 consumes the least. This is because of the difference in total use time.

Table 10 summarizes the features of each cluster.

Cluster 1 can be called middle-use in terms of total usage time. however, long usage in every category is not confirmed. To stretch a point, the usage of the Tools category is long. Its particular feature is that there is little usage of the major and the long-used applications such as Messenger and SNS1 in the top 25 in Table 4.

Cluster 2 is middle-use as is cluster 1, which is characteristic of Messenger and SNS1 from the viewpoint of application usage compared with cluster 1.

Cluster 3, as mentioned above, is characterized by the usage of games and display brightness.

Although both cluster 4 and 5 are characterized as heavy-use, they have different features from the viewpoint of application categories. cluster 4 is characterized by the usage of Home, Mail and Social categories, while cluster 5 is characterized by the usage of Game_Casual/Arcade/Action and Communication categories. Moreover, the battery usage of cluster 5 is higher than that of cluster 4. Therefore, the difference in the categories could affect the battery usage.

Cluster 6 can be called the lightest-use group as mentioned above. To stretch a point, the Comics category is relatively long usage.

4.2.1 Demographics of each cluster

We present the demographic features of each cluster extracted by the cross tabulation analysis. As mentioned above, the result of the cluster analysis is the classification of daily usage, that is to say dates of use. Therefore, to derive the demographic statistics of each cluster, we count the demographic values of each user according to the dates of use.

Figure 13 shows the age distribution of each cluster. Focusing on clusters 1 and 6, users over the age of 40 were active 70% of the dates. In the case of clusters 2, 3, 4 and 5, users under the age of 40 were active in each cluster on more than half the dates. Therefore, the age distributions of the clusters differ little.

Figure 14 shows the sex distribution of each cluster. In the cases of clusters 1and 6, male users were active on the majority of use dates, while female users were most active for clusters 2, 3, 4 and 5. The sexes exhibit different distributions in the clusters.

Figure 15 shows the residential area distribution of each cluster. As shown in Figure 15, there is no unique association between clusters and residential areas because the distributions of each cluster are almost the same. To stretch a point, Kinki (area) has slightly higher ratio than the others in cluster 2.

Figure 16 shows the occupation distribution of each cluster. As is true for the residential area distribution, there is little difference among the clusters. A small feature is Student has lower ratio for clusters 1 and 6 than the others. This is because that the ratio of users under the age of 30 is very low in clusters 1 and 6 as shown in Figure 13.

4.2.2 Diversity in Usage Patterns of each user

As mentioned at the beginning of Section 4, we assumed that even the same user may use his/her smartphone in different ways (patterns) depending on the day. In order to validate this hypothesis, we study the number of usage patterns of each user, that is the number of clusters to which each user belongs (described Nc in the following). Figure 17 shows the user distribution of each Nc. We count Nc of each user and assign them to two groups if they meet two conditions: the number of the days using the same cluster is 1) one day or more, and 2) four days or more.

Focused on the distribution of 1), the number of usage patterns is at most three for 90% users. However, Nc in this case may be indicative of extremely-exceptional patterns which occur infrequent in the logging period, about one month. Therefore, we also study the distribution of 2) to confirm only typical usage patterns. Here, the number of usage patterns is at most two for 90% of the users.

Next, we confirm how many combinations of the typical patterns are adopted by each user. Table 11 shows the features of the top five combinations of the typical usage patterns. Take, as an example, the combination of cluster 1 and 2, both of them are middle-use pattern; they differ as in whether the Messenger application is used or not. In this way, we can see that there exist specific common features in the major combinations.

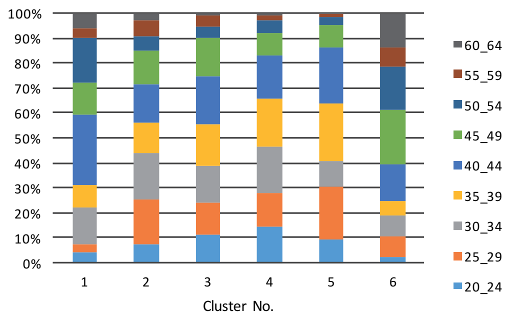

Figure 13: Age distribution of each cluster

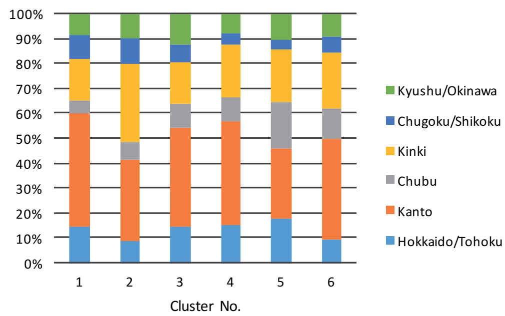

Figure 15: Residential area distribution of each cluster

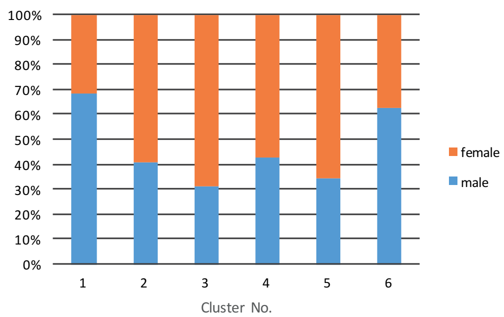

Figure 14: Sex distribution of each cluster

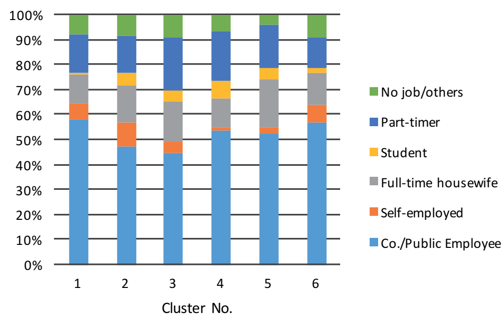

Figure 16: Occupation distribution of each cluster

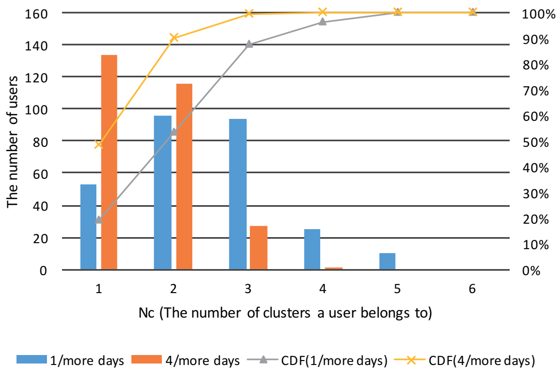

Figure 17: The number of usage patterns for each cluster

Table 11: The common combinations of the typical usage patterns

No.	Combo[1]	UU[2]	Common features	Difference
1	C1 & C2	24	Middle-use	<Messenger usage> C1:rarely,C2:often
2	C1 & C6	22	Light-use	<Browser usage> C1:1hour,C6:1min.
3	C1 & C4	14	<Browser usage> 50min.	C4:long usage of Mail &Social categories C4:100min. longer total usage time
4	C3 & C5	10	Heavy-use Long usage of Puzzle game	C5:long usages of the other kinds of Game
5	C2 & C3	7	Middle-use	C2:long usage of Browser&Messenger C3:long usage of Puzzle game

[1] combination of usage patterns(cluster#s)
[2] The number of unique users

5. CONCLUSIONS

In this paper, we presented a smartphone usage model based on a survey of actual users and usage log collection from the devices of the actual users. Based on cluster analysis using 58 explanatory variables such as categorized application usage time, we derived six daily usage patterns and studied their features including demographics. Moreover, focusing on the diversities of typical usage patterns of each user, we found out almost all users employ their smartphone in at most two patterns depending on the day.

6. REFERENCES

[1] M. Dong and L. Zhong. Chameleon: A color-adaptive web browser for mobile oled displays. In *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services*, MobiSys '11, pages 85–98, New York, NY, USA, 2011. ACM.

[2] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin. Diversity in smartphone usage. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*, MobiSys '10, pages 179–194, New York, NY, USA, 2010. ACM.

[3] H. Li, X. Lu, X. Liu, T. Xie, K. Bian, F. X. Lin, Q. Mei, and F. Feng. Characterizing smartphone usage patterns from millions of android users. In *Proceedings of the 2015 ACM Conference on Internet Measurement Conference*, IMC '15, pages 459–472, New York, NY, USA, 2015. ACM.

[4] X. Lu, X. Liu, H. Li, T. Xie, Q. Mei, D. Hao, G. Huang, and F. Feng. Prada: Prioritizing android devices for apps by mining large-scale usage data. In *Proceedings of the 38th International Conference on Software Engineering*, ICSE '16, pages 3–13, New York, NY, USA, 2016. ACM.

[5] I. Manotas, C. Bird, R. Zhang, D. Shepherd, C. Jaspan, C. Sadowski, L. Pollock, and J. Clause. An empirical study of practitioners' perspectives on green software engineering. In *Proceedings of the 38th International Conference on Software Engineering*, ICSE '16, pages 237–248, New York, NY, USA, 2016. ACM.

[6] Ministry of Internal Affairs and Communications. *2015 White Paper on Information and Communications in Japan*. The Government of Japan, 2015.

[7] D. Mohapatra and S. B. Suma. Survey of location based wireless services. In *2005 IEEE International Conference on Personal Wireless Communications, 2005. ICPWC 2005.*, pages 358–362, Jan 2005.

[8] M. Nagappan and E. Shihab. Future trends in software engineering research for mobile apps. In *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, volume 5, pages 21–32, March 2016.

[9] M. Noro, T. Murakami, T. Kamiwada, and T. Ishihara. Gpu state control method with estimating termination of drawing. *IPSJ Journal*, 57(2):394–405, feb 2016.

[10] S. Patil, Y. Kim, K. Korgaonkar, I. Awwal, and T. S. Rosing. Characterization of userâĂŹs behavior variations for design of replayable mobile workloads. In *Proceedings of the 7th International Conference on Mobile Computing, Applications, and Services*, MobiCASE '15, pages 51–70. Springer, 2015.

[11] E. M. Rogers. *Diffusion of Innovations, 5th Edition*. Free Press, 5th edition, 8 2003.

[12] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux. Quantifying location privacy. In *2011 IEEE Symposium on Security and Privacy*, pages 247–262. IEEE, 2011.

[13] T. Tomita, Y. Igarashi, M. Yamasawa, K. Chujo, M. Kato, I. Iida, and H. Mineno. A study of the product development process through strengthened fundamental r&d - based on a case study of mobile phone businesses for senior users -. In *Proceedings of International Workshop on Informatics(IWIN2014)*, pages 69–78. Informatics Society, 2014.

[14] K. Zickuhr. Location-based services. *Pew Research*, pages 1–25, 2013.