Smartphone Usage Pattern based on Geographical Location and Addiction Issue

Rong Zhou, Dianxi Shi*
National Laboratory for Parallel and Distributed Processing
National University of Defense Technology
Changsha, China 403105
Email:dxshi@nudt.edu.cn,shilx__123@sina.com

***Abstract*—The outcome of smartphone greatly changes the life of people. It is of importance to understand the usage pattern of varieties of applications in social life. Many researches have centered on this topic through statistics of the log files. However, most of them treat the location of the user as an element of the statistical result but not the cause for the usage pattern. In this paper, by analyzing the time and frequency of APP in different location, we find the location of the user is one of crucial factors for usage pattern. The migration takes up a great proportion to change the habit of users, implying that it is an effective cure for addicted people .**

***Index Terms*—Smartphone usage, user behavior, location, smartphone addiction, DJ-cluster, semantic translation**

I. Introduction

The eMarketer [1] predicts there would be more than one third of the population worldwide using smartphone and this growing trend would continue for a long time. After traditional phone extends the range of our ears, the touch screen and integrated sensors (GPS, accelerometer, gyroscope and so on) in smartphone build a more convenient bridge to connect our brain with the outside world. The small volume of such hand-held device makes it portable and its continuously enhanced functions, such as obtaining news, watching videos, playing games, learning and so on, make it more like a new "organ" instead of a tool.

Although smartphone plays an important role in our life, high dependence on it also draws the attention of the public. In one aspect, the developers of smartphone always try their best to attract users because the huge commercial profit is proportional for the time that user spends. In another aspect, similar to over-use of our eyes, overdependency on smartphone can also bring harm to ourselves. Once stickiness evolves into addiction, the degree of hazard might be like drug addiction. At least, there are a rather large number of phubbers all over the world that not only hurt their own health but also the emotions with their families or friends. Therefore, we should be careful to understand the behavior of users, balance the commercial profit and users' health, clarify the boundary between normal use and addiction and make intervention if necessary.

Lee *et al.* [2] have developed the Smartphone Addiction Management System (SAMS) for the mental health research due to the pathological use. It is difficult to define behavior addiction for it is related not only to physical but also to psychological and social aspects, the researchers define smartphone addiction as the excessive use of the smartphones that interferes with the daily life of the users, which is accompanied by tolerance, withdrawal symptoms, salience, mood modification, craving, loss of control and similar clinical features. Ahn *et al.* [3] presented the comparison between smartphone addict and non-addict people through popularity analysis of applications, usage pattern analysis of weekly and hourly patterns and usage pattern-based categorization methods based on SAMS system.

However, these works only recorded the tracking information about location. There is no connection between APP usage patterns with geographical entities (such as workplace, classroom, shopping mall and so on). Therefore, this paper further studies such relationship. We center on usage pattern of student and combine geographical entities with the usage habit by associating the semantics of GPS data. We develop the analyzer: LUHA (Location-based User Habit Analyzer) and our findings are mainly as follows.

- Time might be superficial phenomenon if we divide a day into several stages and make statistics of behaviors per stage. The habits of using APP are highly dependent on the geographical environment. In short, the place where the users are staying mainly determines the behavior pattern of most people.

The remainders of this paper are organized as follows. In section II we present preliminary background and related work about usage pattern and addiction research. In section III we analyze the difference of existing strategies, build a compatible framework based on location and describe our model. Experiment and result evaluation are provided in section IV. Section V draws conclusion and future work.

II. Related Work and Background

In the past decade, substantial diversity in smartphone usage has been studied. At early stage, scholars studied usage behavior by asking people to record their SMS (Short Message Service) [4]. Questionnaire survey was designed to study the difference of massage and call usage [5]. By means of reporting people's own situation of using call, massage, APP and playing games, Butt and Philips [6] showed the relationship between phone usage and five personality models of phone user. However, reporting is subjective, leading to inaccurate

978-1-5090-2129-1/16/$31.00

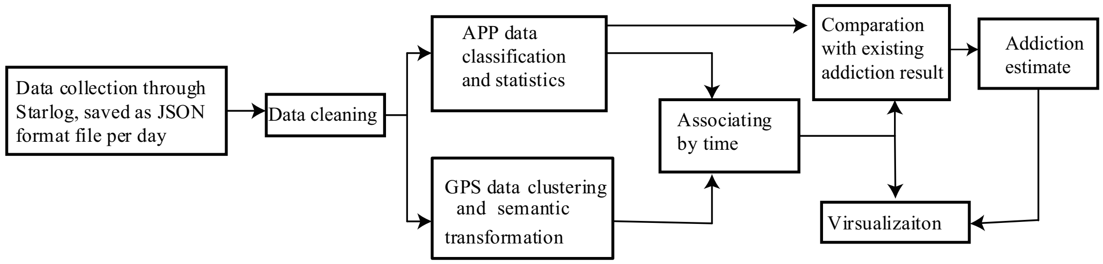


Fig. 1. The data processing framework of LUHA.

or wrong results. Therefore, using sensors in smartphone to capture information about usage behavior is popular for it can be more objective and accurate. Scholars in literatures [7] [8] tended to collect data continuously by smartphone and combined the data with other factors, such as questionnaire survey, to analyze usage behavior. Apart from smartphone, Anderson *et al.* [7] focused on several mobile devices, including notebook computer, ipad, mobile phone, Internet notebook and so on. They provide a "racial mining" algorithm to study a large number of people. The paper [9] centered on the usage pattern of large-scale mobile phone applications, but it took no consideration about the location information. Battestini *et al.* [10] collected mobile phone software usage records of 70 college students for more than 4 months and recorded the location where they sent and received text messages. They analyzed the relationship between the usage behavior of short message and the location. The paper [11] studied the relationship between social context and location. They adopted a large-scale experiment in a natural and unobtrusive manner of daily life and analyzed multiple phone applications (SMS included) with location anchors automatically estimated from multiple sensor data types (GPS, GSM, Wifi, motion), which results in higher accuracy in location estimation. In addition, they presented an analysis of usage with a proxy for social context derived from Bluetooth. The behavior patterns of diverse usage were studied in [12]. The authors characterized intentional user activities interactions with the device and the applications used and the impact of those activities on network and energy usage. The level of diversity suggests that mechanism to improve experience or energy consumption will be more effective if they learn and adapt to behavior. The SAMS can be the forerunner who researched the relationship between addiction and smartphone usage pattern. Based on these previous work, we combine location as one factor to study its influence on addict and non-addict behavior.

III. THE FRAMEWORK OF THE ANALYZER

As we all know, iOS and Android are the two mainstream smartphone operation systems in the world. Because Android has an overwhelming advantage in terms of market share and it is open, we chose Android mobile phone system as a data collection experiment platform. In our framework, the data include call, SMS, Bluetooth, Wi-Fi hotspot, location, activity, acceleration, luminosity, gyroscopes, magnetometers, draw the screen and many other APPs. We call it LUHA (Location-based User Habit Analyzer). The framework of LUHA is depicted in Figure 1.

A. Data Collection and Storage

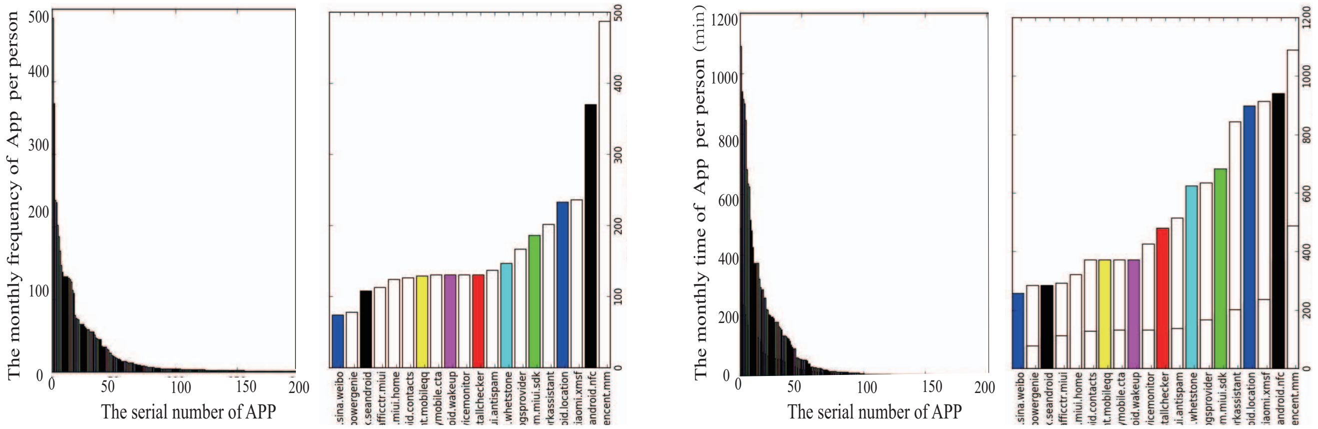
The collection tool in this paper is Starlog, which is based on the open software–Ubiqlog [13] by expanding the sample data. Since we need to constantly scan the state of senors, power and memory will be largely consumed. In order to balance the power and memory consumption with the accuracy of experimental data, the Starlog accesses the latest data every 5 seconds and records both start-time and end-time of one process, which is in chronological order of start-time. We continuously sample location data every 10 seconds. Starlog automatically collects data background and does not affect the normal use. The collected data are saved in a memory card and different types of daily data are recorded in JSON format and put in the same folder. All collected data are classified as a class according to the category, namely, we separately build a class for the GPS data, call, APP and other same type of data.

B. Data Cleaning and Classification

In the beginning, we need to clean the log file. By scanning all the files, we delete all the empty value and repeated items. If the start time is equal to the end time of a certain APP, we treat it as an invalid record. Indeed, we also ignore the items whose duration is less than 1 second. For abnormal values, such as the GPS records of Antarctica, we revise them depending on context or directly delete them. Similar to [14], the name of the recorded APP is the process name rather than human-readable name. For instance, Wechat uses *com.tencent.mm* as URL. We converted these names to the human-readable ones through the same method.

C. GPS Data Clustering

Intelligent mobile phone can provide real-time location of latitude and longitude and altitude information, but it is clear that the semantic position information of the specific meaning is only convenient for people to understand. We can use the semantic information for detailed analysis of different behavior between different locations. Therefore we need to label the physical location information with the semantic meaning. After clustering process, the GPS data sequences are mapped



(a) (b) (c) (d)

Fig. 2. The power-law distribution of APP monthly usage pattern in both frequency and duration and the average top-20 APP with monthly usage per person

to position sequence and the center location of the cluster is associated with the geographical semantic information.

1) DJ-cluster Approach: We adopt both the traditional *K-Means* and the *DJ-Cluster* (**D**ensity-and-**J**oin-based Cluster) algorithms [15] to make a GPS-data cluster. K-means algorithm needs to fix k before clustering, which might bring inconvenience to the unknown distribution of the actual data. DJ-cluster is more intelligent than K-means algorithm. The basic idea of DJ-Cluster algorithm is to extend the DBSCAN [16]. We define $N(p)$ as the set of all the neighbors of point p, then

$$N(p) = \{q \in S \mid dist(p, q) \leq Eps\} \tag{1}$$

where $dist(p, q)$ refers to the Euclidean distance between p and q. The number of elements in the $N(p)$ is the density. We calculate the density of each point. It will be treated as an abnormal point if the density values of one point and all of its neighbors are less than a given value, *minPt*. Otherwise, we make a new category or merge it into the adjacent existing category if all the neighbors of the point do not belong to one existing category. Although *DJ-Cluster* algorithm can get very good results, the performance of the algorithm might be affected when the Eps value is large. The reason is that the different value directly affects density value of every point which would increase the number of comparison with *minPt*.

2) Semantic Translation: We use simple manual markers for the most visited locations. We ask participants to do an online survey and report the name of the place they go to every day. We predefine fined places for people to choose. Each place is represented as a rectangle, the size of which is determined by the grid clustering algorithm. Because we want to study usage behaviors of the students based on the location of mobile phone APP, we show them on a rectangular campus around the map, allowing users to label different positions of predefined tags.

D. App Usage Statistics

By scanning the log files, we mainly obtain the frequency, duration and location information of the process of APPs. Since we have no idea about the exact applications in different smartphones, a dictionary is built. After accessing every line, we firstly check whether it existed in the dictionary. If there is, we add the duration and frequency to the existing items. While if not, we make a new items. Then, we summarize the statistical information of the processes in the same category as the final result of the APP.

1) Asscoiating Location by Time: For the starlog separately records the data of APP and location, the starttime and endtime of certain APP are recorded in log-files. In location log-files, there is only one time point for writing the GPS logs. So we connect the time point of location with the time segment in APP log-files. We allocate the locations to corresponding APP and then make a GPS-clustering. After obtaining the labeled place, we focus on the distribution of APP on different place in daily, weekly and monthly usage pattern. For certain place, we also research the hourly distribution in one daily.

2) Visualization: We plot the trace of certain app in the map and show the main places where people use the app.

IV. Experiment Result and Discussion

A. Overall Result

We collect the log data of 16 people for more than 6 months. During the whole experiment, each student has a set of 10 candidates, among which there are 7 locations that users most frequently visited and 3 randomly selected one. These 10 locations will be randomly allocated to participants. For a more comprehensive understanding of the behavior, we analyzed the data from different respects. Firstly, we only consider the frequency and duration of using APP per people. As depicted in Figure 2, we find the monthly usage frequency and duration of more than 300 APPs present a power-law distribution, suggesting that people spend most of their time on extreme few APPs. We present the top-20 of them in Figure 2(c)(d). We can find that Wechat is the most popular APP among participants. The NFC (Near Field Communication), location and network assistant are the most useful function in smartphone.

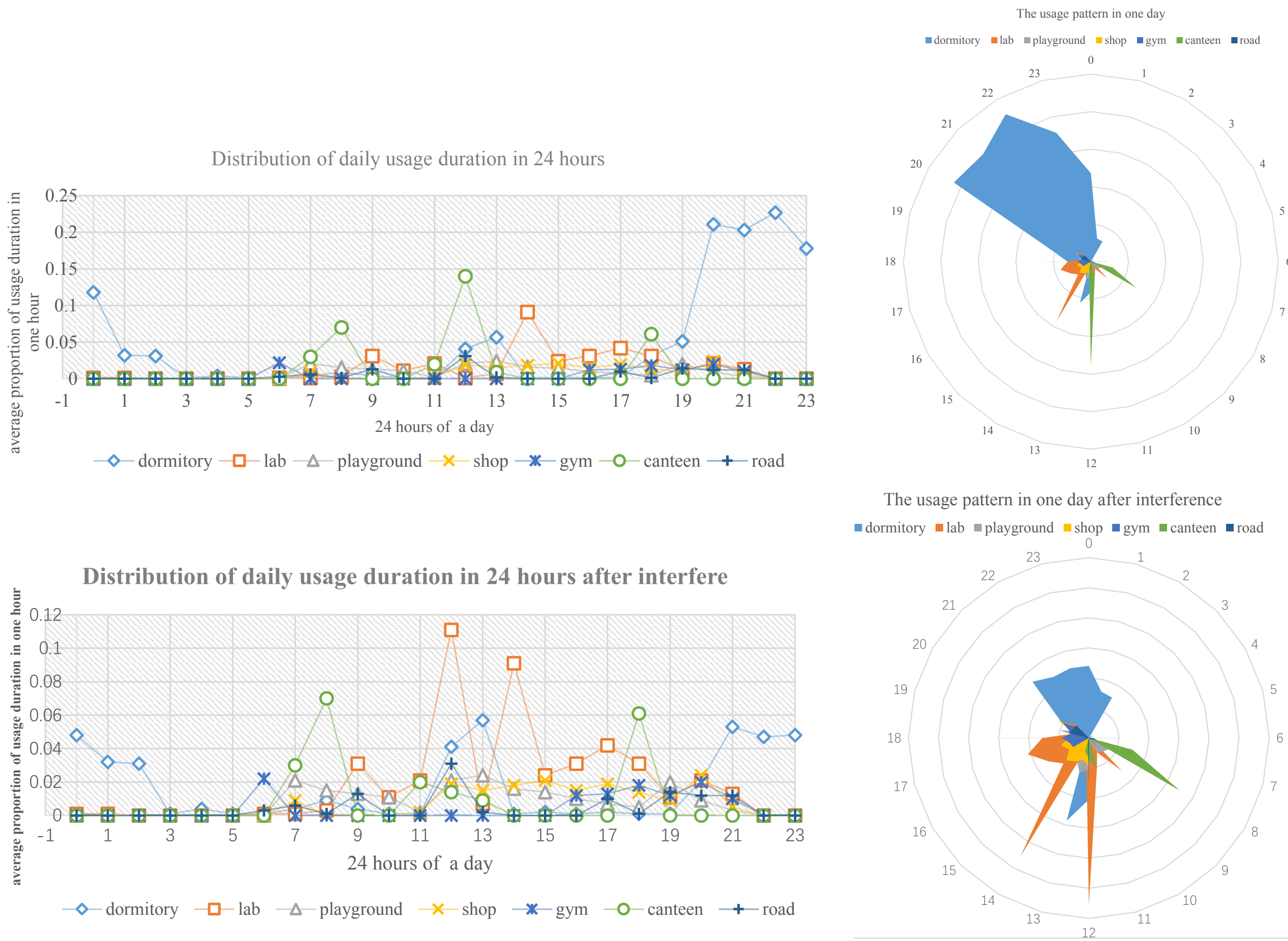


Fig. 3. The daily usage frequency of the participant in Social Network and Entertainment aspects

Secondly, we associate APP with location using the time component. We find that students frequently use Wechat at school and on the road. We classify the whole APPs into about 5 categories illustrated in Table I by inheriting the same way of literature [3]. We find the average duration proportion of these categories is very similar to the result of that paper, although the participants are from different places. We separately record the frequency and duration of each participant. The results present large difference. By further studying the location of these behavior, we find the two people spend lot of time in dormitory on chating with people or playing game. The location plays an important role in determining the behavior. Dormitory takes up most proportion for people using smartphone. In the common place, such as laboratory, they almost do not use smartphone probably due to the studying environment. Refer to the addict and non-addict result proposed in Table I [3], there are only 2 people presenting addict behavior in two aspects. As depicted in Figure 3, the duration of using social network APPs of one person and both the frequency and duration of using Game APPs of another person exceeds the threshold of addict behavior.

B. *Addiction Analysis*

By referring the result of SAMS, we can judge the addict, potentially addict and non-addict usage pattern of people. SAMS fails to fine the time-space distribution of the app-addicted people's behavior. Through hourly usage pattern learning, we find that the time takes control of the daily routine life while it is not the direct reason that leads to the addiction of people. The location environment of people plays more important roles in determining the habit of people. As depicted in the figures, the gap between the proportion of people in different places justifies the restrictions on the freedom of behavior because different places still impact the degree that addicted people immersed in the virtual world. To further prove it, we use interference by only asking people spend double time in lab or gym without telling them they have addiction to smartphone. After two months, the duration and frequency of usage declined about 1/3 time, which is under the addiction standard in Table I.

V. Conclusion and the Future Work

In this paper, we studied the usage pattern of smartphone user by combining the location of people. We study the

TABLE I
REFERENCE FOR DAILY USAGE STATISTIC BETWEEN ADDICTS (A) AND NON-ADDICTS (N)

Category	Common App	Frequency		Duration (min)	
		A	N	A	N
Soical Network (SN)	QQ, Wechat, Weibo, Momo	98.2	62.1	111.8	86.0
Game (G)	Tencent Games, Chess, Rhythm Master, sport game, popstar	7.2	3.7	46.3	22.6
Entertainment (E)	qqmusic, pps, kugou music, leTV	21.0	12.0	89.3	88.2
Browsing (B)	News, Amazon, Taobao, JD, Map, VIP.com	37.8	33.7	77.5	61.9
Other (O)	tracer, camera, locker, timer	44.0	15.4	17.9	15.7

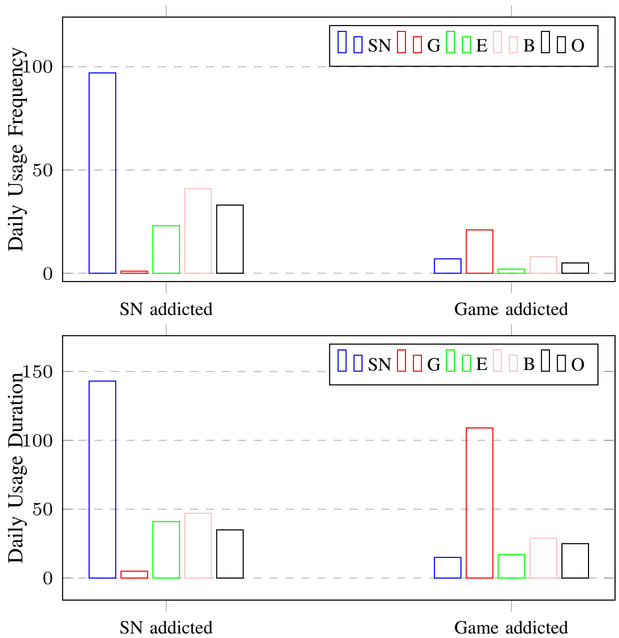


Fig. 4. The time and frequency distribution of the addicted people

problem of "when and where the people use what kinds of APP". By recording the frequency and duration, we find that place plays an important role in determining the behavior of people. The difference of addicted people's performance at different place suggests that we can force a change or an intervene to the environment of people to change the bad habits.

ACKNOWLEDGMENT

The authors would like to thank the support of the National Natural Science Foundation of China(No.61202117).

REFERENCES

[1] "emarketer: Worldwide smartphone usage to grow 25% in 2014," http://www.fiercewireless.com/press-releases/emarketer-worldwide-smartphone-usage-grow-25-2014.

[2] L. Heyoung, A. Heejune, C. Samwook, and C. Wanbok, "The sams: Smartphone addiction management system and verification." *Journal of Medical Systems*, vol. 38, no. 1, pp. 1–10, 2014.

[3] H. Ahn, M. E. Wijaya, and B. C. Esmero, "A systemic smartphone usage pattern analysis: Focusing on smartphone addiction issue," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 9, no. 6, pp. 9–14, 2014.

[4] R. E. Grinter and M. A. Eldridge, *y do tngrs luv 2 txt msg?* Springer Netherlands, 2010.

[5] D. Reid and F. Reid, "Insights into the social and psychological effects of sms text messaging," *University of Plymouth*, 2004.

[6] S. Butt and J. G. Phillips, "Personality and self reported mobile phone use," *Computers in Human Behavior*, vol. 24, no. 2, pp. 346–360, 2008.

[7] A. Ken, N. Dawn, R. Tye, and A. Ryan, "Numbers have qualities too: Experiences with ethnomining," in *Ethnographic Praxis in Industry Conference Proceedings*, 2009, pp. 123–140.

[8] M. Raento, A. Oulasvirta, and N. Eagle, "Smartphones: An emerging tool for social scientists," *Sociological Methods and Research*, vol. 37, no. 3, pp. 426–454, 2009.

[9] T. M. T. Do and D. Gatica-Perez, "By their apps you shall understand them: mining large-scale patterns of mobile phone usage," *Perez*, pp. 1–10, 2010.

[10] A. Battestini, V. Setlur, and T. Sohn, "A large scale study of text-messaging use," in *Proceedings of the 12th Conference on Human-Computer Interaction with Mobile Devices and Services, Mobile HCI 2010, Lisbon, Portugal, September 7-10, 2010*, 2010, pp. 229–238.

[11] T. M. T. Do, J. Blom, and D. Gatica-Perez, "Smartphone usage in the wild: a large-scale analysis of applications and context," in *Proceedings of the 13th International Conference on Multimodal Interfaces, ICMI 2011, Alicante, Spain, November 14-18, 2011*, 2011, pp. 605–627.

[12] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin, "Diversity in smartphone usage," in *International Conference on Mobile Systems, Applications, and Services*, 2010, pp. 179–194.

[13] R. Rawassizadeh, M. Tomitsch, K. Wac, and A. M. Tjoa, "Ubiqlog: a generic mobile phone-based life-log framework," *Personal and Ubiquitous Computing*, vol. 17, no. 4, pp. 621–637, 2013.

[14] J. Rana, J. Bjelland, T. Couronne, P. Sundsy, D. Wagner, and A. Rice, "A handset-centric view of smartphone application use," *Procedia Computer Science*, vol. 34, pp. 368–375, 2014.

[15] C. Zhou, D. Frankowski, P. Ludford, S. Shekhar, and L. Terveen, "Discovering personally meaningful places: An interactive clustering approach," *Acm Transactions on Information Systems*, vol. 25, no. 3, p. 2007, 2007.

[16] J. S. X. X. Martin Ester, Hans-Peter Kriegel, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96*, 1996.