Research Article

The Study of Digital Identity of Mobile Users Based on Android Dataset

A. R. Mohd Shariff[1*], M. F. Abdul Jalil[2], S. I. Fadilah[3]

[1,2,3]School of Computer Sciences,Universiti Sains Malaysia,11800 Pulau Pinang, Malaysia
azizulrahman@usm.my[1]

Article History: Received: 10 November 2020; Revised: 12 January 2021; Accepted: 27 January 2021; Published online: 05 April 2021

Abstract: A smartphone is said to be an extension of a mobile user's oneself. It is the ultimate mobile device today for humans, the one device that humans cannot avoid using, leave behind when on the move and becomes a necessity to own one. It allows humans to communicate and perform many daily personal tasks. Smartphones continue to evolve embedding advanced device technologies, sophisticated processes and applications. This symbiotic combination generates immense data on a temporal and spatial scale for a mobile user. These data represent the behaviors of a mobile user, which is unique. In the digital mobile realm, there exists some digital identity about a mobile user, what can be thought as the mobile user's digital pheromone. This data mining study uses Android Device Analyzer dataset consists of Battery Level, Battery Temperature, Battery Volt and App variables to understand the levels of volatility and their correlation measures. The finding shows there are wide dispersions for Battery Level, Battery Temp, Battery Volt and no statistical relationship exists for Battery Volt/App and Battery Temp/App correlation. Using a simple inequality that only a limited number of mobile users are contained within one standard deviation from the population mean, a smaller set of users exists. The results show that 'uniqueness' percentage of a mobile user increased from 61.7% and 66.7% for single variable to 81.7% in multi-variate condition, to 100% using the simple inequality. Ultimately, this uniqueness indicates that a mobile user can have a unique identity, in the digital mobile world from the daily mobile data generated.

Keywords: digital identity, mobile user, device analyzer, android data set

1. Introduction

Mobile devices are embedded with a host of sensor technologies and some advanced hardware functions[1,2]. They can perform complex computational tasks such as multimedia rendering and gaming due to its integration with a high-performance processor and the graphical processing unit. Mobile users generate immense amount of data using their ultimate mobile device, a smartphone, daily everywhere. A wealth of mobile data can be collected from the sensors and hardware built in the smartphones to understand user behavioral patterns, exploring and extracting its intrinsic 'hidden' features. Ultimately the hidden features can be utilized to develop some intelligent functions or mathematical models that may well improve the users experience and services.

A review by Khan et al.[1] reported comprehensively on the use of smartphones and mobile phone sensors to achieve a better interaction between humans. As in Jia et al., user preferences are captured to provide video recommendations[3] based on some textual information of Android applications. Ren et al.[4] mined user's trajectories history to predict whether users moves or stays at a location for a more extended period. This information offers precise marketing to a user based on the profile of the location. Chittaranjan et al. use mobile devices data from Nokia N95 to study the relationship between the behavioral characteristic and self-reported Big-Five personality traits[5]. User behaviors can be predicted based on smartphone activities and usage. The study done by Ren et al. seeks unique behavior pattern from users' trajectory data[2]. The authors are able to provide intelligent and personalized services to achieve accurate marketing services based on individual behaviors.

Studies by Huang et al. and Liao et al.[7,8] are directed on predicting the tendencies on what type of application in mobile devices that are most inclined to be opened or used by the user. The authors explored contextual information such as the last application used, time, location, user profile, and application usage by using Mobile Data Challenge (MDC) dataset. Ramos et al.[9] studied a diverse group of smartphone datasets that are based on user's usage behavior. The authors were able to find 382 distinct types of users in a population of 106,762 Chinese smartphone users, based on application usage. Quercia et al.[10] investigated the relationship of preferences of social event and geography in a metropolitan area to recognize movement patterns by the users and recommend social events based on a user's location data[10]. Chen et al. used location data from a mobile device to predict the population at large, a spatiotemporal scale of a city[6]. The authors estimated the future inflow population with the current inflow pattern, and then apply the spatial correlation of the population using neural network. Digital identity construction is not restricted only to social networking as pointed out by Larsen[11]. Belk[12] studied the digital identities existence from a digital world perspective and suggested three important features to consider, dematerialization, re-embodiment, and co-construction of self[12]. As for mobile devices application, the mobile activities' usage can create a digital personality that individually unique[22]. According to studies by Vallina-Rodriguez et al.[13] and other authors in[9,14], the authors selected variables from a mobile device dataset and suggested that bat-

tery capacity, battery mode and battery voltage is correlated independently to a user. The selected variables[16] could be tested simultaneously to explore the degree of relationship between them. The main objective in this study is to prove a simple hypothesis using some statistical methods that some mobile data captured from a mobile user is unique, varying and orthogonal compared to other mobile users.

To prove that uniqueness exists; high volatility or wide dispersions for a given variable is needed, and independent statistical relationship in a multi-variate condition must be found. The uniqueness is denoted specifically as the 'digital identity' or a concept promoted as 'digital pheromone' that can be employed to develop user-behavioral based solutions[17,18], biometric or identity-based authentication and encryption mechanisms[23-26]. As suggested by Vallina-Rodriguez et al.[13], each independent user is highly correlated with battery capacity, battery mode, and battery voltage. Hence, *Battery Level*, *Battery Temperature*, *Battery Voltage*, and *Application* are chosen as the variables in this study. Three stages are involved, which includes Android Device Analyzer Dataset, Data Ingestion & Partitioning, and Information Mining Process. The data for this study is based on the Android device analyzer project[15].

2. Digital Identity: The Concept of Digital Pheromone

Mobile users are highly dependent on the mobile applications they use on a day to day basis and locations they domicile, work, leisure and travel. Thus, there are variations in the applications usage and charging patterns individually, from one mobile user to another mobile user. The variation patterns if exists and able to be captured, can be used to model one mobile user to another. How much a variable value varies (being volatile) from one another is the level of orthogonality. However, if only one variable is studied such as *Battery Level*, it is extremely difficult to un-earth the variations between one user to all other users, since there are millions of mobile users that may hold same values for variable *Battery Level*. It is argued that the more random the variations exist for a chosen variable, the higher the volatility and closer to orthogonality.

A mobile user's unique 'digital identity' model can be captured and developed given the conditions;

(a) wide range of variables and its values (data) can be captured temporally and spatially
(b) each variable values tend to vary or be volatile from other mobile users
(c) each variable dependencies and independencies to other variables and their properties is known – correlation measure
(d) the composition of all variable values and properties for a given mobile user can be modelled
(e) for each mobile user, a model that is unique representing that mobile user is always orthogonal to another mobile user's model. The work extends further the concept of 'digital pheromone'.

Pheromone is some chemical substance that is released, excreted or secreted by animals or insects that affects the behaviour of receiving individuals or itself. In the digital mobile realm, the pheromone can be seen as mobile data that is continuously generated by a mobile user and transacted in the mobile network across multiple Internet end hosts. The data from a mobile user to another is unique, a depiction and extension of the mobile user's behaviours. There are many areas where a mobile user's digital identity model is important and beneficial, such as in AI-based privacy and security of mobile applications[17-21]. In this paper, the study relates to part (a), (b) and (c) to prove there exists variations in the mobile user's data, and is fundamentally related to intrinsic behaviours using similar variables of *Battery Level*, *Battery Temperature*, *Battery Voltage*, and *Application.*

a. Device Analyzer Dataset

The Device Analyzer dataset[15] comprise of multiple variables from Android mobile devices. It comprises 1,900 years of aggregate trace duration generated by mobile activities information from 1277 different devices, which comprise of 16,000 contributors in 175 countries dated up to 2014. In this paper, we used the latest dataset provided by the University of Cambridge comprising of 100 billion records of smartphone usages from 17,000 devices worldwide. The huge dataset (big data) may contain different variables, which are grouped into different hardware profiles, software versions and multiple types of data. Processing huge dataset is expensive in terms of computation resource and time. Large files will incur an extended period to complete the task. To overcome computational and storage requirements, uncompressed data between 10MB to 300MB are selected. Thus, the initial raw dataset size decreased to approximately 83GB comprised of 1828 files in the compressed state taken from the primary dataset.

b. Data Ingestion and Partitioning

```
INPUT: raw data directory path, output directory path

initialize variable
InputDir ← raw data files directory
OutputDir ← outputfiles directory

check output directory
if ProcessDir! = exist then
    create ProcessDir directory

for i = 1 to lenght(InputDir) do
    InputFile ← listofFileFrom(InputDir[i])
    for j = 1 to lenght(InputFile) do
        data.frame ← LoadFromFile(InputFile)
        data.frame ← GetFormattedDataset(data.frame)
        VariableInfo ← GetVariableInfo(data.frame)
        DeviceInfo ← GetDeviceInfo(data.frame)
        ManufacturerDir ← file.path(ProcessDir, DeviceInfo.Manufacturer)
        if ManufacturerDir! = exist then
            create ManufacturerDir directory
        outFilename ← CombineString(file.path( OutputDir),DeviceInfo.Model,DeviceInfo.Version)
        if outFilename == exist then
            increment outFilename
        CreateXlsx(VariableInfo,outFilename)
```

Figure 1. Data Ingestion and Partitioning Algorithm

The dataset contains unformatted, raw data, which included some noise information or garbled data. The dataset consists of data collected from the various manufacturers such as Samsung, LG, Motorola and Sony. Thus, the dataset must be 'cleaned' and arranged into the desired format. R-software tool is used to implement the algorithm for the purpose of Data Ingestion and Partitioning phase. In Figure 1, during initial state, input directory path is configured using raw files location taken as a string datatype, while output directory path is configured to store processed data. Both input and output path directory is then stored as *InputDir* and *Output-Dir*, respectively. Next, the output folders are checked for its existence based on the output directory provided. This process is due to the output is written into a path directory provided at initial state after a single execution cycle of the algorithm. An error will trigger if the script cannot write into nonexistence directory folder.

The algorithm comprises of two loops; the first loop iterates on the number of folders exist in any given input directory path. While an inner loop is used to iterate on the number of files contained in each directory captured by the first loop. Then, a variable with the type of *data.frame* is loaded with data from the raw file. The *data.frame* is formatted by calling method *GetFormattedDataset()* that accepts an object with the type of *data.frame* as its value. Inside the method, several steps are taken to eliminate the noise in the dataset and returns a *data.frame* contained formatted data. Next, data is trimmed by removing unnecessary variables and columns, and the only variables needed for data mining are *Battery Level*, *Battery Voltage*, *Battery Mode*, *Battery Temperature* and *Application* while *DateTime, Variables*, and *Values* being the columns.

The data is written into *data.frame* object before proceeding to the next step, which is partitioning the data according to their respective variables. The *Application* variable resides inside *GetVariableInfo()* method. A regular expression applied to propagate into the dataset to find a result that meets the requirements based on the string of expression passed in short execution of time. All the results from each variable are stored as an array of object in the *VariableInfo* object. *DeviceInfo* object stores device information consists of a device manufacturer and device model used for the naming convention of an output file. Output files are grouped based on manufacturer's name; thus, *ManufacturerDir* stores the combination of a string from output's path directory and manufacturer's name and uses it to create a folder inside an output file directory. For the output file's name, the aggregate of the output directory path with the device model and device version is used. If the name existed in the corresponding manufacturer folder, the current string stored in *outFilename* is incremented by one, thus creating a new unique name. Finally, formatted data is written into Excel file extension (.xlsx) as shown in Figure 2 sorted into respective manufacturer's name.

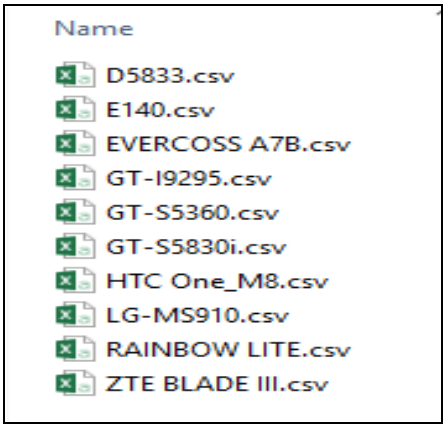

Figure 2. Output files sorted according to manufacturer's name

c. Information Mining

A sample size of 60 mobile users (devices) data from Samsung manufacturer to keep the homogeneity of the device characteristics analyzed. Data for each variable is collected within the same 24-hour period, simultaneously. For each *Battery Level*, *Battery Temperature* and *Battery Voltage* variable, the discharging rate (the slope) is calculated for each mobile user and the mean discharging rate overall for all mobile users. For *Application* variable, which is the total number of running processes noted at discharging intervals in the 24-hour period, only the mean value over all mobile users is calculated. It is argued that mobile users are extremely attached to their mobile applications of choice, its types, the temporal and spatial usage patterns. Therefore, we hypothesize that the running processes for using these mobile applications may vary from one mobile user to another. Other system performances directly related to the usage behaviors are *Battery Level*, *Battery Temperature* and *Battery Voltage*, specifically also their discharging patterns. To assess the volatility between mobile users for each variable, a measure of Coefficient of Variation (CV) is performed to understand the statistical dispersion between mobile users for each variable. The variables performances are compared to investigate whether the variables are volatile (or varying) from one mobile user to another. Correlation analysis are then performed crucially to understand if one value's value changed, what will happen to other variables, and how significant are the changes.

3. Results

Table 1. Comparisons of Statistics

	Battery Level	*Battery Temp*	*Battery Volt*	*App*
Mean	2.359	0.449	0.089	53.31
VAR	6.822	0.638	0.144	248.863
Std Dev	2.612	0.799	0.380	15.775
(CV) (%)	1.11 (111%)	1.78 (178%)	4.25 (425%)	0.296 (29.6%)

Table 2. Correlation Results

		Level	**Voltage**	**Temperature**
Voltage	Pearson corr.	0.652		
	P-value	0.000		
Temperature	Pearson corr.	- 0.386	- 0.476	
	P-value	0.002	0.000	
App	Pearson corr.	0.287	0.162	- 0.180
	P-value	0.026	0.216	0.168

From Table 1 and with reference to standard deviation plots in Figure 3, the statistics for *Battery Level*, *Battery Temp*, *Battery Volt*, and *App* are given. Based on CV, *Battery Volt* has huge variations followed by *Battery Temp*, *Battery Level*, which is above 100% variations about the mean. The *App* variable has the smallest percentage of volatility, which is 29.6% indicating less dispersion of the data. Fundamentally, *Battery Level*, *Battery Temp* and *Battery Volt* provide very high data dispersions or spread that potentially indicating wide volatility between mobile users. For any mobile user, it is highly likely that the values (*Battery Level*, *Battery Temp* and *Battery Volt*) collected may not exhibit 'closeness' with values collected for other mobile users. This is related to

the behavioral patterns of each mobile user. It can be hypothesized that any variable that provides some small dispersions about its mean may not reveal mobile users intrinsic behaviors. This is because the values between the mobile users for the variable are tightly adjacent. Pearson's correlation between variables was examined and are shown in Table 2. Correlation value for *Battery Level* and *Battery Volt* is 0.652 (indicating strong correlation). The relationship between *Battery Temp* to *Battery Level* and *Battery Volt* are also considered to have strong negative correlation at -0.386 and -0.476 respectively. This potentially indicates that as battery temperature continue to rise due to prolonged usage of the mobile devices, both *Battery Level* and *Battery Volt* values reduces. There can be many factors attributed to this such as increased application running processes, multiple running sensor or hardware applications all related to how a mobile user uses his device. It can be seen that of all the variables, *App* has weakest correlations to all other variables, the lowest with a value of 0.162 for *Battery Volt*.

The *p*-value measured must be less than 0.05, inferring the relationship between any two variables as statistically significant. Only *App* and *Battery Volt*, and *App* and *Battery Temp* exhibit a p-value > 0.05, which means no statistical significance between them. This is another important characteristic that indicates 'uniqueness' between variables, needed to build a rich independent multi-variate [16] profile of a mobile user. The former being the 'volatility' measure of a given variable based on its CV as found in Table 1. It can be seen that for any variable that has a low dispersion (low CV), it potentially would have weak significance and not strongly correlated with other variables as shown by *App* variable in this small study. A measure of volatility between mobile users data hence their behaviors can be detected from a measure of the data variable's coefficient of variation (CV) and the *p*-values. To unearth mobile users identity from their usage of mobile devices, it is imperative to choose a data variable or a group of data variables that exhibit wide statistical dispersions.

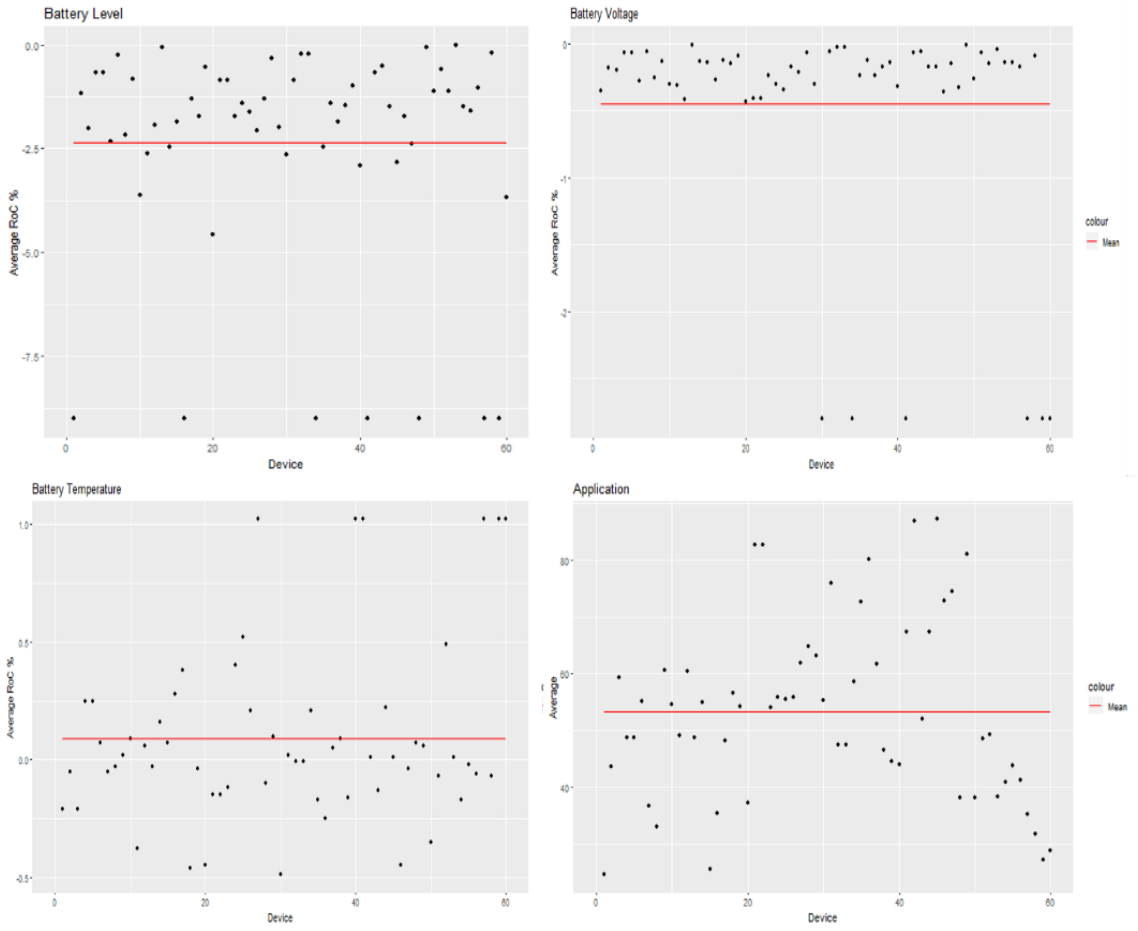

Figure 3. Standard deviation plots from the mean for *Battery Level*, *Battery Temp*, *Battery Volt* and *App* data

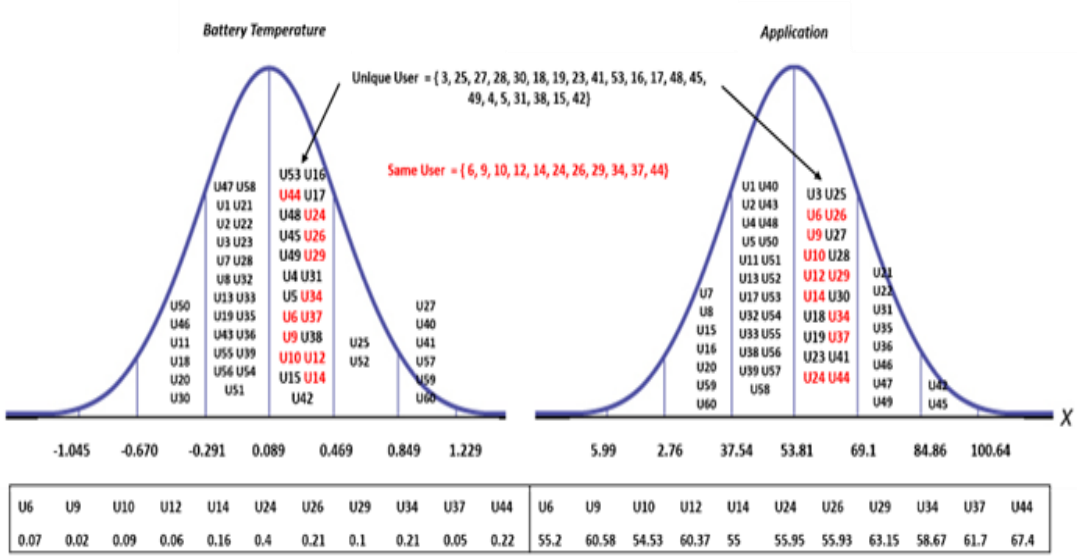

Figure 4. The normal curve showing mobile users distribution within one, two and three standard deviations from the population mean

For any mobile user where data are collected for a wide selection of variables over some long periods of duration, the variables mean value may reach statistical stability, that is its mean error is small. It is very unlikely that when a variable has shown positive characteristic over some long temporal scale, will change the behavior that reverses it characteristic to negative due to few instantaneous events or data values. It is safe to say that, for any mobile user, the data values collected for any given variable may converge to one standard deviation from its mean. Given a large population sample, if the CV for that variable is large as in Table 1, therefore $(\mu_{varA,user} + \sigma_{varA,user}) \ll (\mu_{varA,pop} + \sigma_{varA,pop})$. Ultimately this simple inequality says that only a limited number of mobile users are contained within one standard deviation from the mean for a given population. If the rule above is applied, a smaller set of users exists. The other mobile users outside $(\mu_{varA,user} + \sigma_{varA,user})$ can be considered as outliers. A special case of one positive standard deviation from the mean is considered, the list of mobile users in that interval for *Battery Temp* and *App* variable data is shown in Figure 4. There are 23 and 20 mobile users contained in that interval for *Battery Temp* and *App* respectively, which also denotes to 61.7% and 66.7% uniqueness. This crucially indicates that for any mobile user contained in that interval $(\mu_{varA,pop} + 1 * \sigma_{varA,pop})$ for variable *Battery Temp* and *App*, other mobile users outside one standard deviation interval are considered as outliers and very unlikely to have similar identities. They are 61.7% and 66.7% likely to be unique. Consider a multi-variable condition for *Battery Temp* and *App*, the total number of mobile users contained in both intervals is only 11, which is {6, 9, 10, 12, 14, 24, 26, 29, 34, 37, 44}**.** Out of the 60 mobile users in this study, a total of 49 users can be considered as outliers since they are either, not contained within the said interval $(\mu_{varA,pop} + 1 * \sigma_{varA,pop})$ or, they are contained in $(\mu_{varA,pop} + 1 * \sigma_{varA,pop})$ but only in one of the two variables of *Battery Temp* and *App*. The uniqueness percentage grows to 81.7%. For any of the 11 mobile users, there can be only 10 more mobile users that may potentially have same identities, that is they may have close mean values or sits in the same intervals for the multi-variate condition of *Battery Temp* and *App*. In Figure 4, the mobile user 24 has mean value for *Battery Temp* and *App* at 0.4 and 55.95, assume that *Battery Temp* and *App* standard deviations for mobile user 24 are about 10% away from its mean, i.e. $\sigma_{BattTemp,24}$ = 5.595 and $\sigma_{App,24}$ = 0.04, then mobile user 24 has unique intervals from 0.4 – 0.44 for variable *Battery Temp* and from 55.95 – 61.55 for variable *App*. This ensures inequalities $(\mu_{Batt\ TempA,24} + \sigma_{Batt\ Temp,24}) \ll (\mu_{Batt\ Temp,pop} + \sigma_{Batt\ Temp,pop})$ and $(\mu_{App,24} + \sigma_{App,24}) \ll (\mu_{App,pop} + \sigma_{App,pop})$ are always met. Based on mobile user 24 unique intervals for *App* variable, only mobile user 9, 12 and 34 are contained in the same interval. However, when multi-variate condition is considered for both *Battery Temp* and *App*, mobile user 9, 12 and 34 are outside from mobile user 24's *Battery Temp* interval (0.4 – 0.440), and they are considered as extreme outliers. Mobile user 24 is said to be 100% unique based on its personal intervals. It can also be strongly argued that potentially for a mobile user and from his temporal and spatial mobile data captured, a unique digital identity can be mined and modelled using some mathematical and statistical methods that represent himself distinctively in the digital world. This is the digital pheromone of a mobile user.

4. Conclusion

There is an ever-increasing growth of smartphone usage and advancement in mobile device technologies. Today, smartphone is vital to a person, the ultimate device that the person must have, carry and use continuously everywhere. A smartphone has become an extension of a person's behavior and identity. With all the device technologies embedded in a smartphone, a multitude wealth of data can be collected from them, and from the processes and applications that run with them. These data collected from various variables are temporally and spatially influenced and when mined, analyzed either as a single variable or in multi-variate forms can exhibit 'unique' behavioral patterns of a person. It is explicitly denoted in this work as having some unique 'digital pheromone'. There are various benefits for using human-based or behavioral-based digital identity such as in Identity-based authentication and encryption. This study is composed of a data ingestion and partitioning process, and data mining strategies onto an Android-based smartphone data-set of 60 mobile users. The statistics of four different types of data variable are studied, *Battery Level*, *Battery Temp*, *Battery Volt* and *App* to understand its variability or volatility. The more dispersed the data, therefore higher volatility can be expected for data values between different mobile users. *Battery Level*, *Battery Temp* and *Battery Volt* exhibit over 100% variations about their population mean. Only *App*/*Battery Volt* and *App*/*Battery Temp* exhibit p-value > 0.05, which means no statistical relationship, another important characteristic that indicates 'uniqueness' between variables. The two factors of high volatility for a given variable and independent statistical relationship in a multi-variate condition ultimately points to the existence of wide variety of behaviors in the use of mobile devices between mobile users. This is a desired phenomenon in the ability to build a rich digital identity of a mobile user, hence his or her digital pheromone. Using a simple rule $(\mu_{varA,user} + \sigma_{varA,user}) \ll (\mu_{varA,pop} + \sigma_{varA,pop})$, which says that only a limited number of mobile users are contained within one standard deviation from the mean, a smaller set of users exists. The results show that 'uniqueness' percentage increased from 61.7% and 66.7% for single variable to 81.7% in multi-variate condition, to 100% when the simple rule above is applied. Finally, this study is limited as only 60 mobile users and 4 different types of variables are tested. However, it has been shown that some unique identity exists from one mobile user to another in this simple study based on the two factors above and a simple inequality rule proposed. Future studies will involve higher number of users with larger number of variables.

5. Acknowledgement

The work is funded by a Bridging Research Grant 304/PKOMP/6316015 under USM. The authors would also like to acknowledge that the data for this work is provided by the Device Analyzer Project of the University of Cambridge.

References

1. Khan WZ, Xiang Y, Aalsalem MY, Arshad Q. Mobile phone sensing systems: a survey. *IEEE Comms Surv & Tut*. 2013; 15(1): 402–427.
2. Han Q, Cho D. Characterizing the technological evolution of smartphones. *Procs 18th Intl Conf Elect Commerce*. 2016 Aug; 32: 1–8.
3. Jia X, Wang A, Li X, Xun G, Xu W, Zhang A. Multi-modal learning for video recommendation based on mobile application usage. *Procs IEEE Intl Conf Big Data*. 2015 Dec; 837–842.
4. Ren M, Yang F, Zhou G, Wang H. Mining individual behavior pattern based on semantic knowledge discovery of trajectory. *J Comp & Info Tech*. 2015; 23(3): 245–254.
5. Chittaranjan G, Blom J, Gatica-Perez D. Mining large-scale smartphone data for personality studies. *Pers & Ubi Comp*. 2013; 17(3): 433–450.
6. Chen J, Pei T, Shaw SL, et al. Fine-grained prediction of urban population using mobile phone location data. *Intl J Geog Info Sci*, 2018; 32(9): 1770–1786.
7. Huang K, Zhang C, Ma X, Chen G. Predicting mobile application usage using contextual information. *Procs ACM Conf Ubi Comp*. 2012; 1059–1065.
8. Liao ZX, Pan YC, Peng WC, Lei PR. On mining mobile apps usage behavior for predicting apps usage in smartphones. *Procs 22nd ACM Intl Conf Info & Knwl Mgmt*. 2013 Oct; 609–618.
9. Zhao S, Ramos J, Tao J, et al. Discovering different kinds of smartphone users through their application usage behaviors. *Procs ACM Intl Jnt Conf Perv & Ubi Comp*. 2016; pp. 498–509.
10. Quercia D, Lathia N, Calabrese F, Lorenzo DG, Crowcroft J. Recommending social events from mobile phone location data. *IEEE Intl Conf Data Mining*. 2010 Dec; 971–976.
11. Larsen MC. Understanding social networking: On young people's construction and co-construction of identity online. *Political Sci*. 2008; 18–20.
12. Belk RW. Extended self in a digital world. *J Cons Res*. 2013 Oct; 40(3): 477–500.

13. Vallina-Rodriguez N, Hui P, Crowcroft J, Rice A. Exhausting battery statistics: understanding the energy demands of mobile handsets. *Procs 2nd ACM SIGCOMM Wksh*. 2010 Aug; 9–14.
14. Falaki H, Mahajan R, Kandula S, Lymberopoulos D, Govindan R, Estrin D. Diversity in smartphone usage. *Procs 8th Intl Conf Mob Sys, App, & Serv*. 2010 June; 179–194.
15. Wagner DT, Rice A, Beresford AR. Device analyzer: understanding smartphone usage. *Springer LNICST*. 2014 Sep; 131: 195–208.
16. Trivedi KS. Probability & statistics with reliability, queueing and computer science applications. 2nd ed. *Wiley*; 2008.
17. Shafique U, Sher A, Ullah R, et al. Modern authentication techniques in smart phones: security and usability perspective. *Intl J Adv Comp Sci & App*. 2017; 8(1): 331–340.
18. Tseng YM, Tsai TT, Huang SS, Huang CP. Identity-based encryption with cloud revocation authority and its applications. *IEEE Trans Cloud Comp*. 2018 Dec; 6(4): 1041–1053.
19. Mohammed SMZ, Shariff ARM, Singh MM. An authentication technique: behavioral data profiling on smart phones. *Springer LNEE*. 2018; 488: 88–98.
20. Liang X, Zou F, Li L, Yi P. Mobile terminal identity authentication system based on behavioral characteristics. *Intl J Dist Sensor Nets*, 2020 Jan; 16(1).
21. Mohammed SMZ, Shariff ARM, Singh MM. A secure mobile app solution using human behavioral context and analytic hierarchy process. *Elsevier Proc Comp Sci*, 2015; 72: 434–445.
22. Mønsted B, Mollgaard A, Mathiesen J. Phone-based metric as a predictor for basic personality traits. *J Res Personality*, 2018 June; 74: 16-22.
23. Benzekki K, Fergougui AE, Elalaoui A. A context-aware authentication system for mobile cloud computing. *Proc Comp Sci*. 2018; 127: 379-387.
24. Gulati H, Huang C. Self-sovereign dynamic digital identities based on blockchain technology. *IEEE SoutheastCon*, 2019 Apr; 1-6.
25. T. Padmapriya & S.V. Manikanthan, "Retracted: Security and Routing protocol for 5G wireless mobile networks". IJIMT, 2020.
26. Westerlund M. The emergence of deepfake technology: a review. *Tech Inno Mgmt Rev*. 2019 Nov; 39-52.