A Smartphone-based System for Personal Data Management and Personality Analysis

著者	Guo Ao
出版者	法政大学大学院情報科学研究科
journal or publication title	法政大学大学院紀要．情報科学研究科編
volume	11
page range	1-6
year	2016-03-24
URL	http://doi.org/10.15002/00012874

A Smartphone-based System for Personal Data Management and Personality Analysis

Ao Guo
Graduate School of Computer and Information Sciences
Hosei University
Tokyo 184-0003, Japan
guo.ao.gi@stu.hosei.ac.jp

***Abstract*—The data from or about an individual, called personal data, is continuously increasing due to popularity of smart phones, wearables and other ubiquitous devices. Such personal data can be used to model a user and even digitally clone a person, e.g., Cyber-I (cyber individual) that aims at creating a unique and comprehensive description for every individual to support various personalized services and applications. Due to heterogeneity and sensitivity of personal data, one important issue is how to effectively collect and manage person data with sufficient security protection. Another important issue is how to figure out an individual's character, i.e., personality from personal data. Therefore, this research is focused on personal data management and personality analysis in a smartphone based client-server system. The smartphone functions as not only a source of personal data but also a gateway to manage other wearables and communicate with a server that keeps personal data in a larger amount and a longer period. A multi-security mechanism is implemented to ensure data security in collection, transmission and storage. Personality analysis is made from data normalization, feature extraction and clustering, to personality computation based on sociological personality theories.**

Keywords—personal data; personality; system security; cluster algorithm; smartphone; mongo

I. INTRODUCTION

A lot of big data is from or about users because of the popular uses of the PCs, smartphones, wearables and other ubiquitous devices. Such user's related data is generally called personal data, which exhibits two basic features, the heterogeneity, i.e., data from diverse sources in various forms, and the sensitivity, i.e., private data with high security concerns. Therefore, two fundamental issues are (1) how to effectively collect and manage personal data with sufficient protection of data security and user's privacy, and (2) how to utilize personal data to know more about users and further provide better personalized services and applications.

Cyber-I (Cyber Individual) is our lab's continuous research to tackle the two fundamental issues since 2009, and a Cyber-I is a counterpart of a real individual (Real-I) in cyberspace [1]. It aims at a unique and comprehensive personal description as a digital clone for every real individual and various personalized services. In our previous research, a PC based client-server system was developed for a general management framework of heterogeneous data and multiple Cyber-Is, especially in their life cycle of birth, growth and death [2]. In the PC-based system, the data collections are basically done manually or semi-automatically, and data security function has not been developed.

Hence, the first research objective in this study is to develop a smartphone based client-server system that can automatically manage the collection of data from the smartphone itself and nearby wearables, the storage of data in a SQLite local database inside the smartphone, and the upload of data from the local database to a remote Mongo database in the server. The dynamic scheduling, prompt adjustment and instantaneous data pre-process are key techniques necessary for the automatic data management.

Our next research objective is to ensure personal data security in the process of data collection, storage as well as transmissions between the smartphone and the corresponding server. Although our previous work has been carried out on privacy preservation in the accesses of Cyber-I's information [3], our current work is mainly for necessary security techniques to protect personal data. A multi-security mechanism with using a hash algorithm and an asymmetric encryption algorithm is designed and implemented to guarantee the data integrity and data safety.

The third research objective is to analyze an individual's basic character, i.e., personality, automatically with utilizations of personal data collected and accumulated by the smartphone-based system. Previously, we have done relatively more research on Cyber-I modeling that could be growable along with the increasing personal data [4]. Since the personality is human's essential character and would be useful to guide the growable modeling, we have conducted preliminary personality analysis consisting of data normalization, feature extraction, feature clustering and personality computation.

The remainder of this paper is organized as follows. The next section is about others' work and their relations with our work. Section III describes our system architecture and key functions including the security mechanism. Section IV illustrates the management of personal data and related operations. Section V explains the process and algorithms in personality analysis. In Section VI, system performance tests and representative case studies are presented. Conclusions and future work are drawn in the last section.

Supervisor: Prof. Jianhua Ma

II. Related Work

Lifelog systems, inspired by Vannevar Bush's concept of "MEMory EXtenders" (MEMEX), are capable of collecting and storing a person's lifetime experience as a multimedia database [5]. As defined by Dodge and Kitchin, lifelogging is conceived as a form of pervasive computing consisting of a unified, digital record of the totality of an individual's experiences, captured multi-modally through digital sensors and stored permanently as a personal multimedia archive [6]. Lifelog, as a special human logging technology, is basic in our research that needs to capturing and storing personal data from a smartphone and the nearby wearable devices.

Another primary task in lifelog is to manage personal data from various sources and in different forms or media. Ahmed and his lab proposed a lifelog framework "SemanticLIFE" [7], which was aimed at building a systematic personal information management mechanism over a human lifetime with using ontologies for the representation of semantics according to the living characteristics. The corresponding system consisted of a personal data acquisition layer and an elementary data process layer. Our system is built based on a smartphone that manages the data acquisitions from both the smartphone and wearables, and further upload the data to the server with high more integrated storage and management. We further apply a series of security techniques in personal data collection, storage and transmissions in our system.

User modeling that is a subdivision of human-computer interaction (HCI) and aims at customization and adaptation of systems to the user's specific needs [8]. A generic user modeling server for adaptive Web systems (GUMSAWS) was proposed by J. Zhang, et al [9]. This system is able to act as a centralized user modeling server to support adaptations of various Web systems concurrently and offer multiple user modeling functions. Accordingly, associated Web sites can be automatically adapted to users' needs for personalization. Our Cyber-I is a more general one for human modeling, which is beyond of a specific application dependent user modeling and expected to be used for various personalized services and applications. The unique feature of Cyber-I modeling is the growing capability that a Cyber-I can grow to become bigger, higher and closer for successive approximation to its Real-I [4]. The personality analysis, which is essential in general user modeling and also important in our Cyber-I modeling.

Clustering techniques are certainly necessary in personality analysis. The K-means cluster algorithm, proposed by Stuart Lloyd [10], is a method of vector quantization for data classification widely used in data mining. Besides using a clustering algorithm to get user's commonness, the personality analysis further needs to figure out special features of a person. Personality refers to individual differences in characteristic patterns of thinking, feeling and behaving [11]. Currently, there are various kinds of personality research based on different sociological personality test theories [12]. Among them, the Big Five personality trait theory is the most authoritative one, widely used to obtain personality traits [13]. Similar to the previous research in Cyber-I's birth stage [2], our present personality analysis is also based on the Big Five trait theory but using the data collected automatically during the Cyber-I's growth stage.

III. Overview of the Smartphone-based System

Our smartphone-based system has adopted the client/server mode where the smartphone acts as a client and Tomcat acts as a server. The general system architecture and functions in both client and server sides are shown in Fig. 1. The local DB is one embedded in the smartphone, and the server DB is for persistently keeping all data from many users.

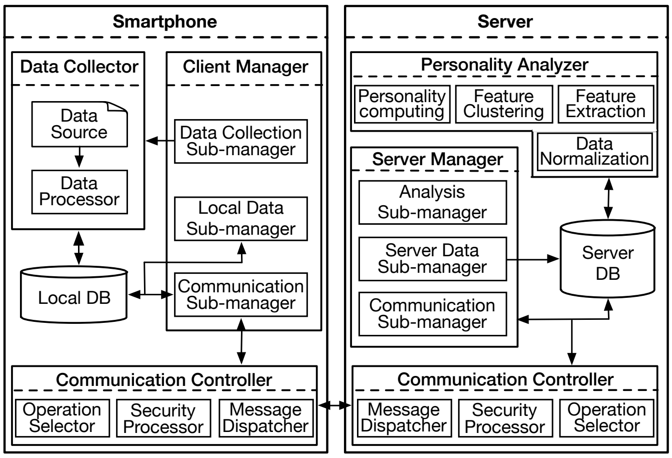


Fig. 1. Architecture of the smartphone-based client/sever system

A. Function Modules in Smartphone

In the smartphone side, there are three modules, a client manager, a data collector and a communication controller. The client manager contains three sub-managers to manage data collection, local data and data communications, respectively. The data collector includes a data processor to collect data from various sources. The communication controller consists of an operation selector, a security processor and a message dispatcher, which acts as a bridge to provide a stable and safe communications between the client and server. For the features of small storage, swift operation and high compatibility, the SQLite database is more suitable to be a local DB. The client manager will also control the local DB, which will be discussed in the next section.

B. Function Modules in Server

In the server side, there are also three modules, a server manager, a communication controller and a personality analyzer. The server manager contains three sub-manages to data operation, communication and analysis, respectively. The server data sub-manager provides a data operation interface for a user to manipulate his/her personal data in sever DB and for a system administrate to manage the whole data of all users. The Mongo DB is chosen as the server DB due to its feature of no-SQL, swift index-ability and simultaneous access support. The detail about the data operations and the server DB are discussed in the next section. The personality analyzer provides a series of functions to compute a user's personality, which is discussed in Section V.

C. Data Collector

The data collector in a smartphone is in charge of collecting data from sensors and toolkits equipped inside the smartphone as well as from outside wearables devices and APPs on the Internet as shown in Fig. 2.

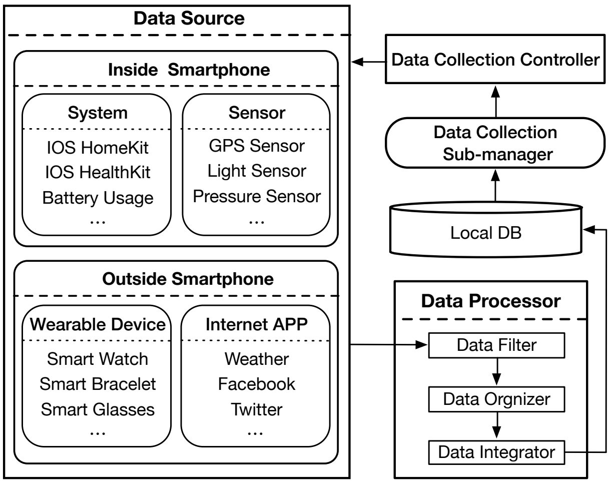


Fig. 2. Personal data collector and its function modules

The smartphone itself can provide a plenty of usage data and sensing data. For instance, an iPhone has two integrated interfaces, HomeKit and HealthKit. In addition, sensors embedded in a smartphone can also be used to get GPS, light, acceleration, pressure and other data. Currently many wearables offer a network such as Bluetooth to connect a smartphone. So a smartphone can take the data from wearables such as smart watch, bracelet, ring, etc. Further, a smartphone can be used to get information such as weather from outside APPs and SNS via the Internet. When data is collected, there may exist more or less redundancy in the data. So, the data filter shown in Fig. 2 is responsible for detection and removal of redundant data. The personal data after screened may be disorganized, and thus needs to be reorganized chronologically by the data organizer shown in Fig. 2. Data from different sources may also need to be combined, and this job is done by the data integrator shown in Fig. 2.

D. Communication Controller

Stable and safe communications between a smartphone and a server are essential in the system. The communication controller consists of three modules, an operation selector, a security processor and a message dispatcher as shown in Fig. 3. The operation selector is to provide some concise but comprehensive operations including data synchronizations and commands, such as basic client functions, user personal settings and system administration instructions. The security processor is for validation of a smartphone user, verification of data integrality, and generation of security information. The security info generator contains four parts: (1) key generator for personal data and operation, (2) dynamic private key generation by the special key rotation table, (3) information discretization processor with hash algorithm, and (4) personal data digital signature integrator for encrypting and combing data with security information. The message dispatcher includes the data protocol normalization and de-normalization. It supports client-server communications using standard Internet protocols such as HTTPS and FTP, and some special protocol such as DSR (Data Segment Resolution) protocol to distinguish operation, personal data and security information, and EKE (Encrypted Key Exchange) protocol for password authentication.

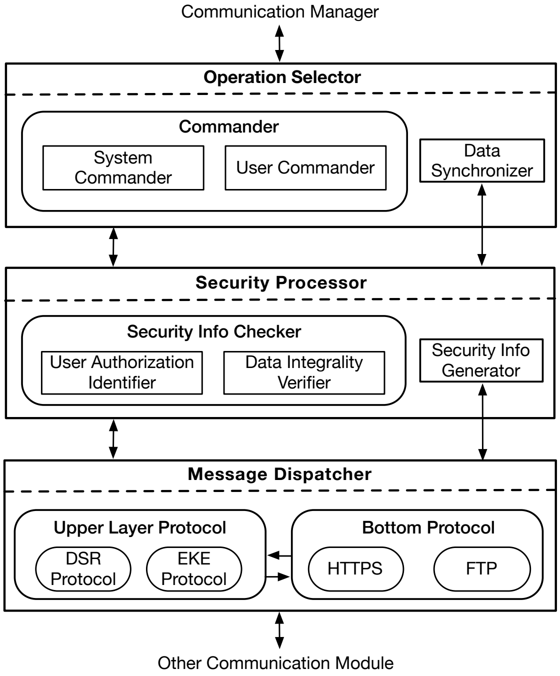


Fig. 3. Communication Controller and its function modules

IV. Management of Personal Data

Management of personal data is one of core issues in general personal information systems. As described in the last section, there are a client managers and a server one in our system. As shown in Fig. 4, the client management functions are completed by the three sub-manages of data collection, local data and communication, and the sever management functions are carried out by the three sub-managers of server data, analysis and communication. The communication sub-managers in client and server are very similar, but one running on a smartphone and the other running on a backend machine, which keeps and manage all personal data from many users. The concrete functions of these sub-mangers as well as data types in the local and server DBs are described in the following.

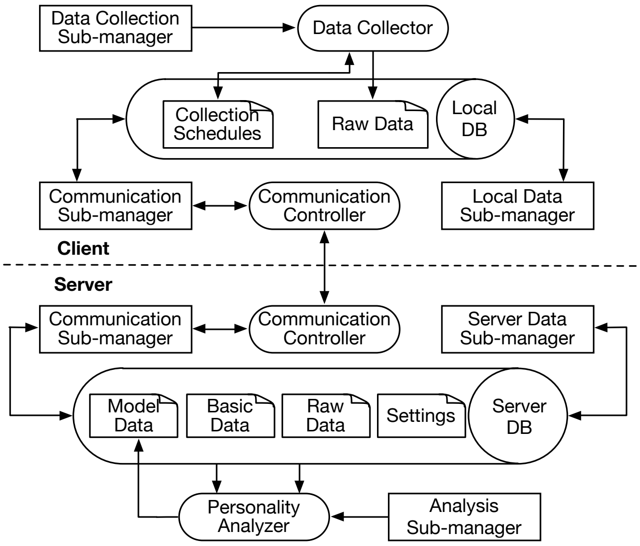


Fig. 4. The client and server managers

A. Data Collection Sub-Manager

As mentioned previously, the personal data in our system is from various sources in different forms. It is thus necessary to control and coordinate the data collections from the multiple sources. The data collection control is according to a file of the collection schedules that is stored in the local DB. The schedules specify that when data will be captured or acquired for every data source. Due to the data heterogeneity in different sources, the schedule for a data source is decided with considerations of data change rate, collection condition, local DB capacity, upload situation, etc. Another factor to be considered in setting the schedules is the energy consumption of a smartphone battery. It is necessary to make some tradeoff between the frequency of data collection and the reduction of power usage.

B. Local Database Sub-Manager

The local DB is a tiny database that is embedded in a smartphone. Considering the small size and high compatibility, it is a SQLite database. The local DB sub-manager is responsible of the database initialization and temporary personal data storage before the data is uploaded to the server DB. The sub-manger also provide an interface with basic data operations of CRUD (Create, Retrieval, Update and Delete) for a user to access and manipulate personal data in local and server DBs. Another CRUD interface is also made for administrators.

C. Communication Sub-Manager

The communication sub-manager is needed for both the smartphone and the server, and it connects directly to a database and a communication controller. The sub-manager is aimed at the fast data synchronization between a client and a server, and the rapid response during the client-server communications. The above functions are actually done by the operation selector, security processor and message dispatcher, whose working flow is automatically coordinated by the sub-manager.

D. Server Database Sub-Manager

In the server side, the No-SQL Mongo DB is adopted as a server database with considering its high index-ability, large data amount from many users, and heterogeneous data types from various sources. Once a user registers to the system, the user's ID and login information will be created and kept in a table including all users. Each user's personal data is cataloged in four types, i.e., basic data, model data, raw data and settings, which are in four correspond tables of the Mongo DB. Because of the no-structure feature of Mongo DB, all raw data to a user can be placed into a single table in the database no matter what data forms and media are

E. Personality Analysis Sub-Manager

The personality analysis sub-manager is resided in the server side as a controller to schedule and handle procedures in the personality analysis. For different users, the corresponding procedure may be different according to the user's personal data types, data amount, increasing rate and so on. The detailed procedure and related processing in the personality analysis is discussed in the next section.

V. Personality Analysis

The personality analysis is a process consisting of data normalization, feature extraction, feature clustering and personality computing, as shown in Fig. 5. First, the raw data will be taken from the server database and the data normalization will be carried out. Next, the normalized data will be used to extract features, which will be further clustered. Finally, the personality computation will be conducted based on feature extraction and clustering results.

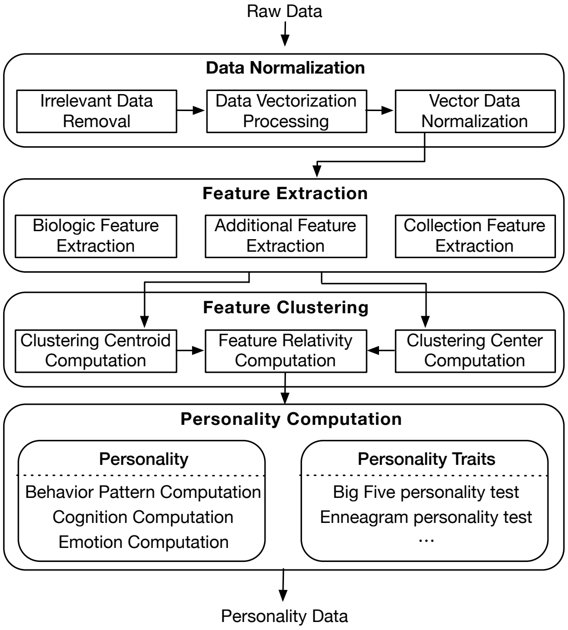

Raw Data
Data Normalization
Irrelevant Data Removal
Data Vectorization Processing
Vector Data Normalization
Feature Extraction
Biologic Feature Extraction
Additional Feature Extraction
Collection Feature Extraction
Feature Clustering
Clustering Centroid Computation
Feature Relativity Computation
Clustering Center Computation
Personality Computation
Personality
Behavior Pattern Computation
Cognition Computation
Emotion Computation
Personality Traits
Big Five personality test
Enneagram personality test
...
Personality Data

Fig. 5. Personality analysis process

A. Data Normalization

Data normalization contains three main functions that are irrelevant data removal, data vectorization processing and vector data normalization. The irrelevant data removal is to remove unnecessary information such as data ID and some flags. The vectorization processing is to take data values with different attributes from a single data source. For instance, a user's movement routes and often staying areas can be taken from the same GPS records. The vector data normalization can be done with using three different algorithms: min-max scaling, standard deviation normalization, and logarithmic normalization. The advantage of min-max scaling is that all data values can be scaled to the maximum of "1" and the minimum of "0". The standard deviation normalization is an algorithm to normalize data according to the data's average value and compute the deviation with this average. The logarithmic normalization is a nonlinear procedure to deal with nonlinear data. It is widely used especially when data analysis involves a large amount of data. In our research, the min-max scaling and the standard deviation are used more often.

B. Feature Extraction

Data features are extracted according to three aspects: (1) the features during a user' data collection process, called data collection features, (2) the features associated with a user's existence states including biologic states such as the user's heartrate and location information, which is called biologic features, and (3) the features associated a user's additional information such as the user's preference according to his/her browsing history, which is called additional data features. The latter two features called data value feature. The collection features are mainly extracted from the data sources, the data collection schedule and the system usage. The biologic features can be described as the basic properties and natural behaviors of a living being. The regularity computing and burst computing are applied to the biologic feature extraction for different types of data. The additional data feature extraction is mainly for getting statistics of data type, usage and proportion in order to discover the behavior characteristics.

C. Feature Clustering

The feature clustering is for clustering the commonness with the appropriate data features from the majority of persons. The K-means clustering algorithm is embedded in the system as the basic clustering algorithm, and its formula is below.

$$C_\sigma = \arg\min_{S} \sum_{i=1}^{k} \sum_{x \in S_i} \| x - \mu_i \|^2 \qquad (1)$$

where each element x belongs to the element set S_i, the k is a number of centroid μ_i , and the C_σ belongs to the set of arguments ($\arg$) that minimize ($\min$) the function norm in set S . With the K-means clustering algorithm, we can separate the data into two or more different cluster. The centroid computation is shown in formula (2),

$$r_\sigma = \frac{\sum_{i} m_i r_i}{\sum_{i=1}^{n} m_i} \qquad (2)$$

where the r_σ is a centroid in cluster of data, the r_i and m_i are each element's influence factor and value, respectively. A centroid is relied on a clear data distinction. The calculation of feature relativity R is given in formula (3).

$$R = \frac{\sum_{i=1}^{n} C_i}{n \sum_{i=1}^{n} r_i} \qquad (3)$$

where the C_i is the K-means center. We can just apply the above the clustering calculations to any of data features and take the same processing. Each process has own cluster centers and the feature relativity. Further, completed cluster centers will be integrated by features of relativities and their cluster centers.

D. Personality Computation

We compute the personality from four different perspective. (1) The behavior pattern computation is based on the Type A and Type B personality theory, in which human's behaviors is classified to two contrary types. (2) The cognition computation is to calculate the set of mental abilities such as human's preferences in this research. (3) The emotion computation is about a person's attitude, which is a complicated and stabled in biological evaluation and physiological experience. We only extract and verify human's daily emotion based on the Robert Plutchik's theory of a wheel of emotions, mainly by counting emotion changes associated with behaviors. (4) The personality traits computation in our research is based on the Big Five personality test that is the most popular one. And a factor is approached by integrating each features and its relativity. The personality feature value F can be computed by formula (4).

$$F = \frac{\sum_{i=1}^{n} R_i \frac{t_i}{C_i}}{\sum_{i=1}^{n} R_i} \qquad (4)$$

where, the R_i is the relativity, the t_i is the extracted data, and C_i is the cluster center in each feature. Personality can be separated into several different levels or parts according to different personality theory.

VI. Evaluation and Case Study

To evaluate performance of our system, a series of experiments have been conducted for testing its data communication and security. A case study about the personality analysis is also shown with using actual personal data samples.

A. Network and Energy Usages

In the experiments, an iPhone6 was used as a client to manage data collections and uploads to a server PC running a Windows 2003 OS. Fig. 6 shows the smartphone's network activity and energy usage in 24 hours. It can be seen that the network activity is not continuous but periodic, and the same is for energy usage. This is because the data collections and uploads are done periodically with following the pre-specified data collection and synchronous schedules. Also, experiment has shown that the energy consumed in running our system is only about 2% in the smartphone's total energy consumptions.

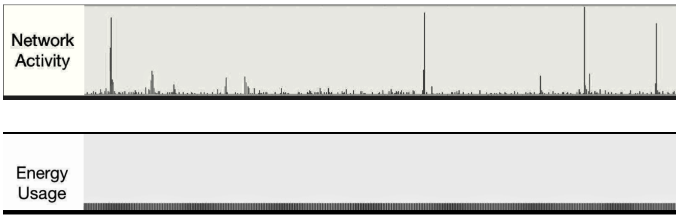
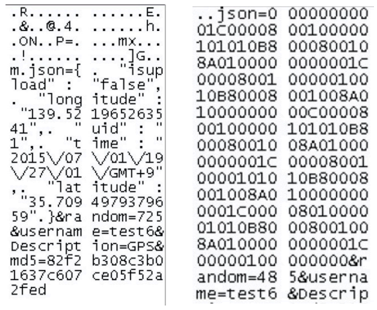


Fig. 6. Network activity and energy usage

B. Data Integrity and Security

Each set of collected data has an individual MD5 code for data integrity, and be encrypted with AES as shown in Fig. 7.



```
.R......  ......E.
.&..@.4.  ......h.
.ON..P=.  ...mx...
.!......  ....]G..
m.json={  .   "isup
load" :   "false",
.  "long  itude" :
 "139.52  19652635
41",.  "  uid" : "
1",.  "t  ime" : "
2015\/07  \/01\/19
\/27\/01  \/GMT+9"
,.  "lat  itude" :
 "35.709  49793796
59".}&ra  ndom=725
&usernam  e=test6&
Descript  ion=GPS&
md5=82f2  b308c3b0
1637c607  ce05f52a
2fed
```

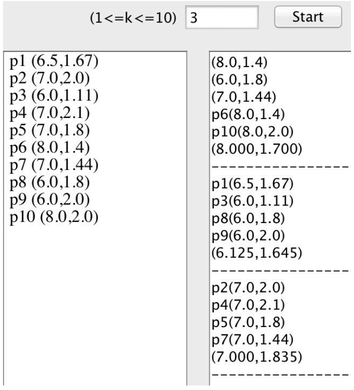
```
..json=0  00000000
01C00008  00100000
101010B8  00080010
8A010000  0000001C
00008001  00000100
10B80008  001008A0
10000000  00C00008
00100000  101010B8
00080010  08A01000
0000001C  00008001
00001010  10B80008
001008A0  10000000
0001C000  08010000
01010B80  00800100
8A010000  0000001C
00000100  000000&r
andom=48  5&userna
me=test6  &Descrip
```

Fig. 7. Communication with and without security process

The figure shows the sample data captured by the Wireshark tool, where the left part is the original data without encryption and the right part is encrypted one, which is just an unreadable symbol string. The security process has prevented illegal modifications and reorganizations in communications between the client and server.

C. Case Study on Personality Computation

In this case study, GPS data and sleep data were chosen as samples for personality computation. Ten tester's data sets were collected and processed with following the four steps, data normalization, feature extraction, feature clustering and personality computation as explained in the last section. Fig. 8 shows the clustered results according to the ten testers' GPS and sleep data features.



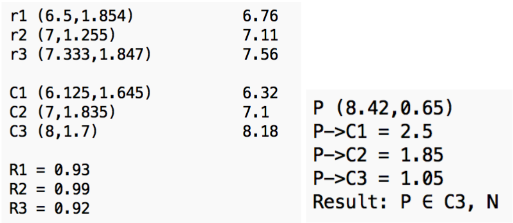
(1<=k<=10) 3 Start

```
p1 (6.5,1.67)
p2 (7.0,2.0)
p3 (6.0,1.11)
p4 (7.0,2.1)
p5 (7.0,1.8)
p6 (8.0,1.4)
p7 (7.0,1.44)
p8 (6.0,1.8)
p9 (6.0,2.0)
p10 (8.0,2.0)
```

```
(8.0,1.4)
(6.0,1.8)
(7.0,1.44)
p6(8.0,1.4)
p10(8.0,2.0)
(8.000,1.700)
----------------
p1(6.5,1.67)
p3(6.0,1.11)
p8(6.0,1.8)
p9(6.0,2.0)
(6.125,1.645)
----------------
p2(7.0,2.0)
p4(7.0,2.1)
p5(7.0,1.8)
p7(7.0,1.44)
(7.000,1.835)
----------------
```

Fig. 8. K-means clustering

Each tester was asked to complete a questionnaire to test his/her behavior personality. We divided the ten testers into three types: 'A' that is ambitious, rigidly organized, highly status-conscious, sensitive and impatient, 'B' that is totally contrasted to type A, and 'N' that is neutral, in-between type A and type B. After the K-means clustering process, the three clusters' centers are as follows: (6.125, 1.645), (8.000, 1.700), and (7.000, 1.835). The centroid and the data reliability are shown in Fig. 9.



```
r1 (6.5,1.854)        6.76
r2 (7,1.255)          7.11
r3 (7.333,1.847)      7.56

C1 (6.125,1.645)      6.32
C2 (7,1.835)          7.1
C3 (8,1.7)            8.18

R1 = 0.93
R2 = 0.99
R3 = 0.92
```

```
P (8.42,0.65)
P->C1 = 2.5
P->C2 = 1.85
P->C3 = 1.05
Result: P ∈ C3, N
```

Fig. 9. Featre relativity computaion and a new sample computation

We can see from the left part that r1 to r3 are centroids, C1 to C3 are cluster centers, and R1 to R3 are data reliabilities in the cluster centers. The right one describes a new sample computation based on the left parameters. The data (8.42, 0.65) indicates that the tester's average sleep time is 8.42 hours and walking speed is 0.65 m/s. A result finally computed shows that the tester's behavior type is N, namely a neutral one. As compared with the questionnaire previously did by the tester, the result shows the tester is indeed a type N, the same as the computed value.

VII. CONCLUSION AND FUTURE WORK

This research is to design and develop a smartphone based client-server system for three main objectives. The first one is to automatically collect the data from embedded sensors and apps in a smartphone and wearable devices worn by a user, and upload the collected data from a SQLite local database in the smartphone to a remote Mongo database in the server. The second one is to guarantee data integrality and security due to the sensitivity of personal data. The third is the personality analysis that is carried out in four steps, namely data normalization, feature extraction, feature clustering and personality computation, based on the Big Five trait theory. A representative example as case study was conducted to show the personality analysis process and the analyzed results.

Future research will be conducted in the following aspects. First, the system should be more scalable so that more sources of data from various apps and wearables can be flexibly added to and integrated into the system. Second, the scheduler needs to be further improved for more coherent management of data collections from extensible data sources. Third, the multi-security mechanism in the system needs to be tested to check its possible security problems. Forth, the effectiveness of current personality analysis should be widely tested for further improvements.

REFERENCES

[1] J. Wen, K. Ming, F. Wang, B. Huang, and J. Ma, "Cyber-I: Vision of the individual's counterpart on cyberspace," IEEE Int'l Conf. on Dependable. Autonomic and Secure Computing (DASC 09), pp. 295-302, 2009.

[2] J. Ren, J. Ma, R. Huang, et al, "A management system for cyber individuals and heterogeneous data", in Proc. of the 11th IEEE International Conference on Ubiquitous Intelligence and Computing (UIC 2014), Bali, Indonesia, December 9-12, 2014.

[3] L. Tang, J. Ma, R. Huang, et al, "Awareness and control of personal data Based on the Cyber-I privacy model", in Proc. of the 11th IEEE International Conference on Autonomic and Trusted Computing (ATC-2014), Bali, Indonesia, December 9-12, 2014.

[4] S. Zhang, J. Ma, R. Huang, et al, "Growable Cyber-I's modeling with increasing personal data," in Proc. of the International Conference on Advances in Computing, Control and Networking (ACCN 2015, Bangkok, Thailand, February 21-22, 2015.

[5] P. W. Cheng, S. Chennuru, S. Buthpitiya, & Y. Zhang, "A language-based approach to indexing heterogeneous multimedia lifelog," International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction. ACM, pp. 26, 2010.

[6] M. Dodge and R. Kitchin, "Outlines of a world coming into existence: Pervasive computing and the ethics of forgetting," Environment and Planning B: Planning and Design, pp. 431–445, March 2007.

[7] M. Ahmed, H. Hanh, H. K. Shuaib, K. Shah, K. Nguyen Manh, et al, " 'SemanticLIFE'–A Framework for managing infromation of a human lifetime," In Proceedings of 6th International Conference on Information Integration and Web-based Applications and Services, pp. 725-734, 2004.

[8] Fischer, Gerhard, "User Modeling in Human-Computer Interaction," User Modeling and User-Adapted Interaction pp. 65–68, 2001.

[9] J. Zhang and A. A. Ghorbani, "Gumsaws: A generic user modeling server for adaptive web systems," in Proc. of the Fifth Confe. Communication Networks and Services Research (CNSR'07), pp. 117– 124, 2007.

[10] MacQueen, J. B. "Some Methods for classification and Analysis of Multivariate Observations," Proc. of 5th Berkeley Sym. on Mathematical Statistics and Probability, U. of California Press. pp. 281–297, 1967.

[11] A, E. Kazdin, "Encyclopedia of psychology," Washington, DC: American Psychological Association, 2000.

[12] D. Funder, "Personality," Rev. Psychol., vol. 52, pp. 197– 221, 2001.

[13] A. Vinciarelli, G. Mohammadi, "A survey of personality computing," Affective Computing, IEEE Transactions on, pp. 273-291, 2014.