A Management System for Cyber Individuals and Heterogeneous Data

Ren, Jun

A Management System for Cyber Individuals and Heterogeneous Data

Jun Ren

Graduate School of Computer and Information Sciences
Hosei University
Tokyo 184-8584, Japan
jun.ren.gy@stu.hosei.ac.jp

Abstract—**The progressive development of information and communication technologies has led us to a new world called hyper world that is composed by the cyber world and the physical world, which, at the same time, brings us the digital explosions of data and connectivity as well as all kinds of smart services. Based on these, Cyber-I was proposed, which aims at creating a unique, digital, comprehensive description for every real person so that the possibility for the person being in the cyber world to get lost in the digital explosions will be reduced. However, the ultimate goal of realizing such Cyber-I needs an endless research process, which will cover multiple areas and disciplines. In order to merge efforts on the study of Cyber-I together as well as provide people with better services utilizing Cyber-I, this paper presents our research and development on a Cyber-I oriented management system where (1) the life cycle of Cyber-I including its birth, growth and death is simulated, (2) with the NO-SQL properties supported by Mongo DB, heterogeneous personal data coming from disparate sources with different formats through varied media can be managed in a scalable way, and (3) apps connected with the Cyber-I can not only fetch personal data but also provide personalized services, while the collected personal data can be used to generate user models from different aspects.**

Keywords—Cyber-I; personal data; model; birth; growth; trait; management; heterogeneous; scalibility; mongo

I. INTRODUCTION

In recent years, various digital explosions including data, connectivity, service and intelligence are emerging in the cyber-physical integrated hyper world [1]. As individuals are facing so many novel services in the digitally explosive world, we may not be aware of what are the most necessary/suitable things. As a matter of fact, in the cyber space composed by various computers and global Internet, we can create not only numerous smart/ubiquitous things, but also "another self" called cyber individual (Cyber-I) for each real individual (Real-I), which may assist the individual in handling the service explosions so as to have a more comfortable and enjoyable life [2] [3]. However, the realization of Cyber-I is a worldwide challenge, which needs aggregation of interdisciplinary and multidisciplinary studies as well as numerous technologies to support the work on the research of individual. Consequently a Cyber-I oriented platform is required for user to create Cyber-I and utilize Cyber-I based applications. Therefore, this research is to develop a system for the management of a number of Cyber-Is including its birth, growth and death as well as the related heterogeneous personal data.

The personal data possesses the great heterogeneity for the reason that such data is largely unknown, unlimited, and always coming from numbers of sources in various formats. As a result, a fundamental requirement imposed to the Cyber-I system is the scalability to manage heterogeneous data. Therefore, MongoDB has been adopted for the storage of personal data while a component is responsible for transforming the heterogeneous data into Json format before being stored into the MongoDB. A Jquery plugin known as Jstree acts as the data display interface, and the data manipulation functions are also provided.

This paper is to present the Cyber-I oriented system and the remainder of this paper is organized as follows. Section II covers related work to describe relationships between our Cyber-I oriented system and similar studies. Section III describes the general functionality and architectures of the system. Section IV and Section V respectively illustrate the key points involved in Cyber-I's birth and growth process. Section VI explains the mechanism for the management of a large number of Cyber-Is as well as their related heterogeneous data. Then we show a set of testing results for evaluation of the system performance in handling large amount of data in Section VII. Conclusions about the system and related issues that we may encounter in the future are presented in the last section.

II. RELATED WORK

Over a period of time, human has been the center of computing. Human-centered computing (HCC) is aiming at bridging the gaps between the various disciplines such as signal processing, machine learning and ubiquitous computing involved with the design and implementation of computing systems that support people's activities [4]. Beyond being a merging area for existing disciplines, HCC also aims at radically changing computing with new methodologies to design and build systems that support and enrich people's lives. Based on the concept of HCC, user modeling and lifelogging are two of the good practices, which will be explained below.

User models as defined by A. Kobsa, are the collections of information and assumptions about individual users, which are needed in the adaptation process [5]. The user models are essential for making personalized services such as adaptive hypermedia system and recommendation system that can assist users in finding the content they preferred. A generic user modeling server for adaptive Web systems (GUMSAWS) was proposed by J. Zhang [6]. This system is able to act as a

centralized user modeling server to support several adaptive Web systems concurrently and offer the user modeling functions, so that Web sites can be automatically adapted to reach the goals of personalization, recommendation, selection, and usage analysis. A. Kobsa also implemented an LDAP based user modeling server (UMS) that allows for the representation of different models [7]. External clients such as user-adaptive applications can submit and retrieve information about users.

A lifelog system can be defined as the one, which involves gathering and exploiting any historical/personal activity records. The objective of the lifelog technology is to enable to trace the threads of an individual's life in terms of events, states, and relationships. S. Lee put forward a lifelog system named LifeLogOn where users are allowed to easily build and exploit lifelog ontology by importing various available data sources [8]. R. Rawassizadeh in his doctor dissertation proposed "A Holistic Multi-Purpose Life Logging Framework" which is open to extension and configuration of various sensors, and at the same time enables users to share their lifelog information and supports required features for life logging [9].

Our Cyber-I system differs from other user modeling systems from the point that it is neither domain specific nor application dependent. And instead of serving as a middleware for user applications, the priority of our Cyber-I oriented system is to provide a platform for the overall studies and analyses of the individual as a human being. Moreover, compared to life log systems, our system also contains log data for individuals, but we place more emphasis on processing and presenting the collected data in a human related concept and manner. From this perspective, the Cyber-I oriented system can be seen as the combination of user modeling and lifelog, which is the inevitable outcome of human centric computing.

III. THE GENERAL SYSTEM DESCRIPTION

The system serves as a common place for Cyber-Is' residence with a set of Cyber-I life control functions, by which Cyber-I is able to (1) born up to its Real-I's request, (2) grow up with the aggregation of personal data, dynamically changing applications and more comprehensive models, and (3) be terminated on its Real-I's demand in various ways. The general functions of the system are shown in Fig. 1. In the first place, a birth process is required before Cyber-I's entering into the system, during which Cyber-I's core elements are captured. Thereafter, with the help of the Data/App/Model Management Interfaces, (1) users are capable of viewing and manipulating their personal data; (2) applications are able to be dynamically associated with Cyber-Is; (3) models can be generated from the personal data in different aspects of the users. The details about Cyber-I's birth process as well as its growing process will be explained in the next two sections.

The system is based on the Browser/Server architecture where Tomcat acts as the server. Mongo DB is selected as the database for the persistence of personal data. The details about Mongo DB and the reason why it can meet the requirements in management of Cyber-I related heterogeneous data, will be discussed in Section VI. In the development of the system, Java is utilized to implement the business logic, and the Struts2 Framework is adopted to simplify the manipulation of Servlet actions. In the foreground of the system, HTML5, CSS3,

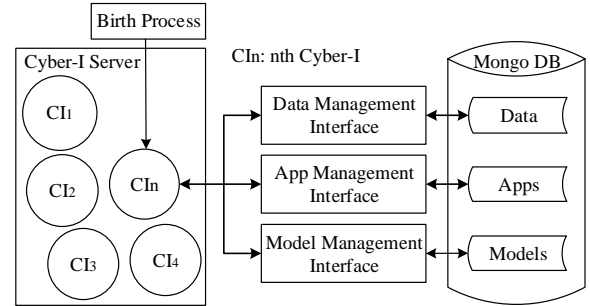JavaScript/Jquery/Ajax cooperate with each other to represent the Cyber-I in a lively and impressive way.

Fig. 1. The general system description

IV. MANAGEMENT OF CYBER-I'S BIRTH

Cyber-I's birth process is of great significance. In this section, the detailed procedures involved in Cyber-I's birth process are described.

In order to start a "new life", a Cyber-I has to experience the stage of incubation, during which the core data of a human that will act as the foundation of the sustainable development in the subsequent growing process is captured. Figure 2 shows the concrete procedures for the Cyber-I's birth. A user is asked to go through 5 steps before his/her Cyber-I's entering into the system. The first three steps are mandatory while the last two steps are optional, and details will be discussed in the following.
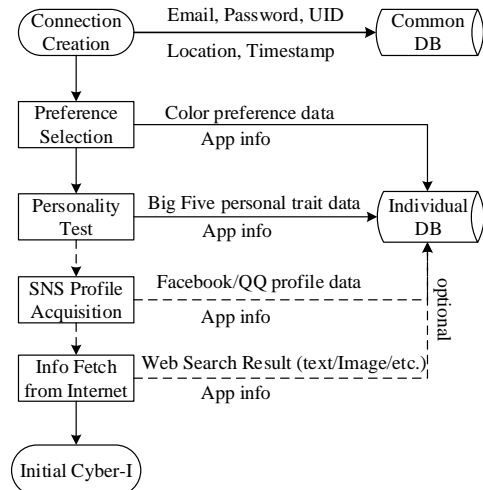

Fig. 2. Cyber-I's birth process

Just like we human beings need an ID card in the physical world, Cyber-Is also require a unique identifier to mark them in the cyber world. In addition, a bridge should stand between a Real-I and a Cyber-I for th eir interactions. In this process, a user is asked to input his/her email address as well as his/her login password. Besides, the current timestamp and the IP-address based location information are automatically recorded. In this way, a combination of the email, timestamp and location will comprise the unique identifier while the email can serve as the connection between a Real-I and a Cyber-I. Meanwhile, related data generated in this process is stored in the common database, which is in charge of the overall management of all the Cyber-Is resided in the system.

Preferential choice has long been of interest in the area of psychology, and psychological research provides evidence that the recognition of user preferences could lead to his/her decision-making improvement. A user's preferences are able to influence the selection and instantiation of the actions that achieve his/her goal. In our system, we started from the simple color preference (CP) since (1) CP may not so sensitive to user's privacy concern; (2) it is relatively easy to make color choice and users may be happy to know their own CP; (3) CP modeling can be extended from such the simple one at the moment to sophisticated ones (state, behavior and mind levels) in the future. The data generated in this process is stored in Cyber-I's individual database, and moreover, this step along with the steps that will be described in the following are treated as Cyber-I's applications/apps. As a result, application information including app name, app URL, added time and location will be stored in an individual's apps database as well.

Trait theory is the biggest area of research in personality development, which states that human beings possess wide varieties of characteristics that are constant over time. Many evidences accumulated have proved that virtually all personality measures can be reduced and categorized into five basic dimensions, often referred to as the "Big Five personality traits", which was first proposed by E. Tupes and R. Christal in 1961 [10]. The Big Five traits are *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness* and *Neuroticism*. These factors could be viewed as the secrets or keys in one's lifetime so that they are also dominative in Cyber-I's growth. A questionnaire is implemented to obtain the Big Five personality traits. In this questionnaire users are required to answer some questions and then the values that covered the five dimensions are inferred from users' answers to the questions. The values are transformed to the percent style and stored in the database. One part of a questionnaire (left side) and a sample result (right side) are shown in Fig. 3.
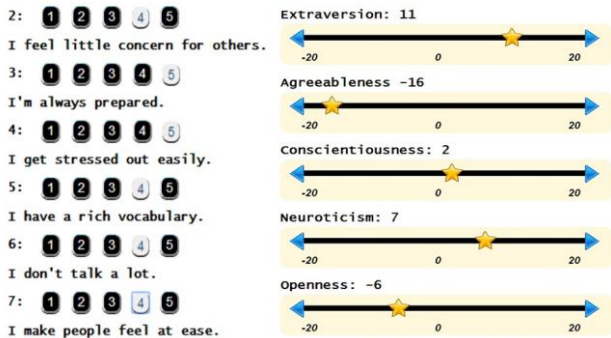

Fig. 3. Big Five questionnaire and results

Social network sites (SNSs) have become the essential components permeating our daily lives as they not only enable us to communicate with lots of people but also are capable of holding our self-presentation information. From the SNSs such as Facebook, QQ, LinkedIn, etc. users' profiles could be obtained. For some SNSs like Twitter that doesn't support a comprehensive user profile, users' preferences and emotional moods and so forth could also be extracted from their history posts. A user's profile plays as an important role during Cyber-I's growth process since it contains some context information that can be utilized when Cyber-I is offering services to Real-I.

Ways to get Facebook profile and QQ profile are implemented in our system where a user has to provide his/her ID and access token. This step is designed to be optional since some persons may not have any SNS at all, while some persons are unwilling to share their personal profiles.

There are always scattered personal pieces of information existing on the Internet especially persons who got some fame. By observing a person's traces on the Internet, we can grab some information about him/her as much as possible. It is reasonable to fetch such kind of data since (1) information available on the Internet are pubic; (2) it is useful in covering a person's features in the cyber world; (3) such data is able to grow along with the user's growth. In order to precisely gather a user's information, not only his/her name but also some other personal information is suggested to input into our search interface. Besides, the profile obtained in the last step could also help to locate a person on the Internet. Our system is able to fetch related Web pages as well as images. Figure 4 shows an example of the fetched results (Web pages & images). It is up to the users whether to store this information or not as the fetched information may have been mistaken. This step is also optional since some persons didn't leave any traces on the Internet at all.

Fig. 4. Web page & image search result

Cyber-I is said to be born after the above five steps (three mandatory ones and two optional ones). During this process, Cyber-I is registered in the cyber world, its connection between the Real-I is built, and the obtained core data is going to dominate Cyber-I's further growth while the information collected from the SNSs or Internet will enrich Real-I's features. The last four steps are also viewed as applications that could be designed to get the same target information from users through different forms in the future. After born, the Cyber-I is well prepared for its growth.

V. MANAGEMENT OF GROWING CYBER-I

Cyber-I's growth is a process of approximating to Real-I's essentials along with collecting more personal data as well as related analyses, models, uses, etc. The details involved in this growth process will be explained in this section.

A. Key Features/requirements in Cyber-I's Growth

The core of Cyber-I's growth depends on a well-defined mechanism for the management of the relationships between Cyber-I and its related personal *data*, *applications* and *models*.

The main elements and the dependencies between them are depicted in Fig. 5. The interactions between a Real-I and its Cyber-I are mainly based on their associated applications/apps. Normally, the personal data is yielded from the applications, some of which is going to be utilized to generate user models in some specific aspects. Thus, with the help of a user's models, applications will get a better understanding to the user when providing services to him/her. Users have full control over their personal data and they are also capable of terminating their Cyber-Is in different ways. The Cyber-I's growth functions are explained in the following sub-sections.
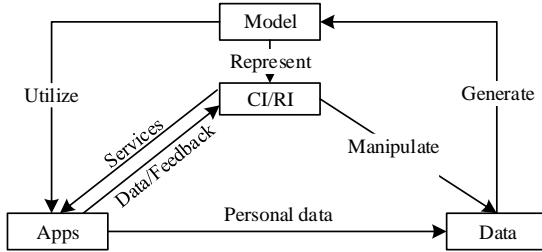

Fig. 5. Cyber-I's growth process

B. Apps Associated with a Cyber-I and Their Functionalities

The apps are a Cyber-I's applications (CI-Applications) that contain various expressions of perceiving the world and performing consequences of Cyber-I's computing. A CI-Application usually can take one of the three functions: (1) To make Cyber-I born and/or grown; (2) To serve a Real-I based on its Cyber-I; (3) To integrate the above 1 and 2. Metadata of the applications that reflects the relationship between a Cyber-I and its related applications is preserved in a collection named app metadata in each Cyber-I's individual database. With the help of App Management Interface, applications reserved in the App Library could be added dynamically to a Cyber-I. Conversely, a connection between a Cyber-I and an application could also be deleted on a user's demand. An app named "AppUsage" exists in our App Library that is focused on offering services to users. In order to implement the AppUsage, we took advantage of the software called "Manic Time" that is able to generate the logs about a user's activities on a PC, which will be automatically or manually uploaded to the user's individual database later. After analyzing the log data reserved in database, AppUsage is able to give a general description about how much time has been spent on each kind of software, and show the results in a pie chart. Users can have a brief understanding of their habits of using the software with services provided by AppUsage, which will also help them make better use of time to some extent.

C. Increasing Personal Data

Personal data refers to the information related to a living individual who can be identified. In our system, the steady increase in personal data relies mainly on user's interactions with applications. In addition, users can manually upload their files (almost all kinds of format are supported) to their individual database using fileuploader. The possibilities of a higher leveled and more precise abstraction of a person can increase along with the increase of personal data. A tree structure is used for the representation of personal data, which also enable users to manipulate their related data. Details involved in the data management will be discussed in the next section.

D. Model Generation

User modeling has been recognized as an essential technique in accessing, reforming and storing information for the purpose of effective recommendations and personalized services. The model in our system is composed of a three-layer structure, i.e., basic layer, feature layer and intrinsic layer as described in [11]. The bottom basic layer contains user's basic information such as user name, age, profession, etc. The middle feature level covers the features of a user, including his/her interests, preferences, goals, etc. And the top intrinsic layer describes user's intrinsic characteristics that may consist of his/her way of thinking, personality traits, etc.

We can make an initial Cyber-I's models corresponding to each of the above three layers with the data obtained in the birth process. The initialization of the basic layer mainly relies the SNS profiles, from which user's basic information is extracted. The basic layer will be more general and more accurate with more available SNS profiles. Color preference as one of the typical preferences is used to initialize the feature layer since it can reveal a person in some aspects to a certain extent. The feature layer is able to generalize user's features both deeply and broadly if more interests/preferences are captured in the future. The intrinsic value of a person is dominative in his/her growth while the personality traits play a significant role. The top intrinsic layer is originally based on the Big Five personality traits. As we have illustrated in the preceding section, the Big Five contains five aspects while each aspect has both a positive and negative side. A set of words has been adopted in order to express the level in each aspect on both sides. Take the agreeableness as an example. On the positive side, the words gullibility, submissiveness and selflessness orderly express the level from low to high. On the negative side, words like deceitfulness, grandiosity, manipulativeness and suspiciousness are used to express the level in the same order. With the growth of the intrinsic layer, the model will become closer and closer to a user's essence. Making use of these models, personalized applications such as a Cyber-I based searching engine would provide better quality of service. On the other hand, feedback or data generated by users during their interactions with the applications will be utilized to reappraise the models.

E. Cyber-I's States When Terminated

A Real-I and its Cyber-I keep a symbiotic relationship that could be broke by the Real-I, which means the Cyber-I's death. However, a Cyber-I could have different forms of existences in the cyber world after being terminated. As a result, our system has supported three ways for Cyber-I's death. The most easily understood way is Cyber-I's complete disappearance where all the Cyber-I's data is fully removed. Another way is to suspend a Cyber-I's life, and during the period when the Cyber-I is temperately halted, all the activities in which the Cyber-I is involved are stopped. When a Cyber-I is suspended, its identification information including the contact email is still kept, which will be utilized for the Real-I to resume his/her Cyber-I again. The last way of death may sound a bit lofty. Just like some self-giving people who are willing donate their organs for others or medical research after death, a Cyber-I is also able to donate parts of its data/model to the cyber world for further research on human. A Real-I has a full control on which parts of data are to be donated while the other parts will be removed.

VI. Management of Multiple Cyber-is and Heterogeneous Data

This section is mainly focused on the illustration of a flexible mechanism for the management of a number of Cyber-Is as well as the huge amount of heterogeneous data that is associated with these Cyber-Is.

A. Key Features in Cyber-Is and Data Management

In order to cope with the problems of supporting efficient integration of heterogeneous personal data and enabling the coexistence of the large amount of Cyber-Is, MongoDB is exploited as the data persistence tool. As shown in Fig. 6, each Cyber-I is assigned an individual database while a common database is holding some global information of all Cyber-Is' status on the whole. Every database maintains a set of collections for the storage of document-based data while all kinds of files are preserved by GridFS collections (abbreviated as GridFS in Fig. 6). The personal data yielded by applications is required to be converted to the Json format before being stored into the collection. And a data processing module is used to transform each document in the collection to a Json string that is recognizable by the Jstree, which is the main data display component. With Ajax, any modifications done on the Jstree could be represented in the database. Files uploaded to the system by the file uploader are kept in GridFS collections while users are also able to download the files through the downloader tool.
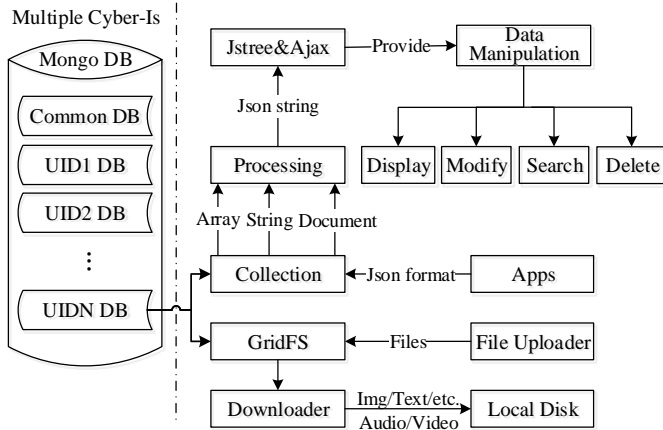

Fig. 6. System general data management

B. Advantages of Mongo DB

MongoDB is a leading NoSQL database that possesses strong capability and high performance in processing massive data [12], which is well suitable for the management of multiple Cyber-Is and their related personal data that is always increasing without limit. Moreover, with the schema less property provided by Mongo, there is no need to define a fixed collection structure in advance, and it thus enables to immensely improve the flexibility of the system in dealing with heterogeneous data. The GridFS collection is effective in storing files and it also exhibits special advantages over relational database in files management. Unlike the relational database that only maintains the file path but conserves the real files in a file system, Mongo directly stores the files in the database so that the file consistency can be easily kept. Besides, we don't have to worry about the size limitations existing in file systems like limited maximum number of files or directories.

C. Management of Multiple Cyber-Is

Although our system is a place for holding a possibly large number of Cyber-Is, every Cyber-I exists as an independent entity whose related data ought to be separately preserved. As a result, an individual database named as Cyber-I's unique id is assigned to each Cyber-I residing in our system. Thus, the related personal data is separated so that there is no need to worry about scoping it out. In addition, a common database is also required to maintain all these Cyber-Is' most basic information as well as the Cyber-I's applications whose information is specified in the App Library. A set of functions have been also made to allow a system administrator to check the general status about the Cyber-Is existence and do some system management work.

D. Management of Heterogeneous Personal Data

Personal data is generally classified as document oriented data and files, which is dealt with different mechanisms. BSON is the basic data storage in MongoDB that is analogous to Json. Thus, in the data persistence layer, a data conversion component is implemented with utilizing the Gson library for translation of the heterogeneous data into the Json format. In the view layer, a powerful and extendable Jquery plugin named Jstree is exploited as a visualization tool to display personal data. A component plugged between the data persistence layer and the view layer is responsible for translating the three forms of data, attribute-value pairs, document and document array that exist in the collection, into the Json strings, which are recognizable by the Jstree as the contents of the tree. Except from data display, other data manipulation functions such as data modification, deletion and search are implemented by combining the Jstree with Ajax. Figure 7 shows the procedure when a user is manipulating the data on the left part, and the right part is a piece of detailed data represented by a tree structure. Personal files are handled by GridFS, which is a specification for storing and retrieving files that exceed the BSON-document limit of 16MB. Files in GridFS collections are split into small chunks while each chunk is stored as a separate document in a chunks collection, usually 256k sizes while metadata about the file, including the filename, content type, and other optional information needed, is stored as a document in files collections. Theoretically, all kinds of files with any size are supported by GridFS. Users could download files conserved by the databases freely so long as they have logged into the system. Moreover, files could be directly inspected on the browser if they are in the supported formats.
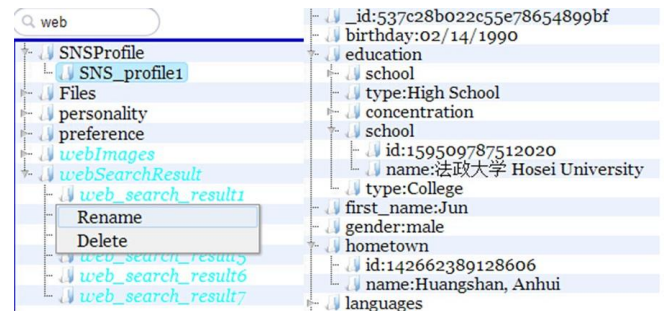

Fig. 7. Data shown in Jstree

VII. System Evaluation

A set of experiments was conducted in assessing our system's performance in dealing with large amount of heterogeneous data. And the performance differences between MongoDB and MySQL were also tested. The experiments were made under a 2.80GHz, Intel(R) Core(TM) i7 CPU with 8GB of memory, and 1TB of disk, for evaluating performance of the operations of insert, update, and selection on MongoDB 2.4.9 and MySQL 5.0 with different rows of data. In practice, we have found that the time cost on doing insert and update is nearly the same on both databases, therefore, insert and update are viewed as the same operation in our experiment. The storage spaces occupied by the data were also recorded. The data sample is picked from our system, and is shown in Fig. 8. The left side of the figure is the sample document in MongoDB, and the right side is the schema of the table in MySQL. The performance experimental results are shown in Fig. 9.

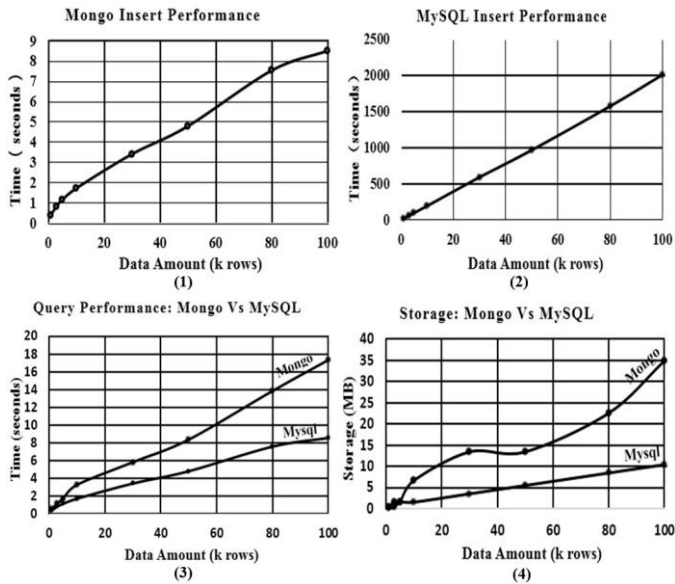Fig. 8. Data sample used in experiment

Fig. 9. Experimental results

In Fig. 9, it is very apparent from (1) and (2) that Mongo performs hundreds times faster than MySQL for insert operation. From (3), we can see that the difference of query performance between Mongo and MySQL is not obvious when dealing with small amount of data and Mongo performs a little better. However, with the growing of the data, the difference becomes bigger and bigger, which suggests that Mongo is more suitable in querying large amount of data than MySQL. The (4) shows that Mongo usually occupies two to three times more disk space than MySQL. Given all the above, it is clear that Mongo is much better in handling humongous data than relational database such as MySQL. And it has shown good performance when the data is in million level. The evaluation results partially show that our Cyber-I management system is capable of dealing with big personal data for many co-existing Cyber-Is.

VIII. Conclusion and future work

This research has been focused on the research and development of a management system for Cyber-Is and heterogeneous data. The system developed is capable of holding a large number of Cyber-Is and each Cyber-I's life cycle is completely simulated. Mongo DB has been adopted to achieve a high performance system with great scalability and flexibility for effectively managing Cyber-Is related heterogeneous personal data. A set of experimental results has shown Mongo DB's promising future in handling big personal data.

This system is only the first step in creating an environment for basic management of Cyber-Is, and the future work is to be done in the following aspects. (1) The general descriptive schemes of Cyber-I, data, models and applications should be defined. (2) In order to help other researchers to build new personalized applications in the system, more APIs and tools should be provided for easy development. (3) Three clouds of Cyber-Is, apps and models will be built so as to make the system as an open platform to share services via the clouds.

References

[1] J. Ma, L.T. Yang, B.O. Apduhan, R. Huang, L. Barolli, and M. Takizawa, "Towards a smart world and ubiquitous intelligence," International Journal of Pervasive Comp, pp. 53-68, March 2005.

[2] W. Jie, M. Kai, W. Furong, H. Benxiong, and M. Jianhua, "Cyber-I: Vision of the individuals counterpart on cyberspace," IEEE International Conference on Dependable Autonomic and Secure Computing (DASC 09), pp. 295-302, 2009.

[3] J. Ma, J. Wen, R. Huang and B. Huang, "Cyber-Individual meets brain informatics," IEEE Intelligent Systems, vol. 26, No.5, pp. 30-37, September/October 2011.

[4] J. Alejandro, N. Sebe and D. Gatica-Perez. "Human-centered computing: a multimedia perspective." In Proceedings of the 14th annual ACM International Conference on Multimedia, pp. 855-864. ACM, 2006.

[5] J. Kay, B. Kummerfeld, and P. Lauder, "Personis: A server for user models," In Proceedings of Adaptive Hypertext 2002, Springer-Verlag, pp. 203-212, 2002.

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

[7] A. Kobsa and J. Fink, "An LDAP-based user modeling server and its evaluation," User Modeling and User-Adapted Interaction, vol. 16, no. 2, pp.129–169, 2006.

[8] S. Lee, et al. "Lifelogon: A practical lifelog system for building and exploiting lifelog ontology," In Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC), 2010.

[9] R. Rawassizadeh, "A holistic multi-purpose life logging framework," Ph.D. Dissertation, University of Vienna, Vienna, Austria, 2012.

[10] E.C. Tupes and R.E. Christal, "Recurrent personality factors based on trait ratings," Technical Report ASD-TR-61-97, Lackland Air Force Base, TX: Personnel Laboratory, Air Force Systems Command, 1961.

[11] G. Zou, B. Zhang, J. Zheng, Y. Li and J. Ma, "MaaS: Model as a service in cloud computing and Cyber-I space," IEEE 12th International Conference on Computer and Information Technology (CIT), pp.1125-1130, 2012.

[12] J. Han, G. Le and J. Du, "Survey on NoSQL database," The 6th International Conference on the Pervasive Computing and Applications (ICPCA), pp.363-366, 2011.