MHA: a Life Logging Data Aggregation Platform for Interdisciplinary Healthcare Research and Collaboration

Zhikun Deng, Youbing Zhao and Feng Dong
Centre of Computer Graphic and Visualisation
Department of Computer Science and Technology,
University of Bedfordshire, Luton, United Kingdom
zhikun.deng@beds.ac.uk

Po Yang
School of Computing and Mathematical Sciences
Liverpool John Moores University Liverpool, UK
p.yang@ljmu.ac.uk

ABSTRACT
Recently, the popular use of wearable monitoring devices and mobile apps makes the effectively capture of life logging personal health data come true. The effective collection of these data in long term is beneficial to interdisciplinary healthcare research and collaboration between citizens (patients), clinicians and researchers. However, most of wearable devices and mobile apps available on market focus on a specific aspect of personal fitness and lack of the compatibility or extensibility of each other. From literatures, no platforms are reported to have successfully aggregated the heterogeneous data from different resources. This paper explores the possibility of collecting and aggregating life logging data with the integration of wearable devices, mobile apps and social media for interdisciplinary healthcare research and collaboration. It reviews existing personal health data collection techniques and identifies the requirements of implementing a life-logging data aggregator. An integrated data collection solution is designed and implemented on MyHealthAvarter (MHA) [14] platform with the integration of five well-known tools: Fitbit[4], Moves[8], Withings[7], Twitter[10] and Facebook[11]. The preliminary experiment demonstrates that the platform can successfully record, store and reuse the unified and structured personal health information in a long term, which including activities, location, exercise sleep, food, heat rate and mood. This demo receives highly positive feedbacks from interdisciplinary partners in EU VPH project MHA [14].

Categories and Subject Descriptors
J.3 [HEALTH]: H.2.8 [DATA MINING]: H.3 [INFORMATION STORAGE AND RETRIEVAL]

Keywords
Life-logging data, wearable computing, interdisciplinary

1. INTRODUCTION
Owing to the highly fragmented health system in many countries, gaining access to a consistent personal health record of individual citizens is very difficult. Recently, the popular use of wearable monitoring devices and mobile apps [1] [3] makes the effectively capture life logging personal health data come true. The effective collection of this long-term health-status information is valuable for clinical decisions and leading to strengthened interdisciplinary healthcare research and collaboration in supporting innovative medical care [2]. However, due to the commercial perspective, nearly all of well-known wearable device and mobile apps in the market focus more on personal fitness plan and lack of the compatibility and extensibility. From literatures, no platforms are reported to successfully aggregate the heterogeneous data from different resources.

In this paper, we investigate the possibility of proposing a semi-automatic life logging data aggregator with the integration of wearable devices, mobile apps and social media for general patient rehabilitation use. It firstly gives a comprehensive review of existing health data collection techniques and identifies the requirements of implementing a life-logging data aggregator. An integrated data collection solution is designed and implemented under MyHealthAvarter (MHA) [14] platform with the integration of five well-known tools: Fitbit[4], Moves[8], Withings[7], Twitter[10] and Facebook[11]. The preliminary experiment demonstrates that it can successfully record, store and reuse the unified and structured personal health information in a long term, including activities, location, exercise sleep, food, heat rate and mood. This demo receives highly positive feedbacks from interdisciplinary partners in EU VPH project MHA [14].

2. RELATED WORK
This sections reviewed existing health data collection techniques, including mobile apps based health data collection, device based health data collection, health data mining in social medial and platform based health information sharing.

Mobile applications recently are turning out to be a great source of user empowerment. The most well-known mobile apps are based on GPS information for tracking user moving activities. The type of health data includes the location, distance, speed and other manual recorded health data.

- ***Endomondo*** is a popular GPS based mobile application for tracking route, distance, duration, split times and calorie consumption. It offers a full history with previous workouts, statistics and a localized route map for each work out.

- **Google MyTracks** is also based on the use of GPS to record users' path, speed, distance and elevation while they are walk, run, and bike or do any activities outside.
- **Cardiio** is a simple touch-free heart rate monitoring mobile application. It can measure heart rate from a distance by simply having users' face straight into the front camera of smart phone for few seconds.
- **Emotionsense** is an Android application that lets user explore how they mood relates to the data that their smart phone can invisibly capture.

Table 1: Pros and Cons of health data collection technique

	Programs	*Data*	*Pros*	*Cons*
Mobile Apps	*Moves*	route, steps	Android and iOS	Moderate batter longevity
	Endomondo	route, distance, speed	community sharing, Android and iOS	Short battery longevity, not work indoor
	MyTracks	route, distance, speed	view data live, API support, only Android	Short battery longevity, not work indoor
	Cardiio	heart rate	Touch free, support iOS	Low accuracy, no API support
	Emotionsense	mood	Context-awareness by sampling strategies	No API support, only research purpose
Device	*Fitbit*	steps, calorie, food	low cost, Android and iOS, long battery life	Limited API
	Withings	steps, SpO2, sleep	reasonable cost, Android and iOS	Limited API
	Nike+	steps, calorie, food	reasonable cost, Android and iOS	Variations on accuracy
	Jawbone Up	steps, distance, calorie	reasonable cost, Android and iOS	No API
Social Media	*Twitter*	mood	Large volume of dataset	Privacy and legal issue
	Facebook	mood	Large volume of dataset	Privacy and legal issue
Platform	*HealthVault*	multiple types of data	many devices and apps, privacy and security	Focus more on medical record.
	MyFitness	multiple types of data	Real-time visualization	No API

Personal physical activity data has already been shown to be effective in various health and fitness applications. In particular, prior work has shown that wearable sensors can benefit individual patient health and individual personal fitness. We listed the most popular products as below:

- **Fitbit** [4] record steps taken, distance travelled, and calories expended. These devices communicate with a host computer using Bluetooth that sends their data directly to a user's account on the *Fitbit* website.
- **Nike+ Fuelband** is worn on the wrist and records calories, steps, distance, and Nike's own unit of activity terms "Nike Fuel". The device connects via USB to a host machine which syncs the data to a user's account on the Nike+ website.
- **Jawbone Up** is the other high-profile consumer level activity device providing steps, distance, and calories. Currently the *Jawbone up* can only be used with mobile device, drivers for laptop and PCs are not provided.

The popularity of social media allows the websites like Twitter or Facebook become communication hubs where people share life experience. This hub contains a large volume of potential personal health information.

- **Twitter:** Paper [9] explores the potential of mining twitter data to provide a tool for public health specialists and government decision makers to gauge the degree of concern (DOC) expressed in the tweets of Twitter users under the impact of diseases. Also, in paper [10], two semantic-based methods (health-related knowledge or personal relationships) for mining personal health information from Twitter are introduced.
- **Facebook:** Paper [11] reports a study of analysing emotion on the wall messages of Facebook. In their studies, three databases are built to store related emotion terms and symbols for three resources of information on Wall messages of Facebook in Chinese, English and Emoticon. The wall messages of the friends of a user can be analysed for ranking the friends of the user based on the attention index developed in this study.

Due to the great evolution of internet technology, the web-enabled healthcare service is emerging as a new healthcare delivery trend. In these web based healthcare platforms, they provide a multi-functional server for users to store, manage and visualize health data from various third party devices.

- **Microsoft HealthVault** is to enable user to gather, store, use and share personal health information though many medical devices. It enables a connected ecosystem with privacy and security-enhanced foundation including more than 300 applications and more than 80 connected health and fitness devices.
- **MyFitnessCompanion** is another healthcare platform for users to manage their personal health data, including metrics like weight, heart rate & HRV, Blood Pressure, Food intake, blood glucose, insulin, asthma, etc. The functionalities are highly similar to Microsoft HealthVault. But it has a real-time visualization mode, which can keep track and visualize all users' measurement with graph on time and share these graphs with others.

Regard the above technique investigation, it appears that personal user activities based health data is capable of being popularly collected by mobile apps. In designing a life logging data collection aggregator, some essential requirements are identified as below:

1. **Data Type:** Most data related to physical activities such as position, steps, distance and location can be automatic detected and tracked due to the mature GPS and sensor technology. These physical data can be the first identified data type in MHA platform. Some data types like Glucose and Blood sugar are hardly collected regarding the need of extra equipment. These data are probably particularly useful for some patients, but not quite necessary to general publics. They would be not included in our first version of data collection utilities solution, but can be included in the further version. Some data types like food; sleep and mood can be collected though mobile apps, but not be popular used by public. The reason is that they normally need a manual record. Also, heart rate is possibly collected without accessing extra equipment but the accuracy is not reliable. These data are supported into the first version of data collection utilities solution, but the quality and accuracy of these data have to be improved when they are used for prediction or diagnosis in MHA platform.

2. **API support and platform:** Another important issue is the availability of API and support of both iOS and Android system. For most of mobile apps or wearable device, they actually support both iOS and Android system. Google, Fitbit and Moves are all able to provide APIs for third party to retrieve health data from their server. But the preliminary requirement is that users have to register into their server firstly.
3. **Data protection and legal issue:** For the health data from social media, there are some concerns about user data protection and legal consideration. So far there are no commercial applications or platforms practically implementing the extraction. Also, the deep analysis of plain text from social media requires some ontology and semantic techniques. Consequently, in this deliverable, we would design and implement the linking to Twitter and Facebook, but not practically release it until the further clear clarification of data protection and legal issues.

3. Architecture Design and Implementation

3.1 Design

This section proposed an architecture design of life logging data aggregator in MHA project as shown in Figure.1.

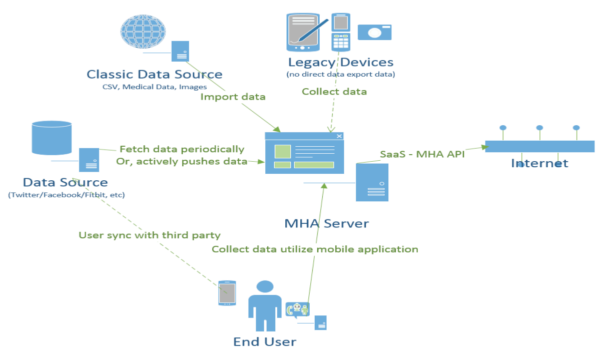

Figure 1 Overview of Data Aggregator in MHA

It is designed in a way that it works with various health related data sources including but not limited to: wearable devices, mobile application, legacy devices and classic datasets. MHA not only collect useful information of users in a lifetime style, it also provides this information to users' other application / services through its API.

MHA acts as a central hub and interacts with other data services automatically or semi-automatically with users' manual assistant. Different data sources can be categories to:

- **Classic Data Source**: this is the data which users can feed in MHA DCU by uploading files, which includes hospital records, clinical trial spreadsheet, CT scan images.
- **Legacy Device**: the data from legacy devices are displayed directly to end-users without the ability to export electronically format. DCU provide easy to use form for user to manually input the data with timestamp to MHA.
- **Mobile Application**: DCU has adapter and interface which interact with popular health related mobile applications, this allows the data automatically collected and stored to MHA central database.
- **Wearable Device**: modern health monitoring device like Fitbit and Nike FuelBand+ SE. if the device provide API for DCU to interact with, the users' data will be collected from the devices' server through their defined security mechanism.
- **Social Network**: User's status update to social networks could include valuable information related to user's location, activity, food, etc. The health related information can be retrieved by DCU for analysis and pass to MHA.

3.2 Implementation

The goal of data collection is to gather as much and appropriate as possible health related information, which is then used to study the correlation of data and combination of different data source which generates synergy for monitor, analyse and visualize participant's physical and emotional health conditions. The internal volunteers are equipped with hardware devices and mobile applications for the real-life data collections, also their social networks accounts are authorized to the platform as social data sources. The hardware, software and social networks include: 1) Fitbit Flex 2) Withings Pulse Ox 3) Moves on Android and iOS 4) Twitter 5) Facebook 6) Google+ 7) Manual input from web UI.

In the sections below we proposed an architecture design of life logging data aggregator in MHA project as shown in Figure.1. It is designed in a way that it works with various health related data sources including but not limited to: wearable devices, mobile application, legacy devices and classic datasets. MHA not only collect useful information of users in a lifetime style, it also provides this information to users' other application / services through its API.

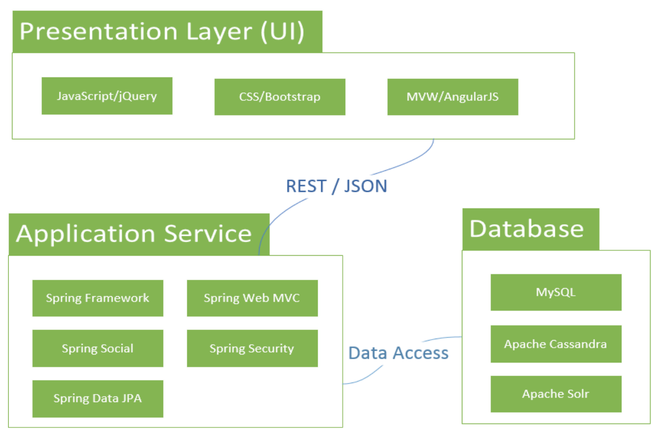

Figure 2

The implementation backend is based on Java programming language and Spring Framework technology stack, the frontend UI is mainly based on Twitter Bootstrap, jQuery and AngularJS. Java runs on all major platforms include Windows, Linux and Mac OS, which makes module based on Java is generally more portable between OSes, also JVM's proven high performance is crucial for our potential large user base. Spring Framework is the de-facto standard in enterprise Java programming, the developer team's high experience in Spring Framework mkes it our pick on implement the project.

Frontend is a relatively separate Grunt project which uses Bower to manage the frontend JavaScript dependancies. The frontend soruce code is unit tested and built on a conituous integration manner. Twitter Bootstrap 3 is used to support both mobile and desktop browser, mobile first responsive design is applied for the web UI. AngularJS and jQuery JavaScript libraries are used to help shape the frontend logic, and interact with the REST service provided by backend. The combination of cutting-edge frontend technology stack does give us a enjoyable developing experience and a great dynamic user-friendly UI.

The project is hosted on Ubuntu 12.04 LTS VPS, Apache is used to server static content while Tomcat 7 is reverse proxied to server the dyamic contents. A HAProxy server is configured to work as load balancer and Radis server is tested to be the cache server for scale to larger user groups.

4. Experimental Results and Discussion

The proposed geometrical correction method has been evaluated using a diverse and representative sample of 10 arbitrarily warped historical document images with complex layout from the IMPACT project dataset [13]. The evaluation methodology used in this paper is based on supervised evaluation with (manually created) ground-truth data. Baselines on both the original warped document image and result document image are marked manually, and then the quality of baselines is compared against the manually entered baselines. The results of two additional geometric correct methods are compared: one is a state-of the-art page-curl correction method designed for IMPACT by NCSR [5] and the leading commercial product Book Restorer™ [14].

5. CONCLUSION

In this paper, a novel grid-based geometrical correction method is proposed for arbitrarily warped historical documents. Our experimental results show that this method can outperform the leading start-of-the-art geometric correction methods on dealing with arbitrary warping effect. Another advantage is that this method can process document images with complex contents, such as graphics, multiple font sizes and multiple column layouts. The method in its current state is designed for book images (as the other methods in the literature), albeit with complex layouts; it has not been evaluated yet on more challenging historical documents with highly dense and intricate layouts, such as newspapers. Future work will focus on this problem.

6. REFERENCES

[1] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, I. Korhonen, “Activity classification using realistic data from wearable sensors” *IEEE Trans. Info. Tech. Bio, vol 10, issue 1, pp. 119-128, 2006*

[2] G. Burdea, V. Popescu, V. Hentz, and K. Colbert “Virtual reality-based orthopedic telerehabilitation” *IEEE Trans. Rehab. Eng., vol. 8, pp. 430-432, 2000*

[3] K. Arai ”Wearable physical and psychological health monitoring system” *2013 International Conference on Science and Information (SAI)*, pp. 133-138, 2010

[4] Fitbit. http://www.Fitbit.com (accessed July 6, 2015)

[5] Nike+ Fuelband. http://www.nike.com/us/en_us/c/nikeplus-fuelband (accessed July 4, 2015)

[6] JawBone Up. https://jawbone.com/up (accessed July 6, 2015)

[7] Withings. http://www2.withings.com/uk/en (accessed July, 16, 2015)

[8] Moves. http://www.moves-app.com/ (accessed July 2, 2015)

[9] F. F. Guo, Y. Li, M. S. Kankanhalli and M. S. Brown, An evaluation of wearable activity monitoring devices. *Proceedings of the 1st ACM international workshop on personal data meets distributed multimedia*, 2013 Oct, pp 31-34.

[10] Sokolova M, Jafer Y, Schramm D, Text mining for personal health information on Twitter. *IEEE Second International Conference on Healthcare Informatics, Imaging and System Biology*. 2012 Sep;pp112-118.

[11] J. Xiang, A. C. Soon, J. Geller, Monitoring Public Health Concerns Using Twitter Sentiment Classification. *IEEE 2013 International Conference on Healthcare Informatics*. 2013 Sep;pp335-344.

[12] H.H. Hsu, Y. F. Chen, C.Y. Lin, C. W. Hsieh, T. K. Shih Emotion Care Services with Facebook Wall Messages. 2012 *26th International Conference on Advanced Information Networking and Applications Workshops*, 2012 Mar;pp875-880.

[13] HealthVault. https://jawbone.com/up (accessed Jan 6, 2014)

[14] MHA: MyHealthAvarter