

# Secure Mobile Automation of Ecological Momentary Assessments (EMA) For Structured Querying

Nikhil Yadav<sup>1,\*</sup>, Mehrdad Aliasgari<sup>2</sup>, Christopher Azzara<sup>1</sup>, Fazel Keshtkar<sup>1</sup>

<sup>1</sup>Division of Computer Science, Mathematics and Science, St. John's University, Queens, New York

<sup>2</sup>Department of Computer Engineering and Computer Science, California State University, Long Beach, California

## Abstract

The ubiquitous nature of mobile devices like smartphones and tablets make them ideal platforms for engaging users in Ecological Momentary Assessments (EMA). In EMA, participants are repeatedly assessed frequently (daily or multiple times per day) through a set of questionnaires. Fluctuations in psychological states, such as cognition and affect can be recorded in real time using mobile devices. EMA results can further be coupled with other physiological sensor data procured through wearables and smartphones, to validate and correlate patient experiences and responses to certain treatments and medications. This can be useful for health care organizations which are interested in the impact of their treatment techniques on patient populations. In this paper, we present an EMA platform developed using Android mobile devices. The collected results are shown and techniques used to query the data are demonstrated. The platform is flexible and can scale up to perform data mining algorithms for sentiment analysis based on the stimulus to a medication or treatment over a prolonged period of time.

Received on 17 January 2018; accepted on 21 March 2018; published on 23 March 2018

**Keywords:** Mobile health, data collection, health surveys

Copyright © 2018 Nikhil Yadav *et al.*, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi:10.4108/eai.23-3-2018.154373

## 1. Introduction

In recent times there has been a prolific emergence of health and wellness mobile applications using wearables and smartphones. There has been an increase in sales for such devices[1]. In addition, there has been an increased adoption of applications performing personal health monitoring of sleep patterns, calories burned, heart rate, etc., using sensors available on mobile devices [2][3]. Though the use of Ecological Momentary Assessments (EMA) while performing physical activities have been studied before, they rely on very targeted and specific sensor technologies with limited subjects [4][5]. The ease of developing applications on mobile devices has meant that it is much more easier to create and deploy EMA studies on a large scale, as opposed to traditional paper and pencil

methods[6]. Levels of engagement for users on devices such as smartphones are typically very high, meaning that they are ideal platforms for high frequency EMA studies [7][8]. Recent work has looked at coupling EMA and sensor data to assess physiological states leading to an abundance of data for the interest of the medical and social science communities [9][10]. Systems such as *psychlog* [11] have been developed for psychological research. Additional work has also been done on performing pattern identification in automating EMA and ESM (Experience Sampling Methods) to classify certain psychological states like stress [12]. Various previous works also focus on how to fuse data from sensors and EMA to arrive at predictions of psychological states[13].

In this paper we focus on the task of building an analytics framework leveraging EMA like surveys on Android tablet devices. The survey we choose focuses specifically on the impact of a drug on a user's mental state using a point based system. The novelty of the work lies in the user experience during the EMA assessment while ensuring the security of the

\*Please ensure that you use the most up to date class file, available from EAI at <http://doc.eai.eu/publications/transactions/latex/>

\*Corresponding author. Email: [yadavn@stjohns.edu](mailto:yadavn@stjohns.edu)

recorded EMA data at rest and in transit. The system allows seamless data storage in a structured format making it suitable for preliminary data mining tasks. Supervised and unsupervised learning to correlate psychological and physiological data is also possible by integrating physiological sensor hardware recording indicators like blood pressure, blood sugar, and heart rate. The current implementation allows a user to record a baseline state assumed to be unaltered by substance usage. Subsequent EMA surveys look at changes in psychological states from the baseline. The data collected over time and analyzed in a human subject study, is valuable for health care organizations studying drug treatment impacts.

This paper is organized as follows: Section 2 gives an overview of the tools and methods used to automate the system using mobile technology. In Section 3 we show results from using the system and how the survey data is organized, stored, and analyzed. We conclude in Section 4, by discussing current drawbacks and future work.

## 2. Implementation

A functional prototype of the automated mobile EMA data collection system (tentatively titled *SJU mHealth*) was implemented with sample runs using mock users. The system allows data about the nature of the disease to be entered together with general information about a patient. This is stored in a patient’s personal health profile. Collected data is stored following stringent security measures and conforms to the current HIPAA (Health Information Portability and Accountability Act) standards. No PHI (personal health information) is stored which can allow for re-identification of a patient.

### 2.1. Mobile platform

Samsung Galaxy Tabs running Android jelly bean (4.2) were used for this project. Android provides multiple open source libraries easily integrable into the main code base, and does not require the tedious software license agreements of other platforms which make it ideal for our prototype.

### 2.2. Data Collected

**Login Information.** The current system uses a standard registration page which captures the first and last name of a user together with a user name and password [Fig. 1]. The names and passwords are stored using the MD5 hash enforced with SALT on a back end cloud service. This ensures uniqueness in the case of duplicate names and passwords. By so doing, we ensure that a user cannot be re-identified. Though the system is not being tested with real human subjects presently, enforcing this requirement was a major concern to ensure HIPAA

compliance in the event that a large scale clinical trial is rolled out for target health organizations.

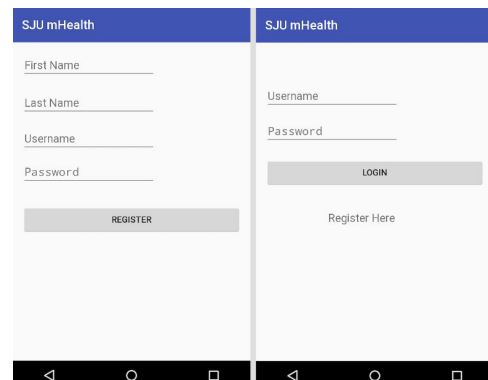
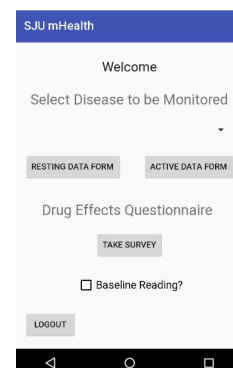
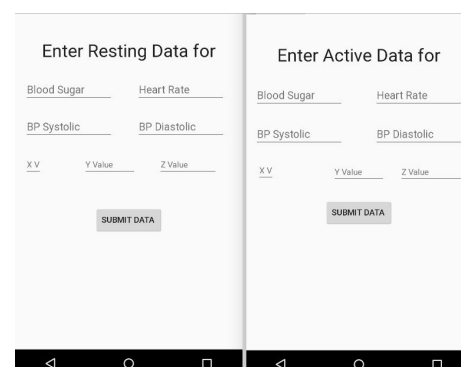


Figure 1. Registration and Login

**Disease and Survey Type.** The application allows a user to select a disease or condition they are being treated for. It also allows the survey being taken to be flagged as either a baseline or post base-line recording as shown in Fig. 2(a).



(a) Disease and Survey Type



(b) Manual Physiological Data Entry

Figure 2. Screens for Disease and Physiological Data Entry

**Physiological and EMA data.** Though the mobile application has no physical integration with external

physiological sensor devices, screens allowing manual entry of parameters such as blood pressure, blood sugar, and heart rate have been provided. These can be instantiated either when the user is at rest, or has been active as shown in Fig. 2(b). The level of physical activity is important to capture as it may influence any physiological indicators being monitored.

A simplified EMA like survey was implemented following the techniques and questions of some existing tests like the PHQ-9 (Patient Health Questionnaire), D.I.R.E. (Diagnosis, Intractability, Risk, Efficacy), P.A.D.T. (Pain Assessment and Documentation Tool) [4][5]. The mobile survey is shown in Fig. 3.

**Figure 3.** Ecological Momentary Assessment for Medication Efficacy

Slider user interface elements calibrated for a score between 1 (not at all) to 10 (extremely) are used. The questions asked pertain to different effects of the medication or drug being administered. These are subjective and specifically focus on whether the user currently *feels* the effect of a drug. Based on the strength of the drug the user could experience negative effects like nausea, fatigue, drowsiness, etc. which they *dislike*. They could also experience more positive effects like high energy levels, less pain, mental alertness, etc. which they *like*. This is also the stimulus for whether they want more of the medication drug or not. The survey effectively captures all these aspects, and can easily be expanded or replaced by researchers.

### 2.3. Cloud Service

The application communicates with the Amazon web service cloud and all data collected from the application is securely stored on servers hosted there. Amazon cloud servers provide the sufficient flexibility needed by the system allowing for the ability to store data in a structured format. The ease of setting up analytical

tools on these servers in anticipation of future data mining tasks was also a motivation to choose them.

### 2.4. Data Flow and Security

We utilize the latest version of Transport Layer Security (TLS) protocol to provide the highest level of protection against the outsider adversaries. TLS guarantees that the connection between the mobile application and the server is secure against any outsider adversary.

The data collected on the mobile devices are encrypted prior to being stored. Each record is encrypted with a fresh new AES 256-bit key. In addition, a hash-based message authentication code (HMAC) of each encrypted data is appended to records prior to storage. The keys are themselves encrypted with a master key. In other words, for a record  $r$ , the following data structure is stored in the mobile application's database:  $\langle c = \text{Enc}_{k_e}(r), t = \text{hmac}_{k_i}(c), \text{Enc}_{k_m}(k_e, k_i) \rangle$ . Note that the encryption key ( $k_e$ ) and the message integrity key ( $k_i$ ) are generated anew for each record. The master key is an AES 256-bit encryption key that is generated when the application is installed. This key is securely stored in a secure storage service provided by the operating system (e.g., Keychain in iOS). The above scheme ensures confidentiality and integrity of user data collected on a mobile device. In order for an insider mobile adversary to compromise the security of user data, it is required to compromise the operating system and access the encryption master key ( $k_m$ ) along with all encrypted records. However, such a scenario is extremely unlikely for operating systems such as Android and iOS.

Similar measures are taken on the server side. In addition, a strict access control policy is implemented on the database of the server to prevent any unauthorized access to data. With this mechanism, sensitive data is always stored encrypted and access is limited only to those entities that require them. Fig.4 shows the data integrity mechanisms just described.

## 3. Results

### 3.1. Structured Data Storage

The format of the survey data collected and retrievable from the server is structured and shown in Table 1. The *User* column is the MD5 hash of the full name of the user (truncated for display purposes). *EMA1* to *EMA5* correspond to the current five EMA questions on the mobile survey application, scored between 1 and 10. *PHY1* corresponds to a physiological indicator like heart rate. These can always be expanded based on the intention of the researcher using the system, and the availability of more physiological sensors. *BSL* in the table indicates whether the data is a baseline reading or not, and *ACT* indicates whether the data has been taken

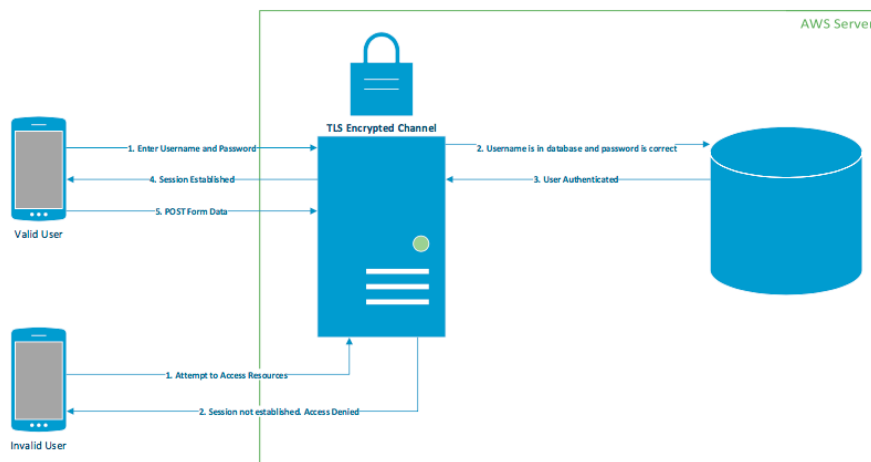


Figure 4. Data Transmission

when the subject is in an active state, e.g., after some physical exercise.

Table 1. Structured EMA and Physiological Data

User	EMA1	...	EMA5	PHY1	BSL	ACT
baa*	10	...	8	5	T	F
Cod*	5	...	1	3	T	T
baa*	10	...	8	5	F	F
Cod*	7	...	1	3	F	T
Dee*	6	...	2	5	T	T
Dee*	7	...	2	6	F	T
Sa0*	10	...	6	4	T	F

A separate table contains demographic information about the user such as age, ethnicity, gender, etc., together with the disease they are being treated for and the treatment drug being administered. The time and date when the EMA and physiological data is collected is also recorded. By performing joins of the tables appropriate structured SQL queries can be used to return results.

### 3.2. Structured Querying of EMA and physiological data

There are various queries that a user can perform using the structured data collected. Some samples in SQL syntax are shown below:

1. *SELECT User, PHY1 from EMA\_PHY\_TABLE, where EMA2<5:* Returns the physiological indicator, e.g., heart rate, for all users who do not want more medication drugs.

2. *SELECT EMA\* from EMA\_PHY\_TABLE, where BSL<> 'T':* Returns the momentary assessment scores which are not baselines.

More complex queries can also be constructed using the structure described.

### 4. Conclusion and Future work

By implementing this system we have laid the foundation for creating a scalable data mining platform which can perform correlational data analysis between physiological and psychological (EMA, ESP) data. Future work will focus on integrating physiological sensor hardware with the mobile devices, and develop an API for researchers and health practitioners to modify (or replace) EMA surveys and analyze the data collected over time using data mining techniques. The self reporting nature of such surveys means that a user has to actively participate to log data. This is likely to pose a challenge if the application is not engaging. However, the popularity of recent social media applications on the smartphone have shown that getting users actively engaged is a possibility with good design, gamification, and the assurance of security.

### References

- [1] International Data Corporation (IDC), *Worldwide Quarterly Wearable Device Tracker*, Jun 18th 2015. (<http://www.idc.com/getdoc.jsp?containerId=prUS41100116>) (accessed October 15, 2015).
- [2] T. Ferguson, A. V. Rowlands, T. Olds, C. Maher, *The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study*, Int. J. Behav. Nutr. Phys. Act. 12 (1) (2015) 1-9.

- [3] J.-H. Park, D.-G. Jang, J. Park, S.-K. Youm, *Wearable sensing of in-ear pressure for heart rate monitoring with a piezoelectric sensor*, *Sensors* 15 (9) (2015) 23 402-23 417.
- [4] S. Shiffman, A. A. Stone, *Ecological momentary assessment: a new tool for behavioral medicine research*, *Technol. Methods Behav. Med.* (July) (1998), 117-131.
- [5] M. Aan het Rot, IC Hogenelst, R. A. Schoevers, *Mood disorders in everyday life: a systematic review of experience sampling and ecological momentary assessment studies*, *Clin. Psychol. Rev.* 32 (6) (2012) 510-523.
- [6] T. J. Trull, U. W. Ebner-Priemer, *Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section*, *Psychol. Assess.* 21 (4) (2009) 457-462.
- [7] S. S. Intille, *Technological innovations enabling automatic, context-sensitive ecological momentary assessment*, in: S. Shiffman, A.A. Stone, A.A. Atienza, L. Nebeling (Eds.), *The Science of Real-time Data Capture*, Oxford University Press, New York, NY, 2007, pp. 308-337.
- [8] S. Kumar, W.J. Nilsen, A. Abernethy, A. Atienza, K. Patrick, M. Pavel, W.T. Riley, A. Shar, B. Spring, D. Spruijt-Metz, D. Hedeker, V. Honavar, R. Kravitz, R. Craig Lefebvre, D.C. Mohr, S.A. Murphy, C. Quinn, V. Shusterman, D. Swendeman, *Mobile health technology evaluation: the mHealth evidence workshop*, *Am. J. Preventive Med.* 45 (2) (2013) 228-236.
- [9] R. M. Chang, R. J. Kauffman, Y. Kwon, *Understanding the paradigm shift to computational social science in the presence of big data*, *Decis. Support Syst.* 63 (1) (2014) 67-80.
- [10] A. Markowetz, K. Blaszkiewicz, C. Montag, C. Switala, T.E. Schlaepfer, *Psychoinformatics: big data shaping modern psychometrics*, *Med. Hypotheses* 82 (4) (2014) 405-411.
- [11] A. Gaggioli, P. Cipresso, S. Serino, G. Pioggia, G. Tartarisco, G. Baldus, D. Corda, G. Riva, *An open source mobile platform for psychophysiological self tracking*, *Stud. Health Technol. Inform.* 173 (January) (2012) 136-138.
- [12] Y. Shi, M. Nguyen, P. Blitz, B. French, *Personalized stress detection from physiological measurements*, in: *International Symposium on Quality of Life Technology*, Las Vegas, NV, September 2015, 2015.
- [13] Ilumivu, *Mobile EMA*, 2016 (<https://ilumivu.com/solutions/> (accessed February 21, 2016)).
- [14] P. C. M. Molenaar, *A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever*, *Meas.: Interdiscip. Res. Perspect.* 2 (4) (2004) 201-218.
- [15] E. L. Hamaker, *Why researchers should think 'within-person': a paradigmatic rationale*, in: *Handbook of Research Methods for Studying Daily Life*, Guilford Publications, New York, NY, 2012, pp. 43-61 (Chapter 3).
- [16] L. Van der Krieke, B. F. Jeronimus, F. J. Blaauw, R.B. Wanders, A. C. Emerencia, H. M. Schenk, S. De Vos, E. Snippe, M. Wichers, J. T. Wigman, E. H. Bos, K. J. Wardenaar, P. De Jonge, *HowNutsAreTheDutch (HoeGekIsNL): a crowdsourcing study of mental symptoms and strengths*, *Int. J. Methods Psychiatr. Res.* 25 (2) (2016) 123-144.
- [17] M. Almalki, K. Gray, F. Sanchez, *The use of self-quantification systems for personal health information: big data management activities and prospects*, *Health Inform. Sci. Syst.* 3 (Suppl. 1) (2015).