

# BeWell: A Smartphone Application to Monitor, Model and Promote Wellbeing

Nicholas D. Lane<sup>†</sup>, Mashfiqui Mohammad<sup>†</sup>, Mu Lin<sup>†</sup>, Xiaochao Yang<sup>†</sup>, Hong Lu<sup>†</sup>, Shahid Ali<sup>\*</sup>  
Afsaneh Doryab<sup>\*\*</sup>, Ethan Berke<sup>‡</sup>, Tanzeem Choudhury<sup>†</sup>, Andrew T. Campbell<sup>†</sup>

<sup>†</sup>Computer Science Department Dartmouth College, <sup>\*</sup>Dartmouth Medical School

<sup>‡</sup>Dartmouth Institute for Health Policy and Clinical Practice, <sup>\*\*</sup>IT University of Copenhagen

{niclane, hong, campbell}@cs.dartmouth.edu, {adoyrab}@itu.dk, {ethan.berke}@tdi.dartmouth.edu

{mohammad.mashfiqui.r.shuvo, mu.lin, xiaochao.yang, shahid.a.ali, tanzeem.choudhury}@dartmouth.edu

**Abstract**—A key challenge for mobile health is to develop new technology that can assist individuals in maintaining a healthy lifestyle by keeping track of their everyday behaviors. Smartphones embedded with a wide variety of sensors are enabling a new generation of personal health applications that can actively monitor, model and promote wellbeing. Automated wellbeing tracking systems available so far have focused on physical fitness and sleep and often require external non-phone based sensors. In this work, we take a step towards a more comprehensive smartphone based system that can track activities that impact physical, social, and mental wellbeing namely, sleep, physical activity, and social interactions and provides intelligent feedback to promote better health. We present the design, implementation and evaluation of BeWell, an *automated wellbeing app* for the Android smartphones and demonstrate its feasibility in monitoring multi-dimensional wellbeing. By providing a more complete picture of health, BeWell has the potential to empower individuals to improve their overall wellbeing and identify any early signs of decline.

## I. INTRODUCTION

Our lifestyle choices have a deep impact on our personal health. For example, our sleep, socialization and exercise patterns are connected to the presence of a wide range of health related problems such as, high-blood pressure, stress [1], anxiety, diabetes and depression [2], [3]. Positive health effects can be observed when these wellbeing indicators (e.g., sleep, physical activity) are kept in healthy ranges. However, people are typically not exposed to these health indicators as they go about their daily lives. As a result, unbalanced unhealthy lifestyles are present in the general population. People demonstrate concern for some aspects of their wellbeing, such as fitness or diet, yet neglect the wellbeing implications of other behaviors, such as, poor sleep, hygiene or prolonged social isolation. We believe this situation is caused by an absence of adequate tools for effective self-management of overall wellbeing and health.

We envision a new class of *personal wellbeing applications* for smartphones capable of monitoring multiple dimensions of human behavior, encompassing physical, mental and social dimensions of wellbeing. An important enabler of this vision are the recent advances in smartphones, which are equipped with powerful embedded sensors, such as an accelerometer, digital compass, gyroscope, GPS, microphone, and camera.

Smartphones present a programable platform for monitoring wellbeing as people go about their lives [4]. It is now possible to infer a range of behaviors on the phone in real-time, allowing users to receive feedback in response to everyday lifestyle choices that enables them to better manage their health. In addition, the popularity of smartphone application stores (e.g., the Apple App Store, Android Market) has opened an effective software delivery channel whereby a wellbeing application can be installed in seconds, further lowering the barrier to user adoption. We believe production-quality wellbeing applications will gain rapid adoption globally, driven by: i) near zero user effort, due to automated sensor based activity inference and ii) universal access, only requiring a single download from a mobile phone application store and installation on an off-the-shelf smartphone.

A number of technical barriers need to be tackled to make this vision a reality. The majority of existing mobile health systems have focused on a specific health dimension (e.g., stress [5], diet [6], physical activity [7]) and consider only one or two types of behavior. Instead, personal management of wellbeing requires applications that monitor a diverse range of daily behaviors, which have broad health related consequences. Unfortunately the small, but growing, number of mobile health applications that consider a wider perspective of user health commonly rely on manual data entry [6]. This type of manual effort is burdensome to users and is unlikely to scale to mass adoption. Rather, the automated and continuous inference of the users behavior and environment based on embedded sensors promotes sustained usage over the long-term. Ultimately, to be effective these applications must understand the impact on the health and wellbeing of the user due to the observed behavioral patterns. Simply reporting behavioral patterns to the user is not necessarily intuitive. It can be difficult for a user to identify which behaviors have a larger impact on wellbeing. Therefore, providing intuitive and interpretable feedback is a key challenge for *wellbeing monitoring apps*.

In this paper, we present BeWell, a personal health application for smartphones that is designed specifically to help people manage their overall wellbeing. BeWell continuously monitors multiple dimensions of behavior and incorporates

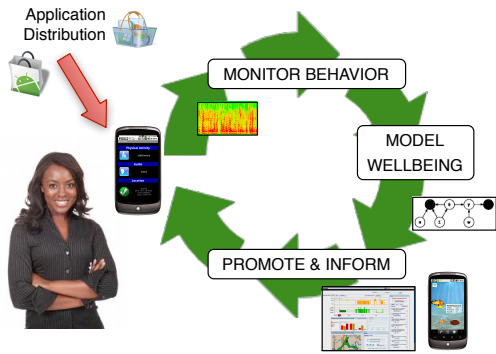


Fig. 1. BeWell approaches end-user self-management of wellbeing with three distinct phases. Initially, everyday behaviors are automatically monitored. Next, the impact of these lifestyle choices on overall personal health is quantified using a model of wellbeing. Finally, the computed wellbeing assessment drives feedback designed to promote and inform improved health levels.

user feedback mechanisms that are able to increase awareness of how different aspects of lifestyle impact the personal wellbeing of the user. BeWell uses commercial, off-the-shelf smartphones (Android Nexus One [8]) to automatically: (i) monitor a person’s physical activity, social interaction and sleep patterns; (ii) summarize the effect of the monitored behavior on wellbeing; and (iii) provide feedback that enables users to effectively manage these three key aspects of their health.

The structure of the paper is as follows. Section II discusses the BeWell architecture and Section III describes the design of our wellbeing model. Sections IV and V present the detailed implementation and evaluation of BeWell, respectively. Section VI discusses the limitations of existing mobile health systems followed by the concluding remarks presented in Section VII.

## II. BEWELL ARCHITECTURAL DESIGN

In this section, we describe the BeWell architecture which includes cloud computing servers as well as smartphones. The operational phases of the BeWell system are shown in Figure 1 and discussed below.

### A. Monitor Behavior

The BeWell application automatically infers behavioral patterns using sensors embedded in smartphones. Rather than tracking a single wellbeing dimension, such as, physical activity, the BeWell system concurrently monitors multiple dimensions (e.g., sleep patterns, social interaction, and physical activity), representing a more complete picture of the user’s overall wellbeing. The current implementation of BeWell is limited to three wellbeing dimensions and does not yet incorporate a number of other important health components such as diet and stress. Users are able to manually add new behaviors by using a diary maintained by the BeWell web portal, which visually show the inferences made by the application and allows users to correct classification errors.

### B. Model Wellbeing

Simply collecting user patterns of behavior is insufficient. Users need to easily understand how their behavior affect different dimensions of their personal wellbeing. Using existing guidelines provided by healthcare professionals, we estimate multi-dimensional wellbeing scores that capture the relationship between behavioral patterns and health outcomes.

### C. Promote and Inform End Users

Armed with the ability to track changes in overall wellbeing, BeWell is able to present users with richer information and make them self-aware of their current wellbeing. BeWell presents information in two ways: i) directly, when the user interacts with the BeWell application installed on their smartphone or from a desktop using the BeWell web portal; and, ii) passively, using an ambient display rendered on the smartphone wallpaper, which is visible as the user performs typical phone operations (e.g., making a call, texting, etc.).

## III. MONITORING AND MODELING WELLBEING

A variety of health outcomes are tightly linked to everyday decisions involving sleep [9], diet [10], exercise [11] and socialization [3] patterns. The BeWell monitoring stage involves sensor-based inferences of user activities (e.g., sleeping, exercising, socialization). The wellbeing model interprets these behavioral patterns and estimates a multi-dimensional wellbeing score to make it easier for the user to understand the impact on overall wellbeing. Behavioral patterns and estimates of wellbeing are used to generate user feedback that are designed to inform users of their current wellbeing, highlighting the behavior changes needed to improve this state.

Although there are still many unanswered questions about the links between behavior and wellbeing, in BeWell we take a pragmatic approach and build an initial wellbeing model, which can be extended and refined.

Under our proof-of-concept model wellbeing is assessed for the three behaviors independently, with three simple wellbeing scores produced that ranges between 0 and 100. Every score seeks to summarize the impact of recent user behavior on overall wellbeing for just one of the three behaviors monitored. A score of 100 indicates the person is matching the accepted guidelines of performance for that behavior (e.g., averaging more than 8 hours sleep per day). Scores of 0 indicate the individual is not even attaining minimum recommended patterns of behavior. The score for each dimension is computed using an exponentially weighted average of daily scores along each dimension. In what follows, we describe the importance of these three behaviors to overall wellbeing and explain how the scores, for individual days, are computed.

### A. Sleep

A body of literature exists that links sleep hygiene to a range of mental and physiological health conditions, including, affective disorders, hypertension, heart disease and the development of diabetes [9]. However, poor sleep behavior (e.g., chronic lack of sleep, erratic sleep cycles) are wide spread

across the general population. People commonly exchange sleep for additional waking hours as a coping mechanism for busy lifestyles.

Research exploring sleep health effects focus both on the quantity and quality of sleep [12]. Although both of these facets are important we focus solely on monitoring sleep duration. We take a simple approach that approximates the sleep duration of users by measuring mobile phone usage patterns, as discussed in Section V-B. Studies show oversleeping (“long sleep”) carries similar negative health consequences to insufficient sleep [9], thus, we penalize both behaviors. BeWell computes a wellbeing score for sleep behavior within a single day using a gaussian function,

$$sleep_{day}(HR_{act}) = Ae^{-\frac{(HR_{act} - HR_{ideal})^2}{2(HR_{hi} - HR_{lo})^2}}$$

in which  $HR_{act}$  is the total quantity of sleep over a 24 hour period,  $HR_{ideal}$  is the ideal hours asleep with  $HR_{hi}$  and  $HR_{lo}$  being the upper and lower limits of acceptable sleep duration. Our sleep function is parameterized using a  $HR_{ideal}$  of 7 hours with a  $HR_{hi}$  of 9 hours and  $HR_{lo}$  of 5 hours, these values are consistent with existing sleep studies [9].

### B. Physical Activity

Benefits of physical activity, such as, lower mortality rates and cardiovascular disease are well-known but health benefits also extend to cancer and sexual dysfunction [11]. Further, a number of studies have linked exercise to improved depression, self-esteem, mood, sleep, and stress [13], [2].

Our automated wellbeing assessment of physical activity begins with common categories of end-user activity (e.g., walking, stationary, running) being recognized by smartphone sensors. The duration the user spends performing these activities is computed, allowing the Physical Activities Compendium metabolic equivalent of task (MET) value [14] to be estimated each day. Being definitive as to the ideal MET levels for an individual is difficult as mental and physical health benefits occur at different levels of activity. These MET levels are also sensitive to user characteristics, such as, existing physical fitness or particular genetic determinants. We currently rely on generic guidelines established by the Centers for Disease Control and Prevention [15] (CDC). Our daily scores of physical activity are simply a linear regression,

$$physical_{day}(MET_{act}) = (MET_{hi} - MET_{lo})MET_{act} + MET_{lo}$$

where  $MET_{act}$  is the actual MET value for a user during that day, with  $MET_{hi}$  and  $MET_{lo}$  being calibrated by the high-end and minimum guidelines for adult aerobic activity set by the CDC. These values range between 300 and 150 minutes of moderate-intensity per week. Such aerobic activity should be accompanied by muscle-strengthening programs, ideal behavioral patterns for these programs are also available from the CDC and are included within the existing physical activity guidelines. However, we neglect this aspect of physical activity due to the inaccuracy in monitoring muscle-strengthening programs without specialized sensors (e.g., on-body sensors).

Currently, BeWell users enter this behavior manually (see Section IV-D). In the future, we will study alternatives such as using coarse inferences based on location or sound.

### C. Social Interaction

The daily social interactions of people have been shown to have impact on many dimensions of wellbeing. The connection between the availability of social support and psychological wellbeing is well established, with low levels being linked to symptoms of depression [3]. Individuals who maintain dense social connections are more likely to have resilient mental health. They tend to be able to cope with stress and often are better able to manage chronic illness.

Medical studies use a variety of measures to capture the social environment of a person. The development of these measures are still an active area of research. BeWell focuses on one of these metrics, social isolation, as it is more easily captured with sensors available in smartphones today. Studies of particular high-risk communities show social isolation is correlated with basic forms of human contact. For example, health deterioration exhibited in the elderly is linked with, amongst others, a decline in the frequency of human interaction (e.g., phone calls and visits with friends and relatives) [16]. In the general population, those with profound acquired hearing loss have been seen to suffer a deterioration of psychological wellbeing due to the associated communication difficulties [17]. We measure social isolation based on the total duration of ambient conversations, which are detected by inferences made using the mobile phone microphone. Insufficient medical evidence exists to parameterize this relationship. At this time we again use a wellbeing score for social interaction with a linear regression,

$$social_{day}(DUR_{act}) = (DUR_{hi} - DUR_{lo})DUR_{act} + DUR_{lo}$$

where  $DUR_{act}$  is the duration of conversation detected relative to the total time the microphone is active during a single day. We determine empirically a value for  $DUR_{hi}$ , 0.35, using the mean conversation ratio of a small 10 person experiment; we also utilize this group to train our classifiers (see Section V). As we lack a population in which poor wellbeing has caused atypical conversation patterns our  $DUR_{lo}$  ratio is simply set to zero.

## IV. IMPLEMENTATION

In this section, we present the BeWell implementation based on the architecture discussed in Section II. As illustrated in Figure 2 the BeWell application and system support consists of a software suite running on Android smartphones and cloud infrastructure. The software components installed on the phones include: i) Sensing Daemon, which is responsible for sensing, classification, data processing (e.g., privacy preserving audio processing) and uploading of sensor data; ii) Mobile BeWell Portal, which displays the user’s behavioral patterns and the wellbeing score associated with these behaviors; and iii) Mobile Ambient Wellbeing Display, which provides an always-on visualization of the user’s wellbeing scores. The

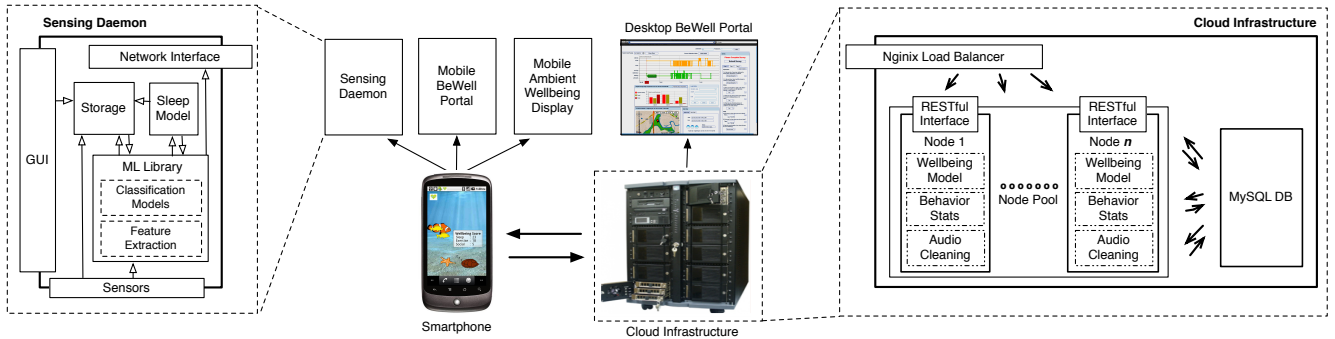


Fig. 2. BeWell implementation, including smartphone components supported by a scalable cloud system

BeWell Cloud Infrastructure provides storage, computation, and web access to user data via the Desktop BeWell Portal.

### A. Sensing Daemon

The BeWell Sensing Daemon combines an internally developed platform-independent machine learning C library with device specific components, written in Java, that are responsible for communication, storage and the user interface. These core components are connected with a JNI bridge. The control flow of the daemon is comprised of two independently operating processes. The first process, the classification pipeline is responsible for the inference of end-user behavior. The pipeline continuously samples the phone sensors and extracts features used by classification models, which also run on the phone. The second process, asynchronously transfers inferences made by the classification models and the raw sensor data to the cloud infrastructure used by BeWell.

The classification pipeline samples three sensors, the GPS, accelerometer and microphone. Using audio sensor data the pipeline recognizes social interaction based on classifying  $\{voicing, non-voicing\}$  audio segments. The accelerometer data is used to classify everyday behaviors necessary to monitor the user's physical activity including  $\{driving, stationary, running, walking\}$ . Inferences are made by applying a combination of feature and classification models developed in previous work, SoundSense [18] and Jigsaw [18]. Specifically, we use time (e.g., mean and variance) and frequency domain features (e.g., spectral roll-off for audio) that are classified using a Naive Bayes classifier based on a multi-variate Gaussian Model for each class. A simple Markov Model is used to apply temporal smoothing to the resulting stream of inferences.

The BeWell sleep model is based on a simple but effective approach to modeling the sleep duration of the user (see Section V-B). The classification process is based on features we extract from the phone, specifically, the frequency and duration of: phone recharging events (since often people recharge their phones overnight); and the periods when the phone is either stationary or in a near silent sound environment. These features are not continuously computed rather they are extracted and computed every 24 hours. We found an effective estimate of

user sleep duration could be achieved using a simple logistic regression model which incorporated a prior based on typical adult sleeping patterns.

All data is stored within independent SQLite files. These files are transferred to the cloud infrastructure with an uploading policy that emphasizes energy efficiency to minimize the impact of using the phone's batteries. Uploading only occurs if the phone has both line-power and WiFi networking available. Sensor data is uploaded to the cloud to make additional context available to end-users so they can verify, and if necessary add or edit, inferences made by the Sensing Daemon. However, people are sensitive to sharing sensor data. To address these concerns users are given complete control of the sensors being used on the phone, and inferences made, through user configuration options. Privacy controls are further augmented with automated voicing filtering. As a result, raw audio is never stored if voicing is detected within  $\pm 5$  seconds, reducing the chance that conversations are captured.

### B. Mobile BeWell Portal

The Mobile BeWell Portal provides a simplified mobile phone version of the BeWell portal allowing users to track their current and historical wellbeing scores and view an automated activity diary (see Section IV-D). Users can view trends in their behavioral patterns as well as their wellbeing scores.

### C. Mobile Ambient Wellbeing Display

In addition to the more interactive mobile portal discussed above, BeWell also supports an ambient wellbeing display to provide constant feedback of a user's current wellbeing state. The BeWell ambient display is an animation rendered on the phone's lock-screen and wallpaper making it visible to the user whenever they glance or interact with their smartphone. Prior examples of effective mobile health applications have found that mobile phone wallpaper is an effective ambient display able to promote changes in user behavior (e.g., UbiFit Garden [7]). The BeWell ambient display represents overall wellbeing as three independent scores each corresponding to sleep, activity and social interactions, as discussed in Section III. Each wellbeing dimension is captured by different characters in an aquatic ecosystem, as shown in Figure 4. The animated activities of the orange clown fish represents the recent activity



Fig. 3. The BeWell web portal provides access to an automated diary of activities and wellbeing scores.

level of the user; the turtle's movement reflects a user's sleep patterns; and a school of fish indicates the sociability of the user. By quickly glancing at the screen at different times during the day the user gets a quick summary of their overall wellbeing. If the user clicks on the star fish character on the ambient display a pop-up dialog box is displayed with the numerical values that drive the animation. The relationship between the ambient display and the wellbeing scores is described below:

1) *Turtle:* Sleep patterns are captured by the turtle. The turtle sleeps on behalf of the user; that is, when the user lacks sleep the turtle sleeps for the user. When the user is getting enough sleep, the dimension score is high and the turtle comes out of its shell and walks around with varying degrees of energy.

2) *Clown Fish:* The clown fish represents the physical activity of the user. The score modifies the speed and movements of the clown fish. At low levels of physical activity the fish moves slowly from across the screen. As physical activity increases the clown fish moves more vigorously and even performs summersaults and back flips when the user has high levels of activity.

3) *School of Fish:* Like the clown fish a school of yellow fish also swims across the screen. The size of the school of fish grows in proportion to the amount of social interaction by the user. In addition, the school and the clown fish swim closer together as social interaction increases.

#### D. Desktop BeWell Portal

BeWell users are able to access a web portal, as shown in Figure 3. The portal provides two primary services: i) to provide an automated diary-like visualization of the behavioral patterns and wellbeing scores of the user, which users are also able to manually edit; and, ii) to collect self-report survey



Fig. 4. Multiple wellbeing dimensions are displayed on the smartphone wallpaper. An animated aquatic ecosystem is shown with three different animals, the behavior of each is effected by changes in user wellbeing.

data using standard validated medical surveys that monitor depression, sleep and overall wellbeing.

The diary-like visualization allows users to both browse and edit the behavioral inferences made by the BeWell Sensing Daemon. Users are able to access any sensor data collected by their phone during the day to assist with their recall of events. Users can access their location, listen to ambient (non-conversational) audio and view the raw time-series graphs of their accelerometer and microphone data. This additional context assists users to recall, for example, when they actually woke up, or identify inference errors, such as a period of walking misclassified as jogging. Users are also able to manually add activities that are unable to be inferred by BeWell, but that may still impact wellbeing. For example, types of exercise such as swimming, which the Sensing Daemon can not recognize, or social interactions, that can be missed if the user forgets to carry their phone or turns it off. For those activities that are forms of exercise users are able to select any physical activity in the Physical Activities Compendium [14]. Wellbeing scores are updated based on both automatic sensor data from the phone and through the manual input by the user via the portal.

#### E. Cloud Infrastructure

The cloud infrastructure supporting the BeWell application is a pool of standard Linux-based servers. Servers offer a variety of support services to BeWell components via RESTful interfaces. These RESTful services perform the following functions: i) to store the SQLite files uploaded from mobile phones along with all user input (e.g., survey responses, activity diary edits) collected by the mobile or desktop BeWell portals; and, ii) to respond to queries for raw data, wellbeing scores or inference data which are made by all three components (viz. Sensing Daemon, Mobile and Desktop BeWell

BeWell Sensing Daemon		
	CPU Usage	Memory Usage
GUI only	0%	13511K
Audio sensor only	2%	14373K
Accel sensor only	2%	13917K
Audio classification	25%	14778K
Accel classification on	11%	14736K
Both Accel and Audio classification	31%	15357K
Benchmark Applications		
	CPU Usage	Memory Usage
MP3 Player	16%	27056K
Web Browser	5%	62376K

TABLE I  
ANDROID NEXUS ONE CPU AND MEMORY USAGE FOR BEWELL AND BENCHMARK APPLICATIONS

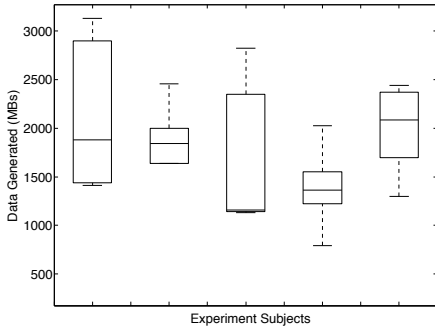


Fig. 5. Daily data generation by subjects during one week experiment

Portal).

In addition, there are continuous background daemons running on each server that perform two key tasks: i) to update the wellbeing scores of users based on incoming inferences; and, ii) to clean all the audio provided by smartphones to further protect user privacy. The cleaning process makes it difficult to inadvertently overhear conversations, which may be accidentally recorded despite the safe-guards implemented on the phone, as discussed in Section IV-A. The cleaning procedure segments each one second of audio into 12 chunks, where every 3rd chunk is zero-ed out. We find this is a simple and effective way to further clean the audio data beyond the voicing based protection implemented on the phone.

## V. EVALUATION

In this section, we perform experiments that validate the design and implementation of the BeWell application. First, we benchmark the resource consumption of BeWell operating on an Android Nexus One phone. We demonstrate that an *automated wellbeing monitoring app*, such as BeWell, can be deployed on an off-the-shelf smartphone. Next, we examine the accuracy of BeWell’s behavior monitoring through a five person experiment. Finally, we show the wellbeing scores collected over one week by a single user.

### A. Benchmarks

We profile the performance of the BeWell prototype application on an Android Nexus One smartphone, for CPU, battery, memory and storage. Phones in this experiment are equipped

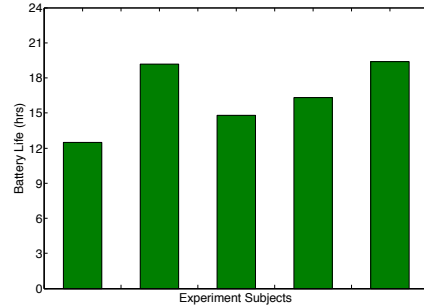


Fig. 6. Smartphone battery life for subjects during experiment

	Voicing	Walking	Stationary	Running
Accuracy	85.3%	90.3%	94.3%	98.1%

TABLE II  
BEHAVIOR CLASSIFICATION ACCURACY

with an extended life battery (3200 mAh) and 4 GB microSD card.

Table I reports memory and CPU usage during different operational phases (e.g., sensing, inference) of the BeWell Sensing Daemon. As the table shows, CPU and memory usage vary significantly depending on the operational phase occurring within the daemon. Not surprisingly, the more burdensome phases involve inference, which encompasses sensor sampling, feature extraction and classification. Under all phases the CPU usage and memory never exceed 31% and 16 MB, respectively. Table I includes a comparison of BeWell with two widely used Android applications (viz. playing MP3 music and browsing the web). A web browser is an example of an application users will use during the day from time to time. Table I shows that BeWell and the web browser can co-exist within resource limitations. Both the MP3 player and the BeWell Sensing Daemon are background processes, and so are designed to be run for long periods of time. Table I shows the MP3 player’s CPU usage when averaged over 5 minutes uses more resources than the BeWell Sensing Daemon. This suggests our daemon is competitive with existing applications in terms of resource efficiency.

We perform a five person experiment which captures battery life performance and data generation rates for BeWell. Subjects are asked to go about their normal daily routines. Figures 5 and 6 present per-subjects values for daily data generation and battery life, respectively. Battery life varies from user to user by 36%. This value increases to 42% for data generation. However, the results for data generation indicate that 4 GB external microSD storage is sufficient. Similarly, the experiment shows that the battery life is consistently above 15 hours, which is sufficient to run BeWell for an entire day if recharged once, very briefly during the day, as well as each night.

### B. Behavioral Inference Accuracy

The accuracy of the classification process is critically important to the overall performance of BeWell. We perform a series

	RMSE	MAE
Linear Regression	2.18 hrs	1.54 hrs
Logistic regression	2.254 hrs	1.56 hrs

TABLE III  
SLEEP DURATION ESTIMATE ERROR

of experiments to test the accuracy of the BeWell classification models (i.e., activity and social interaction classifiers) and sleep model. Models are trained prior to these experiments using training data from 10 people. To maintain consistency across all users for each experiment all subjects position the phone in the same body location, and attach the phone to their hip using a holster we provide.

Our results find that inference accuracy is inline with previous mobile phone sensing experiments [18], [19] conducted with a larger number of subjects. This is expected as our classification models leverage prior work. Table II shows the accuracy for a 5 person experiment where subjects record ground-truth activity diaries for a week.

To validate our sleep model the same 5 users complete a daily survey of sleep duration, using the BeWell web portal. Table III provides the root mean square error (RMSE) and mean absolute error (MAE) when the sleep model uses either logistic or linear regression. The results show that a simple model that correlates phone recharging, movement and ambient sound context is sufficient to predict the amount of hours slept within  $\pm 1.5$  hours. Medical studies suggest that there is little difference in health outcomes for people who have sleep durations that differ by only  $\pm 1$  hour. Therefore, our coarse sleep model which is solely based on a user’s phone, and not on specialized sensors worn while asleep, is adequate for a wellbeing monitoring application.

Raw inference accuracy is only one contributing factor towards how effectively wellbeing is assessed for the physical activity and social interaction of users. In the case of social interaction, for example, our voicing based approach is open to being confused by ambient sound from activities that are not actual conversations (e.g., when the user is watching TV). To test this we perform an experiment involving three people. We randomly select days during the week and require the subject to keep a detailed log of their social interactions for the entire day and evening. From this experiment we find that on average BeWell overestimates true social interaction by 14%. As part of our future work we plan to study more robust techniques for conversation recognition (e.g., using temporal characteristics of conversations such as turn taking).

Similarly, there will be inaccuracies when monitoring user activities that contribute to physical wellbeing. For example, the process of computing METs based on just the “average” energy expended during different categories of activity will introduce noise (e.g., inter-person variation due to weight or sex). However, this is not a new problem. We use the Physical Activities Compendium [14] to compute METs as is standard practice. BeWell potentially introduces an additional error associated with correctly computing the duration of an activity due, for example, to errors in classification. To quantify this error, we compute the difference in ground-truth

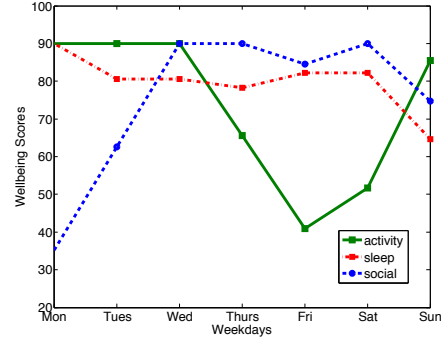


Fig. 7. Automated wellbeing assessments for single representative user

and estimated duration for all BeWell physical activities. We find that duration error averages 22% across all of the users in the experiment. As part of our future work, we plan to study more accurate techniques to compute energy expended during activities; for example, [20] demonstrates that mobile phones can estimate day-long calorie expenditure with 80% accuracy.

### C. Wellbeing Experiment

We perform an experiment in which a single subject uses BeWell to monitor their wellbeing for one week. Figure 7 shows the scores for each wellbeing dimension during the experiment. This is a representative result shown here to illustrate the system in operation. Our preliminary experiment is promising, we observe the scores produced by BeWell correspond to the behavioral patterns of the user for this particular week. Early in the week the social wellbeing score for the user is low, as the subject worked from home. This score climbs later in week as the user interacts with people in the office and again over the weekend. Sleep is fairly consistent throughout the week but dips by Sunday. The subject works out on weekends and this is reflected in a sharp increase in the activity wellbeing score.

## VI. RELATED WORK

There is a growing number of projects in mobile health. However, few apply a holistic approach to the continuous monitoring of wellbeing. Most focus on specific health concerns (e.g., sleep [21]), with the majority of wellbeing tracking systems focusing on physical activity [7], [22]. There is growing interest in systems that take a wider perspective on health. Recent devices (e.g., Fitbit [23]) monitor a broader range of behaviors that impact wellbeing, such as sleep, but require the user to carry a separate dedicated device. BALANCE [24] only monitors one particular aspect of wellbeing, calorie intake and expenditure. It moves beyond only monitoring physical activity, however, it still falls short of truly addressing wellbeing more broadly. Existing wellbeing smartphone applications commonly suffer from an over reliance on burdensome manual user input to gather data, rather than automated sensor-based inferences; and, largely ignore the need to interpret the consequences to wellbeing for the user, based on the data they collect. AndWellness [25] considers a

range of health outcomes but relies on experience sampling to collect user behavior. Although, the monitoring of user behavior is not automatic AndWellness does make use of sensors to intelligently trigger user surveys at more opportune times. In [6], the investigators propose a system that combines self-report with a medical practitioner observing the trends and providing advice via email. This completely manual data entry, and feedback from the practitioners, limits scalability and long-term use. Broad wellbeing has been explored using separate applications that assist in different aspects [26] but this requires users to monitor their own state and select the correct application to address their difficulties. Similarly, [27], [28] consider a comprehensive range of wellbeing but both rely on manual data entry, such as a diary. BeWell focuses on sensing and monitoring wellbeing dimensions just using embedded phone sensors.

## VII. CONCLUSION

In this paper, we discussed the design, implementation and evaluation of BeWell, a personal health application for smartphones capable of automatically monitoring a user's overall wellbeing. BeWell is a real-time, continuous sensing application that provides easily digested feedback that promotes healthier lifestyle decisions. Feedback from BeWell can help users better understand the wellbeing impact of their day to day social interaction, physical activity and sleep patterns. Our prototype BeWell implementation demonstrates the viability of *personal wellbeing applications* using off-the-shelf smartphones. We plan to conduct a large-scale deployment and user study of BeWell to better understand how different users, with different lifestyles, can benefit from this new generation of personal pervasive health applications.

## REFERENCES

- [1] R. Norris, D. Carroll, and R. Cochrane, "The Effects of Physical Activity and Exercise Training on Psychological Stress and Well-being in an Adolescent Population", *Journal of Psychosomatic Research*, vol. 36, no. 1, pp. 55–65, 1992.
- [2] K.R. Fox, "The Influence of Physical Activity on Mental Well-being", *Public Health Nutrition*, vol. 2, no. 3a, pp. 411–418, 1999.
- [3] L.K. George, D.G. Blazer, D.C. Hughes, and N. Fowler, "Social Support and the Outcome of Major Depression", *The British Journal of Psychiatry*, vol. 154, no. 4, pp. 478, 1989.
- [4] Nicholas D. Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T. Campbell, "A Survey of Mobile Phone Sensing", *Comm. Mag.*, vol. 48, pp. 140–150, September 2010.
- [5] Pedro Ferreira, Pedro Sanches, Kristina Höök, and Tove Jaansson, "License to Chill!: How to Empower Users to Cope with Stress", in *Proc. of the 5th Nordic Conference on Human-computer Interaction*, pp. 123–132, Lund, Sweden, Oct 20–22, 2008.
- [6] K. Patrick, F. Raab, M.A. Adams, L. Dillon, M. Zabinski, C.L. Rock, W.G. Griswold, and G.J. Norman, "A Text Message-based Intervention for Weight Loss: Randomized Controlled Trial", *Journal of Medical Internet Research*, vol. 11, no. 1, 2009.
- [7] Sunny Consolvo, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, Ian Smith, and James A. Landay, "Activity Sensing in the Wild: A Field Trial of Ubifit Garden", in *Proc. of the 26th SIGCHI Conference on Human Factors in Computing Systems*, pp. 1797–1806, Florence, Italy, April 5–10, 2008.
- [8] Google Nexus One, <http://www.google.com/phone/detail/nexus-one>.
- [9] GG Alvarez and NT Ayas, "The Impact of Daily Sleep Duration on Health: A Review of the Literature.", *Progress in Cardiovascular Nursing*, vol. 19, no. 2, pp. 56, 2004.
- [10] P. Rozin, "The Meaning of Food in our Lives: A Cross-cultural Perspective on Eating and Well-being", *Journal of Nutrition Education and Behavior*, vol. 37, pp. S107–S112, 2005.
- [11] F.J. Penedo and J.R. Dahn, "Exercise and Well-being: A Review of Mental and Physical Health Benefits Associated with Physical Activity", *Current Opinion in Psychiatry*, vol. 18, no. 2, pp. 189, 2005.
- [12] June J. Pilcher, Douglas R. Ginter, and Brigitte Sadowsky, "Sleep Quality Versus Sleep Quantity: Relationships between Sleep and Measures of Health, Well-being and Sleepiness in College Students", *Journal of Psychosomatic Research*, vol. 42, no. 6, pp. 583 – 596, 1997.
- [13] R.S. Paffenbarger Jr, R. Hyde, A.L. Wing, and C. Hsieh, "Physical Activity, All-cause Mortality, and Longevity of College Alumni", *New England Journal of Medicine*, vol. 314, no. 10, pp. 605–613, 1986.
- [14] B. E. Ainsworth, W. L. Haskell, M. C. Whitt, M. L. Irwin, A. M. Swartz, S. J. Strath, W. L. O'Brien, D. R. Bassett, K. H. Schmitz, P. O. Emplincourt, D. R. Jacobs, and A. S. Leon, "Compendium of Physical Activities: An Update of Activity Codes and MET Intensities", *Medicine and Science in Sports and Exercise*, vol. 32, no. 9 Suppl, September 2000.
- [15] Center for Disease Control and Prevention, <http://www.cdc.gov>.
- [16] D.G. Blazer, "Social Support and Mortality in an Elderly Community Population", *American Journal of Epidemiology*, vol. 115, no. 5, pp. 684, 1982.
- [17] J.F. Knutson and C.R. Lansing, "The Relationship between Communication Problems and Psychological Difficulties in Persons with Profound Acquired Hearing Loss", *Journal of Speech and Hearing Disorders*, vol. 55, no. 4, pp. 656, 1990.
- [18] Hong Lu, Wei Pan, Nicholas D. Lane, Tanzeem Choudhury, and Andrew T. Campbell, "Soundsense: Scalable Sound Sensing for People-centric Applications on Mobile Phones", in *Proc. of the 7th International Conference on Mobile Systems, Applications, and Services*, pp. 165–178, Krakow, Poland, June 22–25, 2009.
- [19] Hong Lu, Jun Yang, Zhigang Liu, Nicholas D. Lane, Tanzeem Choudhury, and Andrew T. Campbell, "The Jigsaw Continuous Sensing Engine for Mobile Phone Applications", in *SenSys '10: Proc. of the 8th International Conference on Embedded Networked Sensor Systems*, pp. 71–84, Zurich, Switzerland, Nov 3–5, 2010.
- [20] Jonathan Lester, Carl Hartung, Laura Pina, Ryan Libby, Gaetano Borriello, and Glen Duncan, "Validated Caloric Expenditure Estimation using a Single Body-worn Sensor", in *Proc. of the 11th International Conference on Ubiquitous Computing*, pp. 225–234, Orlando, Florida, USA, Sept. 30–Oct 3, 2009.
- [21] Shuteye, <http://dub.washington.edu/projects/shuteye>.
- [22] Philips Directlife, <http://www.directlife.philips.com>.
- [23] Fitbit, <http://www.fitbit.com>.
- [24] Tamara Denning, Adrienne Andrew, Rohit Chaudhri, Carl Hartung, Jonathan Lester, Gaetano Borriello, and Glen Duncan, "BALANCE: Towards a Usable Pervasive Wellness Application with Accurate Activity Inference", in *Proc. of the 10th Workshop on Mobile Computing Systems and Applications*, pp. 1–6, Santa Cruz, CA, USA, Feb 23–24, 2009.
- [25] John Hicks, Nithya Ramanathan, Donnie Kim, Mohamad Monibi, Joshua Selsky, Mark Hansen, and Deborah Estrin, "Andwellness: An Open Mobile System for Activity and Experience Sampling", in *Proc. of Wireless Health*, pp. 34–43, La Jolla, CA, USA, Oct. 5–7, 2010.
- [26] Aino Ahtinen, Elina Mattila, Antti Vaatanen, Lotta Hynninen, Jukka Salminen, Esa Koskinen, and Klaus Laine, "User Experiences of Mobile Wellness Applications in Health Promotion", in *Occupational Health*, pp. 1–8, April. 2009.
- [27] E. Lamminmaki, J. Parkka, J. Hermersdorf, J. Kaasinen, K. Samposalo, J. Vainio, J. Kolari, M. Kulju, R. Lappalainen, and I. Korhonen, "Wellness Diary for Mobile Phones", in *Proc. of the 3rd European Medical and Biological Engineering Conference*, pp. 2527, Prague, Czech Republic, Nov. 20–25. 2005.
- [28] Roland Gasser, Dominique Brodbeck, Markus Degen, Jürg Luthiger, Remo Wyss, and Serge Reichlin, "Persuasiveness of a Mobile Lifestyle Coaching Application using Social Facilitation", in *Proc. of 1st International Conference on Persuasive Technology for Human Well-Being*, pp. 27–38, Eindhoven, The Netherlands, May 18–19, 2006.