Unobtrusive Sleep Monitoring using Smartphones

Zhenyu Chen[†‡**], Mu Lin[†], Fanglin Chen[†], Nicholas D. Lane[††], Giuseppe Cardone[‡‡]
Rui Wang[†], Tianxing Li[†], Yiqiang Chen[‡**], Tanzeem Choudhury[*], Andrew T. Campbell[†]
[†]Dartmouth College, [‡]Chinese Academy of Sciences, [††]Microsoft Research Asia, [*]Cornell University
[‡‡]University of Bologna, [**]Beijing Key Laboratory of Mobile Computing and Pervasive Device

***Abstract*—How we feel is greatly influenced by how well we sleep. Emerging quantified-self apps and wearable devices allow people to measure and keep track of sleep duration, patterns and quality. However, these approaches are intrusive, placing a burden on the users to modify their daily sleep related habits in order to gain sleep data; for example, users have to wear cumbersome devices (e.g., a headband) or inform the app when they go to sleep and wake up. In this paper, we present a radically different approach for measuring *sleep duration* based on a novel *best effort sleep* (BES) model. BES infers sleep using smartphones in a completely unobtrusive way – that is, the user is completely removed from the monitoring process and does not interact with the phone beyond normal user behavior. A sensor-based inference algorithm predicts sleep duration by exploiting a collection of soft hints that tie sleep duration to various smartphone usage patterns (e.g., the time and length of smartphone usage or recharge events) and environmental observations (e.g., prolonged silence and darkness). We perform quantitative and qualitative comparisons between two smartphone only approaches that we developed (i.e., BES model and a sleep-with-the-phone approach) and two popular commercial wearable systems (i.e., the Zeo headband and Jawbone wristband). Results from our one-week 8-person study look very promising and show that the BES model can accurately infer sleep duration ($\pm$ 42 minutes) using a completely "hands off" approach that can cope with the natural variation in users' sleep routines and environments.**

***Index Terms*—mHealth; activity recognition; sleep monitoring; smartphone sensing**

I. Introduction

Sleep quality and quantity impacts personal health. For example, poor long-term sleep patterns can lead to a wide range of health related problems, such as, high-blood pressure, high stress, anxiety, diabetes and depression. Quality of sleep can be evaluated partially via physiological measurements, such as respiration rhythm, alternation between deep and rapid eye movement (REM) sleep phases, and partially via personal self-evaluation [1]. The optimal way to monitor sleep quantity and quality is through polysomnographic studies, which use a polysomnogram to monitor sleep. A polysomnogram monitors: brain waves using electroencephalography (EEG), eye movements using electrooculography, muscle contractions using electrocardiography, blood oxygen levels using pulse oximetry, snoring using a microphone, and finally, restlessness using a camera. The complexity and cost of polysomnograms make them impractical for large-scale long-term (i.e., weeks, months) sleep monitoring.

Recently, a number of commercial wearable devices (e.g., Zeo [2], Fitbit [3] and Jawbone [12]) have emerged that make monitoring sleep more accessible to the general public. These technologies use embedded accelerometers and EEG sensors; for example, the Zeo headband uses a combination of inertial sensors and EEG; the Jawbone wristband and Fitbit are based on embedded accelerometers. Many of these devices interact with a smartphone over local bluetooth radio for storage, communication with the cloud and data visualization. However, these approaches are intrusive, cumbersome to use, and place a burden on users to modify their daily sleep related habits in order to gain sleep data; for example, users have to wear devices (e.g., a headband) and in some cases inform the app when they go to sleep (e.g., the Jawbone needs the user to remember to toggle the device to sleep mode or the readings are of little use) and wake up – as a result, the user has to wear the device (e.g., Zeo can irritate the skin and cause headaches) and make sure the monitoring setup is correct (e.g., the phone and device are communicating, the correct switches are toggled).

In this paper, we present a radically different approach for measuring *sleep duration* based on a novel *best effort sleep* (BES) model. BES infers sleep using smartphones in a completely unobtrusive way – that is, the user is completely removed from the monitoring process and does not interact with the phone beyond normal user behavior. The BES model uses a sensor-based inference algorithm on the phone to predict sleep duration by exploiting a collection of soft hints that tie sleep duration to various smartphone usage patterns (e.g., the time and length of smartphone usage or recharge events) and environmental observations (e.g., prolonged silence and darkness).

BES presents a novel approach to sleep monitoring not proposed before. Because BES is a self contained smartphone app it has a number of advantages over other approaches: First, it is unobtrusive and "hands off" for the user – reducing potential usability issues. Next, the user does not have to wear a device nor interact with the phone in any special manner to compute sleep duration. Third, and importantly, it offers a vision of very large scale sleep monitoring well beyond what is feasible with wearables such as the Zeo and Jawbone. The wide-scale adoption of smartphone's in the population make it now feasible to conduct very large studies simply by downloading an app to your phone. Clearly, there are disadvantages of BES over these commercial wearable sleep monitoring systems. The Zeo and Jawbone, for example, can measure sleep quality and patterns to a varying degree. In contrast, BES only estimates sleep duration, which is, however, a coarse indicator of sleep quality; that is, sleep duration

correlates to wellbeing [4] and is influenced by a wide range of health conditions such as affective disorders, depression, anxiety, hypertension, heart diseases and diabetes. Schlosberg and Benjamin report in [5] that heavily stressed people, for example, can suffer from a significant reduction in their total amount of sleep, take longer to fall asleep and wake more frequently during the night.

We present the design, implementation and evaluation of the BES model, which is implemented as part of a broader mobile health app called BeWell [6], [7]. We evaluate four different approaches to sleep measurement: (i) our BES model; (ii) our "sleep-with-the-phone" (SWP) model; (iii) the Jawbone wristband; and (iv) the Zeo headband. We conduct an experimental study with 8 participants who use all four methods for one week – that is, each participant in the study wears the Jawbone and Zeo devices, and, as the name suggests, sleeps with one phone next to their pillow (namely, the SWP smartphone) over 7 nights. Note, that the BES and SWP models represent pure smartphone approaches with no external device interaction. In contrast, the Zeo and Jawbone are commercial products and require the user to wear devices and interact with the phone to collect, store and visualize the sleep data.

This paper is structured as follows. First, we detail the design of the BES model in Section II. Following this, we describe the comparison sleep monitoring systems in Section III. We present results from our study in Section IV that are then discussed in Section V. Section VI describes related work. Finally, we present some concluding remarks in Section VII.

II. Best Effort Sleep (BES) Model

In the following section, we describe the BES model that is studied in this paper and currently implemented as part of the BeWell [6] smartphone app available in Google Play [8]. The BES model unobtrusively estimates user sleep duration by leveraging a series of phone usage and user context statistics collected by BeWell. Specifically, BES keeps track on a daily basis of the total duration of phone-lock, phone-off, phone-charging, phone in darkness, phone in a stationary state and phone in a silent environment, respectively. All the necessary data is gathered solely from sensors embedded in the phone – without user involvement. These features can be grouped in the following four categories; for example, BES assumes that there is a statistical relationship between the sleep duration of the user and the period that the phone experiences darkness. Note, that the BES model is statistical and has multiple features; that is, it is robust to behavior where people might leave the light on at night or days of the year where users experience long daylight hours. BES takes into account all of the features discussed below when inferring the user's sleep duration.

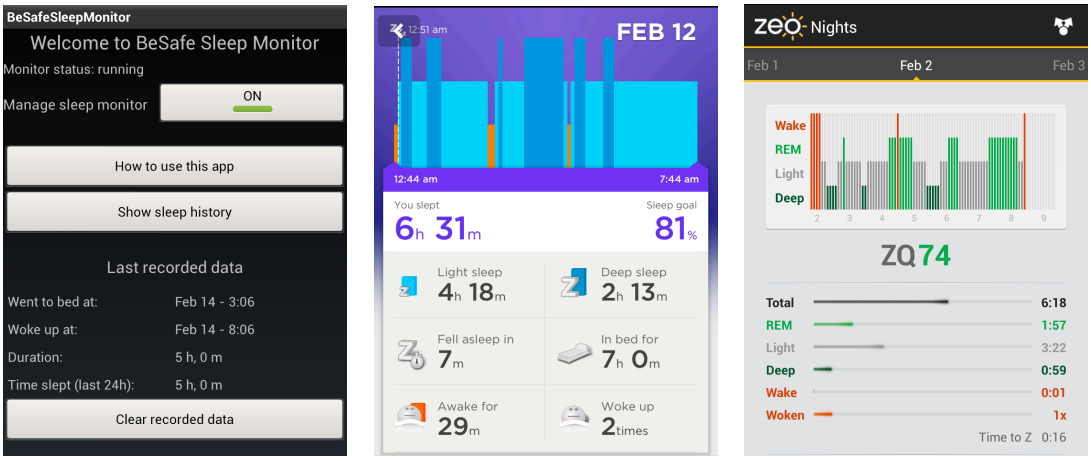
- *Light feature (F_1).* The light sensor on the phone measures the intensity of the light in the phone's context. Therefore, it can robustly detect if the phone is in a dark environment or not. Generally, when people fall asleep at night, they will turn off the lights in their bedroom or the room in the house where the phone has been left. Similarly, the phone transitions from dark to brighter light as the morning arrives or because users wake up and turn the lights on in their house.
- *Phone Usage features*. Phone usage statistics is another rich source for predicting users' sleep duration. For instance, when people go to bed they typically lock the phone (or by default the phone will do this after a period on non-use) and start the recharging process. Or they will choose to simply turn the phone off. Therefore, the duration of phone lock (F_2), phone recharging (F_3) and phone off (F_4) are also good statistical estimators of sleep duration. In BES, we implement a background service using existing Android APIs that can automatically record this duration data.
- *Stationary feature (F_5).* Typically, a user's phone is in a completely static state during sleeping hours. We say typically because the user may use their phone during the night but in general it is mostly in a static state. Therefore the duration while the phone is completely still often correlates with the sleep period of the user. Again, these are not deterministic features but the "stationarity" of the phone represents a rough approximation of sleep duration. Some users may wake up and perform a few activities before picking up their phone; others may pick it up as soon as they wake – for example, to turn off the phone alarm clock. Under BES, the stationarity of the phone is robustly recognized using the Jigsaw [15] activity classification pipeline that processes inertial sensor data (e.g., the accelerometer).
- *Silence feature (F_6).* Similarly, the phone experiences "quiet" at night and the duration of this context correlates with sleep. Again, there are many exceptions that make this an approximation – the noise around people vary with lifestyle and living conditions. BES uses the microphone and the extraction of discriminative audio features to classify "quiet" and "noisy" acoustic environments. This process is performed by the SoundSense [14] audio classification pipeline, embedded in the BeWell App. No audio is recorded during classification for privacy reasons.

As we have indicated, each of these features are individually weak indicators of sleep duration because of the wide variety of phone usage patterns. As a result, BES combines these 6 features to form a more accurate sleep model and predictor. Specifically, BES assumes that the sleep duration of a person (*Sl*) is a liner combination of these 6 features:

$$Sl = \sum_{i=1}^{6} \alpha_i \cdot F_i,\ \alpha_i \geq 0 \tag{1}$$

where the parameters α_i is the weight assigned to each corresponding feature.

As a preliminary study and proof of concept, we conduct a small-scale deployment using 8 subjects for one week to train the BES model. The BES model (as integrated within the BeWell app) is loaded into the primary phone used by each subject. Each phone continuously collects and records all of the above features. Subjects are asked to keep a daily diary



(a) (b) (c)

Fig. 1: Smartphone app for each comparison approach used in our study: (a) Sleep-with-the-phone (SWP); (b) Jawbone Up; (c) Zeo Sleep Manager Pro

of sleep duration as ground truth. BES formalizes the model training process as a nonnegative least-squares regression problem. Specifically, by solving:

$$\min_{\alpha_i} \sum_{j=1}^{4} (Sl^j - \sum_{i=1}^{6} \alpha_i \cdot F_i^j)^2 \quad (2)$$

the optimal set of α_i parameters are identified that minimizes the square error between the ground truth and the predicted sleep duration across all 8 users. The BES model trained by this deployment is evaluated during our experimental study described in Section IV.

III. Comparison Approaches

In this section, we describe three alternative approaches to sleep monitoring that we compare to BES – namely: (i) sleep-with-the-phone model (SWP); (ii) Jawbone Up; and, (iii) Zeo Sleep Manager Pro.

A. Sleep-with-the-phone (SWP) model

We develop an alternative smartphone-based sleep monitoring approach that relies on accelerometer data and a strict usage protocol. Users are instructed to first indicate to the model approximately when sleep begins (for example, via a button press provided to a smartphone app). Afterwards the user must place the smartphone face-down on the bed with them as they go to sleep. During the night the phone monitors the accelerometer, associating stillness with periods of sleep, to estimate the sleep duration of the user. It is important to stress the success of this approach is based on how well the user complies with this usage protocol. Figure 1(a) presents the smartphone implementation of SWP used in our study.

Assuming the user follows the usage protocol, SWP estimates sleep duration as follows. SWP first segments accelerometer data into fixed-length time windows (5 minutes long) from which features are extracted that are representative of two classes of user behavior – either "awake" or "sleeping". Specifically, we extract four time-domain features (viz. average, minimum, maximum, root mean square) from each window for each of the three accelerometer axes (viz. x,y,z); this results in a total of 12 features per time window. Each time window is classified using a C4.5 decision tree [9] as implemented by Weka [10]. The classifier is trained with data collected from three different people; each person uses a different smartphone and sleeps on a different type of mattress – in total we collect more than five hundred hours of data. We use this trained SWP model during our experimental study described in Section IV.

B. Jawbone Up

The Jawbone Up [12] is a wrist mounted device that tracks not only sleep but also physical activity[1]. However, users must explicitly indicate to the device when they are ready for sleep, as well as when they wake up. This is done with a long button press on the device that toggles it between "sleep" and "wake" operational modes. The Jawbone Up infers distinct phases of sleep (e.g., "light" and "deep" sleep) using accelerometer time-series data. Importantly, if the user fails to correctly toggle the device between sleep and wake modes the collected sleep data will be incorrect.

Although the wristband sensor operates independently during data collection to review sleep data the user must connect it with either an iOS or Android smartphone. Data is transferred via the audio connector of the phone and processed within a smartphone app shown in Figure 1(b). The user is able to browse sleep data within this app that includes: total sleep duration, the length of light and deep sleep phases and the start and end times of sleep.

C. Zeo Sleep Manager Pro

The Zeo Sleep Manager Pro [2] relies on sensors embedded inside a headband worn during sleep. These head-mounted sensors monitor the electrical signals of the brain, muscle

[1]Additional health data – mood and eating habits – can be collected with the Jawbone Up through user self-report.

contractions and eye movement. From this data, signal processing techniques are applied to extract features correlated with different phases of sleep. A neural network is then applied to recognize each phase allowing the estimation of total sleep duration (and other sleep-related statistics).

Before a user goes to sleep they must put on the headband and pair it with a smartphone via bluetooth. To guarantee sleep data from the whole-night, the headband must remain paired to the phone which is relied on to both collect and analyze the data. The user must also carefully position the sensors against their forehead and adjust the tightness of the headband so the sensors remain in place during sleep. Failure to do this correctly will result in sleep data not being collected. In the morning, sleep monitoring is ended once the user removes the headband. Because the headband battery only lasts for a single night recharging must be done during the day before the user sleeps again. Figure 1(c) shows the smartphone app included with the Zeo device that allows a user to view the Zeo collected sleep data. The user is able access detailed sleep information including: total duration, the time when sleep began and ended in addition to the length of various sleep phases (e.g., Rapid Eye Movement (REM) sleep).

IV. Evaluation

In the this section, we present the results of a one-week 8-person experimental study that examines the performance of BES and three alternative sleep monitoring systems (viz. SWP, Jawbone, Zeo). In summary, our study findings show: (i) the BES model maintains acceptable levels of accuracy, comparable to alternative systems that use external sensors or require that subjects follow a burdensome usage protocol; (ii) BES is able to cope adequately with routine variations in human behavior – which we call "corner cases" – that may confuse certain BES features (e.g., one night a user may sleep with the light on potentially confusing the light feature F_1 detailed in Section II); and (iii) subjects report a positive user experience when using BES, especially compared to the three alternatives sleep systems. Overall, BES looks promising as a practical technology to unobtrusively track the sleep of large-scale user populations.

A. Study Methodology

Our study population contains 8 males ranging from 23 to 31 years old that are either part of the Computer Science Department or Material Engineering Department at Dartmouth College. Three subjects are visiting scholars, and the remainder are graduate students. Each participant must carry three smartphones during the study with four sleep monitoring apps installed. BES is evaluated using the BeWell app [6] installed on a LG Nexus 4 Android phone. SWP is installed on a Google Nexus One phone. The Jawbone Up interfaces with the "Up" app installed on an iPhone 4S phone, with each participant wearing the Jawbone wristband during sleep. Similarly, at night each subject wears the Zeo headband that is paired via bluetooth to the same LG Nexus 4 Android phone running BeWell. This LG phone also hosts the "Zeo Nights" app that collects data from the Zeo headband. Finally, during data collection each subject records their ground truth sleep activity, including the time they went to bed and woke up.

Feature	Coefficient
Light (F_1)	0.0415
Phone-lock (F_2)	0.0512
Phone-off (F_3)	0.0000
Phone-charging (F_4)	0.0469
Stationary (F_5)	0.5445
Silence (F_6)	0.3484

TABLE I: Weight coefficients for each feature in BES

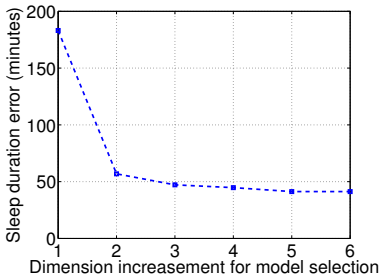


Fig. 2: The reduction in sleep duration error for BES by incrementally adding stationary, silence, phone-lock, phone-charging, light and phone-off features, respectively.

B. BES Model Evaluation

In our first set of experiments we examine the design of the BES model. This model is trained using a preliminary one-week 8-user deployment (see Section II) performed before the start of the experimental study.

Table I shows the feature weight coefficients for BES learned from the preliminary training deployment. We find, for this deployment, that stationary (F_5) and silence (F_6) features are the strongest in predicting sleep duration. Surprisingly, we also find during this training deployment none of the users power their phone off during sleep. As a result, the "phone off" feature (F_4) is zero. In addition, we observe some subjects sleep on occasion with the light on – lowering the weight of the "light" feature (F_1). However, none of these surprises prevent BES from performing well on the training dataset. Under 5-fold cross-validation we find BES estimates sleep duration correctly within $\pm$ 49 minutes.

Figure 2 examines the effectiveness of each of the 6 BES features by incrementally adding each feature to the model – one by one – and presenting the reduction in model error. Features are added into the model in the ranked order of their weight coefficients as listed in Table I. We find model accuracy largely stabilizes once the 4th feature (i.e., phone charging F_4) is added. This suggests the final two features (i.e., light and phone-off features) contribute little to lowering model error. However, these observations regarding model design are inconclusive until larger scale BES user trials are performed.

C. Sleep Duration Accuracy

Figure 3 presents the average per-user sleep duration error for each day across all four sleep systems. We exclude from this result any nights when subjects do not use the Jawbone,

Zeo or SWP correctly. However, we include all nights when subjects use BES. Figure 3 shows BES has a sleep duration error of approximately $\pm$ 42 minutes. In comparison, SWP has an error of nearly $\pm$ 23 minutes while the Jawbone and Zeo devices both have error rates lower than 10 minutes. Although the BES error is larger than comparison systems it achieves this performance without a strict usage protocol (as required by SWP) or cumbersome external sensors (as needed by the Jawbone and Zeo). Furthermore, because most health recommendations related to sleep duration (e.g., [4], [5]) assume hour-level accuracy or often consider just aggregate trends (e.g., Is a person sleeping less than normal?), BES appears sufficiently accurate to help users improve their sleep habits.

D. Coping with Corner Case Behavior

In Figure 3 we report aggregate error during the entire study. However, day-to-day variation in sleep system accuracy due to user behavior is also important to understand. In the following experiments, we examine how each sleep system copes under examples of atypical user behavior – "corner cases" – related to their sleep habits or smartphone usage.

1) BES: Figure 4 reports BES sleep duration error under three examples of unexpected user behavior that occur during our experiments. Each example of corner case behavior is taken from a single day of one user's data that we identify through user interviews. Specifically, "Case1" is a night when a user happens to sleep with the room lights still on; "Case2" occurs when a user has a prolonged nap during the day; and, "Case3" happens when a user fails to recharge their phone. We also provide an average error (i.e., "regular") for all experiment days where it appears the users behave as expected (i.e., in-line with the assumptions of the 6 features used within BES). Figure 4 shows error under BES remains reasonably constant even when unexpected behavior occurs. The largest error is only $\pm$ 62 minutes which is similar to the error under "regular" behavior ($\pm$ 40.5 minutes). In comparison, under SWP errors can jump to four hours or more when users fail to comply with the usage protocol. BES is able to cope with corner case behavior due to the variety of soft sleep duration hints provided by each of the 6 different features it uses. Given a specific situation (i.e., a particular corner case behavior), only a few features are likely to be confused leaving the others to contribute accurate estimates of sleep duration. Figure 5 further illustrates this property of BES by comparing the actual and estimated sleep duration for a single representative user. BES error is shown to be fairly consistent across all days of the experiment.

2) SWP: Figure 6 presents the day-to-day error of SWP for a single representative user. Three days are shown to have noticeably larger errors than the rest – the 3rd, 5th and 6th days. On the 3rd day, we find the user forgot to inform SWP when sleep began, this error is only corrected by the user in the middle of the night. Later in the week – on the 5th day – we find the user stopped the app because the phone had become very hot. Finally on the 6th day, the user reports the

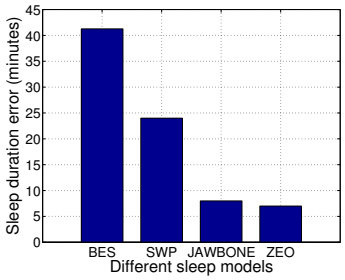


Fig. 3: Overall sleep duration error for BES compared to the three alternative sleep monitoring systems (SWP, Jawbone, Zeo).

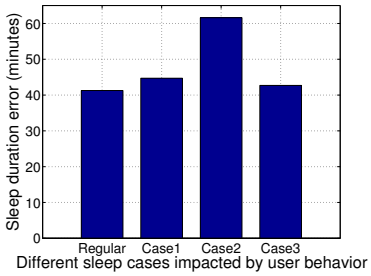


Fig. 4: BES sleep duration error when users behave as expected ("regular") compared to three examples of atypical sleep behavior (i.e., corner cases). Specifically, these behaviors are: Case1 – user sleeps with the room lights on; Case2 – user has a prolonged nap during the day; Case3 – user fails to recharge their phone.

phone shifted position during the night and by the morning the phone was far from their pillow; this limited the accelerometer from accurately detecting the movement (and therefore waking periods) of the person. Collectively, these problems reflect the sensitivity of SWP to the user complying with the usage protocol. Additional reasons for non-compliance may include people with health concerns regarding keeping a phone close to them at all times.

3) Jawbone Up: Figure 7 also presents duration error for a representative subject during the study while using the Jawbone Up. The key difficulty for subjects is that they must remember to toggle the Jawbone device between sleep and awake modes. Two corner case behaviors with high levels of error are shown in Figure 7. On the 3rd day, we find the user

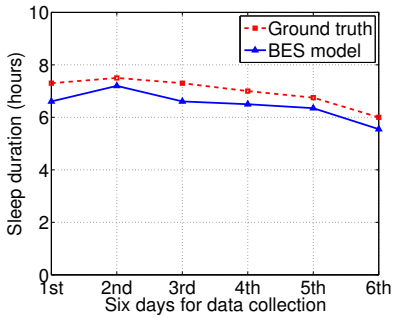


Fig. 5: Comparison of estimated and actual sleep duration under BES for one representative study subject.

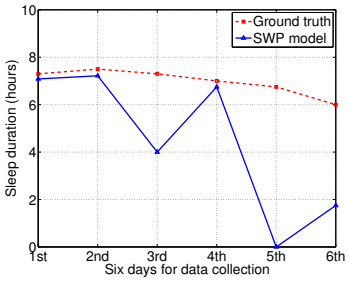


Fig. 6: Comparison of estimated and actual sleep duration under SWP for one representative study subject.

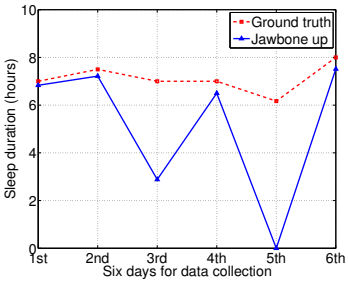


Fig. 7: Comparison of estimated and actual sleep duration for one representative study subject while using the Jawbone Up device.

only remembered to switch the device mode when they woke during the middle of the night after they had already slept for a fairly long time. Similarly, on the 5th night we learn the user simply forgot to wear the device on their wrist. Just as in the case of SWP, we can again see everyday user behavior can greatly influence the accuracy of the Jawbone Up.

4) Zeo Sleep Manager Pro: Figure 8 presents our final case study and illustrates the difficulties of one subject while using the Zeo headband. Overall, the accuracy remains fairly high until the final two days. On the 5th day sleep monitoring is interrupted by a smartphone crash related to a bluetooth connection issue. The user is able to debug the issue and only looses part of the sleep data collected during the night. On the following day, the 6th, we find the user unconsciously took off the headband during the night resulting in little sleep data being collected. By the end of the study this particular subject developed a red mark on their forehead from wearing the Zeo headband. These findings suggest prolonged use of the Zeo device may not be acceptable to all users.

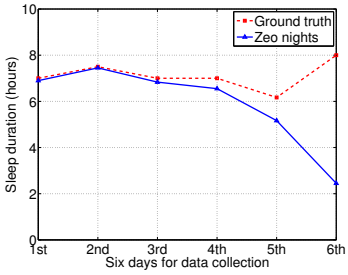


Fig. 8: Comparison of estimated and actual sleep duration for one representative study subject while using the Zeo Sleep Manager Pro device.

Sleep System	Q1	Q2	Q3	Q4	Q5	Q6	Q7
BES	5	5	4.5	4.5	4	2.75	4
SWP	3	2.5	3	3	3.25	2	2.25
Jawbone	3.5	4	4	4	4	5	4
Zeo	1.75	2.75	3.5	2	2.5	3.25	2.75

TABLE II: Summary of User Experience Survey

E. User Experience Survey

To capture the experiences of subjects using each sleep monitoring technology we administer a survey at the conclusion of the study. For each system, all users complete the same 7 questions; specifically:

1) *I found the sleep system was not obtrusive and did not interfere with my day to day life.*
2) *I could use the sleep system for more than six months and still comply with the usage instructions that were given to me.*
3) *Using the sleep system did change my sleep routine. The quality of my sleep was positively impacted.*
4) *The usage instructions of the sleep system are easy to perform.*
5) *I did not notice any change to my phone while using the sleep system. It was as responsive as normal and the battery life was unchanged.*
6) *The sleep system results (i.e., recognition of my daily sleep duration) were accurate enough for me to understand my sleep patterns and any sleep problems I might develop.*
7) *If already available for sale I would download and install the application on my phone (if part of the sleep system) and use it in my daily life.*

Subjects respond to each question using a 5-point likert scale ($1 \approx$ strongly disagree, $2 \approx$ disagree, $3 \approx$ not sure, $4 \approx$ agree, and $5 \approx$ strongly agree) depending on the extent to which they agree or disagree with each statement as it relates to one of the sleep monitoring systems (e.g., BES).

Table II, presents the average user response for each question across all systems. We find subjects strongly prefer BES relative to any of the alternatives. BES receives an average score of 4.25 for all questions. The best performing comparison system is the Jawbone device. In Q5, Q6 and Q7 the Jawbone receives an equal or greater average score than BES. Of particular interest is Q6 that refers to the depth of sleep data presented to the users. Not surprisingly, users rank the Zeo and Jawbone devices – that provide fine-grain sleep data, but require external sensors worn during sleep – ahead of BES and believe they are better able to help them understand their sleep patterns and problems. However, the Neo device receives poor scores in all usability questions (i.e., Q1 thru Q4); and, although the Jawbone device receives strong scores across most questions it still requires that users manually indicate the beginning and end of their sleep activity.

V. Discussion

Based on both the quantitative and qualitative results from our study, we make the following observations and comparisons regarding BES and the alternative approaches we test.

A. On-body Sensors vs. Smartphone Sensing

In the following set of comparisons, we group the four sleep systems we evaluate into two categories: (i) those based on smartphone sensing (i.e., BES and SWP); and, (ii) those leveraging on-body sensors (i.e., the Jawbone and Zeo devices).

1) User Burden: On-body sensor systems place a relatively large burden on the user to ensure sleep data is collected correctly. For example, the Jawbone device must be toggled into and out of sleep mode. Similarly, the Zeo device requires a headband to be put on and the bluetooth connection to be checked. In comparison, approaches based on smartphone sensing have a lower user burden – for example, BES requires no specific user actions.

2) Sleep Data: Our two smartphone-based approaches (i.e., BES and SWP) only provide data on sleep duration. A much wider variety of sleep data is collected by on-body sensor systems, including not only duration but also when the user wakes and falls asleep, the time and frequency of sleep interruptions, and summary statistics of sleep phases (e.g., REM sleep, light sleep etc.).

3) User Feedback: All of the tested systems leverage the smartphone to provide sleep data. On-body sensor systems can also provide feedback via the external sensors – for example, the Jawbone wristband can vibrate and the Neo headband includes LEDs.

4) Cost: Smartphone sensing requires no additional hardware, and therefore monitoring can be nearly free. In comparison, the additional hardware needed for on-body sensor based systems costs between $99 and $130 USD (for the systems we evaluate in our study).

B. Comparison Summary

BES provides acceptable sleep duration accuracy using only the phone's embedded sensors. In comparison, the Zeo and Jawbone devices operate even more accurately largely due to their use of on-body sensors (e.g., head, wrist). However, external sensor driven systems pose different kinds of real-world problems that are likely to limit their suitability for long-term sleep tracking. For example, they require user input to indicate when sleep begins. People often forget to change the Jawbone's operating mode when they are ready to sleep – this problem will only increase with prolonged usage. The Zeo device also has a significant usability problem in that the headband can slip off during sleep due to unconscious user movement. Although the user may tighten the headband this has the side-effect of the headband becoming uncomfortable and potentially causing headaches and discomfort during sleep. Furthermore, the need for continuous wireless pairing between the Zeo device and the user's smartphone is another opportunity for failure. Finally, although SWP – like BES – does not suffer from many of these usability complications related to external sensors it does demand the user follows a specific usage protocol. As we report in our study, if this protocol is not maintained SWP can report inaccurate sleep data.

VI. Related Work

A diverse range of automated and semi-automated technologies for sleep monitoring have been developed for both medical and consumer usage scenarios. Many of these systems require purpose-built sensors that are used to instrument either the user or the sleep environment ([11] provides a survey of home-based sensor-oriented systems). Devices such as the Actigraph [13] and consumer sleep monitors from Phillips [16], FitBit [3] and Jawbone [12] (included our study) require accelerometers to be worn by the user during sleep. Head-mounted sensors are also becoming more popular, for example, iBrain [11] and Zeo [2] (also in our study) exploit brain and muscle signals to track sleep. [17] offers an alternative to user instrumentation and places sensors inside the bed itself. Similarly, [18] shifts the sensor position to above the bed and uses tomography to monitor user body movement during sleep. The strength of this class of sleep monitors is that by relying on carefully placed specialized sensors they can achieve fine-grain sleep monitoring; for instance, tracking aspects of REM sleep or monitoring specific conditions such as sleep apnea. In comparison, BES targets an important complementary part of the design space – sleep monitoring without requiring any additional sensors beyond those present in commodity smartphones. BES seeks to provide coarse "best-effort" sleep monitoring of sufficient quality that it can identify potentially unhealthy sleep behavior that warrants additional investigation (and use of more invasive monitoring techniques). While a number of smartphone apps also enable users to track their sleep patterns without external sensors they have two typical shortcomings. First, many (e.g., [19], [20]) rely on user data entry and so are only effective while the user remains vigilant in their record keeping. Second, apps (e.g., [21], [22]) often require the user to modify their sleep behavior, for example, keeping their smartphone on their bed close to their body during sleep – furthermore, these applications are similarly sensitive to users complying with usage instructions. In contrast, BES is automatic with users completely removed from the data collection process.

In parallel to the advances in sleep monitoring tools, particularly those available to consumers, researchers are developing new technology to promoting healthy sleep behavior (see [23] for a survey of the core problems being studied). One recent example is Shuteye [24], a mobile smartphone app that focuses on improving user sleep-related behavior rather than the process of automated sleep monitoring; in this way, BES and our study are complementary to this system. HealthGear [25] also is a smartphone-based system but is focused on a specific sleep disorder, sleep apnea; and unlike BES, HealthGear requires external sensors. Lullaby [26] seeks to improve sleep conditions in the bedroom by providing comprehensive journaling of a wider range of factors typically considered by automated

tools – although not mobile it presents an interesting future direction for automated sleep assistance.

VII. Conclusion

We presented the design and evaluation of BES – a sensor-based computational model that provides daily automated sleep duration monitoring using commodity smartphones. Unlike existing consumer products (e.g., wearables, smartphone apps) and research prototypes, BES requires zero behavior changes from the user; individuals do not have to change the way they sleep, signal to the phone when they wake up or attach specialized sensors to their body. Rather, BES relies on the collective predictive power of a series of soft sensor-based hints that relate user behavior to sleep duration (e.g., prolonged silence or the smartphone remaining unused and completely still). From a preliminary one-week 8-person study we find BES is able to exploit these hints to estimate sleep duration with a surprising degree of accuracy ($\pm$ 42 minutes) and provides sleep duration estimates close to commercial alternatives requiring wearable sensors. Our results contribute towards the on-going development of mHealth technology able to track key behavioral dimensions that impact overall health and wellbeing; but yet must also remain unobtrusive to users and suitable for daily long-term use.

References

[1] D.J. Buysse, C.F. Reynolds, T.H. Monk, S.R. Berman, and D.J. Kupfer, "The Pittsburgh Sleep Quality Index (PSQI): a New Instrument for Psychiatric Research and Practice", *Psychiatry Research*, vol. 28, no. 2, pp. 193–213, 1989.

[2] Zeo Sleep Manager Pro, http://www.myzeo.com.

[3] Fitbit. http://www.fitbit.com/.

[4] G. G. Alvarez and N. T. Ayas, "The Impact of Daily Sleep Duration on Health: a Review of the Literature", *Progress in Cardiovascular Nursing*, vol. 19, no. 2, pp. 56–59, 2004.

[5] A. Schlosberg and M. Benjamin, "Sleep patterns in Three Acute Combat Fatigue Cases", *Journal of Clinical Psychiatry*, vol. 39, pp. 546–548, 1978.

[6] BeWell, http://www.bewellapp.org.

[7] N. D. Lane, M. Mohammod, M. Lin, X. Yang, H. Lu, S. Ali, A. Doryab, E. Berke, T. Choudhury, A. Campbell. BeWell: A Smartphone Application to Monitor, Model and Promote Wellbeing, In *PervasiveHealth '11*.

[8] Google Play, http://play.google.com/store.

[9] John Ross Quinlan, *C4.5: Programs for Machine Learning. Vol. 1.*, Morgan Kaufmann, 1993.

[10] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten, "The WEKA Data Mining Software: an Update", *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, 2009.

[11] Jessica M Kelly, Robert E Strecker, and Matt T Bianchi, "Recent developments in home sleep-monitoring devices", *ISRN neurology*, vol. 2012, 2012.

[12] Jawbone Up, http://jawbone.com/up.

[13] G. Jean-Louis, D. F. Kripke, R. J. Cole, J. D. Assmus, and R. D. Langer, "Sleep Detection with an Accelerometer Actigraph: Comparisons with Polysomnography.", *Physiology & Behavior*, vol. 72, no. 1-2, pp. 21–28, Jan. 2001.

[14] H. Lu, W. Pan, N. D. Lane, T. Choudhury, A. T. Campbell. Soundsense: Scalable Sound Sensing for People-centric Applications on Mobile Phones, In *MobiSys '09*.

[15] H. Lu, J. Yang, Z. Liu, N. Lane, T. Choudhury, A. T. Campbell. The Jigsaw Continuous Sensing Engine for Mobile Phone Applications, In *SenSys '10*.

[16] Fitbit, http://www.fitbit.com.

[17] Juha M Kortelainen, Mark van Gils, and Juha Parkka, "Multichannel bed pressure sensor for sleep monitoring", in *Computing in Cardiology (CinC), 2012*. IEEE, 2012, pp. 313–316.

[18] Pierre Maquet, "Functional neuroimaging of normal human sleep by positron emission tomography", *Journal of sleep research*, vol. 9, no. 3, pp. 207–232, 2000.

[19] Sleep Cycle Alarm Clock, http://www.sleepcycle.com/.

[20] Sleep as Android, https://sites.google.com/site/sleepasandroid/.

[21] "Sleep Tracker", http://www.sleeptracker.com/.

[22] SleepBot Tracker - Sleep Suite, https://play.google.com/store/apps/details?id=com.lslk.sleepbothl=en/.

[23] Eun Kyoung Choe, Sunny Consolvo, Nathaniel F Watson, and Julie A Kientz, "Opportunities for computing technologies to support healthy sleep behaviors", In *CHI '11*.

[24] Jared Bauer, Sunny Consolvo, Benjamin Greenstein, Jonathan Schooler, Eric Wu, Nathaniel F Watson, and Julie Kientz, "Shuteye: encouraging awareness of healthy sleep recommendations with a mobile, peripheral display", In *CHI '12*.

[25] N. Oliver and F. Flores-Mangas, "Healthgear: automatic sleep apnea detection and monitoring with a mobile phone", *Journal of Communications*, vol. 2, no. 2, pp. 1–9, 2007.

[26] Matthew Kay, Eun Kyoung Choe, Jesse Shepherd, Benjamin Greenstein, Nathaniel Watson, Sunny Consolvo, and Julie A Kientz, "Lullaby: A capture & access system for understanding the sleep environment", In *Ubicomp '12*.