Using Time Series Techniques to Forecast and Analyze Wake and Sleep Behavior

Jennifer Williams
Washington State University
Pullman, WA 99163, USA
jen_williams@wsu.edu

Diane J. Cook
Washington State University
Pullman, WA 99163, USA
cook@eecs.wsu.edu

ABSTRACT

Smart home technologies provide numerous benefits for providing healthcare to individuals in a non-invasive manner. Our goal of this research is to use smart home technology to assistance people recovering from injuries or coping with disabilities to live independently. In this paper, we propose an algorithmic method, Behavior Forecasting (BF), to model and forecast both the wake and sleep behaviors that are exhibited by an individual. The BF method consists of (1) detecting wake/ sleep cycles, (2) defining numeric values that reflect sleep behavior and numeric values that reflect wake behavior, (3) forecasting the numeric wake and sleep values based on past behavior, (4) analyzing the effect of wake behavior on sleep by using wake behaviors when forecasting for the next sleep behavior observed, and vice versa, and (5) improving the performance of score prediction by using both past wake and past sleep scores. We evaluate the performance of our BF method with data collected from CASAS smart homes. We found that incorporating time series techniques such as a periodogram improves the detection of regular sleep and wake cycles. We also found that regardless of the utilized forecasting method, we can model wake behavior and sleep behavior with the minimum accuracy of 87%. These results suggest that we can effectively model wake and sleep behaviors in a smart environment. Furthermore, that future wake behavior can be determined from sleep behaviors and vice versa.

Categories and Subject Descriptors

Time series analysis; Ubiquitous and mobile computing/Ambient intelligence

Keywords

Machine learning; smart environments; cycle detection; behavior forecasting; sleep analysis

1. INTRODUCTION

Getting a good night's sleep is important to all but elusive for many. Problems that occur during sleep are particularly common for individuals who are experiencing stress or are managing chronic health conditions. The relationship between wake behavior and sleep quality has been investigated for years. However, only with the maturing of pervasive computing technologies and machine learning is it possible to quantify sleep quality and quantitatively relate wake behavior to sleep quality. A smart home offers the capability to monitor sleep and wake behavior in naturalistic settings. While a large portion of the smart home research to date has focused on analyzing behavior patterns for health monitoring, less attention has been given to anticipating or forecasting upcoming behaviors. In the context of sleep monitoring and assistance, both behavior monitoring and forecasting are valuable for anticipating and circumventing sleep difficulties.

Sleep is an important component in our everyday lives, and thus should not be considered just another activity within the smart home. We postulate that behavior during wake periods can affect sleep behavior and vice versa. Further, we hypothesize that these behaviors can be predicted based on prior wake and sleep patterns. To validate these hypotheses, we analyze data collected from CASAS smart homes; we evaluate the effectiveness of the forecasting methods with two evaluation metrics, mean absolute error and root mean square error.

In this paper, we explore an algorithmic method, *Behavioral Forecasting (BF)*, in which we forecast wake and sleep behavior using smart home data. Our BF algorithmic method consists of (1) using time series techniques to detect wake/sleep cycles, (2) defining numeric values that reflect wake behavior and numeric values that reflect sleep behavior, (3) forecasting the numeric wake and sleep values based on past behavior (independent prediction), (4) analyzing the effect of wake behavior on sleep by using previous wake behavior when forecasting for the next sleep behavior observed, and vice versa (cross prediction), and (5) improving the performance of value prediction using both past wake and past sleep values (joint prediction). We consider this be a univariate forecasting problem, as we are considering how the individual is behaved previously to predict how they we behave in the future. We evaluate of our BF method with data collected from actual smart home testbeds.

2. SLEEP BACKGROUND

In the majority of smart home research, sleep is viewed as just another activity to recognize and track. However, this is not the most effective use of the sleep data that is being gathered from smart home environments, since sleep plays a fundamental role in a person's overall health and general wellbeing throughout life. Our goal is to analyze and forecast sleep behavior. Accomplishing this goal will assist researchers in identifying sleep problems from sensor data, which often arise in conjunction with problems that occur during an individual's waking hours. To understand the relationship between sleep and wake behaviors, we first review the common components that comprise a single night's sleep and discuss the impact that sleep has on wake behavior.

2.1 Stages of Sleep

There are two main types of sleep: rapid eye movement (REM) sleep and non-REM sleep. The non-REM sleep can be further broken down into four stages of sleep. When a person sleeps, he or she cycles through non-REM and REM sleep, spending the majority (approximately 75%) of the time in non-REM sleep. The sleep cycle begins with non-REM sleep. In stage 1, at the initial onset of sleep, people will often believe that they have not actually fallen asleep; if the person is currently in a sleep cycle, then this is the stage in which he or she transitions out of REM sleep. From stage 1, the person then progresses into stage 2 of sleep, in which he or she is no longer aware of the surroundings and his or her breathing and heart rate become regular. As the person transitions into stages 3 and 4 of sleep, these are the deepest and the most restorative stages of sleep. In these stages, blood pressure drops and breathing becomes slower, the person's muscles relax, the blood supply to the muscles increases, and hormones are released, promoting tissue growth and repair. These are also the stages in which energy is restored. After going through the stages of non-REM sleep, we reverse back through the stages to enter into REM sleep. During REM sleep, a person's body is immobile and his or her eyes move. This is also the stage when the sleeper's brain is active and when dreaming happens. In this stage, energy is provided to the brain and the body to ultimately support wake performance [1]. Figure 1 illustrates the stages of sleep that take place throughout the night [2].

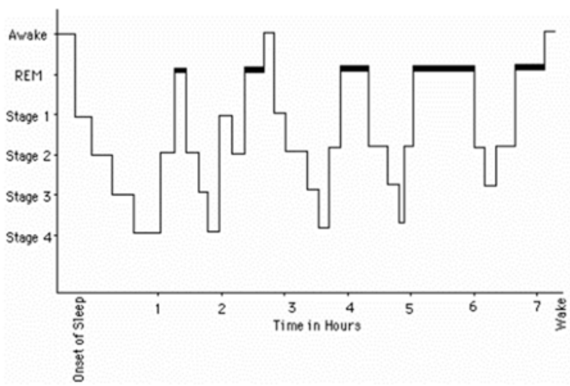

Figure 1: Stages of sleep throughout the night. The graph shows the stages of sleep that take place during the night as the hours of the night progress.

2.2 Sleep Quantity vs. Sleep Quality

There are two common areas that are evaluated with sleep: sleep quantity and sleep quality. Sleep quantity is the amount of sleep a person gets. Sleep quantity is subjective to an individual; thus, the amount of sleep required is whatever is needed for that individual to make them feel rested. The other aspect of sleep assessment, which is often considered more influential, is sleep quality. Sleep quality indicates how well a person sleeps throughout the night. While there are many ways to determine sleep quality, there is not a single standard for evaluating sleep; however, a widely utilized approach is the Pittsburgh Sleep Quality Index [3].

2.3 Pittsburgh Sleep Quality Index (PSQI)

The PSQI contains 19 self-report questions and 5 additional questions that are rated by either a partner sharing the bed or a roommate. The questions are classified into seven categories that combine to develop the PSQI score. These categories are: (1) subjective sleep quality, (2) sleep latency, (3) sleep duration, (4) habitual sleep efficiency, (5) sleep disturbances, (6) use of sleep medications, and (7) daytime dysfunction. Since the PSQI does not limit the self-report questions to highlight just one category in sleep analysis, it has become a common tool for measuring sleep and it has been shown to provide an acceptable score across multiple participation groups (i.e., insomnia) [4]. Therefore, in this research, we focus on the PSQI sleep analysis when developing sleep scores from the smart home data.

2.4 Impacts of Sleep

Both sleep quantity and quality have major impacts on mental health and physical health. In the case of physical health, sleep maintains the body's circadian rhythm. Circadian rhythms run a large number of biological processes that occur throughout the body during the day including body temperature, sleep-wake cycles, and hormone release [5], [6]. Not getting the proper amount and quality of sleep can throw off the circadian rhythm which will then throw off the biological processes in the body, greatly impacting how a person performs throughout the day.

Furthermore, not only does poor sleep impact the circadian rhythm, which in turn influences biological processes throughout the day, but as we saw with the stages of sleep, poor sleep also impacts the restorative stages of the sleep cycle, as the person does not spend as much time in those stages as necessary. It has been shown that poor sleep quality can cause people to be less productive during work [7]. This is in part due to the fact that sleep deficiency can alter brain activity, which immediately affects how a person will think, react, and behave overall. With sleep impacting daily behavior and wellbeing, lack of sleep has repercussions on how people behave throughout the day.

3. Defining the Wake/Sleep Cycle

Since sleep can have a dramatic impact on how a person performs throughout the day, we postulate that sleep must be analyzed separately in a smart home system, rather than being treated as just another activity when monitoring or predicting health and wellbeing. Thus, instead of creating an overall daily score, we create separate wake and sleep scores that will be used in our univariate forecasting method to predict the behavior that the house occupant is experiencing.

As seen in Figure 2a, using the typical 24-hour day can be problematic. This is due to the fact that with the typical 24-hour day, starting and ending at midnight, normally places the end of the cycle change in the middle of a sleep period. As a result, using this standard day notion will not provide a ready basis for measuring the corresponding day's sleep quality. Therefore, in the BF method, rather than break an individual's routine into daytime and nighttime periods, we propose to divide the routine into wake and sleep periods. We then look for cycles that end with the completion of the *end-of-cycle sleep activity* (last sleep activity before a cycle change) and the next cycle begins with the first wake activity after the end-of-cycle sleep activity

Since the cycle is dependent on the individual's sleeping pattern, we are no longer constrained to a 24-hour cycle duration. Instead, we investigate the use of time series techniques to automatically detect the sleep/wake cycle and adapt it to each person. In this paper we consider two cycle detection methods that we describe in the next sections: **RU**le based **C**ycle **D**etection (RUCD), and **PE**riodogram **I**nfluenced **R**ules **C**ycle **D**etection (PEIRCD).

3.1 Rule based Cycle Detection (RUCD)

The first cycle detection framework considered is a rule based cycle detection (RUCD) method. In the RUCD method, the current cycle ends with the completion of the *end-of-cycle sleep activity*, which is determined by taking into consideration two rules: (1) the time since the cycle began (cb) and (2) the duration of the sleep activity (sdur).

To determine an appropriate value for *cb*, we need to first determine an approximate cycle length. As previously stated, all organisms have a series of recurring physiological changes that

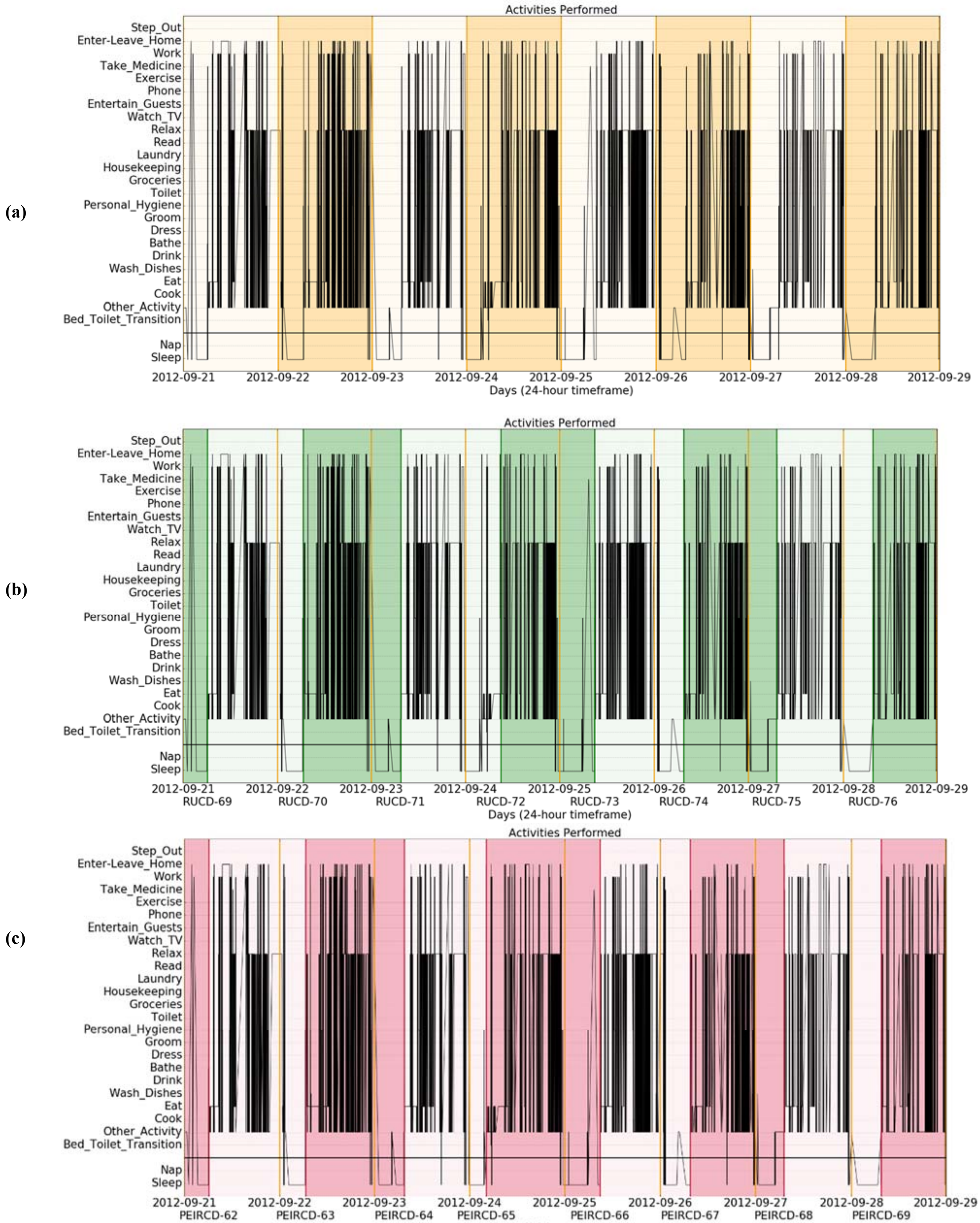

Figure 2: Illustrates the cycle detection methods utilized in this work. Each graph shows 8 days of data collected from a smart home. The black lines indicate the activity that is taking place at that time, when the activity line falls below zero (horizontal line), a sleep activity (i.e., Sleep, Nap) is taking place; otherwise a wake activity is taking place (i.e., Eat, Relax, Work). (a) Shows the cycle changes using when the typical day (changes at midnight). The days are distinguished by alternating white and yellow backgrounds. (b) Shows the cycle changes using the RUCD method. The days are distinguished by alternating green and white background. (c) Shows the cycle changes using the PEIRCD method. The days are distinguished by alternating pink and white background.

take place during their day; these are called circadian rhythms. In humans, the circadian rhythm is roughly on a 24-hour time frame, since it is influenced by daylight hours; furthermore, a well-known circadian rhythm is the sleep-wake cycle [8]. Thus, we still consider roughly a 24-hour time frame. Additionally, studies have shown that adults typically report sleeping between 6 and 8 hours [9]–[11]. Therefore, we determine that an end-of-cycle sleep activity begins at least 18 hours since their cycle began.

The second rule that is considered is the duration of the actual sleep activity. Studies show that people report typically sleeping between 6 and 8 hours. However, a long nap can last as long as 1 to 2 hours [12]. We split the difference between the maximum nap time (2 hours) and the minimum sleep time (6 hours), to determine that an end-of-cycle sleep activity duration is at least 4 hours.

The RUCD method considers two rules when evaluating whether a cycle change should take place: an end-of-cycle sleep activity occurs at least 18 hours after the individual's cycle began (cb=18 hours) and ends with a sleep activity that lasts at least 4 hours (sdur=4 hours). The cycles found by RUCD are shown in Figure 2b.

It is worth noting that in the RUCD method the cycles are changed only on a sleep activity. Therefore, if the house occupant does not sleep at home, a cycle change will not be occur.

3.2 Periodogram Influenced Rules Cycle Detection (PEIRCD)

As previously stated, sleep is subjective to each individual. Since the RUCD method is dependent information found in sleep research based on the average person, it is likely that this method will not work for everyone, especially someone with sleep disorders. Therefore, we also consider a data-driven cycle detection approach. Specifically, we investigate using a periodogram for cycle detection. We then combine the rule-based and periodogram-based cycle detection approaches, resulting in a periodogram-influenced rules cycle detection (PEIRCD) method.

A useful time series technique for this challenge is spectral analysis. Spectral analysis decomposes time series data into a sequence of sine and cosine waves. We utilize a periodogram to identify the strength of importance or strength of the frequencies (or periods) to explain the variations in the time series data. By examining the strengths of the frequencies, we take the frequency that has the highest strength as the frequency that is most representative of the data. Once we have the most important frequency, the *cycle length* is $1/\text{frequency}$.

For the houses that were explored in this work, the periodogram mainly highlighted 3 frequencies, 8 hours, 12 hours and 24 hours, as illustrated in Figure 3. Since 24 hours had the highest value, we found that the cycle length was 24 hours. It is worth noting that there were minor variations in cycle length on a month to month basis; these variations are dependent on whether the participant consistently slept at home during the month.

While the periodogram identifies the cycle length, we did not use just the periodogram when determining cycle lengths. This is because the periodogram relies strictly on the duration of the cycle without using any other outside information. Therefore if the participant has any outside influences that causes their schedule to change slightly, the cycles detected are going to be slightly off as well. This is problematic since using just the periodogram detected cycle lengths may trigger a cycle change in the middle of a sleep activity that will impact the sleep analysis, or cause a cycle change to be after the *end-of-cycle sleep activity* has already taken place.

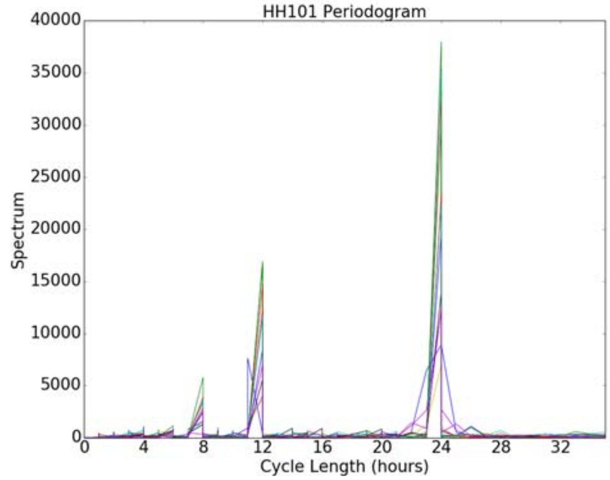

Figure 3: Cycle information gained by the peridogoram. Illustrates the important cycle length distinguished by the periodogram. The graph shows the strength of that cycle length for cycle lengths between 0 and 34 hours.

Thus, we need a method that combines both RUCD and the periodogram cycle detection; this is the PEIRCD method. To incorporate both the rules from RUCD method and the information gained from the periodogram cycle detection, we begin by slightly adjust the rules from the RUCD method. In the PEIRCD method we still define a cycle based on: (1) cb, the time since the cycle began and (2) sdur, the duration of the sleep activity. To incorporate all information we gained from the periodogram, we add cycle length, cl, as an additional periodogram-influenced parameter.

To add the information from the periodogram into the *cb* rule, we consider the cycle length found by the periodogram and determine at what point in the cycle the participant will start a sleep activity. We use the sine wave from the periodogram to roughly indicate when the individual will start a sleep activity. The cycles found by periodogram are illustrated in Figure 4.

From the sine wave, we see that sleep typically starts in the last quarter of their wake/sleep cycle. Therefore, in regard to the time since the time began rule, we consider the cycle length found by the periodogram (typically 24 hour cycle), and we begin looking for a cycle change points during a sleep activity that occurs in the final quarter of their current cycle (cycle length $-$ (cycle length$/4$)).

To determine sdur, the duration of the sleep activity rule, we compare the duration of the detected sleep event against multiple sleep durations. We begin with the sleep duration that was used in the RUCD method of 4 hours, since that is the standard time that is used to distinguish between a long nap and a short sleep activity. However, there are instances in which people sleep less than 4 hours in the smart home during one activity occurrence. Therefore we need to develop a more robust method. Since we use the sleep activity as the pivotal activity to change the cycle; the PEIRCD change preference is to change the cycle on a sleep activity rather than forcing a cycle change without sleep. We still need to be aware of the differences between a sleep activity and a nap activity. Therefore, if the periodogram has indicated that a cycle change is needed and a sleep activity has taken place, we only initiate a cycle change if the participant has slept at least 1 hour (the duration of a long nap).

The new rule we have incorporated is that if the periodogram cycle detection has indicated that a cycle change point has occurred,

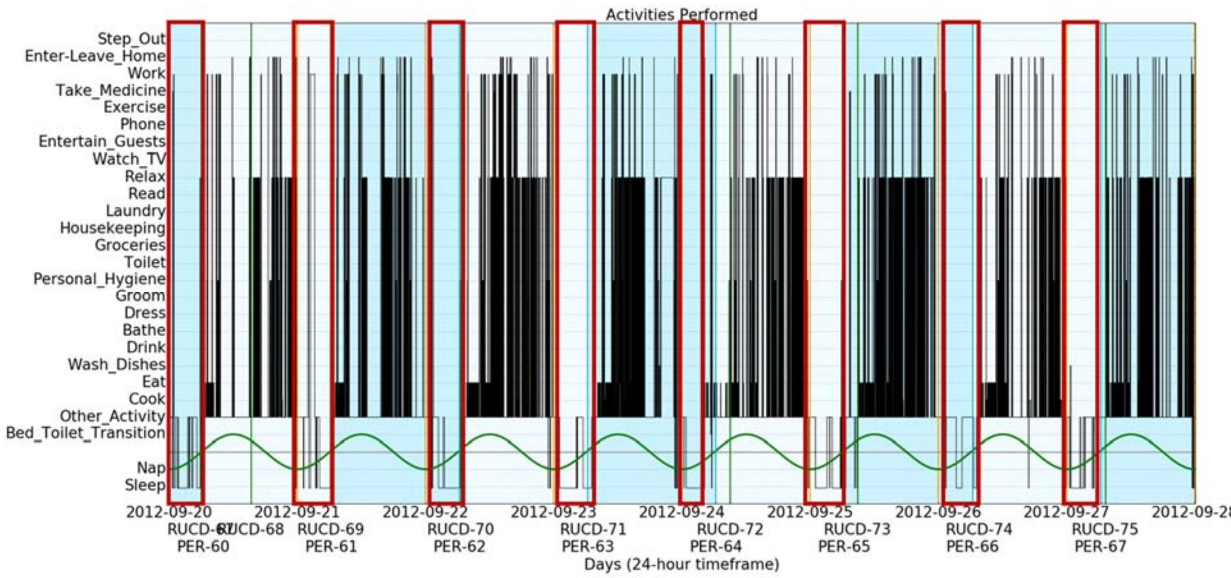

Figure 4: Illustrates 8 days of data collected from a smart home. The black lines indicate the activity that is taking place at that time. When the activity line falls below zero (horizontal line), a sleep activity (i.e., Sleep, Nap) is taking place; otherwise a wake activity is taking place (i.e., Eat, Relax, Work). Shows the cycle changes using a typical 24-hour day (midnight to midnight). Shows cycles found by the periodogram and highlights approximately when during the cycle that the *end-of-cycle sleep* activity takes place. A sine wave is shown to illustrate the cycle length, with the red boxes indicate when an *end-of-cycle sleep* activity; the days are distinguished by alternating blue and white backgrounds

and there is not a sleep activity occurrence, we force a cycle change. This allows us to have appropriate cycle lengths when the house occupant sleeps somewhere other than in the home.

Therefore, the PEIRCD method utilizes the modified rules from RUCD that utilizes information gained by the periodogram cycle detection, in which a cycle change point is tested for a sleep activity if the sleep activity has started in the final fourth of the participants found cycle (cb=cycle length − (cycle length/4)), and the participant has at least 1 hour of sleep (sdur=1). While the PEIRCD method sets a cycle change point on the sleep activities whenever possible, if there is not a sleep activity when the cycle should change the PEIRCD method forces a cycle change point (cl=cycle length found by periodogram). The cycles found by PEIRCD are illustrated in Figure 2c.

4. UNIVARIATE FORECASTING

The goal of this work is to forecast wake and sleep behavior in order to anticipate potential sleep problems and to understand the relationship between wake and sleep. There are numerous approaches that can be taken to achieve this goal. We choose to draw from time series literature in which statistical forecasting techniques are used to predict the value of a single numeric parameter (e.g., daily temperature, stock market values). Time series forecasting techniques are typically univariate, which means that they forecast values for only a single variable. In order to utilize these techniques we need to compress all of wake behavior into a single numeric parameter and all of sleep behavior into a single numeric parameter.

We note that the wake and sleep values themselves may not be easily interpreted in terms of wake behavior quality or sleep quality. However, if the wake/sleep numbers are reflective of actual behavior then they can be predicted. In addition to creating predictable numeric values, or scores, for wake and sleep behavior, we also want a methodology for scoring that is consistent with the literature on wake and sleep behavior. Once these scores are defined we can use time series techniques to forecast them based on past wake and sleep behavior.

4.1 Sleep Score

For our sleep score, we use the PSQI as a reference. Below, we explain which components of the PSQI were incorporated. It is worth noting that the PSQI asks the participant about their sleep behavior during the previous month, but here we adapt the scores to a daily schedule. We also note that scores are generated based on information collected from our CASAS smart home system [13].

Component 1: Subjective Sleep Quality was not included.

Component 2: Sleep Latency was not included.

Component 3: Sleep Duration. We use the total time the participant has spent in an automatically recognized sleep activity as the duration.

Component 4: Habitual Sleep Efficiency. To determine the habitual sleep efficiency, we consider the ratio of the duration in the sleep activity (not including interruptions) over the total duration in the sleep activity including interruptions.

Component 5: Sleep Disturbances. We measure how often the participant gets out of bed during the sleep activity for any reason (i.e., getting out of bed to use the bathroom). Additionally, we can potentially monitor temperature disturbances. We correlate sensor-based temperature data with the sleep measures of duration, interruptions, and habitual sleep efficiency. If a strong correlation ($r > 0.5$) is found between temperature values and one or more of these measures, then that temperature sensor is included in the sleep disturbance component.

Component 6: Use of Sleep Medication was not included.

Component 7: Daytime Dysfunction was not included.

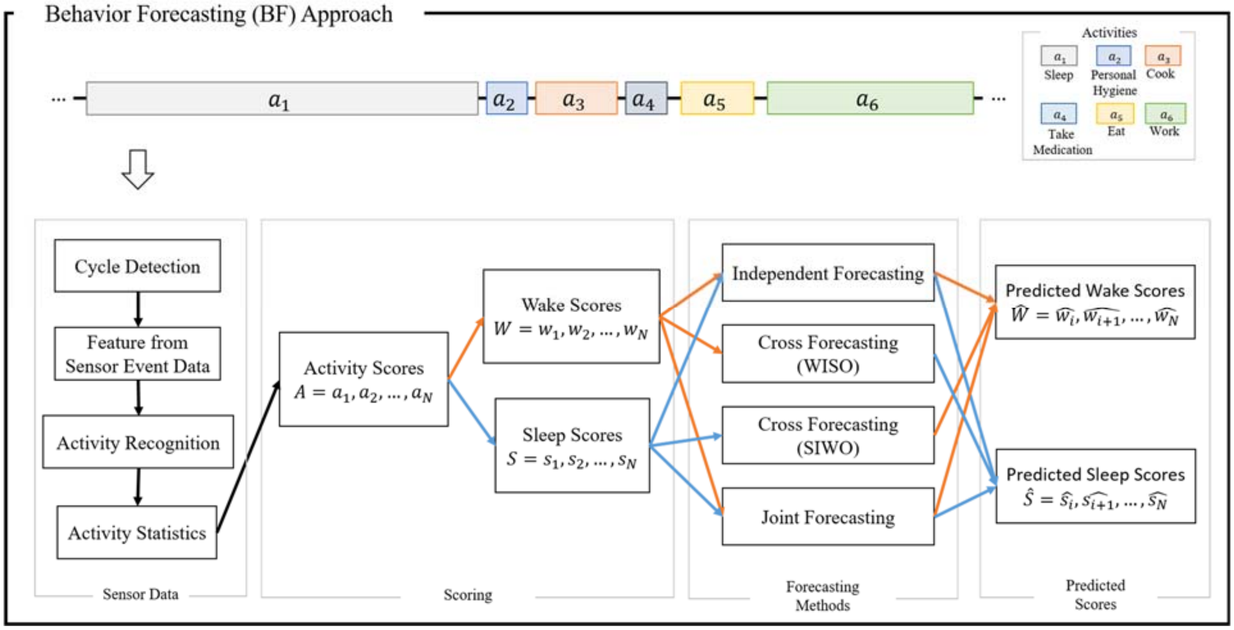

Figure 5: Behavior Forecasting approach framework. WISO: wake scores input and sleep score output. SIWO: sleep score input and wake scores output.

Each of the component receive a score between 0 and 3, the components are then combined to develop the overall sleep score for that wake/sleep cycle.

4.2 Wake Score

Unlike sleep scoring, there is no single theory or measure that is commonly used to assess the quality of wake behavior or to quantify it. In this paper, we are not attempting to quantify the quality of wake behavior. We are defining a single numeric value that is reflective of the activity-based sensed behavior for the purpose of studying the correlation between wake behavior (and changes in wake behavior) and sleep behavior (and changes in sleep behavior). Because our wake score is built upon information that can be detected by smart home sensors and labeled by activity recognition software, our wake score is the combination of the activities seen throughout the wake period in the cycle. Each activity that is seen in the cycle is scored individually (i.e., if the Eat activity takes place three times a day, there will be three Eat scores). Our activities score is built from three activity-centric components: the time since the same activity last occurred, activity performance statistics, and activity level statistics.

Time Since Last Activity. This is the amount of time that elapsed between when the current activity was started and when the activity was previously performed.

Activity Performance Statistics. We consider two activity performance statistics: (1) the duration of the activity without including any interruptions and (2) the number of interruptions that occurred during the activity.

Activity Level Statistics. We consider two additional statistics for each activity. (1) How often the occupant set off the overhead sensors while completing this activity (movement). (2) The ratio of the time spent in the activity to the total time it took to complete the activity.

Each wake activity is scored as follows:

$$\text{wake activity score} = \text{TSL} + \frac{\sum \text{APS} \ + \sum 0.5 * \text{ALS}}{|\text{APS}| + |\text{ALS}|}$$

To develop the overall wake score, the individual wake activity scores are summed.

A troublesome activity that takes place throughout the day is the ‘Other’ activity; this is assigned when the activity does not fall into a predefined category. As this is such a widely used activity label, we put a weight of 0.1 on all the ‘Other’ activities so the overall score is not over-fitted to the ‘Other’ activity category.

4.3 Wake and Sleep Score Summary

Since the sleep score follows the PSQI method, lower sleep scores indicate better quality of sleep. However, with the wake score, a higher score or a lower score is not better, the wake score is simply a score that represents the behavior pattern exhibited during that particular cycle.

It is important to note that creating the wake score for that cycle, the individual wake activity scores were simply summed; similarly creating the sleep score for that cycle, the individual sleep activity scores were summed. This is not the only way to create the wake and sleep scores for the cycle, this is simply one way to accomplish it. The focus of this paper is on creating scores that reflect wake and sleep behavior and designing techniques to forecast these scores.

5. FORECASTING METHODS

The overall hypothesis of this work is that we can predict behavior using novel independent, cross-component, and joint prediction techniques. Forecasting for both wake behavior and sleep behavior is a univariate forecasting problem, since we are using past behaviors to predict future behavior. Therefore, we consider our Behavior Forecasting (BF) approach in which sensor event data is used to create scores that illustrate the behaviors of the participant while the participant is awake and asleep; ultimately the scores will be used predict the next behavior in the sequence. Our BF approach is illustrated in Figure 5.

We use the BF approach to predict the next wake behavior and the next sleep behavior in the sequence. To do this, features are extracted from the sensor event data. These features are input to an activity recognizer, in which the sequence of sensor events is classified into activity categories. With activities labeled, statistics for the individual activities (i.e., duration, interruption count, sensor event count) are computed and the individual activities are scored. The individual activity is scored with the wake scoring method if the activities take place while the participant is awake (i.e., Cook, Eat, Work), or with the sleep scoring method otherwise (i.e., Nap, Sleep). The wake behavior score for that cycle is the sum of the individual activity scores for the activities that take place while the participant is awake; the sleep behavior score for that cycle is the sum of the individual sleep activity scores that take place.

The wake scores and the sleep scores are used for forecasting; each data point corresponds to one cycle. To model and test the forecasting method we utilized a sliding window validation approach, in which a window is moved through entire the dataset. A window consists of t training points, which is the training set, and one test data point. Each training point has a fixed lag length of l, where l is the number of previous cycle scores. As we move through the dataset, we create a model with the training set and predict the next value after the training set, or predict the test data point. After the prediction is made, the window is moved forward one data point. This is repeated until the training window has run through the entire dataset. The only data points that are not predicted are the data points in the initial training window.

Since we are considering both wake behaviors and sleep behaviors, we expand the possible forecasting methods to utilize the behavioral data that we have. Thus, we are interested in exploring three different types forecasting methods for this research: (1) an independent forecasting approach, (2) a cross forecasting approach, and (3) a joint forecasting approach. For the independent forecasting problem, we are interested in determining whether the scores created are a valid representation of the behaviors exhibited such that we can accurately predict for the next score. In independent forecasting, previous wake scores are used to predict the next wake score and previous sleep scores used to predict the next sleep score. For the cross forecasting problem, we are interested in whether we can get enough information from sleep patterns to predict how the participant is going to perform during the next wake period, and vice versa. In the cross forecasting, the previous sleep scores are used to predict the next wake score, and the previous wake scores are used to predict the next sleep score. For the joint forecasting problem, we are interested in determining whether adding all the behavioral score information to the model will increase the accuracies when predicting for the next score. In joint forecasting, both the previous wake scores and the previous sleep scores are used together to predict the next wake score and/or the next sleep score.

6. CASAS SMART HOME DATA

For all the experiments presented in this research, we used data that was collected in actual smart home systems deployed in communities. The data was collected by the CASAS smart home system, developed at Washington State University, which was then installed in the participant's home. We analyze data collected from

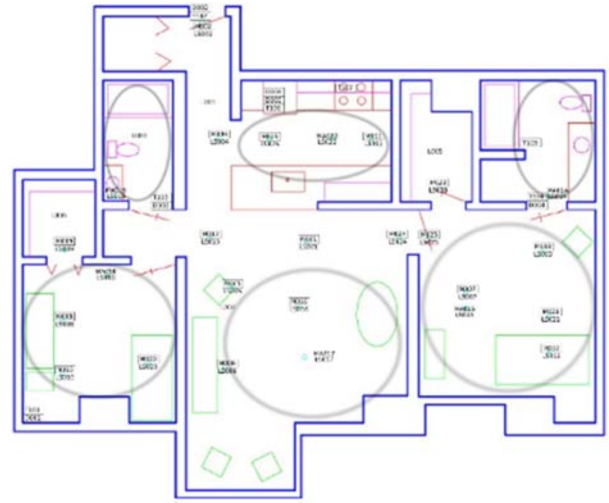

Figure 6: An example CASAS smart home floor plan and sensor layout. The example layout has a bedroom, a kitchen, a dining area, a living area, and two bathrooms.

10 CASAS smart home test beds, each with one resident. A sample floor plan is illustrated in Figure 6.

Each of the smart homes used in this research have least one bedroom, a kitchen, a dining area, a living area, and at least one bathroom. All of the CASAS smart homes have different sizes and layouts, yet they all include the standard sensor setup. There are four sensor types that contribute to the CASAS smart home system: (1) *narrow-field motion sensors*, these sensors reports an ON label when motion is detected, followed by an OFF label when the movement stops. (2) *Wide-field motion sensors*, these are an overall area sensor determine whether there has been movement in the room, not where the participant is. (3) *Door sensors*, these sensors use a magnetic switch to determine whether the doors are opened or closed. (4) *Temperature sensors*, these sensors create an event when the temperature changes.

As the participants are going through their daily routines, when any of the sensors are triggered a sensor event is created and that event is then recorded into the database. The recorded sensor events is then run through an activity recognition method (either through an automated activity recognition software or through human annotation) to label the sequence of sensor events.

6.1 Activity Labeling

Sensor events are mapped to activity labels using activity recognition. The activity recognition algorithm learns a general mapping based on training data labeled with ground truth activity labels. These ground truth activity labels are provided for one month of data in each home using human annotation. Human annotators reference the participant's floor plan with a sensor layout and interview the participants to determine their daily routines. To ensure consistency, multiple people are used to annotate activities for the CASAS smart home sites; the inter-annotator agreement is $\kappa=0.80$ for the data used in this research.

The activity recognition utilized by the CASAS smart home system is AR [14], [15]. AR recognizes activities of daily living including cooking, working, and sleeping, from the sensor data collected from the smart homes. There were 40 activity labels distinguished by AR that were used in this research.

Table 1: Shows the top forecasting methods used with each house to forecast the wake scores. Each house shows the NMAE score for the RUCD method and PEIRCD. If the NMAE error was less than 2% we highlighted with blue, and if the NMAE error was between 2-4% we highlighted with green.

Forecasting for Wake Scores													
	Wake Predictions						Sleep Predictions						
	Independent Prediction		Cross Prediction		Joint Prediction		Independent Prediction		Cross Prediction		Joint Prediction		
	RUCD	PEIRCD	RUCD	PEIRCD	RUCD	PEIRCD	RUCD	PEIRCD	RUCD	PEIRCD	RUCD	PEIRCD	
Home1	0.049	0.030	0.049	0.030	0.049	0.030	0.049	0.054	0.059	0.054	0.059	0.054	Home1
Home2	0.018	0.021	0.018	0.021	0.018	0.021	0.018	0.078	0.106	0.078	0.106	0.078	Home2
Home3	0.017	0.041	0.017	0.041	0.017	0.041	0.017	0.082	0.041	0.082	0.044	0.080	Home3
Home4	0.024	0.017	0.024	0.017	0.024	0.017	0.024	0.108	0.097	0.108	0.097	0.108	Home4
Home5	0.028	0.083	0.023	0.065	0.029	0.083	0.029	0.041	0.104	0.000	0.100	0.000	Home5
Home6	0.025	0.073	0.025	0.073	0.025	0.073	0.025	0.125	0.116	0.125	0.117	0.125	Home6
Home7	0.045	0.069	0.045	0.069	0.045	0.069	0.045	0.108	0.125	0.108	0.121	0.108	Home7
Home8	0.03	0.024	0.03	0.024	0.03	0.024	0.03	0.066	0.091	0.066	0.091	0.066	Home8
Home9	0.023	0.083	0.023	0.085	0.023	0.084	0.023	0.032	0.069	0.036	0.069	0.055	Home9
Home10	0.008	0.026	0.024	0.026	0.019	0.026	0.019	0.127	0.000	0.130	0.000	0.130	Home10

7. METHODS

7.1 Machine Learning Algorithms

In our BF algorithm, the goal is to predict the next wake score or sleep score. Because the scores are continuous values, this can be viewed as a regression problem. We focused on five machine learning methods, (1) regression tree, (2) random forest regression ensemble (2) linear regression, (4) logistic regression, and (5) support vector machine (SVM) with the radial basis function (RBF) kernel.

7.2 Performance Measures

We used two evaluation measures: mean absolute error and the root mean squared error. For all equations used, y_i are the ground truth values that we are comparing all the predictions against, $\hat{y_i}$ are the predicted values, and N is the number of instances being evaluated; each instance for evaluation contains the pair of predicted value and ground truth value, $\{y_i, \hat{y_i}\}$.

The first evaluation measure that we utilize is the mean absolute error (MAE). MAE computes the average of the absolute value difference between the predicted value and the observed value. MAE is defined as:

$$MAE = \frac{\sum |y_i - \hat{y_i}|}{N}$$

MAE measures the average scale of errors without considering the direction of the error (negative or positive error). As only the difference is taken between the predicted value and the ground truth value, all the examples maintain an equal weight; therefore, MAE simply measures the accuracy of the prediction. We also explored a normalized MAE (NMAE), using the maximum score value found in that particular house, defined as:

$$NMAE = \frac{MAE}{\max(y_i)}$$

Another common evaluation measure that we utilize is the root mean squared error (RMSE). RMSE is defined as:

$$RMSE = \sqrt{\frac{\sum (y_i - \hat{y_i})^2}{N}}$$

Since the difference between the predicted value and the ground truth value is being squared, RMSE also is indifferent to the direction of the error (whether the error is negative or positive), and measures the average magnitude of the error providing an emphasis on the higher errors. In other words, RMSE will uncover and highlight when to find if any large errors occurred. We explored a normalized RMSE (NRMSE), using the maximum score value found in that particular house, defined as:

$$NRMSE = \frac{RMSE}{\max(y_i)}$$

For both MAE and RMSE (as well as NMAE and NRMSE), the best possible score is zero, meaning there was no error. Additionally, there is no upper bound on these errors. Using the normalized scores, NMAE and NRMSE, shows the percentage of the error; therefore, the smaller the percentage the better our methods performed.

7.3 Sliding Window Validation

As we are forecasting for wake scores and sleep scores, using a standard cross validation methods is not feasible. This is due to the fact that the wake and sleep scores have a degree of dependence on how the occupants behaved the previous cycles. With that in mind, we use a sliding window validation method. In this approach, we choose a fixed training set size and test on a single data point. The training set starts at the beginning of the dataset and is tested on the next data point after the training set. The window is then shifted forward one data point in order to preform predictions for the next point. This is continued until the sliding window has moved through the whole dataset. In this method, the only data points that are not used for testing are the data points in the initial training set.

8. FORECASTING

8.1 Day Elimination Criteria

Before running the forecasting methods, we need to consider any elimination criteria for the cycles. Because we define a cycle framework there is the potential that the cycles are shorter than or longer than the standard 24 hours since we are using the participant's sleeping patterns to indicate a change of cycle. However, the participant will still have roughly a 24-hour time frame for wake/sleep cycles because of their circadian rhythms, also shown by the periodogram. Therefore, we remove cycles that have fewer than 20 hours or more than 26 hours.

8.2 Results

Table 1 shows the results for forecasting wake scores (left side) and sleep scores (right side) for 10 smart homes. The results in Table 1 show the best NMAE for the both the RUCD method and the PEIRCD method.

In all the methods used for forecasting both wake scores and sleep scores, the maximum NMAE error was 13%, representing 87% accuracy. A noticeable result found was the differences in accuracies when forecasting for wake and sleep scores; forecasting for wake scores provided higher accuracies. When forecasting for wake scores, none of the houses yielded more than 10% error. When forecasting for sleep scores, with the independent method there were no houses with a greater than 10% error; however, when using both the cross method and the joint method there were 5 houses with error greater than 10%. Therefore, when forecasting for both wake and sleep scores, the majority of the results produce less than 10% error, or a greater than 90% accuracy. With the wake and sleep scores, we found that predominantly SVM was out performing the other methods.

We observe that when forecasting for sleep scores, the independent forecasting method yields better accuracies in general over the other forecasting methods. This is not the case when forecasting for wake scores, as each method is preforming within a few percentage of the forecasting methods.

The results for the cross forecasting methods showed that the accuracies were only a few percent different then the accuracies for the independent forecasting method. This is a promising result, since it shows that wake behavior can be modeled from previous sleep behavior and that sleep behavior can be modeled from previous wake behavior, with an acceptable accuracy.

In the case of joint forecasting, the accuracies were only a few percent different then the accuracies for the independent forecasting method; additionally, the results for the joint forecasting methods were very close (if not the same) to the results for the cross forecasting methods. This shows that adding all the behavior score information from the cycle does not improve the accuracies or have much of an effect compared to using just one cycle period.

9. RELATED WORK

As smart home environments are outfitted with various sensors, there is a greater potential of assistive care for participants with either a cognitive or a physical impairment [16], [17]. Assistive healthcare technology that has been researched includes prompting, where the smart home reminds the participant when an action or activity should take place [18], [19].

There is also a wide range of healthcare monitoring available in a smart home. Cognitive health monitoring in a smart home can also be analyzed based on the frequency and quality of activities completed throughout the day [20]–[22]. A related area to both physical and cognitive monitoring is behavioral monitoring. There has been some research regarding behavior predictions [23], in which the behavior is determined by the usage of the household appliances. In this case, forecasting has been utilized for predicting appliance usage durations to ultimately predict the behavior of the participants. Detecting behavior anomalies [24] in a smart home environment has also been explored; when detecting anomalies, a sequence of the events occurring in each particular room is analyzed based on the start time and duration of typical sequences. Clusters of behavior patterns are created for each room and provide the basis for identifying and predicting anomalies in the home. In both of these cases, research has focused on the older adult population, often with some form of cognitive impairment.

10. CONCLUSIONS

We have explained the importance of sleep in a person's overall health and wellbeing; thus illustrating the need to have the sleep component be more important than it has typically been given in other research with smart home environment. Through the experiments that were performed, we found that adding information gained from a periodogram to rule based cycle detection can drastically improve effectiveness of cycle detection. Additionally, we found that regardless of the forecasting method utilized, we can effectively model wake behavior and sleep behavior within a smart home environment.

As the results for the cross forecasting methods and the joint forecasting methods were only a few percent different than the results for the independent forecasting methods, we found that adding all the behavior score information into the forecasting models did not provide enough information to make a noticeable difference in accuracy. However, the results from cross forecasting methods are promising since it shows that wake behavior can be modeled from previous sleep behavior and that sleep behavior can be modeled from previous wake behavior, with an acceptable accuracy. This provides the opportunity to explore the relationship between wake behavior and sleep behavior, which will be explored as the research continues.

11. ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation grant 0900781. The authors would like to thank Maureen Schmitter-Edgecombe, Aaron Crandall, and Brian Thomas for their help in collecting smart home data.

12. REFERENCES

[1] D. L. Walcutt, "Stages of Sleep," *Online*, 2009. [Online]. Available: http://psychcentral.com/lib/stages-of-sleep/.

[2] K. Lovell and C. Liszewski, "Summary: Normal Sleep Patterns and Sleep Disorders." .

[3] D. Buysse, C. Reynolds, and T. Monk, "The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research," *Psychiatry Research*, vol. 28. pp. 193–213, 1989.

[4] J. Backhaus, K. Junghanns, a Broocks, D. Riemann, and F. Hohagen, "Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia," *J*

Psychosom.Res., vol. 53, no. OCTOBER, pp. 737–740, 2002.

[5] E. Maury, K. M. Ramsey, and J. Bass, “Circadian rhythms and metabolic syndrome: From experimental genetics to human disease,” *Circ. Res.*, vol. 106, no. 3, pp. 447–462, 2010.

[6] Y. Dagan, “Circadian rhythm sleep disorders (CRSD),” *Sleep Med. Rev.*, vol. 6, no. 1, pp. 45–55, 2002.

[7] L. M. Swanson, J. T. Arnedt, M. R. Rosekind, G. Belenky, T. J. Balkin, and C. Drake, “Sleep disorders and work performance: Findings from the 2008 National Sleep Foundation Sleep in America poll,” *J. Sleep Res.*, vol. 20, no. 3, pp. 487–494, 2011.

[8] A. Germain and D. J. Kupfer, “Circadian Rhythm Disturbance in Depression,” *October*, vol. 23, no. 7, pp. 571–585, 2009.

[9] A. N. Vgontzas, D. Liao, S. Pejovic, S. Calhoun, M. Karataraki, and E. O. Bixler, “Insomnia With Objective Short Sleep,” *Diabetes Care*, vol. 32, no. 11, pp. 1980–1985, 2009.

[10] A. N. Vgontzas, D. Liao, E. O. Bixler, G. P. Chrousos, and A. Vela-Bueno, “Insomnia with objective short sleep duration is associated with a high risk for hypertension.,” *Sleep*, vol. 32, no. 4, pp. 491–497, 2009.

[11] D. J. Gottlieb, S. Redline, F. J. Nieto, C. M. Baldwin, A. B. Newman, H. E. Resnick, and N. M. Punjabi, “Association of usual sleep duration with hypertension: the Sleep Heart Health Study.,” *Sleep*, vol. 29, no. 8, pp. 1009–1014, 2006.

[12] M. BURSZTYN, J. MEKLER, N. WACHTEL, and D. BEN-ISHAY, “Siesta and ambulatory blood pressure monitoring : comparability of the afternoon nap and night sleep,” *Am. J. Hypertens.*, vol. 7, no. 3, pp. 217–221.

[13] D. J. Cook, A. S. Crandall, B. L. Thomas, and N. C. Krishnan, “CASAS: A smart home in a box,” *Computer (Long. Beach. Calif).*, vol. 46, no. 7, pp. 62–69, 2013.

[14] D. J. Cook, N. C. Krishnan, and P. Rashidi, “Activity discovery and activity recognition: a new partnership.,” *IEEE Trans. Cybern.*, vol. 43, no. 3, pp. 820–8, 2013.

[15] N. C. Krishnan and D. J. Cook, “Activity recognition on streaming sensor data,” *Pervasive Mob. Comput.*, vol. 10, no. PART B, pp. 138–154, 2014.

[16] G. Acampora, D. J. Cook, P. Rashidi, and A. V Vasilakos, “PROCEEDINGS OF THE IEEE 1 A Survey on Ambient Intelligence in Health Care.”

[17] Morris, Meg E, A. Brooke, K. Ozanne, E. Hansen, Ralph, Pearce, and Alan J, “Smart-Home Technologies to Assist Older People to Live Well at Home,” *J. aging Sci. Aging Sci*, vol. 1, no. 1, pp. 1–9, 2013.

[18] B. Das, C. Chen, N. Dasgupta, D. J. Cook, and A. M. Seelye, “Automated Prompting in a Smart Home Environment,” in *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on*, 2010, pp. 1045–1052.

[19] C. Chen, B. Das, and D. J. Cook, “A data mining framework for activity recognition in smart environments,” *Proc. - 2010 6th Int. Conf. Intell. Environ. IE 2010*, no. July 2015, pp. 80–83, 2010.

[20] M. Yu, A. Rhuma, S. M. Naqvi, L. Wang, and J. Chambers, “Posture Recognition Based Fall Detection System For Monitoring An Elderly Person In A Smart Home Environment,” *Inf. Technol. Biomed. IEEE Trans.*, vol. 16, no. 6, pp. 1274–1286, 2012.

[21] P. N. Dawadi, D. J. Cook, M. Schmitter-Edgecombe, and C. Parsey, “Automated assessment of cognitive health using smart home technologies.,” *Technol. Health Care*, vol. 21, no. 4, pp. 323–43, 2013.

[22] P. Dawadi, M. Schmitter-Edgecombe, and D. J. Cook, “Smart home-based longitudinal functional assessment,” *ACM UbiComp Work. Smart Heal. Syst. Appl.*, pp. 1217–1224, 2014.

[23] P. Dawadi, D. Cook, C. Parsey, M. Schmitter-Edgecombe, and M. Schneider, “An approach to cognitive assessment in smart home,” *Proc. 2011 Work. Data Min. Med. Healthc. - DMMH ’11*, p. 56, 2011.

[24] N. K. Suryadevara, S. C. Mukhopadhyay, R. Wang, and R. K. Rayudu, “Forecasting the behavior of an elderly using wireless sensors data in a smart home,” *Eng. Appl. Artif. Intell.*, vol. 26, no. 10, pp. 2641–2652, 2013.