Activity-Based Sleep-Wake Identification in Infants

NA Sazonova, ES Sazonov, SAC Schuckers, CHIME study group

West Virginia University, Morgantown, WV, USA

Abstract

Actigraphy offers one of the best known alternatives to polysomnography for wake-sleep identification. The advantages of actigraphy include high accuracy, simplicity of use and low intrusiveness. These features allow of use actigraphy for determining wake-sleep states in such highly sensitive groups as infants. Both logistic regression and neural networks were tested as predictors. The accuracy of predicted wake-sleep states were established in comparison to the sleep/wake states recorded by technicians in a polysomnograph study. Both prediction methods provided good accuracy of prediction and agreement in results with other studies, thus validating the suggested methodology.

1. Introduction

Identification of sleep/wake states is used in several areas of medical science. For infants, sleep state identification in combination with other parameters may be useful in prediction of life-threatening events such as the Sudden Infant Death Syndrome (SIDS) [1]. Polysomnography (PSG) which includes an electroencephalogram (EEG), electrooculogram (EOG) and electromyogram (EMG), it is the most accurate procedure and is considered to be the "gold standard" in determining sleep states. The largest shortcoming of PSG is that it is rather expensive and too complex to be used by an untrained person. Relatively high intrusiveness of the PSG method is also the cause of its low tolerance by nursing-home patients and infants. An appealing alternative is presented by the actigraphic methods. Actigraphy does not require complex equipment that has to be serviced by a trained technician and is perfectly suited to be used in home conditions. Another advantage of the actigraphy is its low intrusiveness on the patient. An actigraph is a wireless portable device usually worn on a wrist or an ankle. It includes a motion sensor (an accelerometer), a microprocessor with analog/digital circuitry and a memory chip. The motion patterns are recorded throughout the day and analyzed for the information of interest. Usually, actigraphy doesn't aim at identifying sleep states, rather it was traditionally used for determining sleep-wake patterns. Actigraphy provides sleep detection results comparable to those of polysomnography and behavioral response monitoring [2] when applied to different population groups like adults [3], demented nursing-home patients [4], young children and infants [5, 6], etc. Such a wide spectrum of subjects can be covered due to the actigraphy's non-invasiveness. Sleep/wake identifications made in adults by using actigraphy have shown 85-95% agreement rates between actigraphy and polysomnography [7]. In infants agreement rates varied from 54% to 87% at different ages [6].

This study aims at validating the use of actigraphy when a multi-axial accelerometer (normally used to determine infant's position) is also used to provide motion data for actigraphic analysis. The described method also features a different position of the accelerometer on the subjects (diaper instead of an ankle). Accuracy of the sleep/wake state prediction by a statistical method (logistic regression) is compared to that of a soft computing method (neural network). The multi-axial accelerometer was used as a part of Collaborative Home Infant Monitoring Evaluation (CHIME) NIH study, which studied home infant monitors for apnea and bradycardia for over 1000 infants. Sleep state information would be helpful in analyzing approximately 100,000 minutes of data recorded on the home monitor.

2. Data

Data used in this study were collected from infants as part of the CHIME study. Each infant had a standard, monitored 8-hour PSG performed with EEG, EOG, and EMG for calculation of sleep state. Additionally electrocardiogram, respiratory volume, pulse oximetry and an accelerometer were recorded from the same equipment as used in the home study [8]. Although the data set did not include specialized actigraphic measurements, a multi-axial accelerometer (ACC) attached to the diaper was utilized to determine the infant's position in the crib. The placement of the accelerometer on the diaper allowed recording of the infant's position (such as on the belly or on the back) by assigning different DC levels to each of the accelerometer's axes. Such a placement of the sensor is significantly different from traditionally used wrist or

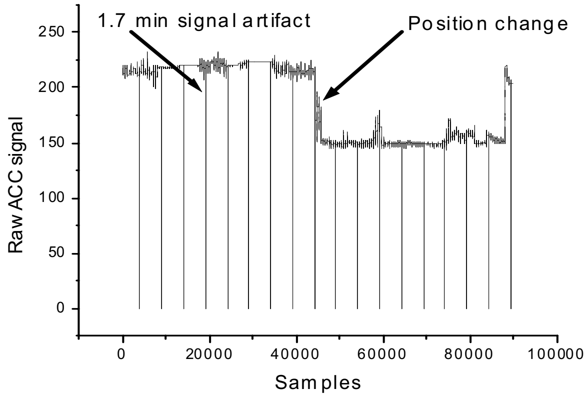

Figure 1. Raw accelerometer signal

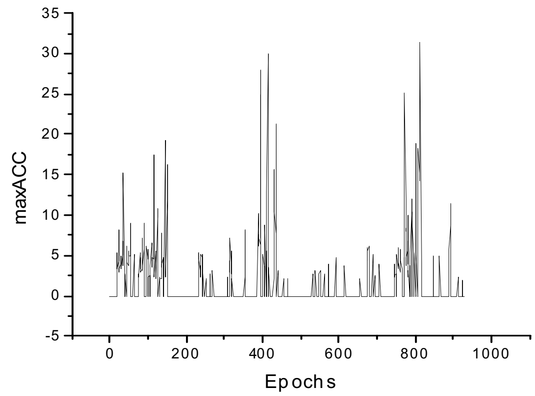

Figure 2. Extracted maxACC measure

ankle and may cause lesser sensitivity to motion (motion of the body is less pronounced than motion of the limbs) and lower noise tolerance (the body has motion artifacts originating from breathing, heart beat and gastro-intestinal processes). At the same time, the accelerometer may pick-up enough motion to identify sleep/wake states and thus eliminate the need for another sensor on a limb.

The raw signal from the accelerometer sampled at 50 Hz (Figure 1) is the primary source of the motion data in this study. Simultaneously with the acquisition of the physiological indicators, independent PSG-based sleep state identification was performed by trained technicians on the infants with 30-second intervals (epochs) [9]. The PSG-identified sleep states (Awake, Active sleep, Quiet sleep and Indeterminable) were used as a baseline for building the regression model and training of the neural network.

The raw accelerometer signal was preprocessed in the following manner:

1. Both the beginning and the end of the accelerometer recording were synchronized in time with the PSG data on an epoch (30 seconds) boundary.
2. An artifact of unknown origin (Figure 1) with a period of 1.7 min was removed from the signal.
3. Position-related DC levels in the ACC signal were removed by applying Discrete Fourier Transform (DFT) to an epoch of the signal (50Hz x 30 seconds = 1500 data points), zeroing the DC harmonic and applying the inverse DFT to the data.
4. The following measures were extracted from the signal: maximum and average accelerometer reading (maxACC and avgACC), and frequency of changes in the ACC reading (frqACC) were computed for each epoch. As an example, the maxACC measure is illustrated in Figure 2.

3. Methods

The methods chosen to address the problem were logistic regression and neural networks as opposed to rather broadly used discriminant analysis [2, 3, 6]. Unlike discriminant analysis logistic regression does not require normally distributed data while neural networks are capable of non-linear mapping of the data. The data clearly showed high degree of positive skewness (Figure 3) and, therefore, the use of discriminant analysis was not validated.

The following variables were used in the analysis:

- response variable is represented by the PSG records of Wake, Active or Quiet sleep during the 30 second epochs. Indeterminate states were ignored for the purpose of the analysis and model building.

- predictors: maxACC measure taken for the last n consecutive periods, thus giving respectively ACC_{-n} , … , ACC_{-2} , ACC_{-1} , ACC_{0} as separate predictors.

Utilization of lagged metrics as predictors was based

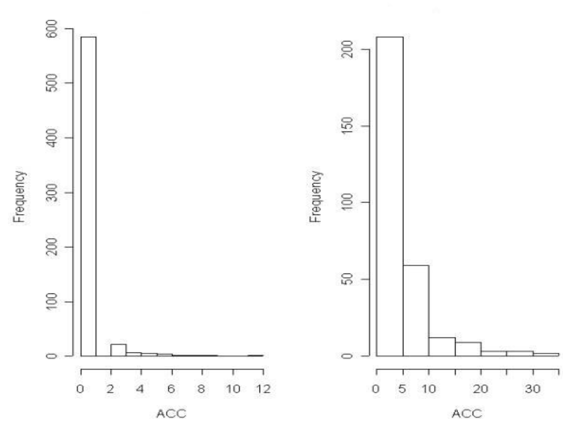

Figure 3. Histograms of the maxACC measure for Sleep (left) and Wake (right) states. Note the scales are different.

Table 1. Agreement and validation rates (%) for the combined model (logistic regression)

Group (rate)	Accuracy of sleep prediction	Accuracy of awake prediction	Weighted Average
Training (agreement)	80.8	78.9	79.4
Validation (validation)	71.3	77.1	76.0
Average	76.8	78.0	77.8

on a simple reasoning that there could be no sudden change in the wake/sleep state. Similar lagged models were used in related studies on adults [7]. All sleep states (Active and Quiet) were combined so that the response vector contained only "Wake" and "Sleep" states.

The first part of the analysis consisted of applying logistic regression with described response and predictor variables. The following form of the regression was assumed:

$$h = \log(p/(1-p)) = b_0 + b_1 ACC_0 + b_2\, ACC_{-1} + \ldots + b_i\, ACC_{-(i-1)} + \ldots + b_{n+1}\, ACC_{-n}, \quad (1)$$

where n represents number of previous 30-second periods for which the maxACC measure is included in a predictor set and p is the probability of sleep. The model predicts a state to be "Sleep" if $p>0.5$ and "Wake" otherwise. Initially the model was developed separately for each subject by taking approximately half of the observations as a training (calibration) set and another half as a validation set. The combined model built on the data pooled from a subset of 4 (out of 8) infants that formed the training sample and the remaining 4 infants were used as a validation sample.

A Learning Vector Quantization (LVQ) neural network [11], a subclass of so-called Kohonen networks, was used to build the neural predictor. The LVQ networks are primarily used as non-linear classifiers, which makes this method a perfect candidate for the task at hand. The experiments with LVQ networks used the same datasets as the logistic regression. The full individual data sets were used both for training and validation of an individual subject while 4 subjects were used for training and 4 for validation in the combined model. The neural predictor was initially trained utilizing the LVQ-1 algorithm and fine-tuned with LVQ-3.

Classification by both LVQ-1 and LVQ-3 is based on a codebook of vectors mi, where each codebook vector belongs to a certain class. The input vector X is compared to each code book vector mi and assigned to the same class to which the closest codebook vector belongs. The distance d between X and mi is calculated according to the following formula:

$$d = \min_i \|X - m_i\| \quad (2)$$

Table 2. Agreement and validation rates (%) by the LVQ predictor

Group (rate)	Accuracy of sleep prediction	Accuracy of awake prediction	Weighted Average
Training (agreement)	95.1	65.0	84.2
Validation (validation)	93.5	44.1	77.6
Average	94.3	54.6	80.9

More information on the LVQ neural networks can be found in [12].

4. Results

First, individual models for each of the 8 infants were built using logistic regression and analyzed for significance of three types of predictors (maxACC, frqACC and avgACC). Among other variables up to 4 lagged maxACC measures showed the highest degree of significance in every individual model and therefore were included in the final model. The coefficients for the same predictor agreed in sign but were quite different in magnitude among all subjects indicating on the presence of pronounced individual effects.

Agreement rates computed on the training set for the individual models were slightly higher than validation rates computed on the validation set: average rates varied from 67.8% to 91.3% for the training set and from 63.3% to 89.9% for the validation set. Moreover, low agreement rates normally corresponded to low validation rates, thus giving indication of the difficulty of a particular subject.

Second, the combined model was built by pooling data from 4 infants in a training set. The following is the linear part of the combined model:

$$\eta = \log(p/(1-p)) = 1.727 - 0.256ACC_0 - 0.154ACC_{-1} - 0.136ACC_{-2} - 0.140ACC_{-3} - 0.176ACC_{-4} \quad (3)$$

All coefficients in (3) have shown high significance in influencing the response variable. The overall differences between individual and combined models were not large. Despite the significant individual effects, the combined model showed good results. Average agreement rates varied from 64.4% to 89.4% and validation rates varied from 68.8% to 86.5%. Overall rates were 79.41% for the training group and 76% for the validation group (Table 1). The largest difference between an individual and the combined models is approximately 10%. Thus, the combined model can be used in place of an individual one with high degree of confidence.

Prediction results by an LVQ neural network with 200 codebook vectors are summarized in Table 2. Prediction performance of the neural network is similar to that of the logistic regression, thus effectively validating the statistical model. Average improvement in the prediction results over the logistic regression is approximately of 4%, while the average prediction rate is about 80%.

5. Discussion

Previous studies showed prediction rates of sleep/wake identification given by actigraphy of 85%-95% [7]. Both statistical and neural predictors of this study provide approximately 72%-92% of accuracy which is comparable to those in other studies, thus confirming validity of using a multi-axial accelerometer for determining the infant's sleep/wake state and position.

In case of logistic regression the results showed high level of agreement and validation rates when compared to similar rates for studies done on adults using actigraphy or other methods. Moreover, the combined model showed almost identical quality of prediction compared to the individual models. Subsequent analysis of the data may improve the accuracy of the combined model by including other explanatory variables in the model to specify individual subjects characteristics.

The neural predictor offered a moderate improvement in the results, displaying the same trends as the statistical predictor. This result may be considered as an indicator of the fact that the utilized methods achieve prediction performance close to the maximum possible. A promising direction of research would be utilization of a self-organizing neural network that could bring the prediction performance close to the performance of an individual model by adapting to the individual trends displayed by each subject.

Obtained results indicate that positioning the motion sensor on the infant's diaper yields approximately the same quality of sleep/wake state prediction as traditionally used placement on the ankle. Using one sensor to perform both position detection and acquisition of the motion data drives down the cost of the system, minimizes intrusiveness on the subject and simplifies collection of the data. Future work may include addition of more subjects in this model and determination of sleep state with more physiological information.

References

[1] Xueyan X. Prediction of Life-Threatening Events in Infants Using Heart Rate Variability Measurements. Ph.D. dissertation submitted to West Virginia University, 2002.

[2] Blood ML, et. al. A Comparison of Sleep Detection by Wrist Actigraphy, Behavioral Response, and Polysomnography. Sleep 1997;20:388-95.

[3] Sadeh A, Sharkey K, Carskadon MA. Activity-Based Sleep-Wake Identification: An Empirical Test of Methodological Issues. Sleep 1994;17:201-7.

[4] Ancoli-Israel S, Clopton P, Klauber MR, Fell R, Mason W. Use of Wrist Activity for Monitoring Sleep/Wake in Demented Nursing-Home Patients. Sleep 1997;20:24-7.

[5] Sadeh A, Lavie P, Scher A, Tirosh E, Epstein R Actigraphic Home-Monitoring Sleep-Disturbed and Control Infants and Young Children: A New Method for Pediatric Assessment of Sleep-Wake Patterns. Pediatrics 1991;87:494-9.

[6] Sadeh A, Acebo C, Seifer R, Aytur S, Carskadon MA. Activity-Based Assessment of Sleep-Wake Patterns During the 1st Year of Life. Infant Behavior and Development 1995;18:329-37.

[7] Cole RJ, Kripke DF, Gruen W., Mullaney DJ, Gillin JC. Automatic Sleep/Wake Identification From Wrist Actigraphy. Sleep 1992;15:461-69.

[8] Neuman MR. Cardiopulmonary Monitoring at Home: the CHIME monitor. Physiol. Measurements, 2001;22(2):267-86.

[9] Crowell DH, Brooks LJ, Colton T, Corwin MJ, Hoppenbrouwers TT, Hunt CE, et.al. Infant polysomnography: reliability. Collaborative Home Infant Monitoring Evaluation (CHIME) Steering Committee. Sleep, 1997;20:553-60.

[10] Haddard GG, et. al. Determination of Sleep State in Infants Using Respiratory Variability. Pediatr Res. 1987; 21: 556-62.

[11] Kohonen T, Hynninen J, Kangas J, Laaksonen J, Torkkola K. LVQ PAK: The Learning Vector Quantization Program Package. Laboratory of Computer and Information Science, Helsinki University of Technology, Finland; 1995.

[12] Kohonen T. Self-Organizing Maps. Springer Series in Information Sciences, Vol. 30, Springer, Berlin, Heidelberg, New York; 1995, 1997, 2001; Third Extended Ed., 501 pages.