Published in final edited form as:

Proc Hum Factors Ergon Soc Annu Meet. 2013 September ; 57(1): 1859–1863. doi:

10.1177/1541931213571415

Highway Healthcare: How Naturalistic Driving Data Index Adherence to CPAP Therapy in Obstructive Sleep Apnea

Anthony D. McDonald¹, John D. Lee¹, Nazan S. Aksan², Jeffrey D. Dawson², Jon Tippin², and Matthew Rizzo²

¹University of Wisconsin-Madison, Madison, WI, USA

²The University of Iowa, Iowa City, IA, USA

Abstract

Drowsy driving is a major factor in many vehicle crashes around the world. Sleep disorders, such as obstructive sleep apnea (OSA), underpin many of these crashes. Continuous positive airway pressure (CPAP) therapy is an effective treatment for sleep apnea but it requires consistent use and is often rejected by OSA patients. Rejection of CPAP treatment creates a dangerous on-road environment for both OSA sufferers and the general public. Algorithms capable of detecting CPAP use and its effects on driving are integral to identifying and mitigating this danger. This work uses naturalistic kinematic driving data to develop an algorithm which can detect nightly CPAP abstinence and adequate CPAP use. Speed and lateral acceleration data were collected using a data recorder in participant's primary vehicle and CPAP data were collected by downloading adherence data from participant CPAP machines. The speed and acceleration data were reduced to a set of symbols using Symbolic Aggregate approximation (SAX) time-series analysis. The symbols were converted into a sequence frequency dataset using sliding windows of size 1 to 10 s with a 1 Hz sampling rate. A Random Forest classifier was trained on the data to create a classification algorithm. On a held aside testing set, the Random Forest algorithm correctly identified 71% of the instances and had an area under the receiver operating characteristic curve of 0.76. The variable importance of the algorithm suggested that kinematic patterns associated with common drowsy driver crash types were key features in the algorithm's prediction performance.

Introduction

In 2011 the National Highway Traffic Safety Administration (NHTSA) reported results from a five year study which showed drowsiness causes more than 100,000 automobile crashes, 40,000 injuries and 1,550 deaths per year (National Highway Traffic Safety Administration, 2011). Sleep disorders, such as obstructive sleep apnea (OSA) are a major factor in many of these crashes (Brown, 1994). OSA affects at least 2–4% of middle-aged adults (Young et al., 1993) and is particularly pervasive in professional drivers (Moreno et al., 2004). A recent large meta-analysis showed that OSA drivers have a mean crash risk ratio of 2.72, having a 172% greater chance of a crash relative to the general population (Sassani et al., 2004; Tregear, Reston, Schoelles, & Phillips, 2009).

OSA is a chronic disorder associated with repeated episodes of complete or partial collapse of the upper airway during sleep. These episodes lead to fragmented sleep which results in excessive daytime sleepiness and increased cardiovascular morbidity and mortality. Continuous positive airway pressure (CPAP) is the treatment of choice for patients with clinically significant OSA (Kushida et al., 2006). CPAP reverses the upper airway collapse in OSA patients by delivering pressurized air through a nasal or an oral–nasal mask sufficient to “splint” open the airway. By providing a patent air-way, CPAP prevents obstructive respiratory events (Kushida et al., 2006). Proper use of CPAP has been shown to improve both crash risk and simulated driving performance in as little as two weeks of treatment (Barbe et al., 2007; Findley, Smith, Hooper, Dineen, & Suratt, 2000; Orth et al., 2005; Turkington, Sircar, Saralaya, & Elliott, 2004). Specifically Tregear et al. (2010) found that crash risk declines by 72% with CPAP.

Unfortunately, CPAP treatment is often uncomfortable leading to reduced acceptance and adherence to CPAP therapy in many patients (Kribbs et al., 1993). In one study, only 72% of patients who started on CPAP agreed to continue it after the first night (Rauscher, Popp, Wanke, & Zwick, 1991). Patterns of use vary widely in patients who accept CPAP; about half use it an average of 6 hours in 90% of nights, while the rest use it less than 4 hours in 2–79% of nights (Weaver et al., 1997).

The accident risk induced by abandonment of CPAP treatment creates a dangerous environment for both drivers with OSA and the general driving public. At the same time, the link between vehicle crashes and untreated OSA suggests that driving data may provide insights on the degree to which CPAP use modulates crash risk. Additionally driving data may help identify drivers' CPAP adherence. This identification is advantageous because it can be used to facilitate interventions prior to a crash caused by CPAP abandonment. Furthermore algorithms developed to detect CPAP adherence may be used to evaluate new clinical interventions for treating OSA.

The development of such algorithms is difficult because it involves searching through data from multiple vehicle sensors in an effort to identify similarities between drivers. More specifically it involves a search for common patterns between drivers. Such patterns are difficult to describe and analyze with continuous time-series data, however the problem is greatly simplified by transforming the data into symbols. This transformation enables the application of Machine Learning tools developed for natural language processing and other domains involving symbolic analyses.

This work explores the development and analysis of an algorithm following this paradigm. The remaining sections describe the data utilized in this work, process to convert this data into symbolic form, conversion of the symbolic data to a form amenable to machine learning, algorithm development, algorithm results, and conclusions that can be drawn from this analysis.

Naturalistic Driving Study

Study data originate from an ongoing NIH funded study focused on evaluating the effects of CPAP therapy. Currently 65 participants have begun the 3.5-month data collection process.

The 65 participants represent two groups: participants diagnosed with OSA, and healthy control participants. The work here focuses solely on the participants with OSA. These patients represent 45 of the 65 drivers, 31 of which are male with a mean age of 47 (SD = 7.36). Drivers are compensated for their time and effort at the completion of data collection.

An in-vehicle data acquisition system (IV-DAS) was installed in each participant's own vehicle to record GPS information, speed, three-axis accelerometer, and accelerator input at 10Hz. The IV-DAS also collects intermittent video of the driver's face and the road. Data were collected for a period of two weeks prior to OSA participants receiving CPAP therapy and for three months following the start of CPAP therapy. The data were partitioned into individual drive files defined by ignition engagement and disengagement. In addition to driving data, participants' CPAP usage was monitored daily. The CPAP data included nightly usage and several sleep quality measures.

The present analyses focus on CPAP usage, speed, and accelerometer data. CPAP usage was collected nightly and measured in minutes. Speed and accelerometer data were selected because they are the most robust data in the dataset. They are inexpensive and unobtrusive to collect and are likely diagnostic of the main crash and near crash types associated with drowsy driving, lane departures and rear end collisions (MacLean, Davies, & Thiele, 2003). Lane departures are often preceded by a period of inactivity and followed by a large lateral acceleration as a driver attempts to quickly return to his or her lane to avoid crashing. Rear-end collisions are often preceded by large changes in longitudinal acceleration and speed due to drivers applying the brake pedal strongly to avoid crashing. Alert drivers should show more gradual deceleration and fewer peak lateral acceleration events. In order to capture these phenomena acceleration data were analyzed in three ways. Lateral and longitudinal acceleration measures were included separately in the data and then combined together in a vector sum (combined acceleration).

Converting Continuous Driving Data to Symbolic Data

The driving data set contains 15,953 drives with speed and acceleration collected at 10 Hz. This level of granularity is computationally prohibitive with respect to pattern recognition and thus prevents thorough analysis of the data. In order to make the data amenable to machine learning they must be reduced, converted to symbols, and coerced into a set of measures (a feature vector) with a binary label. The reduction and symbolic conversion can be accomplished through Symbolic Aggregate approximation (SAX) time-series analysis. The SAX method reduces continuous data by replacing small segments of data with the mean over the segment. The means are converted to symbols via bins of the measure on the y-axis, which are defined by quantiles of the normal distribution (Lin, Keogh, Wei, & Lonardi, 2007). Thus the input to SAX is a continuous source of data and the output is a series of letters.

The concept of applying SAX to driving data was introduced in McDonald et al. (2012). The steps for a segment of speed data are illustrated in Figure 1.

Figure 1 shows the steps of SAX reduction for a ten second segment of speed data. The top plot displays the raw data at 10 Hz. The middle plot displays the mean reduction step to

reduce the data to 1 Hz (1 letter per second), and introduces the labeled bins on the y axis. The bottom plot illustrates the final symbolic conversion overlaid on the raw data. The letter sequence, “dfghhhhhii,” is the final output of SAX for this segment of speed.

SAX time-series analysis was applied separately to speed, lateral acceleration, longitudinal acceleration, and the combined acceleration measure. The application process used 1s sized, non-overlapping windows and was consistent for all four measures except for the number of bins (alphabet size). Nine bins were used for speed, as depicted in Figure 1, and four bins were used for lateral acceleration, longitudinal acceleration, and combined acceleration. The change was made for two reasons. First acceleration measures were noisy and thus a smaller alphabet reduced the effects of this noise. Second the distributions of acceleration data had less variance so a smaller alphabet better captured qualitative differences in acceleration, and in essence helped parse out rare, potentially dangerous accelerations from more common values observed in normal driving. The result of the four applications of SAX was a set of four words corresponding to speed, acceleration, lateral acceleration, and longitudinal acceleration, which characterized each drive. Thus the reduced dataset contained the date of each drive, the CPAP usage from the previous evening, and the set of four words.

Converting Symbolic Driving Data to Algorithm Input

In any symbolic sequence analysis the goal is to find a way of measuring the similarities between sequences. The simplest way to accomplish this goal is to compare two sequences directly and assess their differences based on each letter. Unfortunately this technique is not robust to sequences of different lengths and small differences. For example consider two drivers who live next door to one another driving to a traffic signal at the end of their street. One driver will have to drive for an extra few seconds before she reaches the other driver's home. After this point the drives will be exactly equivalent. If compared directly the drives may look very different because of this small offset. A more informed approach would break the drive into smaller sequences and analyze their frequencies in the respective datasets (Leslie, Eskin, & Noble, 2002). In the previous example this conversion would make the two drives very similar except for the frequency of the sequence which defines the drive between the two houses.

This sequence frequency method was employed here. One underlying assumption of the example above is that the duration of each meaningful sequence is apparent in the data. In the OSA case this assumption fails because specific patterns of data have not been associated with CPAP use or CPAP disuse. In order to compensate for this a range of sequence lengths, from 1 to 10 letters, and their respective frequencies were examined for each of the four words. The result of this process is a dataset which contains columns for the date of the drive, the CPAP use in minutes, and one column for each sequence observed in the data.

Figure 2 demonstrates the process of creating this dataset for a sequence length of 3. At each step of the process a sliding window is moved one letter to the right. If the sequence in the

window is already in the dataset, then the frequency of that sequence is incremented by one. If the sequence has not been observed, it is added to the dataset with a frequency of one.

Additional data processing

The sequence frequency dataset contained one row for each of the 15,953 drives in the data, labeled with the CPAP usage from the previous night. This arrangement is somewhat problematic because often multiple drives occurred in one day for each driver and therefore the drive rows are not consistent with the daily CPAP use labels. To compensate for this differential the sequence frequency data was aggregated to the day level. Thus if the sequence, “Speed_aaa,” was observed once in three different drives on a given day, the aggregate data would show “Speed_aaa” with a frequency of three.

Each row of aggregated data was labeled with a continuous measure of CPAP use during the previous evening specified in minutes. This continuous specification does not integrate well with most machine learning algorithms, which perform best with binary labels. Therefore it was necessary to discretize the CPAP use measure into two levels, use and lack of use. Intuitively this process consists of defining a threshold for each condition and then simply re-labeling the data according to the thresholds. In this case the two thresholds used were more than 360 minutes (6 hours) of use and 0 minutes of use respectively. Nights where participants used their CPAP, but did so for less than 360 minutes were removed from the dataset. The goal of the bounds and data removal was to create a clear definition between the two labels that ensured nights classified as positive CPAP use complied with recommended CPAP practices. In the removed cases it is likely that participants did not get a sufficient amount of sleep which could contribute to general drowsiness related driving incidents and would make the conclusions drawn by the algorithm more ambiguous. Exploration of these “in-between” cases is left for future work. After this reduction and the day aggregation, the final dataset contained 1,799 instances (days), corresponding to 725 days where people used CPAP and 1,074 when they did not.

Algorithm Development

The data were partitioned into stratified training and testing datasets, so that the generalizability of the algorithms could be assessed. The training set contained approximately 80% of the original data (1440 instances) and the testing set contained the remaining 20%. The testing set was withheld from all feature (measure) and algorithm selection decision-making processes.

Feature Reduction

One limitation of relying on sequence frequency process of measure creation is that the process produces many irrelevant and uninformative features for describing the data. For example an unexpected and rare occurrence such as a squirrel running into the road may cause a unique speed pattern in one of the driving days. Since the pattern is unique it will be associated with CPAP use or disuse alone, and may be used to differentiate between the two cases by an algorithm trained on the data. Obviously the pattern is a consequence of the squirrel and thus the algorithm would be inaccurate. In contrast, events such as a single letter

or second of 0 speed may occur in nearly all drives. These features are a hindrance to learning and cost processing time. In order to compensate for these two limitations sparse features, those found in less than 1% of days, were removed from the data and then an information gain based feature selection process was applied to the remaining data.

Information gain based feature selection involves ranking features by the amount they reduce entropy (uncertainty) in the original data labels and then selecting the best k features (Yang & Pedersen, 1997). Entropy reduction can be envisioned as a process where knowledge of certain information affects conclusions drawn based on other information. For example the probability of a particular result in a die roll may be uncertain, however knowing that the die has the same number on all sides would reduce this uncertainty to 0. In contrast knowing that the die is red or blue would provide little insight and thus no information gain.

In this application the 500 most informative features were retained. Thus the final algorithm training data set contained 1440 days (576 days of CPAP use) each described by 500 different features corresponding to segments observed in speed, lateral acceleration, longitudinal acceleration, and combined acceleration lasting between 1 and 10 seconds.

Algorithm selection and training

The primary goal of this work is to create an algorithm that can identify CPAP usage from driving data. However the interpretability of this algorithm is also important because it enables one to gain insight into the phenomenon of OSA induced drowsy driving. At the same time the measures here are inherently noisy and may undermine algorithms that are sensitive to noise such as k-nearest neighbor and rule based methods (Kotsiantis, Zaharakis, & Pintelas, 2007). The goal of interpretability and presence of noise eliminate a large portion of potential modeling types including a Support Vector Machine with a string kernel, which would generally be associated with sequence frequency features (Leslie et al., 2002). One modeling approach, which satisfies these requirements, is Random Forests.

Random Forests, developed by Breiman (2001), are a series of decision trees trained on random bootstrapped samples of training data, and randomly selected feature subsets. Classification is performed by a majority vote amongst the predictions of the trees for each new instance. A key feature of Random Forests is that they naturally produce a measure of variable importance. This measure is based upon the amount of times a given feature appears in the algorithm and its predictive power in each appearance. The variable importance compensates for the loss of interpretability of a multi tree algorithm by demonstrating key variables associated with classification.

The Random Forest algorithm for OSA data was trained and its parameters were optimized using a ten-fold internal cross validation with the training set. The resulting algorithm contained 126 trees.

Results

This work sought to explore the feasibility of predicting CPAP usage with driving data using an algorithm that provides explanatory power for the underlying phenomenon of OSA induced drowsy driving. These two goals are evaluated separately below.

Algorithm Evaluation

The primary evaluation criterion for the algorithm is area under the receiver operating characteristic curve (AUC). The receiver operating characteristic (ROC) curve is a plot of True Positive Rate (sensitivity) by False Positive Rate (1-specificity). The AUC is a measure of classifier performance that is insensitive to distributions of positive and negative examples in the data set. An AUC of 0.5 (the line with a slope of 1 in the ROC plot) represents the performance of a random classifier and an AUC of 1.0 represents a perfect classifier (Fawcett, 2004). Figure 3 shows the ROC plot and AUC for the test data. The algorithm AUC and the lower bound of the confidence interval are higher than 0.50, which suggests that the algorithm is significantly better than random.

Moreover this result indicates that the algorithm is generalizable and that it is detecting meaningful associations in the data. This assertion is supported by secondary measures including accuracy and negative prediction. The algorithm predicts approximately 71% of cases correctly which is significantly better than a random classifier or a null classifier, one that predicts all true or false, which would have accuracies of 50% and 60% respectively.

Variable Importance Analysis

The quality of the algorithm predictions demonstrated by the evaluation criteria suggests that this algorithm identifies meaningful associations between driving data and CPAP use. Thus exploring the variable importance of the features used to train the algorithm can provide valuable insight. Figure 4 shows a plot of the variable importance for the ten most important features in the Random Forest CPAP use algorithm. The features in Figure 4 are identified by their associated variable followed by the appropriate sequence. Thus the most important variable is the sequence, “cccaaa” for the combined acceleration variable. Note that speed is on a scale from ‘a’ to ‘i,’ where ‘a’ represents speeds over 79 kph and ‘i’ represents 0 kph, and that the four acceleration values are on a scale of ‘a’ to ‘d,’ where ‘a’ represents accelerations over 0.1g and ‘d’ represents 0g of acceleration.

Several interesting observations emerge from Figure 4. The most prominent feature, “acc[combined]_cccaaa” represents a 6 second event of accelerations with a 3 second peak acceleration over 0.1 g. This may occur due to quick braking or when a turn or exit ramp is negotiated at high speed. The second most important feature, “speed_ghhhhhhi,” suggests a gradual transition into a stop over 8 seconds. This pattern theoretically should be more common amongst normal drivers as compared to sleepy drivers who may frequently need to brake quickly. Two other intriguing patterns in Figure 4 are “lat-acc[lateral acceleration]_dddaaa” and “latacc_ddddddddda,” both of which represent one second transitions from 0 g acceleration to lateral accelerations over 0.1 g. These patterns may indicate swerves to avoid a lane departure following a micro-sleep.

The patterns observed in Figure 4 lend credence to the hypothesis that events that likely surround common accident types amongst drowsy drivers such as hard swerves and hard (verses gradual) braking can differentiate between OSA patients using their CPAP device and those who are currently not using their device.

Conclusions

The goal of this work was to explore the efficacy of identifying CPAP adherence using kinematic data from naturalistic driving in patients with OSA. The results here suggest that this concept is valid using a sequence frequency feature vector and a Random Forest. The variable importance results from the Random Forest algorithm also demonstrate that driving patterns such as hard swerves and gradual stops are important identifiers in CPAP usage. This observation appears to support the hypothesis that events surrounding common drowsy driving crashes are the main differentiating features between patients who use CPAP and those who do not.

This work demonstrates an innovative approach to detecting effects of CPAP usage on safety critical behavior in the real world. If refined the approach could be used to lessen the burden on healthcare providers charged with monitoring CPAP use, as well as improving patient safety and reducing traffic danger for the general public. Moreover this approach could be used as a preliminary screening tool to identify those suffering from sleep apnea and to evaluate new therapy methods for OSA by analyzing driving data from patients undergoing treatment, in essence employing a personal vehicle as a medical diagnostic device. The patterns identified as important through this algorithm also warrant further investigation as they could provide key insights into drowsiness-related crashes and may be applicable to a wider range of impairments.

References

- Barbe F, Sunyer J, de la Pena A, Pericas J, Mayoralas LR, Anto JM. Effect of continuous positive airway pressure on the risk of road accidents in sleep apnea patients. *Respiration*. 2007; 74(1):44–49. [PubMed: 16804289]
- Breiman L. Random Forests. *Machine Learning*. 2001; 45:5–32.
- Brown ID. Driver Fatigue. *Human Factors*. 1994; 36(2):298–314. [PubMed: 8070794]
- Fawcett. ROC Graphs: Notes and Practical Consideration for Researchers. *ReCALL*. 2004; 31(HPL-2003-4):1–38.
- Findley LJ, Smith C, Hooper J, Dineen M, Suratt PM. Treatment with nasal CPAP decreases automobile accidents in patients with sleep apnea. *American Journal of Respiratory and Critical Care Medicine*. 2000; 161(3):857–859. [PubMed: 10712333]
- Kotsiantis SB, Zaharakis ID, Pintelas PE. Supervised Machine Learning: A Review of Classification Techniques. *Informatica*. 2007; 31:501–520.
- Kribbs NB, Pack AI, Kline LR, Smith PL, Schwartz AR, Schubert NM. Objective measurements of patterns of nasal CPAP use by patients with obstructive sleep apnea. *American Review of Respiratory Disease*. 1993; 147:887–895. [PubMed: 8466125]
- Kushida CA, Littner MR, Hirschowitz M, Morgenthaler TI, Alessi CA, Bailey D. Practice parameters for the use of continuous and bilevel positive airway pressure devices to treat adult patients with sleep-related breathing disorders. *Sleep*. 2006; 29(3):375–380. [PubMed: 16553024]
- Leslie C, Eskin E, Noble WS. The spectrum kernel: A string kernel for SVM protein classification. Paper presented at the Pacific Symposium on Biocomputing. 2002

- Lin, Jessica; Keogh, Eamonn; Wei, Li; Lonardi, Stefano. Experiencing SAX: a novel symbolic representation of time series. *Data Mining and Knowledge Discovery*. 2007; 15(2):107–144.10.1007/s10618-007-0064-z
- MacLean AW, Davies DRT, Thiele K. The hazards and prevention of driving while sleepy. *Sleep Medicine Reviews*. 2003; 7(6):507–527. [PubMed: 15018093]
- McDonald, AD.; Lee, JD.; Aksan, NS.; Dawson, JD.; Tippin, J.; Rizzo, M. The Language of Driving: Advantages and Applications of Symbolic Data Reduction for Naturalistic Driving Data Analysis; Paper presented at the 92nd Annual Meeting of the Transportation Research Board; Washington, D.C. 2012.
- Moreno CRC, Carvalho FA, Lorenzi C, Matuzaki LS, Prezotti S, Bighetti P, et al. Lorenzi-Filho G. High Risk for Obstructive Sleep Apnea in Truck Drivers Estimated by the Berlin Questionnaire Prevalence and Associated Factors. *Chronobiology International*. 2004; 21(6):871–879. [PubMed: 15646234]
- National Highway Traffic Safety Administration. Research on Drowsy Driving. 2011. Retrieved 21 December, 2011, from <http://www.nhtsa.gov/Driving+Safety/Distracted+Driving/Research+on+Drowsy+Driving>
- Orth M, Duchna HW, Leiday M, Widdig W, Rasche K, Bauer TT. Driving simulator and neuropsychological [corrected] testing in OSA before and under CPAP therapy. *European Respiratory Journal*. 2005; 26(5):898–903. [PubMed: 16264053]
- Rauscher H, Popp W, Wanke T, Zwick H. Acceptance of CPAP therapy for sleep apnea. *Chest*. 1991; 100(4):1019–1023. [PubMed: 1914550]
- Sassani A, Findley L, Kryger M, Goldlust E, George C, Davidson T. Reducting motor-vehicle collisions, costs, and fatalities by treating obstructive sleep apnea syndrome. *Sleep*. 2004; 27:453–458. [PubMed: 15164898]
- Tregear S, Reston J, Schoelles K, Phillips B. Obstructive sleep apnea and risk of motor vehicle crash: systematic review and meta-analysis. *Journal of Clinical Sleep Medicine*. 2009; 5(6):573–581. [PubMed: 20465027]
- Tregear S, Reston J, Schoelles K, Phillips B. Continuous positive airway pressure reduces risk of motor vehicle crash among drivers with obstructive sleep apnea. *Sleep*. 2010; 33(10):1373–1380. [PubMed: 21061860]
- Turkington PM, Sircar M, Saralaya D, Elliott MW. Time course of changes in driving simulator performance with and without treatment in patients with sleep apnoea hypopnoea syndrome. *Thorax*. 2004; 59(1):56–59. [PubMed: 14694250]
- Weaver TE, Kribbs NB, Pack AI, Kline LR, Chugh DK, Maislin G. Night-to-night variability in CPAP use over the first three months of treatment. *Sleep*. 1997; 20:278–283. [PubMed: 9231953]
- Yang Y, Pedersen JO. A comparative study on feature selection in text categorization. Paper presented at the Machine Learning-International Workshop then Conference. 1997
- Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. *New England Journal of Medicine*. 1993; 328:1230–1235. [PubMed: 8464434]

Figure 1. Illustration of SAX time-series reduction for a segment of speed data. The three graphs show the raw data, the reduction step, and the symbolic conversion step respectively

Figure 2. Demonstration of three steps of the sequence frequency dataset creation

Figure 3. ROC plot of the test data predicting high versus 0 CPAP use. The gray band is a 2000 replicate bootstrapped confidence interval. The AUC of the curve is shown on the plot along with a bootstrapped confidence interval

Figure 4. Variable Importance pin plot for the 10 most important variables in the Random Forest CPAP prediction algorithm. Note that the raw variable importance has been scaled from 0 to 100 relative to the most important variable