Drowsy Driving Detection System by Analyzing and Classifying Brain Waves

Subha.R , J. Aravind , Vigneshwaran Santhalingam , J. Dhalia Sweetlin
Department of Information Technology
Madras Institute of Technology
Chennai, India.

Abstract — ***The increase in number of accidents is a dangerous situation that needs to be mitigated. If the data from recent past is taken into account, one can see that the number of road accidents has grown exponentially. In fact, roadways record the highest number of accidents in comparison to other modes of transport. Although drowsy driving is not the only contributor, it remains the principal or major issue and can pose a grave threat if it is not averted. This paper thus aims to introduce a method to reduce the number of accidents due to drowsy driving. Using an Electroencephalograph (EEG) headset which consists of multiple EEG sensors and inbuilt Bluetooth transmitter, brain wave data is transmitted to the system. The collected data is fed as input to the prediction model which decides whether the driver is being drowsy or not. The prediction model is trained by using the Auto Regressive and Integrated Moving Average (ARIMA) time series algorithm to predict whether the person is being drowsy. If so , the proposed system can be used to trigger an alarm/warning mechanism.***

Keywords—Drowsy Driving, EEG, Brain Waves, Time Series Analysis, ARIMA model, Classification.

1. INTRODUCTION

Travelling long distances takes a toll on any individual and drivers are no exception. This can lead to exhausted drivers who become drowsy on the road while driving, leading to loss of life and property. The cause of drowsy driving [1] cannot be helped but its effect can still be handled in such a way so as to safeguard valuable human lives. A number of grievous accidents are caused due to this alarming issue. There is a serious call to mitigate these circumstances.

Statistics show that each day a grievous accident happens and most other accidents are deaths of people on two wheelers. Moreover, it has also been shown that Tamil Nadu is the state with the maximum number of road accidents per year [2]. This shows the defects in the current road safety systems. Although there is no denial that the road safety systems in place now do their best bit, it still is not enough considering the current boom in the number of people using personal transport. All these show that the current rules need to be reformatted in favor of new rules and mechanisms to enable more safe travel. Moreover there is no reliable system to help curb such incidents. Though many devices had already been made to prevent drowsy driving and to ensure safety to the driver and the public, they all have some drawbacks and are not completely fool proof. The proposed system uses Electroencephalograph (EEG) [3] to combat accidents by detecting drowsiness.

The rest of the paper is organized as follows: Section 2 presents the existing works in the domain and section 3 deals with the proposed system framework. Experimental results are discussed in section 4 and section 5 presents the conclusion and scope for future work.

2. LITERATURE REVIEW

Some of the articles by eminent scholars on drowsy driving and facts are studied and discussed below :

McIntire et al. proposed a system [4] that estimated eye blink metrics from an eye tracker .The system measured the blink frequency and duration which changed significantly over time during the study and determined eye blinks as an indicator of arousal levels and changes in cerebral blood flow velocities.

Garg et al. proposed an iris recognition system [5] in relation with a heat variation sensor and a camera to determine if the driver is drowsy or not. The driver's eyes are monitored with a camera using image processing system. The variation in the heat in the driver's body is measured with the help of infrared thermal sensor which is used to perceive if the driver is drowsy or not.

Mesharam et al. proposed a system [6] that continuously monitors the head movements of the driver to determine if the driver is falling asleep.The authors propose a technique which utilizes a hybrid geometric and feature based algorithm for head pose estimation, so that both head and eye blink pattern are enough to provide information about the driver's drowsy condition. The proposed method includes face detection, eye region extraction, determining eye blink rate pattern, head postures.

Eskandarian et al. proposed a smart algorithm [7] for determining driver drowsiness detection. The authors describe an experimental analysis of drivers who were subjected to drowsiness conditions in a truck driving simulator and evaluate the performance of a neural network based algorithm which monitors the driver's steering input.

Liu et al. published a review [8] of different methods for estimating driver's drowsiness. The review compared the performance of vehicle based measures like monitoring steering wheel changes and acceleration, subjective measures like driver sleep assessments and physiological methods like head tilt, closing of eyes, Electrooculography (EOG) for determining drowsiness. The review concluded that multiple criteria must be combined for determining if a drowsy or not.

Krajewski et al. proposed a vehicle based measure [9] for estimating drowsiness. The authors describe a steering wheel based monitoring system for determining fatigue. The system monitors the expected changes in steering wheel and its counter correction as a sign of fatigue.

Tawari et al. proposed a system [10] with a distributed camera framework for head movement analysis. The system tracks facial features and their geometric configurations to determine the head pose. The system uses multiple cameras to capture a three dimensional view of the driver's face.

Chutorian et al. proposed a behavioral measure based drowsy detection system [11] by using a head tracking module. The tracking module is used to provide an estimate of the three dimensional motion of the head and uses localized gradient orientation histograms as input for support vector regressors.

Mbouna et al. proposed a system [12] which uses visual analysis of eye state and head pose estimation for monitoring the alertness of the driver. The system uses eye index, pupil activity and head position as estimators for drowsiness. A support vector machine classifier is employed to classify short video segments into alert and non-alert driving events.

Lin et al. proposed a generalized EEG based neural fuzzy system [13] to detect drowsiness. The EEG power spectrum changes are highly correlated with the driver's performance. The authors compared the performance of subject dependent and generalized cross subject prediction models to estimate drowsiness.

Compared to the existing works , the proposed system predicts the state of the driver based on the brain wave frequencies obtained from the collected data. The extracted components alpha , beta , delta ,theta from the EEG signal are analyzed using time series analysis which can help predict if the driver is drowsy by analyzing the historical data extracted from the EEG signal. This can be used to alert the driver even before he gets drowsy.

The manuscript follows a sequential process of data collection from the EEG sensor. The collected EEG data is filtered to remove noise and bad channels. The filtered data is divided into epochs and fast fourier transform (FFT) is applied to obtain the different brain wave frequencies. Time series analysis of the frequencies is used to determine the driver's state. From the literature survey it is seen that there is no similar analysis of brain waves using time series algorithms and hence this article aims to help prevent instances of drowsy driving.

3. SYSTEM FRAMEWORK

The framework of the proposed system is presented in Figure 1. The modules in this framework are data acquisition and sampling, data preprocessing which includes noise/bad channel filtering, maxwell filtering , fast fourier transform on the filtered data and training and prediction using ARIMA time series model.

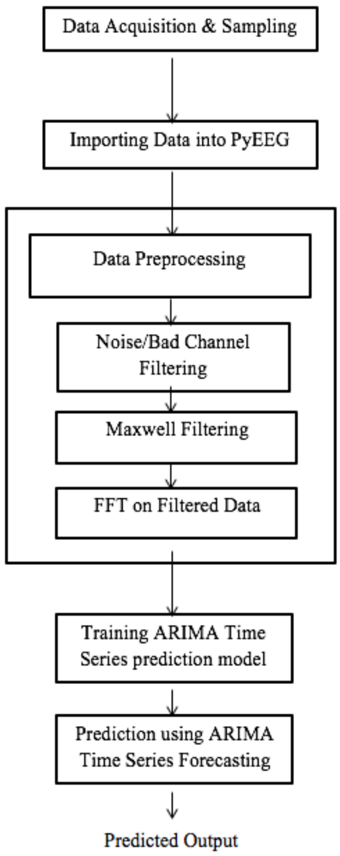

Figure 1: System Framework

3.1 DATA ACQUISITON

The data is collected via an EEG headset called NeuroSky Mindwave which uses the TGAM1(ThinkGear ASIC Module) module, a single AAA Battery with automatic wireless pairing. It uses Bluetooth V2.1 to transfer the sampled data into the system and has a static Headset ID (headsets have a unique ID for pairing purposes). The headset measures raw brainwaves and includes EEG signal quality analysis which is used to detect if the device is off the head or poor contact.

The sensor outputs a single continuous EEG signal is sampled at a rate of 500/s and transferred to the system using the inbuilt bluetooth transmitter in the sensor. The sampled data is stored and the features are extracted using this sampled signal and python libraries MNE Python [14], PyEEG [15] packages. Some of the features include: High Alpha, Low Alpha, Beta, Theta, Delta, Power Spectral Intensity (PSI), Relative Intensity Ratio(RIR), Petrosian Fractal Dimension(PFD), Higuchi Fractal Dimension(HFD), Hjorth mobility and complexity, Spectral Entropy, SVD Entropy, Fisher information, Approximate Entropy (ApEn), Detrended Fluctuation Analysis(DFA) and Hurst Exponent.

PyEEG uses standard NumPy [16] arrays to store data that correspond to EEG. The datasets that are obtained from drowsy, conscious persons are imported into PyEEG using these NumPy arrays. Numpy arrays are tables of elements which are homogeneous, indexed by a tuple of positive integers.

3.2 DATA PREPROCESSING

Data preprocessing is performed on the input data obtained from EEG sensor. The EEG sensor readings are susceptible to noise and head movements which cause changes in the value measured. To remove sources of external interference , data preprocessing is performed which includes noise/ bad channel filtering , maxwell filtering and fast fourier transform on the filtered data.

3.2.1 Noise Removal / Bad Channel Filtering

Filtering can help to select certain frequencies, such that either some frequencies are removed, or possibly that some filters remain. There are a number of types of filters [17]:

Low-pass filter: 'Low' frequencies below a certain value are kept, while high frequencies are removed. This is also known as a high-cut filter. It may help to think of the audio version of this, which would be something that removed all the high notes from a sound.

High-pass filter: The same as above, but only high frequencies remain, and only those below a certain value are removed.

Band-pass filter: This filter selects only frequencies between a lower and upper bound. The band-cut filter which functions opposite to that of the band-pass filter, which removes all frequencies in a particular range.

Notch filter: This is a special type of band-cut filter, that removes a single frequency. It is also possible to combine multiple notch filters, to remove a particular set of single frequencies, useful for things like removing electricity noise.

3.2.2 Maxwell Filtering

Maxwell filtering [18] is used as a pre-processing step in this manuscript. Since the data acquisition source is an EEG sensor connected to a subject who is either conscious or becoming drowsy, head movements and noise are common. However, it is mandatory to remove the effect of these head movements and interference in order to get a good result. In mne-python, Maxwell filtering helps to attenuate external interference and impact of driver's head movements.

3.2.3 Fast Fourier Transform

The Fast Fourier Transform (FFT) algorithm [19] is used to divide a sampled signal into its constituent frequency components. PyEEG is used to perform Fast Fourier Transform on the filtered data to obtain the required alpha, beta, gamma, theta components as shown in Figure 2. The component values in the EEG are normalized to obtain the comparable values.

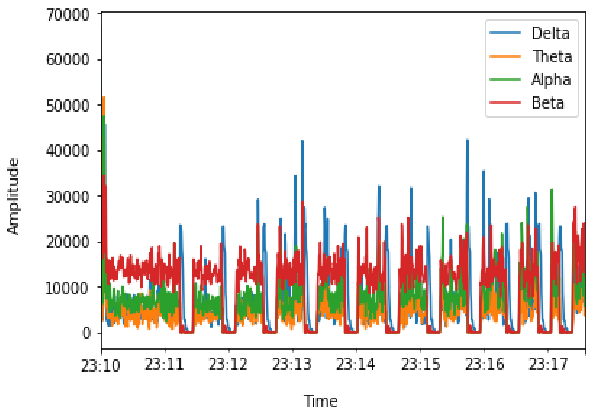

Figure 2: Raw signal split into components

3.3 TRAINING AND PREDICTION USING ARIMA MODEL

A time series is a sequence of data points taken at equally spaced time points and indexed based on time order. Thus it is a sequence of discrete-time data as shown in Figure 3. Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Time series prediction uses a model to forecast values based on historical values.

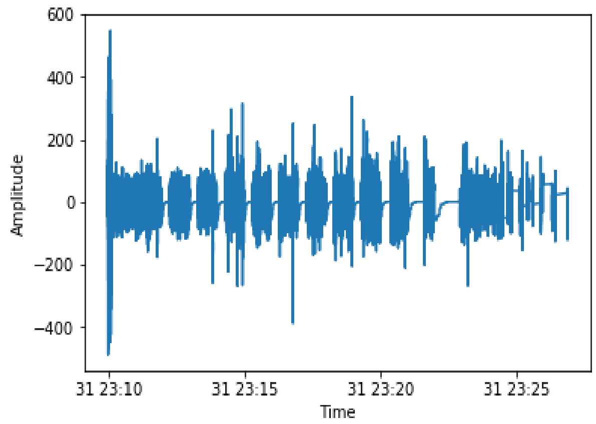

Figure 3: The raw input time series

The Auto Regressive Integrated Moving Average (ARIMA) time series model, which is a standardization of Auto Regressive Moving Average (ARMA) model is applied in some cases where data show evidence of non-stationarity. Since the data from the brain signals are capable of showing unsteady variations, the ARIMA model would be better suited for the application. The ARIMA model has been applied to the dataset and checked for stationarity based on the Dickie-Fuller test [20].

4. RESULTS AND DISCUSSION

The modules proposed in the system include preprocessing raw EEG data (noise/bad channel removal, maxwell filtering, FFT) and time series analysis of the sampled EEG data. The output is the expected values of alpha, beta, gamma, theta for the next epoch which can be used to determine the driver's state. The system focuses on analysis of the EEG data and identifying useful patterns in the EEG data based on time series analysis. The advantage of the system is that it uses EEG to extract driver information, since EEG is unambiguous and cannot be duplicated by any other physiological aspect of the person. Figure 4 shows the plot of the extracted beta wave. The extracted signal values are shown in Figure 5.

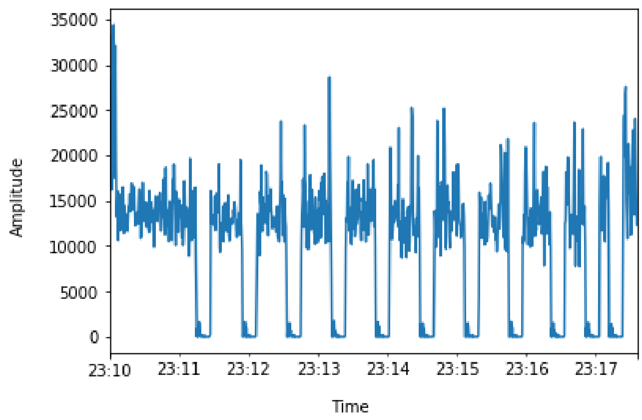

Figure 4: Beta wave plot

Out[24]:

	Delta	Theta	Alpha	Beta
2017-12-19 23:10:00.000	66967.607213	2.629728e+04	2.970435e+04	3.318235e+04
2017-12-19 23:10:00.500	44627.953217	4.292298e+04	2.117908e+04	2.136383e+04
2017-12-19 23:10:01.000	14499.432685	3.110694e+03	7.624004e+03	2.050524e+04
2017-12-19 23:10:01.500	6031.714406	2.531559e+03	7.536532e+03	1.621922e+04
2017-12-19 23:10:02.000	2712.837504	3.843056e+03	6.361775e+03	2.072157e+04
2017-12-19 23:10:02.500	20338.295479	2.694895e+04	2.751465e+04	2.846539e+04
2017-12-19 23:10:03.000	17313.665628	2.992978e+04	4.161394e+04	2.900749e+04
2017-12-19 23:10:03.500	25656.378576	5.162075e+04	4.747597e+04	3.442705e+04
2017-12-19 23:10:04.000	45667.456683	2.679977e+04	3.544084e+04	1.755379e+04
2017-12-19 23:10:04.500	26436.479182	1.433437e+04	8.911995e+03	2.554039e+04
2017-12-19 23:10:05.000	14062.121775	7.856979e+03	1.507807e+04	3.213525e+04
2017-12-19 23:10:05.500	5834.037953	5.797267e+03	7.472385e+03	1.945855e+04
2017-12-19 23:10:06.000	2612.794732	6.886028e+03	8.098902e+03	1.319407e+04
2017-12-19 23:10:06.500	5827.901348	3.670407e+03	1.044599e+04	1.479116e+04
2017-12-19 23:10:07.000	5090.386673	3.104800e+03	1.144903e+04	1.608707e+04
2017-12-19 23:10:07.500	1586.465786	3.548717e+03	7.533779e+03	1.063819e+04
2017-12-19 23:10:08.000	1706.871728	3.369785e+03	6.475098e+03	1.384858e+04
2017-12-19 23:10:08.500	7290.667021	3.239402e+03	4.610035e+03	1.444662e+04
2017-12-19 23:10:09.000	5363.391550	4.295825e+03	1.079469e+04	1.568237e+04
2017-12-19 23:10:09.500	1536.458513	3.495787e+03	8.502761e+03	1.174234e+04
2017-12-19 23:10:10.000	2138.698514	3.078979e+03	5.727216e+03	1.183240e+04

Figure 5: Extracted values from raw EEG signal

Time series analysis is performed on the beta wave extracted, as it is known that these waves show the attentiveness and level of consciousness of a person. The data is stored and contains about 350,000 records out of which about 35,000 records are used for training and the remaining is used for testing purposes. In the initial forward validation testing, the results of the validation testing are shown in Figure 6.

```
>Predicted=34.200, Expected= 44
>Predicted=44.000, Expected= 43
>Predicted=42.600, Expected= 37
>Predicted=37.000, Expected= 35
>Predicted=34.900, Expected= 35
>Predicted=35.100, Expected= 33
>Predicted=33.200, Expected= 29
>Predicted=28.900, Expected= 25
>Predicted=24.900, Expected= 23
>Predicted=22.900, Expected= 20
>Predicted=20.400, Expected= 13
>Predicted=13.200, Expected=  4
RMSE: 5.905
```

Figure 6: Prediction results – Walk Forward Validation

The ARIMA forecasting model is used for the prediction and to get better performance as shown in the Figure 7.

```
[[ 28.60000038  27.64970024]]
[[ 29.70000076  30.33247721]]
[[ 32.09999847  28.98689715]]
[[ 36.59999847  32.16987157]]
[[ 38.70000076  37.90679496]]
[[ 37.29999924  38.26971913]]
[[ 35.70000076  34.53575694]]
[[ 34.59999847  32.87028759]]
```

Figure 7: Prediction and expected values – ARIMA model

The ARIMA gives an RMSE of 0.59 which corresponds to a better prediction than walk forward validation. This is because walk forward validation does not consider the temporal relationship between the time series data and randomly splits the observations into groups without considering the time factor whereas ARIMA models consider the variations in the data due to time and remove non-stationarity in the time series. From the Figure 8, it is seen that there is a high correlation between predicted and expected output for the time series in ARIMA model.

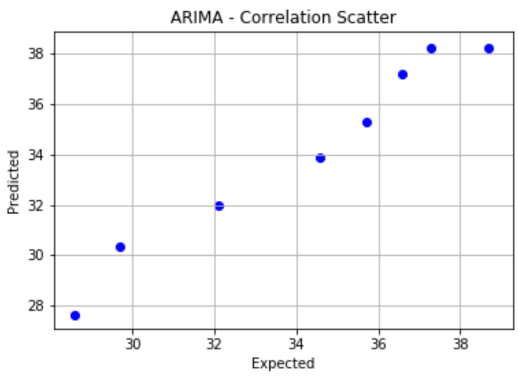

Figure 8: The scatter plot of expected and predicted output

Proposed Method	RMSE
Walk Forward Validation	5.905
ARIMA	0.59

Table 1: Performance Metric – Proposed Method

The table 1 shows the RMSE of the proposed techniques namely walk forward validation and ARIMA forecasting model.

Previous Methods	RMSE
Smart Algorithm	0.2
Mobile phone based	0.65

Table 2: Performance Metric – Previous Methods

5. CONCLUSION

The system would help to tackle drowsiness which is one of the major contributors to road accidents. It uses Electroencephalograph(EEG) to determine the driver's state to prevent accidents caused due to drowsy driving. This can help to reduce/mitigate the destruction to life and property.

Signal Processing has been a widely researched topic and complete accuracy in separating signals has not yet been completely achieved. With better separation procedures, the accuracy of the current system can be improved. In this system, the PyEEG module is used to split the waves into its constituent Alpha, Beta, Delta and Theta components and a time series is used to fit a model on the input data. Since a time series is used, it is only possible to train the model using a single raw signal converted into its constituents.

However, in the future, the signals can be trained using a Recurrent Neural Network (RNN) and Long Short Term Memory (LSTM). The neural network architecture can use multiple raw signals to train the system. The model can then be used to test data. The system can be extended to include solutions for several other road accident causes. Processing of the sampled EEG data can be done over the cloud to ensure a relatively non-intrusive technique. This system can be implemented as a product which can be manufactured and can be mass produced if a low cost sensor can be found to retrieve brain waves.

REFERENCES

[1] Drowsy Driving Effects and Statistics : http://drowsydriving.org/about/facts-and-stats/

[2] Tamil Nadu accident statistics over the years : http://www.tn.gov.in/sta/ra1.pdf

[3] EEG : https://www.nhs.uk/conditions/electroencephalogram/

[4] L. McIntire ,A. McKinley , C. Goodyear and J. McIntire , “ Detection of vigilance performance using eye blinks” , *Elsevier- Applied Ergonomics*, vol 45 ,no 2 , pp .354-362 , March 2014

[5] R.Garg ,V.Gupta and V.Agarwal , “A Drowsy Driver Detection and security system” , *IEEE International Congress on Modern Telecommunications and Control Systems (ICUMT)* , pp 1-8, 2009

[6] P.Meshram, N. Auti and H.Agrawal , “Monitoring Driver Head Postures to Control Risks of Accident” , *Elsevier-Procedia Computer Science* , vol.50 , pp. 617-622 , 2015

[7] A.Eskandarian and A.Mortazavi, “Evaluation of a Smart Algorithm for Commercial Vehicle Driver Drowsiness Detection”, *Proceedings of IEEE Intelligent Vehicles Symposium* , June 13-15, 2007

[8] C.Liu ,S.Hosking and M.Lenne , “Predicting driver drowsiness using vehicle measures: Recent insights and future challenges”, *Elsevier – Journal on Safety Research*, vol .40 ,no. 4, pp. 239-245, 2009

[9] J. Krajewski, D. Sommer, U. Trutschel, D. Edwards and M. Golz,"Steering Wheel Behavior Based Estimation of Fatigue", *Proceedings of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, June 22-25, 2009

[10] A. Tawari, S. Martin, and M. M. Trivedi,” Continuous Head Movement Estimator for Driver Assistance: Issues, Algorithms and On-road Evaluation”, *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 2,pp. 818-830, 2014.

[11] E. Murphy-Chutorian and M. M. Trivedi, “Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 300–311, 2010.

[12] R. O. Mbouna, S.G. Kong and M.G. Chun,” Visual Analysis of Eye State and Head Pose or Driver Alertness Monitoring”, *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3,pp. 1462-1469, 2013.

[13] F.C. Lin, L.W. Ko, C.H. Chuang, T.P. Su and C.T. Lin , "Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system", *IEEE Transactions on Circuits and Systems*, vol. 59, no. 9, pp. 2044-2055, 2012.

[14] MNE–Python : https://martinos.org/mne/stable/index.html

[15] PyEEG : http://pyeeg.sourceforge.net/

[16] Python Toolkit – NumPy : http://www.numpy.org/

[17] Python-Filters : https://docs.scipy.org/doc/scipy-0.18.1/reference/signal.html

[18] MNE-Python Maxwell Filtering : https://martinos.org/mne/dev/manual/preprocessing/maxwell.html

[19] Python-Fast Fourier Transform : https://docs.scipy.org/doc/numpy1.13.0/reference/routines.fft.html

[20] Dickie Fuller Test : http://www.statsmodels.org/dev/generated/statsmodels.tsa.stattools.adfuller.html