Excessive Alcohol Craving Prediction Algorithm Using Smartphone Accelerometer Sensor

Prakan Sornbootnark
Digital Technology Innovation,
Faculty of Science, Maejo University,
Chiang Mai, Thailand
earth.prakan@gmail.com

Paween Khoenkaw
Digital Technology Innovation,
Faculty of Science, Maejo University,
Chiang Mai, Thailand
paween_k@gmaejo.mju.ac.th

***Abstract*— This research presents the development of the algorithm that predict excessive alcohol consumption by using sensors on smartphones in order to solve problems of alcohol drinkers who is unaware of their own alcohol consumption level. Since this incident results in several problems such as, road accidents. Therefore, this algorithm was developed to replace Breathalyzer, which most people do not have. This algorithm will use data from the sensors on smartphones, which most people already have in possession, to forecast alcohol over dosage. The experiment conducted from 6 people of 23 sample groups, which is 100% accurate, revealed that this algorithm could be used instead of the Breathalyzer in a certain extent.**

Keywords— Excessive Alcohol, Intoxicate, Signal processing, Accelerometer sensor

I. Introduction

Excessive alcohol consumption is the leading cause of several deaths worldwide (population decline). Thus, part of it is caused by, drunk driving or having excessive alcohol intake while driving a vehicle, which causes road accidents. So, the problem stems from the alcohol drinkers' inability to recognize that they had consumed an excessive amount of alcohol. Since, other than using the Breathalyzer to measure the level of alcohol through one's breath, there is no other way to find out one's level of alcohol. Therefore, this article presents an algorithm for predicting alcohol over dosage from accelerometer sensor on smartphones. Then, the obtained data will be processed through the signal processing method. This algorithm has been developed using sensor data obtained from the body movement while sitting and drinking alcohol.

II. Literature review

A. *Effects of alcohol on human*

One of the effects of ethanol intoxication is a general impairment of static balance control, characterized by an increased postural sway and the inability to coordinate postural and voluntary activity.

The average sway area and sway path changed in accordance with the Blood alcohol content, especially for the large alcohol dose group. There was a significant positive correlation between sway area increase and Blood alcohol content both. Little change was found in the small alcohol dose groups. Thus, it can be postulated that body sway does not increase significantly when the Blood alcohol content is less than 0.8%. [1]

Acute alcohol intoxication is associated with numerous health risks. For example, impaired driving due to alcohol was implicated in 28% of the 38,000 deaths from motor vehicle accidents in the US in 2016. These consequences largely stem from alcohol's detrimental effects on psychomotor performance. One measure of psychomotor performance that is particularly sensitive to alcohol is gait. Gait requires coordination of multiple sensory and motor systems. Both postural stability and gait are sensitive to blood alcohol concentration (BAC) levels. [2]

Shunpei Ando and colleagues have proposed acute alcohol ingestion is suggested to result in a large sway of low frequency in the medio-lateral direction. Not only neuromotor function but also attention and cognitive function are adversely affected by acute alcohol ingestion, both of which have implications for the ingestion of alcohol before driving a car. [3]

B. *Detecting blood alcohol levels from gait*

Andrew McAfee and colleagues have proposed the anticipated level of alcohol in the blood from walking (gait analysis), in which, the data is stored through accelerometer sensor and Gyroscope sensor on the smartphone of the alcohol drinker. The data is processed by using the method of Machine Learning, which resulted in an accuracy of 89.45%. [4]

Zachary Arnold and colleagues have proposed smartphones could infer the alcohol intoxication levels (how many drinks) of their users based on anomalies in their gait. Time and frequency domain features were extracted from accelerometer data of drinkers and used for classification in a machine learning framework. For a task of classifying the number of drinks consumed by a user into ranges of 0-2 drinks (sober), 3-6 drinks (tipsy) or >6 drinks (drunk), Random Forest yielded 56% accuracy on the training set, and 70% accuracy on the validation set. [5]

The previous research was conducted simply through alcohol detection while walking only. In reality, those alcohol drinkers might not be able to find an open space without any obstructions. Therefore, this research proposes an alcohol detection while the drinker is casually sitting and drinking alcohol instead. So, the magnitude will be measured while the drinkers swing their bodies.

III. Method

The proposed model is based on a sensor exist in smartphone, the overview of the system is shown in Fig.1

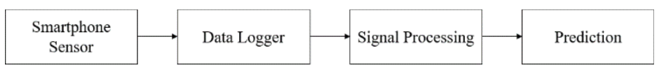

Fig. 1. System Overview

A. Smartphone Sensor

The accelerometer sensor is used in this method, this sensor is used to sense the G-force. The smartphone used this sensor to determine the screen orientation. The shaking of the device also can be measure using this sensor.

The user must keep the smartphone in the pants pocket while the sensor is pulled. The sensor data is pulled collected 20 samples in 1 second (sampling-rate of 20 Hz).

B. Data Logger

The data logger is a process running on a smartphone to collected the data pulled from the accelerometer sensor. The data is stored locally with a timestamp. Since the smartphone only return the sensor data only when the sensor value is changing, then the timestamp data is used to reconstruct the data to the time-series signal. The example of raw sensor data collected in this process is shown in Table I.

TABLE I. SAMPLE SENSOR DATA OF ALCOHOL DRINKERS

Timestamp	X-Axis	Y-Axis	Z-Axis
1532442954	-11.381436	8.072051	0.45130703
1532442954	-11.106103	3.8750155	-7.28376
1532442955	-9.452309	3.803788	-1.4634558
1532442955	-8.194754	2.5833437	-1.3700819
1532442955	-7.425018	1.6059108	-1.9189527
1532442955	-8.280347	3.3333273	-2.9652188
1532442955	-7.367557	1.8112136	-5.768829

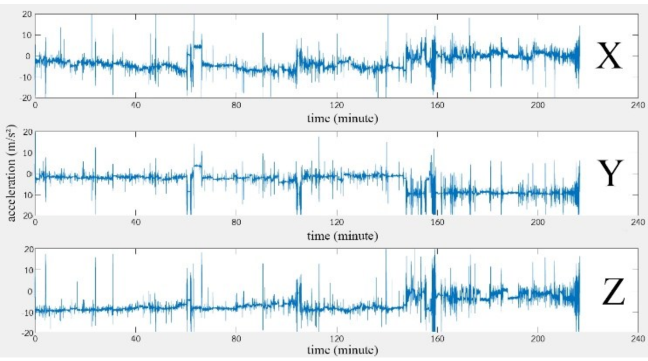

Fig. 2. Sample data sampling from accelerometer sensor.

C. Signal Processing

The signal processing module is used to analyze the raw sensor data collected from the smartphone to determine that the user is in a state of alcohol consuming or not.

The raw sensor data contains 3 axis of x, y, z. This data is varies deepened on the smartphone position, to standardize this signal the normalization process is used. This process combined all 3 axis into one signal that doesn't varies on the position of smartphone. The normalization signal is calculated based on (1), and the sample normalized signal is shown in the Fig. 3.

$$b_t = \sqrt{x_t^2 + y_t^2 + z_t^2} \tag{1}$$

- b is the normalized data.
- t is the time position.
- x is X-axis raw data from the accelerometer.
- y is Y-axis raw data from the accelerometer.
- z is Z-axis raw data from the accelerometer.

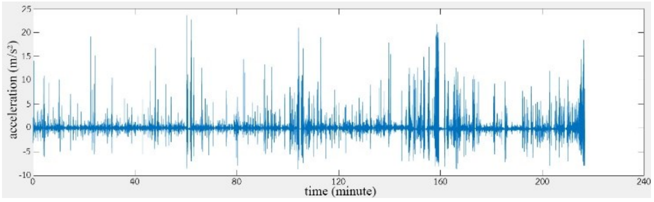

Fig. 3. The example of normalized data.

Then the Rate of Change is computed using (2), the example signal is shown in the Fig. 4.

$$r = |\Delta b| \tag{2}$$

- Δb is the rate of change of b.
- r is the result from computing the rate of change and convert into the absolute value.

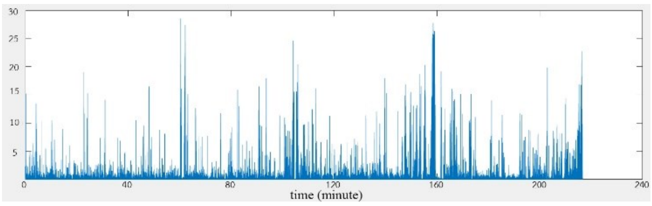

Fig. 4. Data from computing the rate of change and convert into the absolute value.

The rate of change signal is contained many small spike data. This spike is causing by the high sensitive of the sensor, this result in the sensor capturing the unwanted signal. For example, the small vibration causing by the breathing, some disturbance sound (noise) such as natural movement of the body are all result in a small spike in the signal. Thus, the Moving Average was calculated in order to reduce this noise by using (3) as shown in the Fig. 5.

$$m_t = \frac{\sum_{i=t}^{t+w} r_i}{w} \tag{3}$$

- m is the result of removing the noise.
- w is the window size of the Moving Average algorithm.

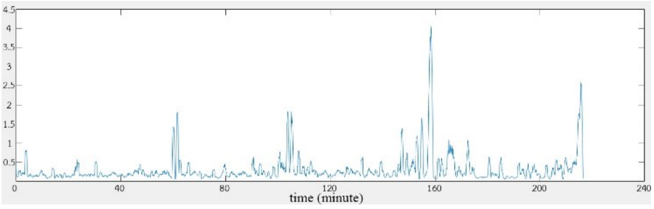

Fig. 5. The example of noise removed data.

According to the result in the Fig. 5, we had investigated the data gathering from the alcohol drinker sample. It was found that through the later of the signal, the over alcohol drinkers show higher value when they consumed more alcohol. In contrast, data from the under alcohol drinkers seemed to be less and stable. This was due to the over alcohol consuming caused the higher value in the graph. Therefore, the threshold l is used which represent the over alcohol consuming. It was the reference line in the graph (red line). In order to identify the alcohol consuming type, the Fig. 6 presented the graph of the over alcohol consuming samples and the Fig. 7 presented the under alcohol consuming samples.

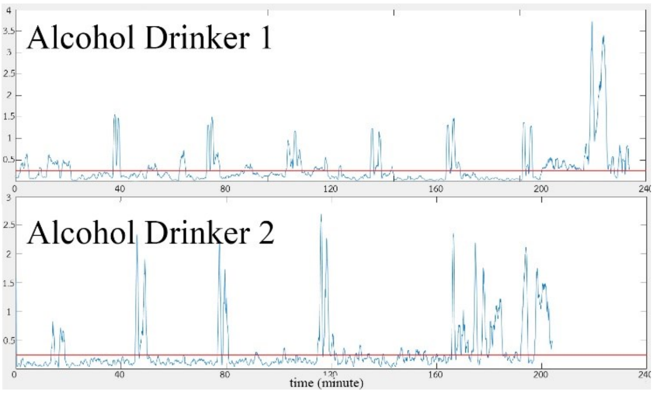

Fig. 6. Data of the over alcohol consuming drinkers.

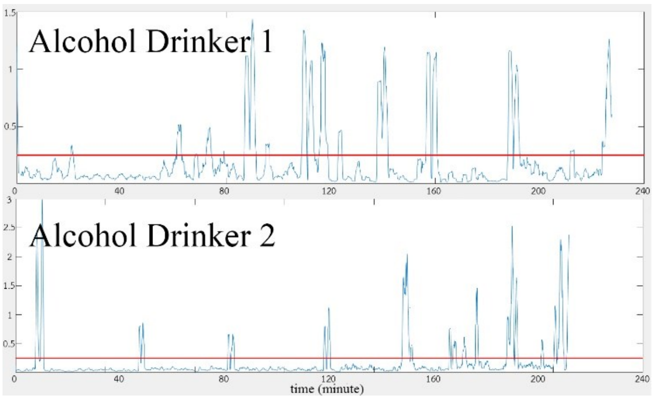

Fig. 7. Data of the under alcohol consuming drinkers.

Finally, the number of sample that exceed threshold l is used to identify the over alcohol consuming. It started from searching from the beginning to the end of the graph by moving 1 minute at a time on the horizontal. this range of this search is defined as m. The exceed values were in the green areas as shown in the Fig. 8, where range of m is shown in vertical green line.

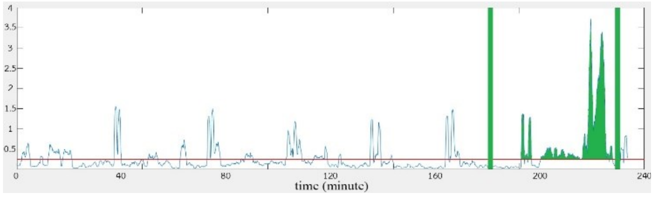

Fig. 8. The example of a search window.

The percentage of sample the exceed l is then calculated, as in (4).

$$p_t = \frac{f}{m} \times 100 \quad (4)$$

- p is the percentage of the value that exceed the l value.
- f is number of samples that exceed l.
- m are the size of search windows.

The p is then classified as over consuming or not base by comparing with the predefined threshold, as in (5).

$$s_t = \begin{cases} 1 & where\ p_t \geq \theta \\ 0 & where\ p_t < \theta \end{cases} \quad (5)$$

Where θ is the threshold to defined the over consuming value.

IV. Experiment and Results

The proposed model is evaluate by varies the θ from 40 to 50, the m is ranging from 12,000 to 36,000 , and l from 0.125 to 0.4. The dataset was acquired from 6 volunteers; the sensor data was recorded from start drinking until overconsuming. The Samsung Galaxy Note 8 used in this experiment. The blood alcohol concentration (BAC) of each volunteer was tested every 30 minute using a breathalyzer. The volunteer that has BAC value greater then 0.15 mg/l is considered as the overconsuming, this data is then used as a ground truth to evaluated the proposed model.

The results is shown in Fig. 9, Fig. 10 and Fig. 11, it is demonstrated that every m value can generated 100% accuracy if using the correct parameters. However, the parameter m is directly effecting the response time of the system, because it is determined the number of sample that required by this prediction model.

In this experiment, the recommend m value is 18,000 because it is using lowest number of samples while can still producing the result with 100% accuracy.

TABLE II. THE ACCURACY OF PROPOSED PREDICTION MODEL ($\theta = 40$)

m \ l	**0.125**	**0.15**	**0.175**	**0.2**	**0.225**	**0.25**	**0.275**	**0.3**	**0.325**	**0.35**	**0.375**	**0.4**
12,000	0.68	0.73	0.77	0.82	0.82	0.82	0.87	0.87	0.87	0.82	0.87	0.91
18,000	0.77	0.77	0.77	0.82	0.91	0.91	0.91	0.91	0.91	0.96	0.91	0.91
24,000	0.77	0.77	0.82	0.87	0.91	0.91	1	0.96	0.91	0.91	0.91	0.91
36,000	0.77	0.77	0.82	0.87	1	1	0.91	0.91	0.91	0.87	0.87	0.87

Fig. 9. The accuracy of proposed prediction model ($\theta = 40$).

TABLE III. THE ACCURACY OF PROPOSED PREDICTION MODEL ($\theta = 45$)

m \ l	**0.125**	**0.15**	**0.175**	**0.2**	**0.225**	**0.25**	**0.275**	**0.3**	**0.325**	**0.35**	**0.375**	**0.4**
12,000	0.68	0.73	0.77	0.82	0.82	0.87	0.87	0.91	0.87	0.91	0.91	0.87
18,000	0.77	0.77	0.87	0.87	0.96	1	0.96	0.96	0.91	0.91	0.91	0.91
24,000	0.77	0.77	0.87	0.91	0.96	1	0.91	0.91	0.91	0.87	0.87	0.87
36,000	0.77	0.77	0.91	0.96	0.96	0.96	0.91	0.87	0.87	0.87	0.87	0.87

Fig. 10. The accuracy of proposed prediction model ($\theta = 45$).

TABLE IV. THE ACCURACY OF PROPOSED PREDICTION MODEL ($\theta = 50$)

m \ l	**0.125**	**0.15**	**0.175**	**0.2**	**0.225**	**0.25**	**0.275**	**0.3**	**0.325**	**0.35**	**0.375**	**0.4**
12,000	0.73	0.73	0.82	0.82	0.87	0.91	0.96	0.96	0.91	0.91	0.91	0.91
18,000	0.77	0.77	0.91	0.96	0.96	**1**	0.91	0.91	0.91	0.91	0.91	0.91
24,000	0.77	0.82	0.91	0.96	**1**	0.91	0.91	0.91	0.87	0.87	0.87	0.87
36,000	0.77	0.82	0.96	0.96	0.96	0.91	0.91	0.91	0.87	0.87	0.82	0.77

Fig. 11. The accuracy of proposed prediction model ($\theta = 50$).

V. CONCLUSION

This research presents an algorithm for predicting excessive alcohol consumption from the accelerometer sensor data that is obtained while the user drinks alcohol. The sensor data is recorded using the smartphone and send to the signal processing unit to analyzed the predicted the over consuming. The recommend parameters are proposed and evaluated and compared the results with the ground-truth obtained using a breathalyzer. The results shown that this algorithm is produced 100% accuracy. However, this method is not able to produce the accuracy results in real-time, the 15 minutes lags are required in order to produce the high accuracy prediction.

ACKNOWLEDGMENT

The authors would like to thank the colleagues at INTNIN Laboratory, Computer Science Department, Faculty of Science, Maejo University for their support.

REFERENCES

[1] M. Nieschalk, C. Ortmann, A. West, F. Schmäl, W. Stoll, and G. Fechner, “Effects of alcohol on body-sway patterns in human subjects,” 1999 International Journal of Legal Medicine

[2] B. Suffoletto, P. Gharani, T. Chung, H. Karimi, “Using phone sensors and an artificial neural network to detect gait changes during drinking episodes in the natural environment,” 2018 Gait & Posture 60, 116-121.

[3] S. Ando, T. Iwata, H. Ishikawa, M. Dakeishi, K. Murata, “Effects of acute alcohol ingestion on neuromotor functions,” 2008 NeuroToxicology 29, 735–739

[4] A. McAfee, J. Watson, B. Bianchi, C. Aiello, E. Agu, “AlcoWear: Detecting Blood Alcohol Levels from Wearables,” 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications,Cloud & Big Data Computing, Internet of People and Smart City Innovation.

[5] Z. Arnold, D. LaRose, E. Agu, “Smartphone Inference of Alcohol Consumption Levels from Gait,” 2015 International Conference on Healthcare Informatics.