AlcoWear: Detecting Blood Alcohol Levels from Wearables

Andrew McAfee, Jacob Watson, Ben Bianchi, Christina Aiello, Emmanuel Agu
Computer Science Dept, Worcester Polytechnic Institute, Worcester, MA 01609
{amcafee, jrwatson, bgbianchi, cjaiello, emmanuel}@wpi.edu

***Abstract*—Alcohol abuse causes 88,000 deaths annually. Alcohol affects neuromotor functions such as walk patterns, making them a reliable bio-measure of intoxication. In this paper, we present AlcoWear, a machine learning based system that passively senses a drinker's Blood Alcohol Content (BAC) by classifying accelerometer and gyroscope data gathered from their smartphone and smartwatch. While gait sensor readings taken from a device attached to the user's trunk (smartphone) are the most accurate, users often do not carry their phones (e.g. leave them on a table) while walking around during their day. Smartwatches are worn continuously but are less accurate due to noisier sensor readings (e.g. confounding hand gestures). AlcoWear extracts and classifies features such as sway area (gyroscope) and cadence (accelerometer) from smartphone sensor data, and features such as Total Harmonic Distortion (accelerometer) and angular velocity (gyroscope) from the same users' smartwatches. On the smartphone, the J48 classifier was the most accurate, classifying user gait patterns into BAC ranges of [0.00-0.08), [0.08-0.15), [0.15-0.25), [0.25+) with an accuracy of 89.45%. For the smartwatch, AlcoWear classifies users as being in BAC ranges [< 0.08 vs >= 0.08] (2 bins) with a 79.8% accuracy using a Random Forest classifier. AlcoWear classifies the smartphone's data when the user carries it, and uses data from the user's smartwatch when the user is not carrying their phone. AlcoWear is the first to combine the smartphone and smartwatch in a collaborative system to sense BAC from gait.**

***Keywords*—intoxication, alcohol, detection, smartwatch, smartphone, machine learning, mobile sensing**

I. INTRODUCTION

Alcohol abuse results in physical harm and mental malfunction [10] and is responsible for 1 in 10 deaths among adults aged 20-64 years in the United States annually [19]. Moreover, binge drinking (defined as 4 or more drinks for women on a single occasion, and 5 or more drinks for men on a single occasion [19]) has been on the rise. Between 2002 and 2005, over a third of college students aged 18-20 reported binge drinking in the prior month [17]. Even though it's well known that alcohol impairs driving ability, many people frequently drive when drunk. In 2010, 47.2% of pedestrian fatalities and 39.9% of vehicle occupant fatalities were caused by drunken driving [18]. However, in many Driving Under the Influence (DUI) cases, the drinker is unaware that they are over the legal driving limit.

Alcohol consumption raises the Blood Alcohol Content (BAC) of the drinker's blood [24], impacting their neuromotor and cognitive functions approximately 20 minutes after alcohol consumption [25]. Gait, or the manner in which a person walks is one of the neuromotor functions affected by alcohol consumption. In fact, aside from direct BAC or Breath Alcohol Concentration (BrAC) testing, neuromotor testing including analysis of gait is the most reliable way to determine intoxication in humans [8]. Leveraging this fact, the walk-and-turn field sobriety test used as a first screen by the police in 70 percent of DUI cases, is based on gait assessment.

In this paper, we present AlcoWear, a system that collaboratively uses a drinker's smartphone and smartwatch to passively sense their intoxication level from their gait (walk). AlcoWear infers the drinker's BAC level by classifying accelerometer and gyroscope sensor features gathered from their smartphone and smartwatch simultaneously using a machine learning approach (See Figure 1). Gait measurements taken from the user's trunk (smartphone) are the most accurate, but users carry their phones on them only about 53% of the time, as they often leave their phones (e.g. on the table or in a bag) [28] while walking around during their day. Smartwatches are worn continuously but have noisier sensor readings (e.g. confounding arm gestures).

Figure 1 AlcoWear Intoxication Sensing (Smartwatch + Smartphone)

AlcoWear infers users' BAC levels by classifying features such as sway area (gyroscope) and cadence (accelerometer) extracted from the user's smartphone, and additional features such as Total Harmonic Distortion that are also extracted from the same users' smartwatches. On the smartphone, the J48 classifier was the most accurate, classifying user gait patterns into BAC ranges of [0.00-0.08), [0.08-0.15), [0.15-0.25), [0.25+) with an accuracy of 89.45%. For the smartwatch, AlcoWear classifies users as being in BAC ranges [< 0.08 vs >= 0.08] (2 bins) with a 79.8 percent accuracy using a Random Forest classifier.

Passive methods to monitor intoxication can be used lto prevent mishaps as well as treat hard drinkers (addictions). *Use in prevention:* Drinkers who are over the legal driving limit can receive just-in-time notifications of excessive alcohol consumption, preventing drunk driving. *Use in treatment:* a frequent drinker's drinking patterns and associated contexts (e.g. time, place, who with) can be logged continuously. Drinkers can reflect on their drinking patterns of abuse and either self-correct or seek treatment. Counselors can use such logs as evidence to prescribe treatment.

II. RELATED WORK

Intoxication detection devices: SCRAM Continuous Alcohol Monitoring [4] is an ankle-worn commercial alcohol detection device that samples the user's perspiration every 30 minutes in order to measure their BAC levels. *Kisai Intoxicated LCD Watch:* is a watch that has a built-in breathalyzer on its side [2]. Users breathe into its Breathalyzer, from which the watch determines and displays graphs of the user's BAC level. SCRAM and Kisai are dedicated devices that must be purchased and worn, unlike smartphones and smartwatches which users already own.

Intoxication detection using heart rate: Heart rate is an alternate bio-measure of intoxication. Gutierrez et al. [1] estimated whether individuals were intoxicated (BAC > 0.065) using heart rates and temperatures measured by a smartwatch worn by the subjects. However, they did not use the accelerometer and gyroscope (gait) to infer the subjects' intoxication level.

Alcohol-related smartphone apps: Existing smartphone alcohol-related applications on the iPhone and Android app markets allow users to manually record their alcohol consumption, or estimate Blood Alcohol (BAC) levels using built-in formulas, offer manual cognition tests to assess users' intoxication levels [5] and encourage positive drinking habits [21, 23] but do not sense intoxication from the user's gait.

Intoxication-detection from gait: Some prior work has utilized just gait data from the smartphone for detecting intoxication [7, 9, 15] but not combined with smartwatches. As mentioned previously, detection using only the smartphone fails when the user leaves their phone (e.g. on the table) while walking around, which occurs about 47 percent of the time. This paper explores combining gait inference from smartphones as well as smartwatches.

III. BACKGROUND: DETECTING INTOXICATION FROM GAIT

In prior work, Ando *et al* [8] found that subjects' postural sway increased on the YZ (anterior-posterior) and XY (medio-lateral) planes in Figure 2, after they ingested alcohol. Nieschalk *et al* [3] determined that sway area (a gait attribute widely used in posturography) was the most sensitive attribute for detecting intoxication. Posturography is a clinically validated approach for assessing balance disorders from gait. Increases in the user's sway area due to alcohol manifests as increases in rotation around the axes of their phone's gyroscope (Figure 2 (b)). AlcoWear, extracts and classifies sway area features to infer its user's BAC level.

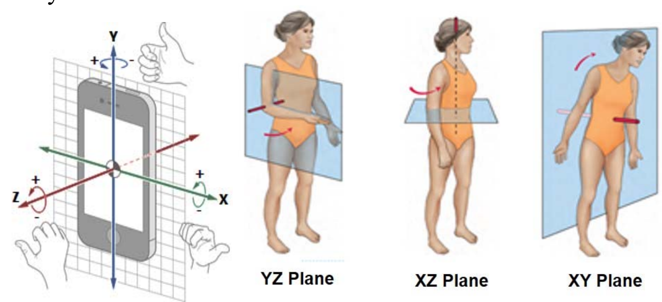

Figure 2 (a) Gyroscope rotation axes in Android smartphones and (b) their relationship with the three axes of the body

IV. ALCOWEAR

We present AlcoWear, a system that senses users' Blood Alcohol Content (BAC) from their walk pattern using data gathered from their Smartphone and Smartwatch collaboratively. For users with only a smartwatch or smartphone, only data from those devices are used. When a user wears both devices, the smartphone readings are utilized (more accurate) unless the phone is kept on the table or not carried by the user. As an initial trigger for initializing gait testing, the user's activity (e.g. using the Android Activity Recognition API on both devices) must be recognized as walking before AlcoWear performs intoxicated gait testing.

Figure 4 shows which steps of the real-time AlcoWear app are performed on each device. Gyroscope and accelerometer data is collected on both the smartwatch and smartphone. The smartwatch data is then sent to the phone where it is segmented along with the smartphone's sensor data, into 5 second segments. The segmented data is sent to the server, where features are extracted since feature extraction can be computationally intense and some signal processing features used are not supported on all Android devices. The extracted features are returned to the phone, where classification is performed. The inferred BAC range is sent to the smartwatch where it is used and displayed.

A. AlcoWatch

Figure 3 AlcoWatch a) Sober Screen b) Drunk Screen

AlcoWatch is the part of the AlcoWear BAC sensing system that runs on the smartwatch. First, the user initializes both the AlcoWatch (smartwatch) and AlcoGait (smartphone) apps, by entering their personal details (gender, weight, height), what emergency action the AlcoWear system should take if they become too drunk to drive (e.g. call a friend or Uber) and records a sample of their sober walk. The sober walk is required as the app determines user BAC as a function of increased sway from their baseline walk. (more on this later). Whenever the user's gait is tested and they are under the limit (< 0.08), the phone shows the sober screen (figure 3a). If the user is over the limit (>0.08), the Drunk Screen is displayed (Figure 3b). At the bottom of screen is a button that allows a user to call a friend, call an Uber, or activate a kill switch in a car based on their preference.

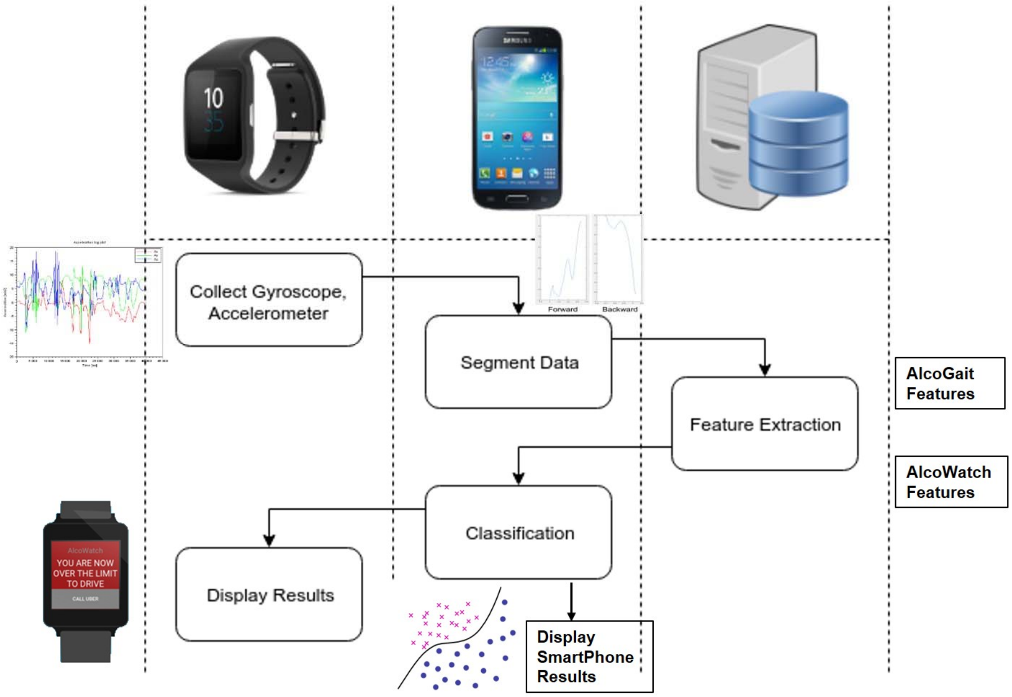

Figure 4 Flow of operations on AlcoWear

B. *AlcoGait*

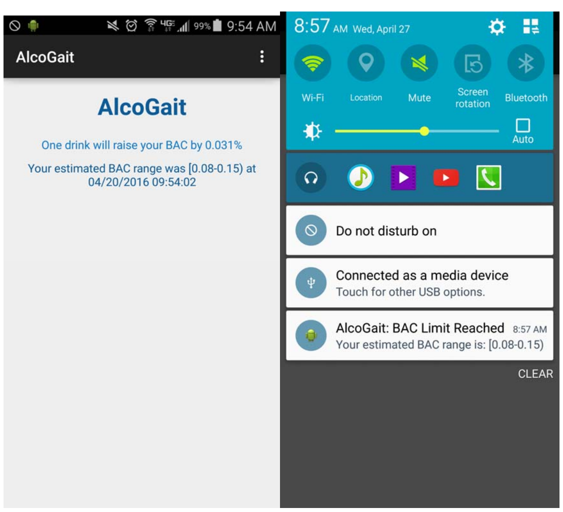

Figure 5 AlcoGait app Main screen (left) Notification sent to user when they have reached their pre-set BAC limit (right)

AlcoGait is the part of the AlcoWear BAC sensing system that runs on the smartphone. AlcoGait uses our classification model trained with data from our 33 subjects to detect the BAC level of smartphone users in real-time based on their gait data. Figure 5 shows the main screen of our AlcoGait app (left) and a notification delivered to the smartphone owner when they have reached their pre-set BAC limit (right).

V. Training AlcoGait and AlcoWatch BAC Classifiers

Our methodology for training the BAC classifiers on the smartphone and smartwatch offline followed a typical machine learning classification approach with a flow diagram illustrated in Figure 6. Smartphone and smartwatch gyroscope and accelerometer data were gathered simultaneously from subjects as they walked while wearing special goggles designed to simulate intoxication. This data was used to train classifiers for both the smartwatch and smartphone. The final classification model was used to create the AlcoGait smartphone app and the AlcoWatch smartwatch app, which can detect the BAC levels from users' smartphones and smartwatches in real-time from their gait while they walked.

A. *Data Gathering Study*

Thirty-three (33) participants (20 males and 13 females) were recruited via a pool of psychology students who receive academic credit for participating in user studies. Subjects were also recruited via email advertisements, social media advertisements, and word-of-mouth.

Sensor-Impairment Goggles: Subjects wore sensor-impairment goggles and walked while accelerometer and gyroscope sensor data was collected from their smartwatch and smartphone. These "Drunk Busters" goggles use vision distortion to simulate the effects of alcohol consumption on the body [6]. Goggles rated at various Blood Alcohol Concentration (BAC) levels simulated the corresponding impairment causing wearers to experience intoxication effects including reduced alertness, delayed reaction time, confusion, visual distortion, alteration of depth and distance perception, reduced peripheral vision, double vision, and lack of muscle coordination [6]. These goggles have been used to educate individuals regarding the effects of alcohol on one's motor skills. It is instructive to note that each alcohol goggle simulates an approximate BAC range (e.g. 0.08 – 0.15).

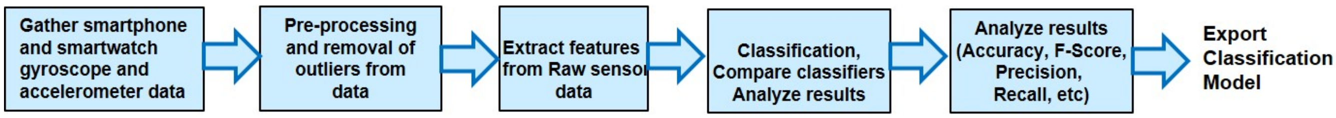

Figure 6 Flow Diagram for sensor data collection, feature extraction and classification

Figure 7 A subject walking while wearing Drunk Busters Goggles. Data is gathered on their Smartwatch and Smartphone

Study Procedure: In an IRB-approved study, participants placed an Android smartphone in either their front or back pants pocket, and wore the smartwatch on their wrist. The participants then walked normally (no impairment) for 60 seconds, while a data gathering app running MATLAB mobile recorded the smartphone and smartwatch's gyroscope and accelerometer data (Figure 7). Subjects then repeated the 90-second walk while wearing goggles rated at BAC of [0.04-0.06, 0.08 – 0.15, 0.15-0.25, 0.25 – 0.35].

A. *Pre-Processing (including outlier removal)*

Gyroscope and accelerometer data was gathered from subjects and stored in segments of 5 seconds. As such no further segmentation was required. However, subjects may trip or fall while intoxicated, which would generate extremal gyroscope and accelerometer data values (outliers). We synthesized a simple outlier removal algorithm by sorting the accelerometer and gyroscope data and removing the top and bottom 1 percent of values on the x, y and z axes.

B. *Feature extraction*

Gyroscope features: In prior work, Nieschalk *et al* [3] found that a gait attribute called sway area, was the most sensitive attribute for detecting increased body sway after subjects ingested alcohol. Kaewkannate [27] showed that smartphone gyroscopes can accurately capture posturography variables (including sway area). Gyroscope signals from our subjects' smartphones and smartwatches were utilized to extract sway features. The smartphone gyroscope sensor returns the rate of rotation around the smartphone's X, Y and Z axes in radians per second. Sway area is calculated by plotting values from two of the gyroscope's axes. For the XZ sway area, all observed gyroscope X and Z values in a segment were projected unto an X-Z plane (see Figure 8 (top)).

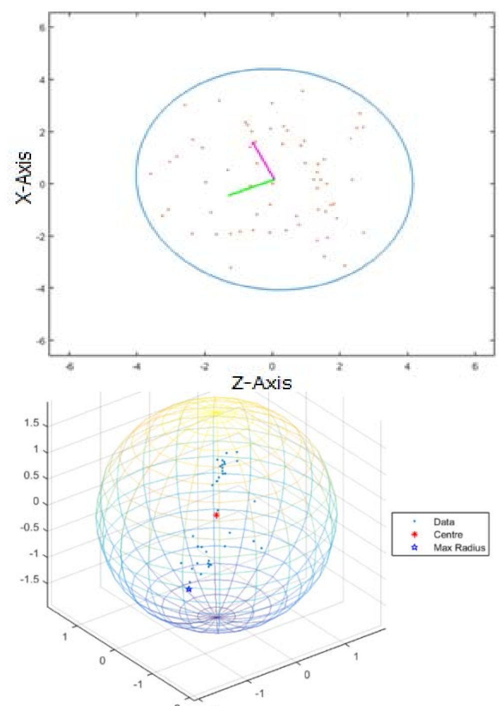

Figure 8 XZ Sway area (top) and Sway Volume (bottom) Plots

The area of an ellipse that encloses the 95 percent confidence interval of all observed points was returned as the XZ sway area. This methodology is similar to that used for calculating sway area using force plate readings in clinics [3]. However, our use of the gyroscope to synthesize these sway areas is novel. As a further contribution, we synthesized gyroscope-based sway volume, the 3D analogue of sway area as a novel feature explored (See figure 4 (bottom)). Sway volume is the sphere that contains the 95 percent confidence interval of all X, Y, Z points in a segment. Accelerometer and gyroscope features were generated from the sensor data gathered from all study participants. Table 5 lists gyroscope features extracted and their formulas, while Table 6 lists accelerometer features extracted and their formulas. In addition to the sway and accelerometer signals in tables 1 and 2, on the smartwatch, additional features that capture wrist rotation and movement were extracted.

Feature Normalization: No two individuals walk exactly the same, especially while intoxicated. To minimize such inter-person differences in gait patterns, users' gyroscope and accelerometer features were normalized per person. This task involves dividing each subject's intoxicated feature value by the value of the same features calculated while they were sober. For instance, the formulas for average sway area and normalized sway area can be seen below (eqns 1 & 2).

$$\textit{Average sway area} = \frac{\Sigma\ (\text{sober sway area calculations})}{\#\ \text{sober readings}} \quad (1)$$

$$\textit{Normalized sway area} = \frac{(\text{current samples's sway area})}{\text{average sway area}} \quad (2)$$

C. Machine Learning Classification

All gyroscope and accelerometer features were imported into the Weka machine learning library to explore classification. The accuracy of various classifiers was compared.

VI. Results

We now present results of analyses and classification of gait data from our 33 participants (20 males and 13 female). Participants' heights ranged from 150cm to 200cm (mean = 172cm, s = 10.22cm), weights ranged from 100lbs to 250lbs (mean = 155lbs, s = 31.96lbs), and ages ranged from 18 to 22 (mean = 20 years, s = of 1.32 years).

A. Feature Selection using Correlation-Based Feature Selection (CFS)

To quantify the predictive value of each extracted feature, we used Correlation-Based Feature Selection (CFS) [13] wherein each feature's correlation with the subject's BAC level and p-value are computed. Features that have statistically significant correlations (p-value < 0.05) with BAC levels have the highest predictive value are used for BAC classification (listed in Table 1 for the smartwatch and Table 2 for the smartphone). Subjects' ages, gender and height, were also included as features as they improved classification accuracy.

A. Results of Classification

Using the Weka machine learning library, the extracted features (normalized vs non-normalized) were classified into the labeled ranges of [0.04 – 0.06, 0.08-0.15, 0.15-0.25, 0.25-0.35] for the Smartphone, and into the [< 0.08, >= 0.08 bins] on the Smartwatch. The accuracy of the J48, JRip, Bayes Net, Random Forest, Random Tree, and Bagging classification algorithms were compared. The classification accuracy of normalized vs. not normalized data were investigated, in addition to investigating different sets of features (all gyroscope features, all accelerometer features, selected gyroscope features, and certain attributes describing the participant such as height, weight, and gender). Table 3 summarizes our classification results for AlcoWatch.

Feature	Correlation	p-Value
Yaw Angular Velocity Variance (Forward)	0.1877	0.00000
Pitch Angular Velocity Median (Backward)	0.1720	0.00001
Total Harmonic Distortion	0.1642	0.00002
Yaw Angular Velocity Variance (Backward)	0.1522	0.00008
Weight	0.1490	0.00011
Y Velocity Median (Forward)	0.1467	0.00015
Roll Angular Velocity Variance (Forward)	0.1265	0.00107
Net Pitch Change (Backward)	0.1247	0.00126
Yaw Angular Velocity Median (Backward)	0.1228	0.00149
Band power	0.1215	0.00169
YZ Sway Area	0.1198	0.00195
Pitch Angular Velocity Variance (Backward)	0.1170	0.00250
XZ Sway Area	0.1146	0.00305
XY Sway Area	0.1130	0.00350
Pitch Angular Velocity Median (Forward)	0.1129	0.00354
Y Velocity Variance (Forward)	0.1026	0.00805
Yaw Angular Velocity Median (Forward)	0.1026	0.00808
Net Pitch Change (Forward)	0.0979	0.01146
Net Yaw Change (Forward)	0.0932	0.01609
X Velocity Median (Backward)	0.0931	0.01628
Net Distance in Y (Forward)	0.0909	0.01893

Table 1 Statistically Significant Features for AlcoWatch (p-value < 0.05)

P-Value	Correlation Coeff	Feature
4.26E-12	-0.1601	Skew
3.71E-10	-0.14496	Kurtosis
0.002902	-0.06918	Gait Velocity
8.22E-06	0.10344	Residual Step Time
2.38E-18	-0.20114	Band power
8.49E-18	-0.19788	XZ Sway
6.71E-13	-0.16596	XY Sway
2.87E-24	-0.23314	YZ Sway
3.59E-08	-0.12763	Sway Volume

Table 2 p-values and correlation coefficients statistically significant features used to train AlcoGait classifiers

Attributes Classified	Classifier	Accuracy
Full-swing, forward and backward swing features, gender, height, age, weight	Random Forest	79.6784
All of the above, except features removed during feature selection	Random Forest	79.8246

Table 3 Classification Accuracy for AlcoWatch

Table 4 shows our results for the AlcoGait (smartphone) classification. Figure 9 summarizes our main results. Our main results were:

i. *All features with normalization:* When classifying all generated features from both the gyroscope and accelerometer (with sway areas and volume normalized), ID, height, weight, and gender, the J48 classifier using percentage split, 99% train and 1% test had the **highest accuracy of 89.45%** and an ROC area of 0.916.

ii. *All features, no normalization:* When classifying all generated features from both sensors (with sway areas and volume not normalized), ID, height, weight, and gender, the J48 classifier using percentage split, 99% train and 1% test had the **highest accuracy of 88.89%**.

iii. *Accelerometer features only:* When classifying all features generated features from just the accelerometer, Random Forest using cross-validation, 10 folds had the **highest accuracy of 64.56%** (24.89% less accurate than all features+ normalization).

iv. *Gyroscope features only:* When classifying all generated features from just the gyroscope, Random Forest using cross-validation, 10 folds had **the highest accuracy of 75.79%** (13.66% less accurate than all features+ normalization) and an ROC area of 0.919.

Including participant gender, weight and height as features: Since intoxication is affected by weight, height and gender we explored including them as features in our classification. Including weight, height and gender as features improved classification accuracy. For example, when classifying the normalized data using Random Forest with percentage split (66% train, 33% test), including gender, height, and weight was 5.3% more accurate than using the features and weight.

Comparison of classifier types: The J48 and Random Forest classifiers were 21% to 39% more accurate in classification compared to the other classifiers explored (Random Tree, Bagging, JRip, and Bayes Net).

Investigating personalization: We investigated the idea of personalization, which involved training classifiers using only single user's gait data. Unfortunately, our results were inconclusive: personalization improved classification accuracy for some subjects, but worsened it for others. As such, personalization was not explored further.

Classifier	Test Set	Accuracy When Normalized	Accuracy (Unnormalized)
J48	Percentage split, 95% train 5% test	73.12%	72.22%
J48	***Percentage split, 99% train 1% test***	***89.45%***	***88.89%***
Random Forest	Percentage split, 66% train 33% test	72.66%	69.18%
Random Forest	Percentage split, 95% train 5% test	81.72%	75.56%
Random Forest	Cross-validation, 10 folds	73.74%	74.79%
Random Tree	Percentage split, 66% train 33% test	66.77%	63.44%
JRip	Cross-validation, 10 folds	50.29%	50.31%
Bayes Net	Percentage split, 66% train 33% test	45.47%	42.46%
Bagging	Cross-validation, 10 folds	67.53%	70.94%

Table 4 AlcoGait (Smartphone) classification results using accelerometer and gyroscope features, height, weight and gender

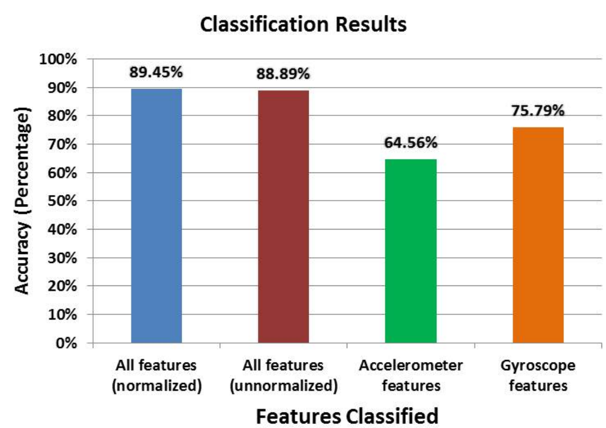

Figure 9 Classification Accuracy of features into [0.04-0.06, 0.08-0.15, 0.15-0.25, 0.25-0.35] ranges

VII. Conclusion and Future work

Alcohol abuse kills 88,000 people annually in the United States. In this paper, we presented AlcoWear, a system that uses a drinker's smartphone and smartwatch to passively sense their intoxication level from their gait (walk). AlcoWear infers the drinker's BAC level by classifying accelerometer and gyroscope sensor features gathered from their smartphone and smartwatch simultaneously using a machine learning approach. Features such as sway area (gyroscope) and cadence (accelerometer) are extracted from the user's smartphone data. Features such as Total Harmonic Distortion and angular velocity are also extracted and used to classify accelerometer data from the same users' smartwatches. On the smartphone, the J48 classifier was the most accurate, classifying user gait patterns into BAC ranges of [0.00-0.08), [0.08-0.15), [0.15-0.25), [0.25+) with an accuracy of 89.45%. For the smartwatch, AlcoWear classifies users as being in BAC ranges [< 0.08 vs >= 0.08]

(2 bins) with a 79.8 percent accuracy using a Random Forest classifier. As future work, we hope to gather intoxicated gait data from a large number of subjects who are also breathalyzed and explore the effects of differences in alcohol tolerances, walking patterns and confounding factors such as fatigue and mood, which also affect gait.

REFERENCES

[1] Gutierrez, M, Fast M, Ngu A, Gao, B, Real-Time Prediction of Blood Alcohol Content using Smartwatch Sensor Data, Volume 9545 of the series Lecture Notes in Computer Science pp 175-186

[2] Tokyoflash Japan, "Kisai Intoxicated LCD Watch," Tokyoflash Japan, 2014. [Online]. Available: http://www.tokyoflash.com/en/watches/ kisai/intoxicated/. [Accessed 09 December 2014].

[3] Nieschalk, M., Ortmann, C., West, A., Schmäl, F., Stoll, W., & Fechner, G. (1999). Effects of alcohol on body-sway patterns in human subjects. Int'l Journal of Legal Medicine, 253-260.

[4] SCRAM Systems, "SCRAM Continuous Alcohol Monitoring," Alcohol Monitoring Systems, Inc., 2014. [Online]. Available: http://www.scramsystems.com/index/scram/continuous-alcoholmonitoring.[Accessed 09 December 2014].

[5] Intoxicheck® Phone App - Buzzed Driving App | Impaired Driving App | Impairment Assessment. (n.d.). Retrieved Dec 1, 2015, from http://fatalvision.com/intoxicheck.html

[6] Lifeloc Technologies - Impairment Goggles. (n.d.). Retrieved Dec 1, 2015, from https://lifeloc.com/c-60-impairment-goggles.aspx

[7] Aiello, C and Agu, E, Investigating Postural Sway Features, Normalization and Personalization in Detecting Blood Alcohol Levels of Smartphone Users, in Proc Wireless Health, 2016.

[8] S. Ando, T. Iwata, H. Ishikawa, M. Dakeishi and K. Murata, "Effects of acute alcohol ingestion on neuromotor functions, *NeuroToxicology,* vol. 29, pp. 735-739, 2008.

[9] H.-L. Kao, B.-J. Ho, A. C. Lin and H.-H. Chu, "Phone-based Gait Analysis to Detect Alcohol Usage," in Proc ACM Ubicomp 2012.

[10] K.-C. Wang, Y.-H. Hsieh, C.-H. Yen, C.-W. You, M.-C. Huang, C.-H. Lee, S.-Y. Lau, H.-L. Kao, H.-H. Chu and M.-S. Chen, "SoberDiary: A Phone-based Support System for Assisting Recovery from Alcohol Dependence," in Proc ACM Ubicomp 2013, Seattle, WA, 2013.

[11] Dietterich, Thomas. Ensemble Methods in Machine Learning, (1-5). Retrieved Dec 4, 2015, http://web.engr.oregonstate.edu/~tgd/ publications/mcs-ensembles.pdf

[12] BACtrack, Long term DUI consequences, http://www.bactrack.com/ blogs/expert-center/35042309-long-term-dui-consequences

[13] Hall, M. A. (1999). *Correlation-based feature selection for machine learning* (Doctoral dissertation, The University of Waikato).

[14] Mathworks. (n.d.). Bandpower. Retrieved April 20, 2016, from http://www.mathworks.com/help/signal/ref/bandpower.html

[15] Arnold, Z, LaRose, D, and Agu, E Smartphone Inference of Alcohol Consumption Levels from Gait, in Proc. IEEE ICHI 2015

[16] Qi, Muxi. "A Comprehensive Performance Comparison of Signal Processing Features in Detecting Alcohol Consumption from Gait Data." WPI Graduate Thesis, April 2016.

[17] Substance Abuse and Mental Health Services Administration (SAMHSA) National Survey on Drug Use and Health. Underage Alcohol Use among Full-Time College Students. Issue 31, 2006.

[18] Chambers, M., Liu, M., Moore, C.: Drunk driving by the numbers. United States Department of Education.

[19] Fact Sheets - Alcohol Use and Your Health. (2015, December 17). Retrieved January 8, 2016, http://www.cdc.gov/alcohol/factsheets/ alcohol-use.htm

[20] Impaired Driving: Get the Facts. (2015, November 24). Retrieved January 8, 2016, http://www.cdc.gov/motorvehiclesafety /impaired_driving/impaired-drv_factsheet.html

[21] Tjondronegoro, D.; Drennan, J.; Kavanagh, D.J.; Zhao, E.J.; White, A.M.; Previte, J.; Connor, J.P.; Fry, M.-L., "Designing a Mobile Social Tool that Moderates Drinking," in IEEE Pervasive Computing, vol.14, no.3, pp.62-69, July-Sept. 2015 doi: 10.1109/MPRV.2015.62

[22] Mathworks. (n.d.). Total Harmonic Distortion. Retrieved April 20, 2016, from http://www.mathworks.com/help/signal/ref/thd.html

[23] Wang, K., Huang, M., Hsieh, Y., Lau, S., Yen, C., Kao, H,. Chen, Y. (2014). SoberDiary. In Adjunct Proceedings ACM Ubicomp 2014

[24] National Highway Traffic Safety Administration. (n.d.). The ABCs of BAC A Guide to Understanding Blood Alcohol Concentration and Alcohol Impairment. Retrieved April 20, 2016, from http://www.nhtsa.gov/links/sid/ABCsBACWeb/page2.htm

[25] S. Demura and M. Uchiyama, "Influence of moderate alcohol ingestion on gait," Sport Sci Health, no. 4, pp. 21-26, 2008.

[26] Smartphone users worldwide 2014-2019 | Statistic. (n.d.). Retrieved April 20, 2016, http://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

[27] Kaewkannate, K, The Correlation Among Body Sway Parameters from the Gyroscope Data Sets, in Proc Biomedical Engineering Conference (BMEiCON) 2013.

[28] Anind K. Dey, Katarzyna Wac, Denzil Ferreira, Kevin Tassini, Jin-Hyuk Hong, and Julian Ramos. 2011. Getting closer: an empirical investigation of the proximity of user to their smart phones. In *Proc UbiComp* 2011.

Table 5: Features Generated from Gyroscope Data

Feature	Feature Description	Formula	Reference
XZ Sway Area	Area of projected gyroscope readings from Z (yaw) and X (pitch) axes	$XZ\ Sway\ Area = \pi r^2$	[3]
YZ Sway Area	Area of projected gyroscope readings from Z (yaw) and Y (roll) axes	$YZ\ Sway\ Area = \pi r^2$	Our contribution
XY Sway Area	Area of projected gyroscope readings from X (pitch) and Y (roll) axes	$XY\ Sway\ Area = \pi r^2$	[13, 14]
Sway Volume	Volume of projected gyroscope readings from all three axes (pitch, roll, yaw)	$Sway\ Volume = \frac{4}{3}\pi r^3$	Our contribution

Table 6: Features Generated from Accelerometer Data

Feature	Feature Description	Calculation	Ref.
Steps	Number of steps taken	calculation of signal peaks above one standard deviation away from mean of gravity corrected magnitude of signal [15]	[15]
Cadence	Number of steps taken per minute	$cadence = \frac{\#\,steps}{minute}$	[15]
Skew	Lack of symmetry in one's walking pattern	$skewness = \frac{\frac{1}{n}\sum(x_i - \mu_x)^3}{\left[\frac{1}{n}\sum(x_i - \mu_x)^2\right]^{3/2}}$ Where x_i is the data sequence, and μ_x is the average of all x_i [16]	[15]
Kurtosis	Measure of how outlier-prone a distribution is	$kurtosis = \frac{\frac{1}{n}\sum(x_i - \mu_x)^4}{\left[\frac{1}{n}\sum(x_i - \mu_x)^2\right]^2}$ Where x_i is the data sequence, and μ_x is the average of all x_i [16]	[15]
Average gait velocity	Average steps per second divided by average step length	$\text{average gait velocity} = \frac{(average\,steps\,/\,sec)}{step\,length}$	[15]
Residual step length	Difference from the average in the length of each step	$\text{residual step length} = \frac{distance}{\#\,steps}$	[15]
Ratio	Ratio of high and low frequencies	$harmonic\ ratio = \frac{\sum_{i=1,3,5,\ldots} V_i}{\sum_{j=2,4,6,\ldots} V_j}$ Where V_i is the amplitude of odd-ordered harmonic frequency and V_j is the even-ordered harmonic frequency [16]	[15]
Residual step time	Difference in the time of each step	$\text{residual step time} = \frac{\sqrt{\frac{1}{n}\sum(interval_i - \mu_{interval})^2}}{\mu_{interval}}$ Where $interval_i$ is a sequence of stride intervals and $\mu_{interval}$ is average of all $interval_i$ [16]	[15]
Band power	Average power in the input signal	$\text{band power} = \text{bandpower(x)}$ Where x is a matrix of the magnitude, and band power calculates the average power in each column independently [14]	[15]
Signal to noise ratio	Estimated level of noise within the data	$snr = \frac{power_{signal}}{power_{noise}}$ [16]	[15]
Total harmonic distortion	"Determined from the fundamental frequency and the first five harmonics using a modified periodogram of the same length as the input signal" [22]	$thd = \frac{\sqrt{\sum_{i=2,3,4,5} V_i^2}}{V_1}$ Where V_1 is energy contained within peak of PSD at the fundamental frequency and V_i are the energy contained within the harmonics [15]	[15]