You're Driving and Texting: Detecting Drivers Using Personal Smart Phones by Leveraging Inertial Sensors

Cheng Bo
Illinois Institute of Technology
Chicago, IL 60616, USA
cbo@hawk.iit.edu

Xuesi Jian
Illinois Institute of Technology
Chicago, IL 60616, USA
xjian@hawk.iit.edu

Xiang-Yang Li
Illinois Institute of Technology
Chicago, IL 60616, USA
xli@cs.iit.edu

Xufei Mao
School of Software and
TNLIST
Tsinghua University, China
xufei.mao@gmail.com

Yu Wang
University of North Carolina at
Charlotte
Charlotte, NC 28223, USA
yu.wang@uncc.edu

Fan Li
Department of Computer
Science
Beijing Institute of Technology
fli@bit.edu.cn

ABSTRACT

In this work, we address a critical task of detecting the user behavior of driving and texting simultaneously using smartphones. We propose, design, and implement **TEXIVE** which achieves the goal of distinguishing drivers and passengers, and detecting texting operations during driving utilizing irregularities and rich micro-movements of users. Without relying on any external infrastructures and additional devices, and no need to bring any modification to vehicles, **TEXIVE** is able to successfully detect dangerous operations with good *sensitivity*, *specificity* and *accuracy*. We conduct experimental study of **TEXIVE** with the help of a number of volunteers using various vehicles and smartphones. Our results indicate that **TEXIVE** has a classification accuracy of 87.18%, and precision of 96.67%.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

Keywords

TEXIVE, Smartphone, Classification

1. INTRODUCTION

One recent study indicates that there are at least 23% of all vehicles crashed and 1.3 million crashes in the US involve using cell phones (especially texting) during driving [2] although over 30 states and District of Columbia has forbidden texting message while driving [1]. Such severe security issue has stirred numerous researches and innovations on detecting and preventing driving and texting operations so that a number of distracted driving behavior detection approaches have been proposed, e.g., mounting a camera to monitor the driver [8], relying on acoustic ranging through car speakers [7], use adapter that provides vehicle speed reference readings to the phone [6], and cloud computing to recognize drivers' operations [3].

In this work, we address this critical task of detecting driving and texting activities. We propose **TEXIVE**, a system leveraging inertial sensors integrated in regular smartphones to distinguish drivers from passengers through recognizing rich micro-movement dynamics of smartphone users, and further detect driving and texting activities of drivers. Our main idea is to let **TEXIVE** recognize micro-movements by fusing multiple evidences collected from inertial sensors in smartphones, e.g., detecting whether a user is entering a vehicle or not, inferring which side of the vehicle he/she is entering, determining whether a user is siting in front or rear seats, or even detecting which side if multiple users in the car collaborate.

We collected the data from motion sensors when users are performing various activities and observed some unique patterns by converting the signal to the frequency domain using DCT and wavelet. To infer whether a user enters the vehicle from left side or right side of the vehicle, or sits in front or rear seats, we exploit the unique patterns mined from the rich dynamics in the accelerometer and magnetometer data observed during respective actions, and make cognitive decision based on machine learning techniques.

Unlike most related solutions, **TEXIVE** is designed based on regular smartphones without supporting from any external infrastructures and additional devices. In addition, **TEXIVE** does not bring any modification to vehicles. We conduct extensive experiments involving in a number of volunteers on various vehicles and smartphones and the results show that **TEXIVE** has a classification accuracy of 87.18%, and precision of 96.67%.

2. SYSTEM DESIGN

To address the challenge of distinguishing between driver and passenger, we will leverage the existing inertial sensors integrated in smartphones and exploit some unique and distinguishable patterns mined from a sequence of sensory data.

2.1 System Architecture Overview

We adopt a three-phase solution to accomplish the task: initial walking detecting, in-vehicle recognition, and evidence fusion respectively. Figure 1 illustrates the basic working flow of the system according to the three phases, and the functionality of detailed components is as follows.

Activity Identification: Generally, most of users get used to carry their smartphones all day long, which facilitate observing multiple activities. One of tasks is to identify related activities from a rich set of potential daily activities, including walking, sitting, standing or even ascending stairs. One thing deserves mention is that **TEXIVE** does **not** require *any* interaction from the user. We
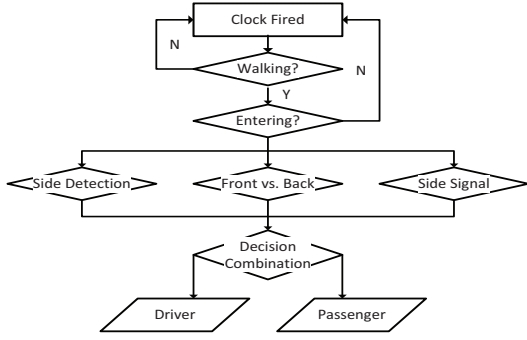
Figure 1: The system overview

study the temporal and spatial distribution of different activities as well, through constructing a Hidden Markov Model (HMM) [5] to personalize the model, and further optimize the energy consumption by carefully adjusting the duty-cycle.

Detecting Boarding Side: Although whether a user entering a vehicle from either side of the vehicle is a driver could be inferred most of time directly, we cannot guarantee the identity of the user precisely. Hence, it is still necessary to judge the boarding side for users. **TEXIVE** recognizes the entering activity based on direction of turning and sequence of lifting leg to judge the boarding side of a user. Turning and lifting actions on different sides of the vehicle could be reflected through inertial sensors, which will act divergent under different side and place of the body.

Detecting Front vs. Back: Our approach relies on the unique and distinguishable patterns reflected on the acceleration between front and back seats when vehicle is crossing a bump or pothole. According to our preliminary tests, the bump signal, although not guaranteed to happen, can always accurately determine whether the phone (user) is in front seats or rear seats.

2.2 Entering Vehicles?
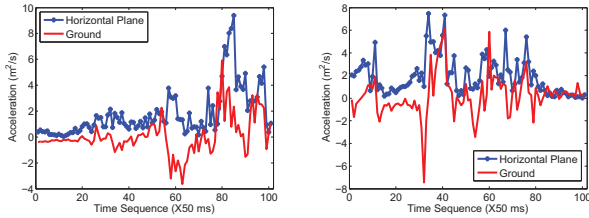
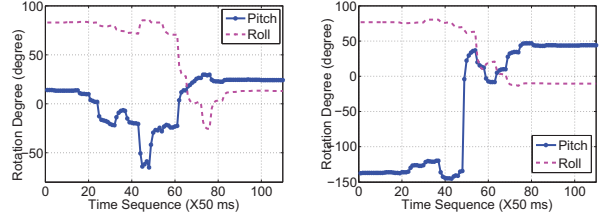(a) Driver Side, Left Pocket   (b) Driver Side, Right Pocket

Figure 2: Data extracted from accelerometer in Horizon plane and ground when people enter vehicles.

A key challenge of this system is to identify specific activities in real-time, especially determining whether a user will enter a vehicle or is just performing other activities, which have similar observable patterns. Empirically, the activity of getting into vehicle consists of five basic actions, including walking towards the vehicle, opening the door, turning the body, entering, and sitting down.

A study [4] indicates that most of the users carry their phone in trousers. We first collected 200 samples of entering vehicle from both driver and passenger sides in the parking lot by a group of volunteers with the smartphones in trouser pockets. We extract the linear acceleration and transform it to the Earth Frame Coordinate, to reduce the impact of irregular and unpredictable positions of smartphones in the pocket. The activity of entering the vehicle in both horizontal plane and ground direction is shown in Figure 2, in which the difference is obvious. With Naive Bayesian classifier, the accuracy and precision of distinguishing 40 cases of entering from

nearly 296 other activities are 84.46% and 45.24%. Unfortunately, several other activities (sitting down mostly) could be misjudged as car-entering activity, which negatively impacts the performance. To overcome this, we propose a more comprehensive ambient-aware filter to improve the accuracy. The filter is based on the observation the sensed ambient magnetic field fluctuates more dramatically when approaching the vehicle than other activities.

2.3 Which Boarding Side?

(a) Left Pocket, Driver Side   (b) Right Pocket, Driver Side

Figure 3: Side detection by observing rotation along BFC.

The system conducts side-detection operations simultaneously with the entering-detection so that the detection delay is minimized. The pure acceleration-based determination may fail in judging the boarding side of a user because of the possibility that a user with his/her phone in left pocket boarding from the right side has similar observing pattern as that of a user with his/her phone in his right pocket boarding from the left side.

We found another key factor determining the direction of body rotation when entering from both sides. Usually, in order to face front, the driver has to turn left while the passenger will turn right before he/she enters into the car, and such small duration of action could be captured by the gyroscope sensor in both *Pitch* and *Roll*, shown in Figure 3. Although the orientation of vehicle is unknown and unpredictable, the turning-based determination is demonstrated to be robust during our evaluation. We also adopt Extend Kalman Filter to eliminate the internal mechanism noise of sensors.

2.4 Front or Back Seats?

Solving the front-or-back problem is inspired by the experience that when we drive through a bump or a pothole, people sitting in the back row feel more bumpy than those sitting in the front. The thought is verified in analyzing our data collected by driving through either bumps or potholes (Fig. 4).
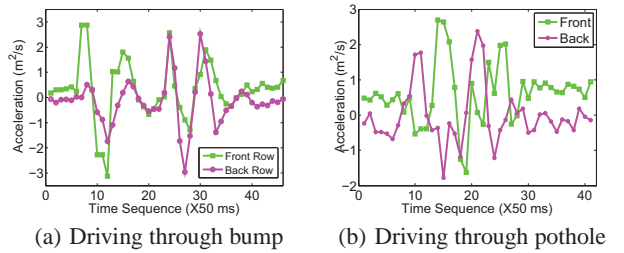

(a) Driving through bump   (b) Driving through pothole

Figure 4: Driving through bumps and potholes.

Due to the special shape of bumps or deceleration strips, one wheel will experience two continuous large vibration when wheels first hit the bump and hit the ground consequently. And smartphone is so sensitive that both front and back wheel bumpy activities could be observed, only in different intensities. The sophisticated analyzed data indicate that in both infrastructures, the front row users experience two jumps with similar intensity, while in the back row, the front-wheel vibration is relatively smaller than the back-wheel, as shown in Figure 4(a) and 4(b) respectively.

2.5 Texting?
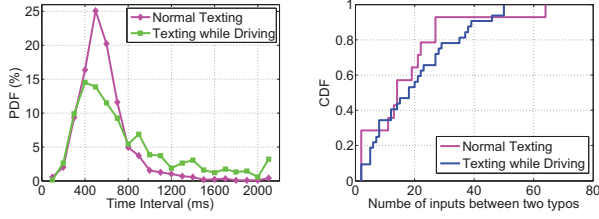
(a) The probability of different time interval between two inputs

(b) The typo frequency

Figure 5: The information extracted from typing.

TEXIVE triggers necessary warning when endangered texting activity is detected, and the key to satisfy the demand is detecting distracted behavior within a small amount of input or before touching the screen. Generally, typing is not a difficult task when user are fully concentrated, with fewer typo and higher typing speed. We conduct a set of experiments by a group of volunteers to compose multiple sentence in smartphone in both normal and driving scenarios, focusing on both the time interval between two inputs, and the frequency of typo.

We plot typing time interval between two inputs in Figure 5(a). We notice that the large difference in average typing speeds and standard deviation in two cases are generated from the fact that driver has to pause and watch the road after one word or phrase to keep alert. Therefore, the typo are more likely to happen in this case, as shown in Figure 5(b).

3. PRELIMINARY RESULTS
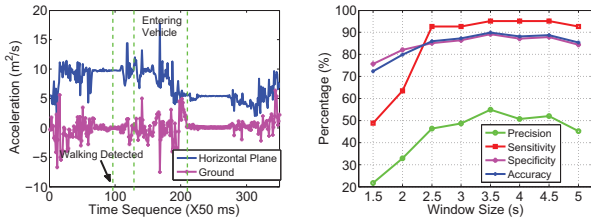
(a) Detecting the first arriving signal

(b) Recognition of entering vehicles

Figure 6: Detecting entering vehicles.

Initially, we evaluate the performance of entering activity detection, more specifically, the capability of both extracting entering pattern in a series of sensory data and distinguish from others.

Figure 6(a) illustrates the successful detection of first signal according to the protocol reflected in acceleration from the perspective of both horizontal and ground direction. The smartphone is put in the pocket, and **TEXIVE** perceives a walking activity from 112th time slot, and the entering signal arrives 2 seconds later(133th time slot). In addition, we evaluate the precision, sensitivity, specificity and accuracy with respect to different window sizes of the action. We set the window size ranging from $1.5s$ to $5s$ and plot the results in Figure 6(b), which indicates that generally the performance improves with the increment of window size. The system reaches a better performance when the window side is around $4s$ to $4.5s$, with both sensitivity are over 90%. After the behavior is determined, the detection of boarding side is followed. For both side, the accuracy reaches 85% and the precision is as high as 95% in the same window size.

We also take a number of comprehensive experiments in both parking lot and local roads to evaluate the efficiency of front-back

distinguish using bumps and potholes. We drive through both one deceleration strip and one bump in the parking lot ten times each with different driving speeds. The test results in Table 1 indicates absolute correctness, 20 bumps are all successfully detected in both locations.

Table 1: Bump in the parking lot

	Bump in Front	Bump in Back
Test in Front	20	0
Test in Back	0	20

In local road test, the smartphone of driver detects 334 samples of readings and 23 of bumps and potholes, while the back seat passenger only detects 286 samples but 58 bumps and potholes. Although the number of bumps and potholes being detected by both smartphones are different, because of the starting time of passenger is behind the driver, both smartphones report they are in the right location with accuracy of 100%.

To detect driving and texting, we sample 20 different typing cases with 8 texting in normal condition and 12 in driving condition in the parking lot. The evaluation in texting detecting is reliable and feasible, the accuracy is 90%.

Based on our experiment, we notice that the performance of **TEXIVE** mainly depends on the first two phases. We test the performance of driver detection based on the fusion of all the phases, the precision is 96.67% and accuracy 87.18%. Meanwhile, according to the real evaluation in Android smartphone, the recognition delay is only 0.2184 second.

4. ACKNOWLEDGMENTS

5. REFERENCES

[1] Insurance institute for highway safety. http://www.iihs.gov/.
[2] National highway traffic safety administration (nhtsa). http://www.nhtsa.gov/.
[3] H. Chu, V. Raman, J. Shen, R. Choudhury, A. Kansal, and V. Bahl. In-vehicle driver detection using mobile phone sensors. In *ACM MobiSys*, 2011.
[4] Y. Cui, J. Chipchase, and F. Ichikawa. A cross culture study on phone carrying and physical personalization. In *UI-HCII*, pages 483–492. Springer, 2007.
[5] L. Rabiner and B. Juang. An introduction to hidden markov models. *IEEE ASSP Magazine*, 3(1):4–16, 1986.
[6] Y. Wang, J. Yang, H. Liu, Y. Chen, M. Gruteser, and R. P. Martin. Sensing vehicle dynamics for determining driver phone use. In *ACM MobiSys*, 2013.
[7] J. Yang, S. Sidhom, G. Chandrasekaran, T. Vu, H. Liu, N. Cecan, Y. Chen, M. Gruteser, and R. Martin. Detecting driver phone use leveraging car speakers. In *ACM MobiCom*, pages 97–108, 2011.
[8] C.-W. You, N. D. Lane, F. Chen, R. Wang, Z. Chen, T. J. Bao, M. Montes-de Oca, Y. Cheng, M. Lin, L. Torresani, et al. Carsafe App: Alerting drowsy and distracted drivers using dual cameras on smartphones. *ACM Mobisys*, 2013.