Automatic Identification of Driver's Smartphone Exploiting Common Vehicle-Riding Actions

Homin Park, DaeHan Ahn, Taejoon Park, *Member, IEEE*, and Kang G. Shin, *Fellow, IEEE*

Abstract—Texting or browsing the web on a smartphone while driving, called *distracted driving*, significantly increases the risk of car accidents. There have been a number of proposals for the prevention of distracted driving, but none of them has addressed its important challenges completely and effectively. To remedy this deficiency, we present an event-driven solution, called *Automatic Identification of Driver's Smartphone* (AIDS), which identifies a driver's smartphone by analyzing and fusing the phone's sensory information related to common vehicle-riding activities, such as walking toward the vehicle, standing near the vehicle while opening a vehicle door, entering the vehicle, closing the door, and starting the engine. AIDS extracts features useful for identification of the driver's phone from diverse sensors available in commodity smartphones. It identifies the driver's phone before the vehicle leaves its parked spot, and differentiates seated (front or rear) rows in a vehicle by analyzing the subtle electromagnetic field spikes caused by the starting of the engine. To evaluate the feasibility and adaptability of AIDS, we have conducted extensive experiments: a prototype of AIDS was distributed to 12 participants, both males and females in their 20 and 30s, who have driven seven different vehicles for three days in real-world environments. Our evaluation results show that AIDS identified the driver's phone with an 83.3-93.3 percent true positive rate while achieving a 90.1-91.2 percent true negative rate at a marginal increase of the phone's energy consumption.

Index Terms—Identification of driver's phone, distracted driving, passenger and vehicle safety, smartphones

1 Introduction

Any activities that could divert drivers' attention away from the road endanger the safety of the driver and others. Of the various activities known to distract drivers, texting is by far the most problematic since it requires visual, manual, and cognitive attentions from the driver. According to the experiments conducted at Virginia Tech Transportation Institute, texting takes drivers' eyes off from the road for an average of 4.6 seconds, which is equivalent to driving *blindly* across an entire football field [1]. Statistics indicate that 15-25 percent of crashes, or approximately 1.3 million crashes leading to 400,000 injuries, in the United States are caused by distracted driving [2].

The menace of distracted driving to public safety has drawn increasing attention from governments as well as mobile and insurance industries. The US National Transportation Safety Board legislated a nationwide ban on texting while driving [3], and almost all US states followed the same. Mobile service providers and device manufacturers have also introduced various services to reduce/prevent distracted driving [4], [5], [6]. Despite the significant amount of efforts and resources invested and legislation, however, accidents related to distracted driving have not decreased thus far [7].

This lack of progress is due to the fact that conventional distracted driving prevention services require users to *manually* identify themselves as the driver in order to activate necessary restrictions on use of the driver's smartphone (DS). Such an approach suffers two practical problems. First, users cannot always be trusted to voluntarily update their status because of their reluctance in restricting use of their favorite mobile services, such as texting and social networking. Second, users may simply forget to set themselves (on their phones) as drivers. Therefore, manual status updates have not been effective in reducing distracted driving. To automatically identify the DS, there have been many proposals that require additional dedicated devices and/or modification of in-vehicle components [8], [9]. However, these incur additional costs to purchase, install, and modify devices, making them unattractive/impractical to the users and the car-makers.

While considering the above deficiencies (e.g., manual interventions and additional devices), several event-driven solutions have been proposed in recent years [10], [11], [12], [13], [14]. They autonomously identify the DS by analyzing the sensory information acquired from specific driving-related events, such as wearing seat belts, pedal pressing, turn signal audio, turning around the corners, and driving over the speed bumps, using a variety of sensors in commodity smartphones. While conventional event-driven solutions successfully automated the DS identification, there are four additional challenges that have not yet been addressed effectively.

First, we do not know if and when driving events of interest will occur. In the worst case, the DS will not be identified, thus failing to protect passengers, pedestrians and vehicles. Even if we assume these events will take place at some time while driving, drivers are in danger for an uncertain period of time until the events are monitored and analyzed. Considering the fact that incidents can happen at any time, we must rely on

- *H. Park and D. Ahn are with the Department of Information and Communication Engineering, Daegu Gyeongbuk Institute of Science and Technology, Daegu 42988, Republic of Korea. E-mail: {andrewpark, daehan}@dgist.ac.kr.*
- *T. Park is with the Department of Robotics Engineering, Hanyang University, Gyeonggi-do 15588, Republic of Korea. E-mail: taejoon@hanyang.ac.kr.*
- *K.G. Shin is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109. E-mail: kgshin@eecs.umich.edu.*

Manuscript received 4 Apr. 2016; revised 22 Dec. 2016; accepted 23 June 2017. Date of publication 7 July 2017; date of current version 5 Jan. 2018. (Corresponding author: Taejoon Park.)
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMC.2017.2724033

the events that occur *before* the vehicle leaves its parked spot (i.e., before making the first movement) to ensure the ultimate level of safety for the driver and others.

Second, there are no guarantees that all passengers will have a smartphone equipped with the DS identification system. Thus, despite the benefits of having a cooperative system design, we should not expect any assistance from the others in the vehicle. Third, smartphones might not have the communication capabilities to connect to the central server for various reasons, such as parked underground. In fact, all feature extractions and decision-making processes must be completed separately.

Finally, smartphones' *positions* (or poses) may change during daily routines based on the users' need and habit. In fact, no matter how accurate the end results are, a DS identification system would be impractical/unuseful if it is designed to operate only in a particular pose. To accurately identify the DS without pose restrictions, we must first define a feasible set of sensors for each smartphone pose, and construct the corresponding feature analyzers. To date, we have not been aware of any solutions that are able to accurately identify DS before the vehicle starts moving without imposing the above system restrictions and constraints.

In this paper, we propose a novel event-driven DS identification system, called *Automatic Identification of Driver's Smartphone* (AIDS). Its core is to fuse heterogeneous sensory information extracted from the common vehicle-riding actions—approaching the vehicle, standing still while opening the door, entering the vehicle, closing the vehicle door, and starting the engine—to identify the driver's phone.

The use of heterogeneous sensors can be divided into three parts. First, entering a vehicle is detected by analyzing electromagnetic field (EMF) fluctuations, significant vertical accelerations caused by sitting-down motion, and vehicle door closing sounds (VDCSs). While each of these can be found in other daily activities, a sequence of these events takes place only when entering the vehicle. Second, vehicle entering directions (left or right) are differentiated by analyzing the body rotations monitored when entering the vehicle. We observed that users turn counter-clockwise (clockwise) when entering from the left (right). Finally, seated (front or rear) rows are differentiated by analyzing subtle EMF changes monitored when starting the engine. The magnitude of such an EMF spike at the front row is likely to be much greater than that of the rear row because electronic devices are densely populated in front of the driver's seat.

For comprehensive evaluation of AIDS, we have implemented a prototype on Android, and distributed it to 12 participants, both males and females in their 20 and 30s. Our experimental results using 7 different vehicles show that entering a vehicle is detected with a 90.0-93.7 percent True Positive Rate (TPR) and a 91.3-93.2 percent True Negative Rate (TNR), while entering directions are identified with an 87.2-95.6 percent TPR and an 84.2-90.4 percent TNR. Moreover TPR and TNR of seated row classification results are found to be 82.8-99.5 and 79.3-95.8 percent, respectively. Finally, AIDS identifies the DS with an 83.3-93.3 percent TPR while the TNR is 90.1-91.2 percent.

This paper makes three main contributions: we

- propose an event-driven DS identification system that achieves accurate results before the vehicle leaves its parked spot with neither smartphone pose restrictions nor external assistance;
- reinforce the proposed system with a unique solution to differentiate the seated (front or rear) rows by analyzing the subtle EMF changes monitored when the vehicle is started; and
- implement and evaluate a prototype of AIDS on Android smartphones.

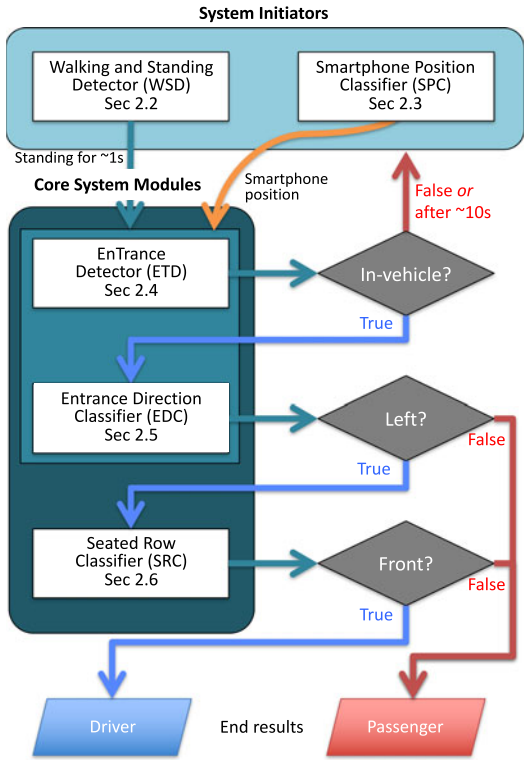


Fig. 1. Decision-making processes and the corresponding system modules in AIDS.

The rest of this paper is organized as follows. Section 2 first provides an overview of system operation with an introduction of key system modules using natural actions taken when riding the vehicle. Each system module is then detailed by describing the algorithms used to analyze monitored sensory data. Section 3 evaluates the performance of AIDS using real sensory data. Section 4 discusses the related work, and the paper concludes with Section 5.

2 System Design

Our main objective is to accurately identify the driver's smartphone without restricting the phone's poses nor requiring additional dedicated devices. Considering limited battery capacity, we also want to minimize energy consumption. To design AIDS with these objectives and constraints, we make the following three assumptions: A1) smartphones would not make significant movements when starting the vehicle; A2) vehicles are running on petrol engines; and A3) remote vehicle door openers are not available. These assumptions can be easily met since the market share of hybrid vehicles in the US is only about 2.75 percent, and remote vehicle door openers are rarely used [15]. The key building blocks of AIDS are detailed next.

2.1 Design Overview

As shown in Fig. 1, AIDS includes 5 system modules that are classified into two groups according to their functions. *System initiators* include Walking and Standing Detector (WSD) and Smartphone Position Classifier (SPC). Their role is to trigger core system modules only when needed with a set of sensors that produce clean data. Considering the importance of the

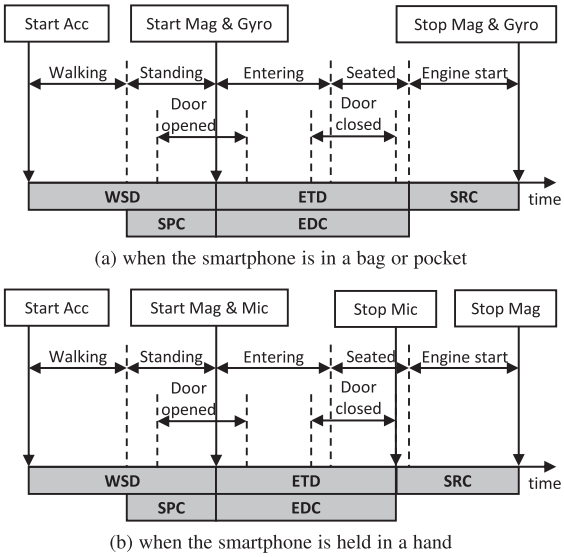


Fig. 2. Sensor use-cases based on common vehicle-riding actions and events. Different sets of sensors are used depending on detected smartphone positions.

energy-efficiency, activated sensors (excluding accelerometers) are put back to sleep after some time if no vehicle-riding actions are detected. The rest—namely, EnTrance Detector (ETD), Entering Direction Classifier (EDC), and Seated Row Classifier (SRC)—are grouped as core system modules with an objective to identify the DS. After AIDS concludes whether the smartphone is the driver's or not, all sensors but accelerometers are deactivated to save the battery energy. Fig. 2 shows the relationships between the common vehicle-riding actions/events and the sensors.

There are two major benefits of this system design. First, energy-efficiency is one of the most important requirements when designing mobile apps. We save as much energy as possible by solely using accelerometer readings for the system initiators while keeping other in sleep. Inactive sensors will be awakened only when the initiators determine that the user is about to enter the vehicle. Second, we improve the accuracy of ETD by using a subset of sensors, determined by SPC results, that are likely to return clean data. For example, microphones cannot accurately detect VDCS when smartphones are placed in a trouser pocket or a bag due to excessive fricative noises caused by surrounding materials. In such a case, ETD only leverages magnetometer and accelerometer readings.

2.2 Walking and Standing Detector (WSD)

WSD aims to detect a sequence of actions—walking and standing—that must be taken before entering the vehicles. Upon detection of standing, the core system modules are initiated by activating a subset of sleeping sensors to confirm whether the user is entering a vehicle or not.

To accurately detect walking and standing states, we use two acceleration features that can always be detected when users are walking. The first feature is a gait cycle. As indicated in [16], walking is a cyclic action where its motion characteristics can be traced by analyzing the frequency components. Based on the observations shown in Fig. 3, the gait cycle is found to be within 1∼10 Hz, depending on the smartphone's position.

The second feature is the presence of significant accelerations toward the horizontal plane. Such a feature is particularly important because it is physically impossible to enter a vehicle if we do not walk toward it. Considering the fact that our body

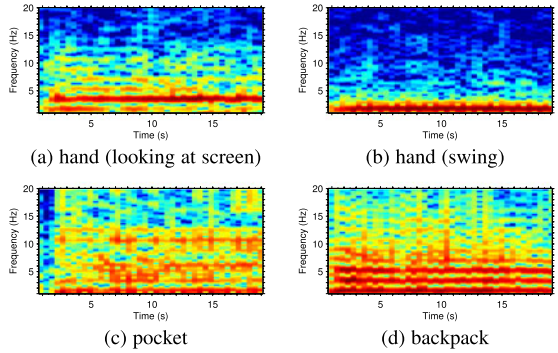


Fig. 3. Frequency spectrum representation of walking with different smartphone positions. Our observations show that walking is a cyclic action where the strongest energy is found in between 1 and 10 Hz.

cannot make *zero* accelerations for various reasons, including heartbeat and intentional body movements, WSD assumes that the user is making significant movements only when the horizontal accelerations exceed 0.5 m/s^2. In conclusion, co-occurrence of these two features strongly indicates that the user is walking, while the absence indicates standing.

To extract the gait cycle, WSD divides incoming triaxial accelerometer readings within a sliding time window into n segments. For each segment W_i^{accel}, where $1 \leq i \leq n$, let $\boldsymbol{a}_{ij} = (x_{ij}, y_{ij}, z_{ij})$ represent the jth acceleration vector within the window. The acceleration magnitude a_{ij} is then computed by taking the euclidean norm of $\boldsymbol{a}_{ij}$. In order to represent the frequency spectrum of these magnitude signals, we take the Discrete Fourier Transform (DFT) after applying a low-pass-filter with cutoff frequency of 20 Hz to remove the noises that may exist in the higher-frequency bands. Finally, the gait cycle is identified by locating a frequency with the strongest energy.

The second feature, the magnitude of horizontal acceleration, is computed as follows. Let $\boldsymbol{g} = (g_x, g_y, g_z)$ represent a unit gravity vector. The magnitude of the vertical acceleration v_{ij} of the acceleration vector $\boldsymbol{a}_{ij}$ is then computed by taking the dot product of $\boldsymbol{a}_{ij}$ and $\boldsymbol{g}$

$$v_{ij} = \boldsymbol{a}_{ij}^T \cdot \boldsymbol{g}. \tag{1}$$

This vertical acceleration can also be used to compute the vertical projection by multiplying the gravity vector

$$\boldsymbol{v}_{ij}^{pr} = v_{ij}\boldsymbol{g}. \tag{2}$$

Then, the magnitude of the horizontal acceleration h_{ij}, which is the magnitude of the horizontal projection $\boldsymbol{h}_{ij}^{pr}$, is computed as

$$\boldsymbol{h}_{ij}^{pr} = \boldsymbol{a}_{ij} - \boldsymbol{v}_{ij}^{pr} \tag{3}$$

$$h_{ij} = \|\boldsymbol{h}_{ij}^{pr}\|. \tag{4}$$

With the above sensory features, WSD concludes that the user is currently walking if the strongest frequency is found in between 1 and 10 Hz, and an average of h_{ij} within $\boldsymbol{W}_i^{accel}$ exceeds 0.5 m/s^2.

2.3 Smartphone Position Classifier (SPC)

SPC differentiates three most frequent positions for holding smartphones, such as including trouser pockets, bags, and hands [17]. The classified positions are then used by ETD to activate a subset of sensors that are most suitable for acquiring clean data for accurate DS identification results. A number of smartphone position classifiers have been proposed, satisfying our needs [16], [18], [19]. Of these, we adopt a method

proposed in [16], which utilizes supervised learning through regularized kernel methods [20], [21].

The feature vector used for SPC consists of: 1) DFTs of the horizontal and vertical acceleration magnitudes, and 2) smartphone orientations. For each time window W_i^{accel}, the horizontal and vertical DFTs, denoted by $\boldsymbol{f}_i^H$ and $\boldsymbol{f}_i^V$, are computed. According to our observations, different smartphone positions generate different oscillation patterns along the horizontal and vertical axes.

For example, smartphones placed in a pocket or a bag exhibit stronger acceleration magnitude toward the horizontal axis than that of vertical axis. In contrast, smartphones held in a hand show different oscillation patterns depending on the gestures taken while walking around. Swinging-like motions are one of the most frequently observed gestures which portray significant accelerations toward both horizontal and vertical axes. On the other hand, smartphone manipulation is another frequently seen gesture which exhibits stronger accelerations toward the horizontal axis than the vertical axis (similar to that of pocket and bag).

Since horizontal and vertical oscillations represented with DFTs are insufficient to clearly distinguish different smartphone positions, an orientation $\boldsymbol{o}_i$ is added to the feature vector. In our daily activities, smartphones held in hands are highly unlikely to stay in the upright position for a long period of time. On the other hand, when smartphones are held in pockets or bags, its screen is rarely facing up or down. Such differences in orientation can be captured by analyzing the combinations of gravity magnitude vectors in one- and two-dimensional spaces

$$\boldsymbol{o}_i = (|g_x|, |g_y|, |g_z|, \sqrt{g_x^2 + g_y^2}, \sqrt{g_x^2 + g_z^2}, \sqrt{g_y^2 + g_z^2}), \quad (5)$$

where $|g_x|$, $|g_y|$, and $|g_z|$ are one-dimensional projection of gravity vector in x, y, and z axis, while the rest represents two-dimensional projects on x-y, x-z, and y-z plane. The combination $|g_z|$ and $\sqrt{g_x^2 + g_y^2}$ represents the degree of a smartphone tilted along the x-y plane invariant from rotation angle along the z-axis. With two other combinations, namely $|g_y|$ and $\sqrt{g_x^2 + g_z^2}$, and $|g_x|$ and $\sqrt{g_y^2 + g_z^2}$, we are able to infer current orientation of smartphones. Finally, the feature vector used for position classification in the time window i is defined as $(\boldsymbol{f}_i^H, \boldsymbol{f}_i^V, \boldsymbol{o}_i)$.

2.4 Entrance Detector (ETD)

When WSD detects the user standing still for some time, ETD is triggered to determine whether the user actually enters the vehicle or not. To make an accurate decision, ETD looks for three specific features of entering the vehicle: 1) the variance of EMF fluctuations, 2) the magnitude of positive vertical accelerations (heading down toward the ground) caused by sitting motions, and 3) Vehicle Door Closing Sound (VDSC).

However, the accuracy of these components depends on the smartphone's position. For example, a sitting motion feature cannot make accurate decisions when the phone is held in a hand because of random swinging of the arm. On the other hand, VDCS cannot be reliable when the phone is placed in a pocket/bag due to the excessive fricative noises generated when the phone's microphone and the surrounding objects are rubbed against each other.

Considering the constraints imposed by different phone positions, ETD selects a subset of components that are likely to yield correct results as follows.

- When phones are *placed in a pocket/bag*, EMF and sitting motion features are used, while VDCS is not.
- When the phone is *held in a hand*, EMF and VDCS are utilized while sitting motion feature is not.

Unlike other features, EMF is always trusted to yield reasonably accurate results since the magnitude of EMF variance monitored during entry of the vehicle is stable, and is not affected much by the phone's position.

Despite the high performance of EMF feature, ETD performs an AND operation between selected features to draw accurate conclusions for the following reasons. First, individual sensory features can sometimes be found from other daily activities. For example, significant EMF fluctuations can be observed if the user—albeit unlikely—swings or shakes the phone wildly, walking by heavy or electrified metallic objects, or placing the phone right next to other electronic devices in a bag. Furthermore, similar sitting-down acceleration magnitudes can be found when users sit on a chair, and VDCS can be monitored even when the user is near another user who is closing the vehicle door. However, simultaneous occurrence of specified features can be found only when the user is entering the vehicle. In what follows, we detail each of these components by illustrating how the sensory features are extracted.

2.4.1 EMF Variance When Entering the Vehicles

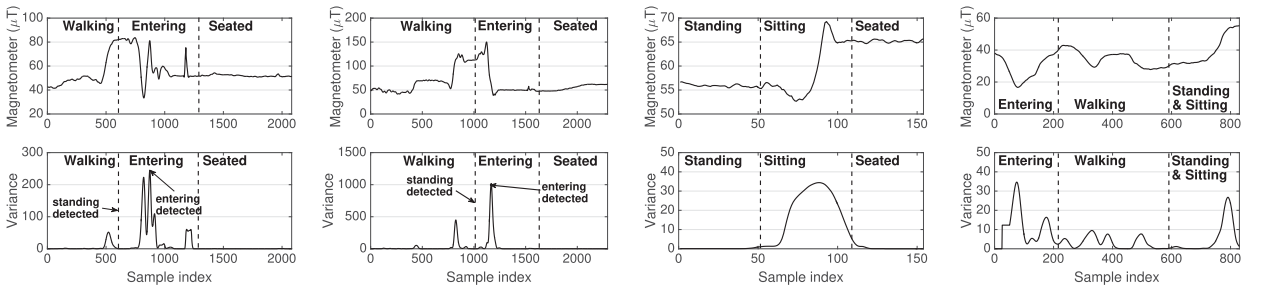
An EMF is a physical field produced by electrically charged objects. Motorized vehicles typically consist of magnetic materials and electronic devices which create a magnetic moment induced by the earth's magnetic field. According to [22], [23], [24], the presence of motorized ground vehicles can be detected by using specially-tuned magnetometers. According to the regulations of International Standardization Organization, every vehicle must pass the electromagnetic compatibility tests which strictly limit the EMF emissions below a certain small level [25]. The EMF emitted from the vehicles can only be detected in close proximity when the magnetometer from commodity smartphones is used.

The main objective is to monitor and analyze the presence of significant EMF fluctuations to verify whether the user has entered the vehicle or not. According to the magnetometer readings shown in Fig. 4a, entering the vehicle generates significant changes in EMF dynamics. To quantify such changes, we use a sliding time window of length m to segment incoming magnetometer readings. For each window W_i^{emf}, let e_{ij}, $1 \leq j \leq m$, represent the jth EMF magnitude within the window i. With these magnitudes, we take the variance

$$var_i^{emf} = \frac{1}{m}\sum_{j=1}^{m}(e_{ij} - \bar{e_i})^2, \quad (6)$$

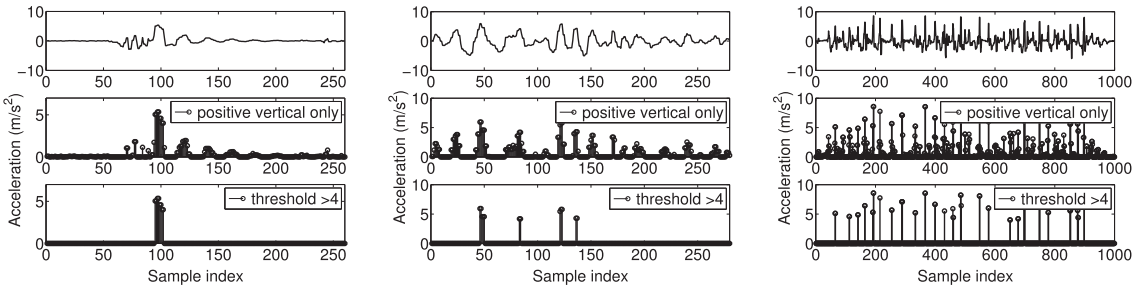
where $\bar{e_i}$ is the average of EMF magnitudes in W_i^{emf}.

The result shown in Fig. 4 illustrates that the variance spikes when the user enters the vehicle and converges to a low level once seated. While other sitting motions also generate similar EMF patterns due to the shifts of smartphone orientation, the magnitude of the EMF variance monitored when entering the vehicle is much greater than the others. Note that walking in between densely parked vehicles also causes significant EMF variance spikes as shown in Fig. 4b, but they are not captured by AIDS since magnetometers are turned on once the user is detected standing. In addition, they do not incorporate simultaneous sitting-down motions. The average variance of entering the vehicle is found to be approximately 300 while other ADLs show the average variance under 100. With an appropriately-chosen variance cut-off threshold, ETD assumes that the user has entered the vehicle if the variance exceeds the threshold.



(a) entering a vehicle (b) entering densely parked vehicle (c) sitting in front of a desk (d) sitting in a public transportation

Fig. 4. EMF dynamics when entering a vehicle versus other daily activities.



(a) sitting, in a pocket (b) walking, in a hand (c) walking, in a pocket

Fig. 5. Vertical acceleration measured with different smartphone poses and activities.

2.4.2 Positive Vertical Acceleration of the Sitting Motions

Sitting-down motions cause our body to accelerate toward the ground without cyclic actions. In fact, ETD preliminarily assumes that the user has entered the vehicle if the following two acceleration features are detected: 1) significant Positive Vertical Acceleration (PVA), and 2) acyclic acceleration patterns.

To compute the vertical accelerations v_t at time t, we take the dot product of acceleration and gravity vectors

$$v_t = \boldsymbol{a}_t^T \cdot \boldsymbol{g}_t. \tag{7}$$

where $\boldsymbol{a}_t = (a_x^t, a_y^t, a_z^t)$ is the acceleration vector at time t, and $\boldsymbol{g}_t = (g_x^t, g_y^t, g_z^t)$ is the unit gravity vector. Once v_t's are calculated, we take the magnitude of PVA to represent the downward movements. The results in Fig. 5—where the top row represents the raw acceleration magnitude, the middle row shows the PVA only, and the bottom row illustrates the acceleration cycles after thresholding—indicate that sitting-down motions produce an acyclic acceleration patterns unlike walking. Furthermore, PVA magnitudes of sitting-down motions are found to be greater than 4 m/s^2. Note that our participants were in late 20s' and early 30s', and hence we did not consider acceleration magnitudes found in elderly and disabled individuals.

Based on thousands of experiments, we have designed ETD to think that a significant vertical movement is present when PVA exceeds 4 m/s^2 while verifying the presence of acyclic acceleration patterns by using the time difference of PVA peaks after applying the threshold. As Fig. 5a shows, PVA peaks for the sitting-down motions are detected within a time window of a second or less (50 samples or less) while others, Figs. 5b and 5c, show larger time differences.

Note that above two features can be manipulated if someone intentionally swings or drops his/her phone down toward the ground, but ETD would not be tricked since EMF features discussed earlier are highly likely to be detected when the user actually enters the vehicle.

2.4.3 Vehicle Door Closing Sounds

To verify whether the user has entered the vehicle or not, we exploit the fact that VDCSs are designed by the manufacturers to have distinct impulsive acoustic features as described in [26]. The presence of VDCSs are detected as follows. First, we use sliding time windows of length b each to segment incoming acoustic signals. For each window W_i^{acst}, let s_{ij}, $1 \leq j \leq b$, represent the jth acoustic amplitude within the window i. With these amplitudes, we verify the presence of impulsive sounds by measuring the maximum amplitude level and the amount of energy enclosed within the window. Enclosed energy is obtained by computing the Short-Term Energy (STE) as

$$en_i = \sum_{j=1}^{b} s_{ij}^2 w(i), \tag{8}$$

where b is the total number of samples in W_i^{acst}, and $w(i)$ is the windowing function used. After conducting extensive experiments with different types of vehicles, we have found that VDCSs generate a maximum amplitude of 0.89 with standard deviation of ± 0.03, and STE above 500. Furthermore, a frequency spectrum analysis using DFT has shown the average duration of VDCSs to be approximately 0.5 second with standard deviation of ± 0.04. An example analysis of VDCS is provided in Fig. 6.

If the monitored sound is found to be impulsive, ETD then computes a feature vector composed of Mel-Frequency Cepstral Coefficient and the duration of the sound. With this feature vector, we use a binary classifier with two sound groups, VDCS and non-VDCS. ETD can, of course, utilize other classifiers, e.g., one might use one-class Support Vector Machine [27], which uses outlier-detection mechanisms when the number of observations are quiet small to clearly define non-VDCSs.

2.5 Entrance Direction Classifier (EDC)

Once ETD confirms that the user has entered the vehicle, AIDS asks EDC to verify the entry directions, left or right. In general, drivers are seated at the left-front (right-front for United

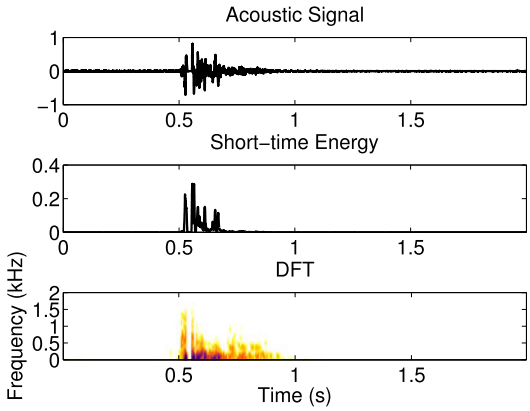


Fig. 6. Vehicle door-closing sound analysis.

Kingdom and Japan) while passengers are occupying the other seats. For accurate identification of the driver's phone, AIDS must identify different entry directions (left or right) and seated rows (front or rear). Considering the fact that entering from the right-hand side eliminates the chance of being the driver, SRC is initiated only if EDC concludes that the user has entered the vehicle from the left (driver-side).

To identify the entry directions, we leverage the body rotations taken when entering the vehicles as shown in Fig. 7. Thus, entering from the left side of the vehicle causes our body to turn counter-clockwise while entering from the right turns clockwise. However, after experimenting thousands of vehicle riding actions with different smartphone poses, we found that the smartphones could experience and detect angular rotations independent of the body rotation generated when entering a vehicle if the user 1) swings his phone, and 2) takes the bag off from his shoulder. For instance, users entering the vehicle from the left-side while taking the bag off from their right-shoulder could be classified as entering from the right-side due to the counter-clock-wise rotations observed when taking their bag off. To handle this problem, EDC is designed to analyze the significant angular rotations, around the global z-axis (yaw), monitored when the user is *entering* the vehicle by identifying the exact moment when the vehicle entering EMF fluctuation is detected, and cropping the collected sensory data to have clean rotation features.

When computing the angles, one must note that local gyroscope readings cannot correctly infer the global yawing due to the difference between global and local sensor axes as shown in Fig. 8. For example, imagine there are two smartphones floating in the air, one facing the sky (local axes perfectly aligned with global axes) and other facing the ground (local z-axis facing the opposite direction from that of global while the other axes are aligned). When we physically rotate them clock-wise, the former will portray an increasing yaw indicating a clock-wise rotation while the latter produces a decreasing yaw indicating a counter clock-wise rotation. As a result, one could draw inaccurate conclusions that the user has sat on the right side of the vehicle when s/he entered the vehicle from the left.

One way to solve this problem is to physically fix the phone's orientation so that the local sensor axes are always aligned with the global sensor axes where y-axis pointing toward the magnetic North Pole. Such an approach will generate accurate sitting trajectories while being highly impractical since handheld devices must be movable without any physical restriction. Since physical constraints are not imposed, we instead use the Quaternion algorithm [28], [29] to virtually rotate the local sensor

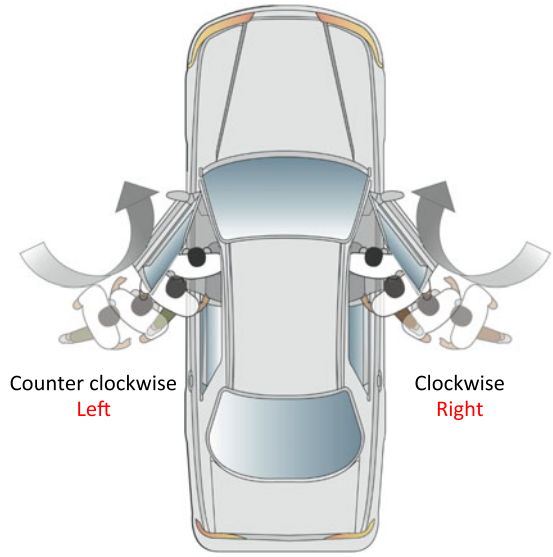


Fig. 7. Different rotations taken when entering a vehicle from driver and passenger sides.

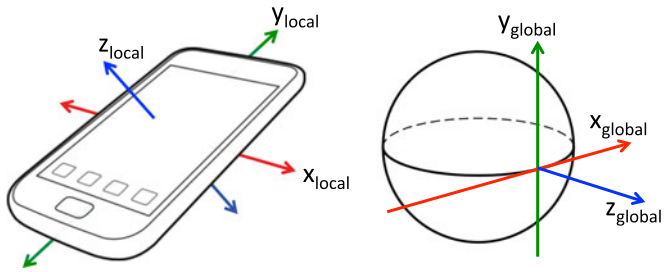


Fig. 8. Difference between local and global sensor axes.

axes by utilizing gyroscope and magenetometer readings. There are two advantages of using Quaternion over others, such as Euler's [30] and orthogonal matrix-based approaches [31]. First, it offers a better chance of avoiding the gimbal lock phenomenon [30] than that of Euler's. Second, it provides a simpler representation of a rotation matrix than that of orthogonal matrices, since Quaternion has 4 numbers while matrix-based approaches have 9.

According to [32], [33], [34], the current smartphone orientation $\boldsymbol{q}$ in a Quaternion 3-dimensional space is defined by the combination of a unit vector $\boldsymbol{u}$ and a scalar θ as

$$\boldsymbol{q} = e^{\frac{\theta}{2}\boldsymbol{u}} = \cos\frac{\theta}{2} + \boldsymbol{u}\sin\frac{\theta}{2}, \tag{9}$$

where $\boldsymbol{u} = (u_x, u_y, u_z) = u_x i + u_y j + u_z k$, and u_x, u_y, u_z are euclidean vector components over three Cartesian axes i, j, k. The extension of Eq. (9) is

$$\boldsymbol{q} = e^{\frac{\theta}{2}(u_x i + u_y j + u_z k)} = \cos\frac{\theta}{2} + (u_x i + u_y j + u_z k)\sin\frac{\theta}{2}. \tag{10}$$

The desired virtual rotations $\boldsymbol{p}$ can be applied to the current smartphone orientation $\boldsymbol{q}$ by taking the conjugation of $\boldsymbol{p}$ by $\boldsymbol{q}$

$$\boldsymbol{p}' = \boldsymbol{q}\boldsymbol{p}\boldsymbol{q}^{-1}, \tag{11}$$

where $\boldsymbol{p} = (p_x, p_y, p_z) = p_x i + p_y j + p_z k$, and $\boldsymbol{p}'$ is the global axes that we aim to be aligned with. For every sensor reading, we compute the current smartphone orientation along the global z-axis by taking the difference between the local axes $\boldsymbol{q}$ and the global axes $\boldsymbol{q}'$.

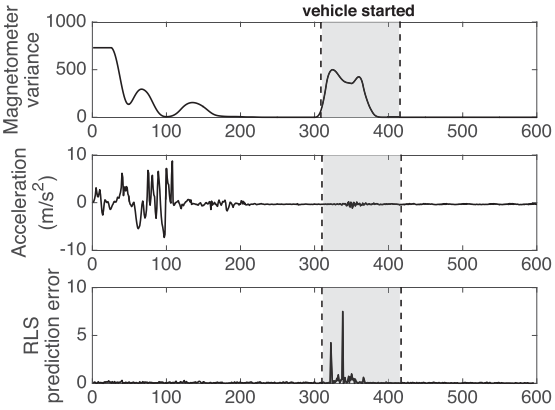


Fig. 9. An EMF spike monitored when a vehicle is starting.

2.6 Seated Row Classifier (SRC)

After AIDS concludes that the user has entered the vehicle from left, SRC is initiated to classify the seated (front or rear) row. The salient aspect of SRC is that the seated rows are differentiated using the subtle EMF fluctuations caused by starting the vehicle, which motorizes the engine, and powers the electronic devices. Such fluctuations are likely to be detectable only at the front row since electronic devices are densely populated in front of the driver's seat while most vehicles have their engines at the front.

Assuming that the phones won't move much, such as relocating the phone to another seat or shaking, while starting the vehicle, EMF features shown in Fig. 9 can be detected. Considering the fact that a user has started the vehicle in between sample index 300 and 400, one can see that there is a subtle EMF fluctuation that peaks at sample index 310, and slowly converges to a normal state.

To achieve accurate classification, we designed SRC to incorporate data-cropping mechanisms to filter out EMF fluctuations caused by the human activities, such as relocating the phone or shaking, before the vehicle is started. While we assumed that the users would not make significant movements when starting the vehicle, there could be an action that causes EMF to fluctuate like starting the car. To handle such a problem, we crop collected sensory data where 1) no significant acceleration is detected, and 2) prediction error of monitored EMF readings is higher than a threshold as shown in Fig. 9, which is computed by using a Recursive Least Squares (RLS) filter [35]. Once the subtle changes are detected, SRC quantifies the phenomena using the variance with Eq. (6).

The RLS filter can be described as follows. At time k, SRC has gathered $k \times 1$ magnetometer readings, $\boldsymbol{e}_k = (e_k, \ldots, e_1)$. With $\boldsymbol{e}_k$, we formulate a least squares prediction problem

$$\hat{e}_t = \boldsymbol{h}_k^T \boldsymbol{e}_{t-1}(M), M < t \leq k, \tag{12}$$

where $k \geq M+1$ and $\boldsymbol{h}_k$ are the $M \times 1$ filter-weight vector at time k defined as

$$\boldsymbol{h}_k = (h_k, \ldots, h_{k-M})^T, \tag{13}$$

and $\boldsymbol{e}_{t-1}(M)$ is the $M \times 1$ past readings as

$$\boldsymbol{e}_{t-1}(M) = (e_{t-1}, \ldots, e_{t-M})^T. \tag{14}$$

Our objective is to find an estimator $\boldsymbol{h}_k$ at time k that minimizes the Sum of Squared Errors (SSE)

$$SSE_k = \sum_{t=M+1}^{k} \lambda^{k-t} |c_t - \boldsymbol{h}_k^T \boldsymbol{e}_{t-1}(M)|^2, \tag{15}$$

where λ (≤ 1) is an exponential forgetting factor, which needs careful tuning for accurate results. A smaller λ places a higher weight toward the most recent information. For the above least squares prediction problem, we develop a recursive algorithm that updates the filter-weight vector upon reception of new magnetometer readings, given the previous filter weights up to $k-1$. The RLS algorithm first calculates *a priori* prediction error based on old filter-weight estimates at iteration k

$$\alpha_k = c_k - \hat{\boldsymbol{h}}_{k-1}^T \boldsymbol{e}_{k-1}(M). \tag{16}$$

The filter-weight vector is then updated as

$$\hat{\boldsymbol{h}}_k^T = \hat{\boldsymbol{h}}_{k-1}^T + \alpha_k \boldsymbol{g}\boldsymbol{v}_k, \tag{17}$$

where $\hat{\boldsymbol{h}}_M = 0$ and an $M \times 1$ gain vector $\boldsymbol{g}\boldsymbol{v}_k$ is computed by

$$\boldsymbol{g}\boldsymbol{v}_k = \frac{\boldsymbol{P}_{l-1}\boldsymbol{c}_{l-1}(M)}{\lambda + \boldsymbol{c}_{l-1}^T(M)\boldsymbol{P}_{l-1}\boldsymbol{c}_{l-1}(M)}. \tag{18}$$

$\boldsymbol{P}_l$ is an $M \times M$ inverse correlation matrix, initialized as

$$\boldsymbol{P}_M = \rho^{-1}\boldsymbol{I}, \tag{19}$$

with a small positive ρ, and recursively updated by

$$\boldsymbol{P}_k = \lambda^{-1}\boldsymbol{P}_{k-1} - \lambda^{-1}\boldsymbol{g}\boldsymbol{v}_k \boldsymbol{c}_{k-1}^T(M)\boldsymbol{P}_{k-1}. \tag{20}$$

A priori prediction error α_k computed with Eq. (16) is used to quantify the prediction errors as

$$p_k = \frac{|\alpha_k|}{max\{c_k, c_{min}\}}, \tag{21}$$

where $k \leq M+1$. The changes of prediction errors are plotted in Fig. 9.

From thousands of experiments, we have found that starting the vehicle causes prediction errors to spike up by more than 0.5 for most of the trials. Once such a prediction error peak is detected, SRC measures the variance of magnetometer readings, and takes the maximum. At last, if both prediction error peak and maximum variance satisfy specific cut-off thresholds, then SRC concludes that the user is seated at the front row, or not.

2.7 Designing Distracted Driving Prevention Services

When designing Distracted Driving Prevention Services (DDPSs), one must carefully consider the following three issues for usability. First, driver detection systems, such as AIDS, sometimes make inaccurate decisions regarding identification of DS. One may thus incorporate additional checking mechanisms, which use sensory features detectible while driving, introduced in [10], [12], [14] for correct determination of the user status.

Second, DDPSs must disable distractive apps and services only when the vehicle is in motion. According to AT&T DriveMode [4], DDPSs are activated when the speed of vehicle estimated by Global Positioning System (GPS) exceeds 25 miles per hour. While GPS provides accurate speed estimation, there are two technical shortcomings we must overcome. First, GPS does not work when the line of sight between the phone and the satellite is blocked, i.e., in-door parking lot, tunnel, and urban area. Second, GPS is energy-expensive due to the communication mechanism requiring the antenna to be powered at all times, preventing the smartphone from moving to sleep state. In order to resolve these issues, one should consider the use of

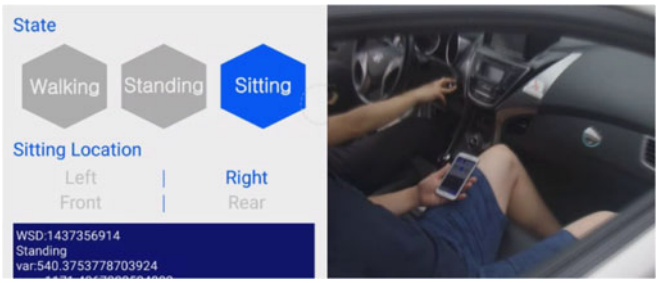


Fig. 10. Snapshot of an experiment conducted with an AIDS prototype application.

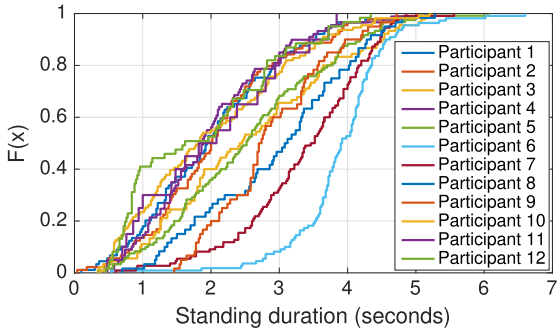


Fig. 11. CDF of the standing duration before entering a vehicle.

accelerometer features generated when the vehicle starts to accelerate from stop state to activate DDPSs. On the other hand, the presence of deceleration can be used to deactivate DDPSs, temporarily allowing the driver to use the smartphone. Moreover, AT&T DriveMode must be activated *manually* by the phone owner, whereas AIDS allows DDPSs to stop distractive activities *automatically*.

Third, DDPSs must be designed to handle the case when the driver asks a passenger to send a text on his/her behalf. In such a case, DDPSs may ask the passenger to simultaneously place two thumbs on the screen for a few seconds to deactivate the proposed solution, which can be detected by using the touch sensors embedded in commodity smartphones. Considering how dangerous it is to take both hands off the steering wheel, unless the vehicle is completely stopped, drivers will not attempt to unlock the device while driving.

3 Evaluation

3.1 Experimental Setup

To comprehensively evaluate the feasibility of AIDS, we have conducted two separate experiments where 1) thousands of vehicle-riding actions are performed with different vehicle types in order to obtain appropriate cut-off thresholds for the prototype application based on the sensory features monitored, and 2) a fine-tuned prototype is deployed in our daily lives to extensively test the AIDS's performance. Throughout this section, we call the former *controlled* experiments and the latter *normal* experiments.

The prototype of AIDS is implemented on multiple Samsung Galaxy S5s running on Android [36] platfrom as shown in Fig. 10. In addition, we evaluated the impact of varying sensor qualities in other smartphones by employing Apple iPhone 6S plus running on iOS platforms [37]. According to our measurements, the EMF readings from iPhone 6S plus show relatively greater magnitudes than those from Galaxy S5 even when both devices are placed at the exact same location. However, the magnitude of EMF variances monitored when the user enters the vehicle and starts the vehicle does not differ much from each other.

These smartphones are equipped with accelerometer, gyroscope, microphone, and magnetometer sensors. The sampling rates of kinematic (accelerometer and gyroscope) sensors and magnetometer are set to 50 Hz (50 samples per second), while microphones are sampled at 44,100 Hz. Note that increasing the sampling rate does not lead to proportional enhancements of the accuracy of the proposed system modules and components due to unstable sampling precisions at higher rates (i.e., the highest sampling rate provided by Android API has a wider range, 100∼120 Hz, than that of a lower one, 48∼52 Hz).

We conducted experiments with 12 male and female participants within their 20 and 30s, and used 7 different vehicles that fall under four different segments all running on petrol engines as follows:

- Kia Pride (B-segment small cars)
- Hyundai Accent (B-segment small cars)
- Hyundai Avante (C-segment medium cars)
- Hyundai Sonata (D-segment large cars)
- Kia K5 (D-segment large cars)
- Hyundai Genesis (E-segment executive cars)
- BMW 520D (E-segment executive cars)

While trucks and sport utility vehicles are not considered in these experiments, the vertical acceleration feature used for ETD can be slightly adjusted to accept the upward magnitudes.

3.2 Evaluation Criteria

Our evaluations use three different performance criteria based on 1) histograms, 2) Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC), and 3) system TPR and TNR. While histograms show the difference between the sensory features extracted from vehicle riding actions and other daily activities with visual illustrations, ROC, AUC, TPR and TNR portray the accuracy of the proposed system modules and its corresponding features.

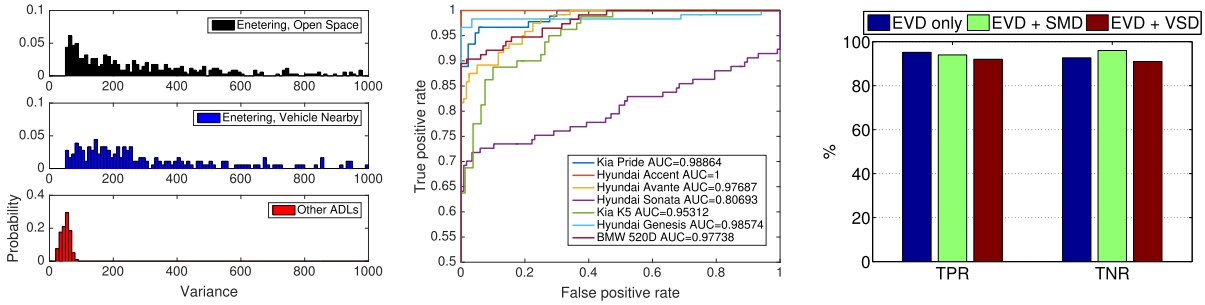
An ROC curve is used widely in statistics to illustrate the performance of a binary classifier where the curve is generated with the TPR against the False Positive Rate (FPR) under various threshold settings. The TPR infers the ratio of true positive for the true case while FPR infers the ratio of false positive for the false case. With an ROC curve, one can evaluate the performance of a classifier by computing the AUC, where AUC close to 1 indicates that the system is able to differentiate the target phenomena perfectly from the others while AUC ≤ 0.5 indicates that the classifier is meaningless.

3.3 Evaluation Results

3.3.1 Performance of WSD and SPC

Since the performances of WSD and SPC have already been evaluated in [16], showing higher than 94 percent accuracy, we will focus on the duration of the standing state monitored prior to entering a vehicle. We found that generating hundreds of standing events throughout a day and initiating the core system modules on all events consume lots of energy on a smartphone. Considering the phone's limited battery capacity, we must therefore carefully set the duration threshold in order to prevent excessive sensor activations.

The cumulative distribution function shown in Fig. 11, generated using data collected from controlled and normal experiments, infers that people stand by vehicles for an average of 2.58 ± 1.23 seconds before entering them. The minimum duration is approximately 0.04 second while the maximum is



(a) Histogram (b) ROC and AUC (c) Controlled TPR and TNR

Fig. 12. Feasibility of ETD under controlled vehicle riding actions and environments.

6.6 seconds. Furthermore, the 5th and 50th percentiles are observed to be 0.66 and 2.58 seconds, respectively. According to these findings, we designed WSD to initiate ETD only when the user is detected standing for more than a second. This cutoff time is approximately the 12th percentile of the observed data. Such a duration threshold initiated ETD 97 times, on average, throughout a day while the total number of standing motions monitored is approximately 250, i.e., filtering out almost 60 percent of unnecessary ETD initiations.

While the actual number of vehicle riding events is much smaller than that of ETD initiations, indicating a high FPR, we are only using accelerometer readings to capture entering-a-vehicle-like events without wasting battery energy on all events. If the accuracy is a primary concern, then we can achieve better TPR and TNR by adjusting the cut-off threshold or using the magnetometer readings to detect the presence of the vehicles at the expense of energy-efficiency.

On the other hand, we can minimize the amount of energy wasted by carefully controlling the activation time of triggered sensors. According to our experiments, ~93 percent of the users with the intention of riding a vehicle completed a set of actions required to enter the vehicle (opening the vehicle door, sitting down and being seated) within next 10 seconds after a standing motion is detected. Considering the importance of energy-efficiency, we, therefore, designed AIDS to deactivate triggered sensors (excluding accelerometers) if no vehicle-riding actions are detected within 10 seconds as shown in Fig. 1.

3.3.2 Performance of ETD

Once WSD concludes that the user is about to enter a vehicle, AIDS initiates ETD using different sets of sensory features based on the SPC result. Boolean values (*true* or *false*) returned from analyzing selected features are *AND*ed to determine whether the user has indeed entered a vehicle or not. To demonstrate the feasibility and robustness of ETD under various conditions, we have instructed the participants to ride the vehicles under two different settings.

In the controlled setting, participants made clean vehicle-riding actions for accurate feature extractions. In contrast, the normal setting allows participants to take common vehicle-riding actions while satisfying the specified assumptions and restrictions, e.g., smartphones do not make significant movements when vehicles are started. In what follows, we first evaluate the performance of each system component under controlled settings, and then discuss the feasibility and robustness of ETD under widely varying normal conditions.

With thousands of EMF variances computed using magnetometer readings captured from controlled vehicle-riding actions and other most commonly found daily activities involving sitting motions (sitting on a desk chair and a public transportation vehicle), we have evaluated the performance of the Entry EMF Variance Detection (EVD) as shown in Fig. 12. According to the histograms shown in Fig. 12a, the EMF variances observed from other daily activities are densely distributed between 0 and 100 while entering the vehicle shows a wider range of variance. Note that entering a vehicle parked in between two other vehicles does not deviate much from the EMF variances monitored from entering a vehicle parked in an open space with no vehicle around it. Despite the small overlaps between entering a vehicle and other ADL distributions, we can differentiate the vehicle-entering events by carefully selecting a cut-off threshold. Fig. 12b shows that AUC for each vehicle exceeds 0.95, except for Hyundai Sonata, indicating that the EMF variance accurately differentiates the action of entering the vehicle from other daily activities. By setting a cut-off threshold to 50, EVD achieved an average system TPR of 96.1 percent while TNR equals 94.6 percent as illustrated in Fig. 12c.

For the performance of Sitting Motion Detection (SMD), we analyzed the collected sensory data to find AUC approximately equal to 0.96. Based on the analysis, we specified a PVA threshold to be 4 m/s^2 as shown in Fig. 5, and set the time difference threshold to be 50 samples (equivalent to 1 second). The results indicate that SMD achieves system TPR of 98.3 percent while TNR is found to be 91.7 percent. Finally, the performance of Vehicle door closing Sound Detection (VSD) was evaluated in two different environments, noisy Costco parking lot and quiet outdoor open space, where different levels of ambient noises were present. As shown in Table 1, the level of ambient noises observed from the Costco parking lot was approximately −25 dB (0 to −90 dB relative to the full scale) while quiet open space showed −42 dB. By entering a vehicle, these ambient noises are drastically attenuated due to sound absorbing materials embedded in various vehicular components. With ambient noises captured inside and outside of the vehicle, we then computed the signal-to-noise ratio (SNR) to illustrate the impact of external noises on the VSD performance. SNR between VDCSs and ambient noises captured outside the vehicle is 13.2 dB, which

TABLE 1
Impact of Ambient Noises on VSD Performance

	Noisy Costco parking lot	Quiet open space
Outside	−25 dB	−42 dB
In-vehicle	−45 dB	−47 dB
SNR,	13.2 dB	29.1 dB
Outside	(21 times greater)	(830 times greater)
SNR,	29.8 dB	30.2 dB
In-vehicle	(980 times greater)	(1,100 times greater)
VSD accuracy	87%	89%

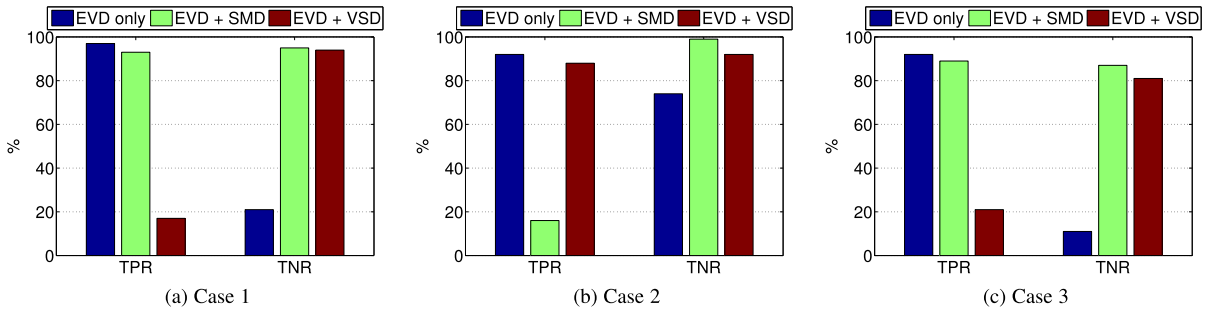


Fig. 13. TPR and TNR of ETD with different combinations of system components under various environmental and behavioral conditions: (a) vehicles parked at an indoor parking lot (case 1); (b) vehicles parked outside and smartphones swung recklessly (case 2); and (c) smarphones kept in a bag (case 3).

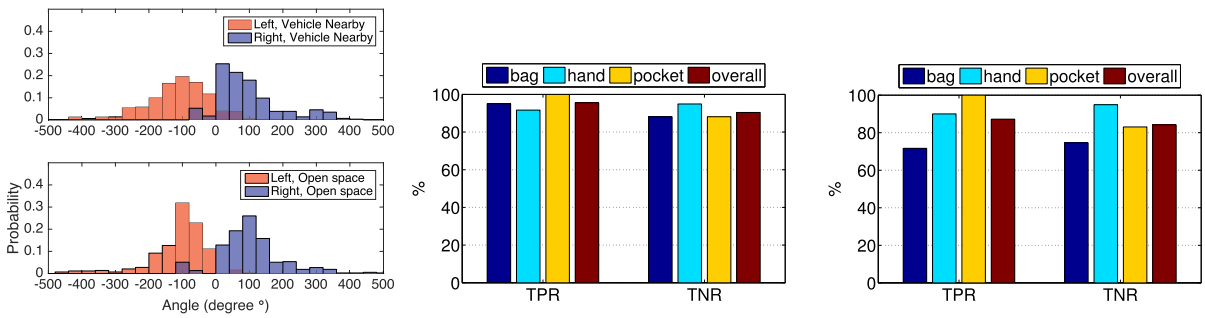


Fig. 14. Angular rotations when entering a vehicle, and performance of EDC for different smartphone positions.

indicates that VDCSs are approximately 21x stronger than ambient noises. In the in-vehicle scenario, SNR is 29.8 dB. Such a ratio indicates that VDCSs are approximately 980x stronger than the ambient noises observed inside the vehicle. Under different environments with different levels of ambient noises, VDCSs were detected with an average of 88 percent accuracy. With a set of sensory features extracted from controlled vehicle riding actions, ETD achieves, on average, 93.7 percent TPR and 93.2 percent TNR.

Since ETD shows reasonably accurate detection performance when specified sensory features are cleanly extracted, the next step is to evaluate the feasibility of ETD under varying environmental conditions that might distort the required information. As Fig. 13 shows, we have identified 3 different cases that distorted required sensory features. For Case 1, vehicles were parked at an indoor parking lot where heavy metallic objects, such as other vehicles, air ventilation control units, etc., are present to distort the EMF feature. The user placed his smartphone inside his pocket, and conducted experiments. For Case 2, vehicles were parked outside, and smartphones are recklessly swung in the users' hands. Finally, vehicles were parked outside for Case 3 while smartphones are kept in a bag along with other electronic devices, such as laptop, tablet and external batteries.

Fig. 13a shows that using EVD only achieves approximately 97 percent system TPR while TNR is below 22 percent. Such low TNR is caused by the fact that even when the user did not enter the vehicle, significant EMF variance spikes are detected due to other metallic objects, leading EVD to conclude that the user has entered the vehicle while s/he has not. To compensate for such shortcomings of EVD in Case 1, SMD or VSD is used. While a EVD+SMD combination is shown to yield satisfactory TPR and TNR results, VSD drops the TPR below 20 percent due to the excessive fricative noises caused inside the pocket. While there is a small performance drop for system TPR when SMD is combined with EVD, it is only a marginal loss for TNR.

For Case 2 shown in Fig. 13b, EVD only achieves 92 and 74 percent system TPR and TNR, respectively. The degradation in system TNR is due to the fact that reckless swinging sometimes replicates the EMF variance monitored when entering the vehicle. To compensate this, a EVD+VSD combination is used for accurate ETD since SMD is infeasible and PVA features can be replicated by swinging the smartphone down toward the ground. At last, Case 3 shown in Fig. 13c, TNR of EVD is lower than 20 percent due to the EMF variance spikes caused by other electronic devices placed right next to the driver's phone. Our experiments illustrate that VSD cannot be used to compensate the performance degradation due to the fricative noises, while EVD+SMD combination yields 89 and 87 percent TPR and TNR, respectively.

Under varying environmental settings, we found that ETD achieves an average of 90 percent TPR and 91 percent TNR. We observed a slight (about 2 percent) performance degradation when AIDS was tested under normal experimental settings, demonstrating the feasibility of the specified sensor features and the proposed ETD design.

3.3.3 Performance of EDC

The performance of EDC was evaluated extensively under two different behavioral settings. First, participants are instructed to take the bag off from their shoulder before opening the vehicle's door, and phone swinging and shaking motions are avoided, thus enabling clear differentiation of the associated body rotations from the others. Second, participants are allowed to enter the vehicle as they would have done in their daily routines.

Under the controlled setting, entering directions are differentiated as Fig. 14a illustrates. Note that entering a vehicle parked in between two other vehicles, providing narrower entrance spaces for the users, shows relatively larger overlaps between two sides. EDC yields the best performance (TPR $>$ 99 percent) when the phone is placed in the pocket while bag and hand

TABLE 2
SRC Performance and EMF Feature Statistics Monitored When Vehicle Engine Starts

	Kia Pride		Hyundai Accent		Hyundai Avante		Hyundai Sonata		Kia K5		Hyundai Genesis		BMW 520D	
	front	rear	front	rear	front	rear	front	rear	front	rear	front	rear	front	rear
EMF mean	8.9	0.4	9.4	1.8	33.1	1.3	48.7	15.4	7.1	6.1	334.3	202.6	161.3	32.0
EMF std	11.8	1.3	21.0	4.3	137.7	4.1	113.1	67.2	33.4	3.3	150.1	239.6	615.8	90.8
EMF min	2.2	0.01	0.03	0.03	2.1	0.02	0.3	0.01	2.7	0.01	114.9	0.1	10.1	0.08
EMF max	26.6	7.8	118.4	22.1	693.5	31.3	785.8	495.7	237.9	10.1	675.9	961.1	3778.2	458.7
AUC	.9681		.8358		.9248		.9414		.9597		.9876		.9090	
accu., contr.	99.1%	95.7%	92.1%	90.3%	95.8%	86.4%	92.6%	93.5%	99.5%	90.3%	98.9%	85.3%	97.2%	87.1%
accu., norm.	93.3%	95.8%	84.6%	84.6%	88.9%	85.3%	82.8%	93.3%	94.7%	87.1%	96.7%	79.3%	96.3%	80.2%

positions show relatively lower performance (TPR >91 percent) as shown in Fig. 14b. The average TNR of EDC is found to be approximately 90.4 percent. While EDC shows reasonably accurate performance when vehicle riding actions are carefully controlled, Fig. 14c shows the overall performance degradation when participants freely enter the vehicle.

EDC yields the worst results when the smartphone is kept in a bag since the angular rotation monitored when taking the bag off from his/her shoulder could override the body rotations taken while entering the vehicle. According to our observations, taking off the bag from the left-shoulder generates clock-wise angular rotations while the right-shoulder shows counter-clock-wise rotations. Since smartphones held in a bag generates independent angular rotations from that of the body, EDC is designed to analyze the significant angular rotations monitored when the user is entering the vehicle by employing the EMF variance fluctuation feature. However, two different angular rotations (one from taking the bag off from the shoulder while the other from entering the vehicle) tend to overlap with each other, degrading the EDC performance. In addition, smartphones held in a hand show relatively lower accuracy than that of pocket since one might swing his arms in a way that the direction of angular changes is reversed to infer incorrect states.

Our performance evaluation has revealed that placing smartphones in a bag significantly degrades the EDC performance, causing an around 7 percent drop of overall system TPR and TNR—the system TPR shifted from 95.6 percent (controlled) to 87.2 percent (normal) while TNR from 90.4 to 84.2 percent.

3.3.4 Performance of SRC

SRC was evaluated by analyzing the magnetometer readings taken when the vehicle is started after the driver entered the vehicle. For comprehensive evaluation of SRC, participants are first instructed to start the engine after placing the phone on a flat surface (on top of the lap, or seat) to avoid unwanted EMF spikes caused by orientation shifts. After completing these controlled experiments, the participants are allowed to start the engine without worrying about the smartphone positions and placements.

According to our analysis in Table 2, the EMF variance monitored at the front row has a higher magnitude than that of rear across different vehicles. Despite the fact that the distributions of EMF variance monitored from the front and rear rows have small overlaps, the overall AUC results indicate that the proposed solution achieves reasonably good classification accuracy. Note, however, that luxurious automobiles, such as Hyundai Genesis and BMW 520D generate significantly higher variance magnitudes, which, in fact, require AIDS to have per-segment calibration for accurate front and rear row classification results. Vehicle segment identification is part of our future work to extend the practicality of AIDS.

By setting the cut-off EMF variance threshold to 1.2 for Kia Pride, Hyundai Accent, Hyundai Avante, Hyundai Sonata, and Kia K5, TPRs of SRC, as shown in Table 2, were found to be >92 percent while TNRs were above 90 percent. For Hyundai Genesis and BMW 520D, TPR was >96 percent while TNR was >85.3 percent with the cut-off threshold of 123. The relatively lower TNR was due to electronic devices at the rear seats in luxurious automobiles, e.g., DVD players and electronically adjustable seats.

Next, we extended our experiments by taking normal vehicle riding actions while starting the engine. The participants were found to have tendency to send a text, browse web, or answer the phone while starting the vehicle. Such actions override specified EMF features, degrading system TPR and TNR by approximately 6 percent, on average, with the collect data. Low-end vehicles suffer from a greater performance degradation than that of luxurious vehicles since the EMF variances monitored from the former are much less than those of latter, making EMF fluctuations caused by the driver's motion difficult to override the EMF feature of starting the vehicle.

3.3.5 Performance of Driver Identification

With sensory features monitored from thousands of vehicle riding motions, we have analyzed each system module. The last step is to evaluate the driver identification performance by combining EDC and SRC results monitored under various conditions (as illustrated in ETD, EDC, and SRC evaluations) with the *AND* logical operator. According to our system design, a smartphone is concluded to belong to the driver if and only if EDC concludes that the user has entered the vehicle from the left while SRC results are found to be the front. The results in Fig. 15 indicate that the driver is identified with TPR (drivers are identified as the driver) of 93.3 percent while TNR (passengers are identified as a passenger) is 91.2 percent under controlled settings. On the other hand, AIDS achieves 83.3 percent TPR and 90.1 percent TNR, respectively, when participants are allowed to ride the vehicle freely. Such performance figures also indicate that 6.7-16.7 percent of drivers will be identified as passengers while 8.8-9.9 percent of passengers will be identified as the driver. As noted in Section 2.7, status of the users can be double-checked by designing DDPSs to employ additional verification methods.

Considering the accuracy of individual seat position classification, the worst performance is identified at the rear-right seat with the controlled TPR (the rate of correctly identified seated positions) of 65.9 percent. Such a low classification accuracy is caused by the subtlety of EMF variances monitored from the rear seats as discussed in Section 3.3.4. Our experiments show that the passengers seated at the rear/right position were frequently identified as seated at front/right. Note, however, that this low classification performance in the rear/right seat position

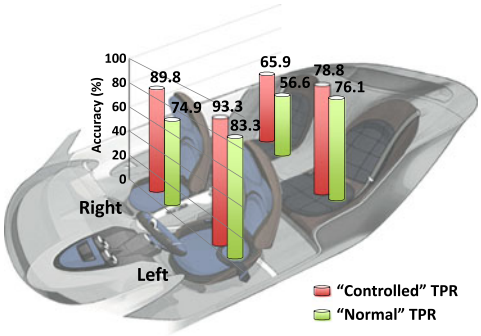


Fig. 15. Performance of AIDS for different seat positions.

(passenger) does not decrease the value of AIDS, since it is designed and tailored to differentiate the driver from the passengers. Nevertheless, in future we plan to improve the accuracy in identifying exact seated positions.

3.3.6 Energy-Efficiency

The last performance metric we consider is the energy-efficiency of AIDS, quantified by the total amount of energy used per day. Energy consumption rate of each sensor is listed in Table 3 based on the studies conducted by [38]. When AIDS is active, accelerometers are used by system modules, consuming 12 mA/day by default. The energy usage of other sensors is determined by the number of times the core system modules are initiated during daily activities. According to our experiments, AIDS requires average of 7.8 μA/s when ETD is initiated. Note that the total energy usage per initiation is determined by how quickly the user finishes his/her riding actions. To quantify the actual energy consumption throughout our daily activities, we have collected the system initiation and battery level logs for three days. Our logs indicate that core system modules were triggered 97 times per day, on average, consuming additional $\sim$140 mA/day compared to the case when AIDS is kept off. Such energy-consumption is equivalent to 10 minutes of GPS navigation, or 5$\sim$10 percent of battery life.

Note that the energy-efficiency of AIDS highly depends on the number of core system module initiations, while the number of actual vehicle riding actions does not have much effect. This phenomenon comes from the fact that activated sensors are used for some amount of time regardless of the presence/absence of vehicle riding actions to detect specified sensory features.

4 Related Work

The latest work on DS detection was reported in [10]. It uses various sensory features extractable from common vehicle riding/driving actions, such as entry swing, wearing a seat belt, pedal-press signature, and turn signal audio, for DS detection. While these features are novel and interesting, there are three critical problems when the objectives of AIDS are considered important. First, sensory features used for front and rear differentiation, pedal-press signature and turn signal audio, cannot guarantee their occurrence before the vehicle enters the traffic. Second, the authors of [10] assumed that all drivers and passengers will have their algorithm installed in their smartphones, which is highly unlikely. Finally, some of the features rely on a cloud server to make accurate decisions.

On the other hand, the authors of [13], [14] proposed sensing of the vehicle dynamics for driver phone use on the basis of the fact that centripetal acceleration observed while making a turn varies with the position in the vehicle. By exploiting the difference between the given smartphone and the reference point, one can decide whether the device is located left or right of the reference location. In order to acquire this reference point, the authors introduced the use of a very small sensory device that can be plugged in a cigarette light adapter which monitors the centripetal acceleration at the center of the vehicle. While incorporating the sensory information observed from a vehicle's turns is interesting, it has two impractical aspects. First, it requires an additional infrastructure. Besides the cost and effort to install the additional devices, the system must be robust when operated as a standalone system. In fact, an infrastructure-free system is much more scalable and easier to deploy. Second, it suffers a long detection latency. Since it requires a number of *good* turns for 95 percent or higher DS detection rate, the time lag between staring the vehicle and making those turns could be very long. The authors also mentioned this latency problem. The driver's safety will be greatly improved if the system can detect the DS before the actual driving.

TABLE 3
Energy Consumption of AIDS

Sensor type	Energy per hour	ETD initiation pose	Energy per initiation
Accelerometers	0.5 mA/h	Riding (hand)	7 μA/s
Gyroscopes	3.2 mA/h	Riding (bag)	9 μA/s
Magnetometers	0.6 mA/h	Riding (pocket)	9 μA/s
Microphones	2 mA/h	Other activities	6 μA/s

Considering the system constraints on scalability and practicality, Bo et al. [11], [12] share by far the closest objectives and results with AIDS, detecting DS without any infrastructure support. The direction of entering a vehicle is identified by fusing the horizontal plane accelerations and 3-axis rotational vectors monitored with commodity smartphone sensors. Their approach to identifying the vehicle entrance direction is very intuitive and limited by the fact that the smartphone must be placed inside a trouser pocket to detect the body, or a leg, rotations. Moreover, they identify the front and rear locations of the device carrier by analyzing the sensory information acquired when a vehicle goes over a bump or pothole on the road. While the DS detection accuracy is above 90 percent, identifying the front and rear using this approach suffers the time lag problem since running over a speed bump or pothole is unlikely to happen when needed. As mentioned earlier, detecting DS before the actual driving is important to the safety of the driver by removing the source of distractions.

While much resources and efforts are continuously being invested to develop feasible DS detection systems, IT companies have also noticed the importance of this problem, and started to make various investments. For example, Apple acquired a patent in the late spring of 2014 regarding driver owned smartphone shutdown methods [6]. This method requires the device owner to take the snapshot of the surrounding environments to go through an image processing phase to conclude the user's whereabout. The patent also specifies the use of external devices which transmit specially forged signals toward the driver's seat to shut down any smartphone located within the targeted area to prevent distracted driving.

While each existing study presents interesting and unique ideas, AIDS is very different from (or even better than) the existing approaches for the following reasons. First, AIDS allows users to carry their smartphones freely while [10] and [12] have a number of impractical constraints on how the phones are carried. The side detection mechanisms introduced

in [10] do not work when the smartphone is held in hand while [12] works only when the phone is placed inside a trouser pocket. The authors of [12] justified these restrictions by stating that 57 percent of male users carry their phone in trouser pockets. However, the referred statistics do not represent the today's actual use of smartphones since it was published in 2005, which was several years before the advent of smartphones. Considering the fact that how to hold/carry a smartphone may change during daily routines based on the user's need and habits, a driver identification system would be less practical/useful if it is designed to operate only in a particular pose.

Second, the features used in AIDS are always present when the user rides a vehicle before the vehicle leaves its parking spot. The side detection mechanisms proposed in [10] cannot always guarantee accurate results when smartphones are not placed in a jacket pocket. While the authors might assume all drivers wear their seat belts, the statistics presented by Centers for Disease Control and Prevention in 2012 show that millions do not buckle up on every trip [39]. Furthermore, the driver detection systems introduced in [12], [14] suffer from indefinite detection delays since we do not know if and when driving events of interest—running over speed bumps or potholes, turn signal lever manipulation and the number of *good* turning movements—will occur. In the worst case, drivers will not be identified, thus not protecting passengers, pedestrians, and vehicles. Even if we assume that these events will take place at some time while driving, drivers are in danger for an indefinite period of time until the events are monitored and analyzed.

Third, AIDS does not need any external assistance. The authors of [10] proposed a front and rear differentiation mechanism that achieves reasonable results by comparing recorded audio with other smartphones. On the other hand, [14] requires additional dedicated hardware to identify the driver. While employing external assistance may increase the accuracy of driver identification, it incurs additional costs to purchase, install, and modify devices, making it unattractive/impractical to the users and the vehicle manufacturers.

While the above approaches are far from meeting the AIDS's design constraints and goals, they indicate the need for an efficient solution to distracted driving as smartphones have become a major source of distraction and thus a serious concern to driving safety.

Most related work focused on detecting and shutting down DS. While such objectives can be met with pre-installed devices, such as blocking the message signal transmitter, the required infrastructure is not available in every vehicle, thus making the solutions less practical. So, we need to detect DS by relying only on commodity smartphone sensors. AIDS meets this important need.

5 Conclusion

Can we accurately identify the DS without external support or unrealistic physical restrictions? AIDS answers this question by effectively extracting, analyzing, and fusing the heterogeneous sensory information on commodity smartphones. Our extensive evaluation of AIDS shows that the TPR and TNR of DS identification are 83.3∼93.3 and 90.1∼91.2 percent, respectively, at the cost of 5∼10 percent reduction of phone battery operation time in a day. At this modest energy-consumption cost, the safety of drivers, passengers and vehicles can be achieved with a high probability by automatically activating necessary distracted driving prevention services. This energy consumption of AIDS can be reduced by smart sensor duty cycling, which is part of our future research.

Acknowledgments

This work was supported by the research fund of Hanyang University (HY-2015-G).

References

[1] Official website for distracted driving. [Online]. Available: http://www.distraction.gov/content/get-the-facts/facts-and-statistics.html, Accessed in 2017.
[2] Governors highway safety association. [Online]. Available: http://www.ghsa.org/html/issues/distraction/index.html, Accessed in 2017.
[3] Distracted driving laws. [Online]. Available: http://www.ghsa.org/html/stateinfo/laws/cellphone_laws.html, Accessed in 2017.
[4] AT&T drivemode app. [Online]. Available: http://www.att.com/gen/press-room?pid=23185, Accessed in 2017.
[5] Bluetooth car kits, bluetooth. [Online]. Available: http://www.bluetooth.com/Pages/Handsfree-Calling.aspx, Accessed in 2016.
[6] J. G. Elias, "Driver handheld computing device lock-out," U.S. Patent 8 706 143, Apr. 22, 2014.
[7] V. K. Lee, C. R. Champagne, and L. H. Francescutti, "Fatal distraction cell phone use while driving," *Canadian Family Physician*, vol. 59, no. 7, pp. 723–725, 2013.
[8] A. Riener and A. Ferscha, "Supporting implicit human-to-vehicle interaction: Driver identification from sitting postures," in *Proc. 1st Annu. Int. Symp. Veh. Comput. Syst.*, 2008, Art. no. 10.
[9] J. Yang, et al., "Detecting driver phone use leveraging car speakers," in *Proc. 17th Annu. Int. Conf. Mobile Comput. Netw.*, 2011, pp. 97–108.
[10] H. Chu, V. Raman, J. Shen, A. Kansal, V. Bahl, and R. R. Choudhury, "I am a smartphone and I know my user is driving," in *Proc. 6th Int. Conf. Commun. Syst. Netw.*, 2014, pp. 1–8.
[11] C. Bo, X. Jian, X.-Y. Li, X. Mao, Y. Wang, and F. Li, "You're driving and texting: Detecting drivers using personal smart phones by leveraging inertial sensors," in *Proc. Annu. Int. Conf. Mobile Comput. Netw.*, 2013, pp. 199–202.
[12] C. Bo, X. Jian, and X.-Y. Li, "TEXIVE: Detecting drivers using personal smart phones by leveraging inertial sensors," arXiv:1307.1756 [cs.NI], Jul. 2013.
[13] Y. Wang, J. Yang, H. Liu, Y. Chen, M. Gruteser, and R. P. Martin, "Sensing vehicle dynamics for determining driver phone use," in *Proc. Annu. Int. Conf. Mobile Syst. Appl. Services*, 2013, pp. 41–54.
[14] Y. Wang, et al., "Determining driver phone use by exploiting smartphone integrated sensors," *IEEE Trans. Mobile Comput.*, vol. 15, no. 8, pp. 1965–1981, Aug. 2016.
[15] U.S. HEV sales by Model (1999–2013). [Online]. Available: http://www.afdc.energy.gov/data/10301, Accessed in 2017.
[16] J.-G. Park, A. Patel, D. Curtis, S. Teller, and J. Ledlie, "Online pose classification and walking speed estimation using handheld devices," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012, pp. 113–122.
[17] H. L. Chu, V. Raman, J. Shen, R. Choudhury, A. Kansal, and V. Bahl, "In-vehicle driver detection using mobile phone sensors," in *Proc. ACM Annu. Int. Conf. Mobile Syst. Appl. Services*, 2011, pp. 397–398.
[18] D.-K. Cho, M. Mun, U. Lee, W. J. Kaiser, and M. Gerla, "AutoGait: A mobile platform that accurately estimates the distance walked," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2010, pp. 116–124.
[19] J. Lester, T. Choudhury, and G. Borriello, "A practical approach to recognizing physical activities," in *Pervasive Computing*. Berlin, Germany: Springer, 2006, pp. 1–16.
[20] T. Evgeniou, M. Pontil, and T. Poggio, "Regularization networks and support vector machines," *Advances Comput. Math.*, vol. 13, no. 1, pp. 1–50, 2000.
[21] A. J. Smola and B. Schölkopf, *Learning with Kernels*. Cambridge, MA, USA: MIT Press, 1998.
[22] S. Ahn and J. Kim, "Magnetic field design for high efficient and low EMF wireless power transfer in on-line electric vehicle," in *Proc. 5th Eur. Conf. Antennas Propag.*, 2011, pp. 3979–3982.
[23] A. Q. Memon, S. L. Lau, and K. David, "Investigation and compensation of the magnetic deviation on a magnetometer of a smartphone caused by a vehicle," in *Proc. IEEE 78th Veh. Technol. Conf.*, 2013, pp. 1–5.

[24] S. Simon, "System and method for identifying vehicle by utilizing detected magnetic field," U.S. Patent 9,333,946, May 2016.
[25] Electromagnetic capability. [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:16750:-2:ed-4:v1:en, Accessed in 2017.
[26] E. Parizet, E. Guyader, and V. Nosulenko, "Analysis of car door closing sound quality," *Appl. Acoust.*, vol. 69, no. 1, pp. 12–22, 2008.
[27] D. M. Tax and R. P. Duin, "Support vector data description," *Mach. Learn.*, vol. 54, no. 1, pp. 45–66, 2004.
[28] S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, 2011, pp. 1–7.
[29] N. Mostofi, M. Elhabiby, and N. El-Sheimy , "Indoor localization and mapping using camera and inertial measurement unit (IMU)," in *Proc. IEEE/ION Position Location Navigat. Symp.*, 2014, pp. 1329–1335.
[30] K. Shoemake, "Animating rotation with quaternion curves," *ACM SIGGRAPH Comput. Graph.*, vol. 19, no. 3, pp. 245–254, 1985.
[31] W. Kabsch, "A solution for the best rotation to relate two sets of vectors," *Acta Crystallographica Section A: Crystal Physics Diffraction Theoretical Gen. Crystallography*, vol. 32, no. 5, pp. 922–923, 1976.
[32] N. Bobick, "Rotating objects using quaternions," *Game Developer*, vol. 2, no. 26, pp. 21–31, 1998.
[33] R. Heise and B. A. MacDonald , "Quaternions and motion interpolation: A tutorial," in *New Advances in Computer Graphics*. Berlin, Germany: Springer, 1989, pp. 229–243.
[34] O. Walter, J. Schmalenstroeer, A. Engler, and R. Haeb-Umbach , "Smartphone-based sensor fusion for improved vehicular navigation," in *Proc. 10th Workshop Positioning Navigat. Commun.*, 2013, pp. 1–6.
[35] T. Park and K. G. Shin, "Attack-tolerant localization via iterative verification of locations in sensor networks," *ACM Trans. Embedded Comput. Syst.*, vol. 8, no. 1, 2008, Art. no. 2.
[36] Galaxy S5, Samsung. [Online]. Available: http://www.samsung.com/sec/galaxys5/, Accessed in 2016.
[37] Apple iOS and iPhone. [Online]. Available: http://www.apple.com/ios/ios-10/, Accessed in 2017.
[38] S. Tarkoma, M. Siekkinen, E. Lagerspetz, and Y. Xiao, *Smartphone Energy Consumption: Modeling and Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
[39] Seatbelt Statistics, Centers for Disease Control and Prevention. [Online]. Available: http://www.cdc.gov/motorvehiclesafety/seatbelts/facts.html, Accessed in 2017.

Homin Park received the BS degree in computing and software systems from the University of Washington, Tacoma, Washington, in 2006, and the MS and PhD degrees in information and communication engineering from the Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu, Korea, in 2016. He is currently a post-doctoral research fellow in the BIGHEART at the National University of Singapore. He jointly conducted numerous researches and projects with the University of Michigan (Ann Arbor), University of Virginia, Hanyang University, South Dakota State University, Hyundai Motors, and Electronics and Telecommunications Research Institute (ETRI). His current research interests inlcude human activity inferencing systems, applied deep learning, cyber-physical systems (CPS), internet of things (IoT), and their applications to smart healthcares, homes, offices, and vehicles. He has authored or coauthored 16 papers and two patents including four prestigious peer-reviewed journals and conference papers.

Daehan Ahn received the BS degree in information and communication engineering from Korea Aerospace University, in 2013. He is working toward the PhD degree in information and communication engineering at the Daegu Gyeongbuk Institute of Science and Technology (DGIST). His current research interests include cyber-physical systems and artificial intelligence, pattern analysis, and algorithmic optimization.

Taejoon Park received the BS degree (summa cum laude) degree in electrical engineering from Hongik University, Seoul, Korea, in 1992, the MS degree in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Taejon, Korea, in 1994, and the PhD degree (under the supervision of Prof. Kang G. Shin) in electrical engineering and computer science from the University of Michigan, Ann Arbor, Michigan, in 2005. He is currently a full professor in the Department of Robotics Engineering, Hanyang University, Gyeonggi-do, Korea. Prior to joining Hanyang University, he was an associate professor with the Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu, Korea, from Feb. 2011 to Feb. 2015, an assistant professor with Korea Aerospace University, Gyeonggi-do, Korea, from Sep. 2008 to Feb. 2011, a principal research engineer with Samsung Electronics, Gyeonggi-do, Korea, from Apr. 2005 to Apr. 2008, and a research engineer with LG Electronics, Seoul, Korea, from Feb. 1994 to Jun. 2000 (promoted to a senior research engineer in 2000). His current research interests include cyber-physical systems (CPS) with emphasis on artificial intelligence and machine learning, and their applications to robots, vehicles, and factories. He has authored or coauthored more than 130 papers/patents including essential patents for the DVD standard, six of which were cited more than 100 times. He is a member of the IEEE and the ACM.

Kang G. Shin is the Kevin & Nancy O'Connor professor of computer science in the Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor. His current research focuses on QoS-sensitive computing and networking as well as on embedded real-time and cyber-physical systems. He has supervised the completion of 79 PhDs, and authored/coauthored more than 900 technical articles, one textbook and more than 400 patents or invention disclosures, and received numerous best paper awards, including the Best Paper Awards from the 2011 ACM International Conference on Mobile Computing and Networking (MobiCom'11), the 2010 and 2000 USENIX Annual Technical Conferences, as well as the 2003 IEEE Communications Society William R. Bennett Prize Paper Award and the 1987 Outstanding IEEE Transactions of Automatic Control Paper Award. He has also received several institutional awards, including the Distinguished Faculty Achievement Award in 2001 and Stephen Attwood Award in 2004 from The University of Michigan (the highest honor bestowed to Michigan Engineering faculty); a Distinguished Alumni Award of the College of Engineering, Seoul National University in 2002; 2003 IEEE RTC Technical Achievement Award; and 2006 Ho-Am Prize in Engineering (the highest honor bestowed to Korean-origin engineers). He has chaired several major conferences, including 2009 ACM MobiCom, 2008 IEEE SECON, 2005 ACM/USENIX MobiSys, 2000 IEEE RTAS, and 1987 IEEE RTSS. He is a fellow of the IEEE and the ACM, and served on editorial boards, including the *IEEE Transactions on Parallel and Distributed Systems* and the *ACM Transactions on Embedded Systems*. He has also served or is serving on numerous government committees, such as the US NSF Cyber-Physical Systems Executive Committee and the Korean Government R&D Strategy Advisory Committee. He has also co-founded a couple of startups.