Sensors Know When to Interrupt You in the Car: Detecting Driver Interruptibility Through Monitoring of Peripheral Interactions

SeungJun Kim¹, Jaemin Chun², and Anind K. Dey³
Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA 15213 USA
{sjunikim, anind}@cs.cmu.edu¹,³, and jaemchun@andrew.cmu.edu²

ABSTRACT
Interruptions while driving can be quite dangerous, whether these are self-interruptions or external interruptions. They increase driver workload and reduce performance on the primary driving task. Being able to identify when a driver is interruptible is critical for building systems that can mediate these interruptions. In this paper, we collect sensor and human-annotated data from 15 drivers, including vehicle motion, traffic states, physiological responses and driver motion. We demonstrate that this data can be used to build a machine learning classifier that can determine interruptibility every second with a 94% accuracy. We present both population and individual models and discuss the features that contribute to the high performance of this system. Such a classifier can be used to build systems that mediate when drivers use technology to self-interrupt and when drivers are interrupted by technology.

Author Keywords
Interruptions; Naturalistic Driving; Sensor Data Mining

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

INTRODUCTION
Human attention is a finite resource [25]. When interrupted while performing a task, this resource is split between two interactive tasks. People have to decide whether the benefits from the interruptive interaction will be enough to offset the loss of attention from the original task. They may choose to ignore or delay dealing with the interruption to a more convenient time. They may alternately choose to immediately address the interruption but this comes with a risk of reduced performance on the primary task [19, 22, 27].

The issue of dealing with peripheral tasks (both self-interruptions and external interruptions) is particularly critical in driving situations. Under normal driving conditions, drivers should have an appropriate following distance (ideally 2-3 seconds behind the car in front of you [e.g., 12]) for safe driving. This following distance provides the driver with enough reaction time to decide whether to stop, slow down, or otherwise react to changing driving conditions.

However, driving guidelines like this assume that the driver is fully attending to the driving task. If a driver is performing peripheral tasks, those not directly related to the driving task (e.g., turning off a smartphone alarm or changing the radio station), she needs to manage when to divide her attention. Depending on the frequency and duration of these additional tasks, drivers need to adapt how they apply safe driving guidelines.

Fortunately, drivers naturally adapt to changing driving conditions when dealing with peripheral tasks, e.g., waiting for a red light to attend to these tasks, and not performing them in heavy traffic. They determine appropriate timings for changing the positions of their hands on the steering wheel, controlling the foot pedals, all while monitoring adjacent traffic. However, if a demand for peripheral interaction arrives at an inappropriate or unexpected time (e.g., phone rings while changing lanes to exit the highway), it can lead to dangerous driving situations which may result in a driving violation, accident or even loss of life. In fact, 25% of car accidents in the U.S. are related to phone use [8].

By identifying situations or patterns when drivers are not able to attend to peripheral tasks due to their current state and driving situation, a workload manager (e.g., Green [9]), may regulate the flow of information to drivers that could otherwise interfere with driving. Intelligent systems will be able to mediate the delivery of external interruptions, or even disable phone/infotainment systems to mediate self-interruptions [e.g., 24], to support safer driving.

In this context, we investigate when and how drivers perform peripheral activities while driving. We identify driver and driving states and road configurations in which drivers choose to deal with interruptions.

Specifically, we explore three overarching questions:
• What are the driver and driving states in which drivers perform peripheral activities?
• What sensors and sensor features best explain the states in which drivers perform peripheral interactions?
• How quickly and accurately can systems assess opportunities for peripheral interaction?
Through an in-car data collection with 15 drivers, we collected a large variety of sensor data to understand driver state and driving situations including vehicle motion states, road conditions, driver motion changes, and physiological responses, while capturing driver interactions and traffic states during naturalistic driving. We demonstrate that this data can be used to build a population model that is 94% accurate at determining the interruptibility of a driver each second. We discuss the most informative features that contribute to this performance, and how such a model can be used in practice for mediating interruptions while driving. We start with an overview of background work motivating the need for sensor-based assessment of driving experiences during naturalistic driving, especially in the situations of divided attention due to in-car interruptions.

**BACKGROUND WORK**

**Driver Experience Sampling**

Traditionally, user experience has been sampled by asking people to stop mid-task and note their experience [17]. The point is for users to record *in situ* aspects of experience like mental effort or emotion, based on their own judgment. When users are engaged in naturalistic and uncontrolled real-world tasks, that approach provides low-resolution data since sampling rates need to be low to avoid disrupting the user too much, but are often too low to track dynamically varying user states [10]. This problem is especially disadvantageous in mobile contexts such as driving.

In automotive contexts, when directed to self-report experience, drivers must divert their attention from the driving task. This can diminish cognitive capabilities for the primary task by drawing attention to interruptive demands for peripheral interaction; while driving, interruptions can negatively impact primary task performance. As well, when sampling user state at the end of a driving session, *in situ* variations of that state can be blurred in relation to the overall user experience. Due to these potential issues, driver experience sampling based on self-reports has typically been explored for evaluating user interfaces (e.g., dashboard designs) or within driving simulations (e.g., [13],[16]) where interruptions are less dangerous.

In our work, we seek to better understand the relationship between a driver’s performance of peripheral tasks and *in-situ* driver states and driving contexts. This necessitates the continuous sampling of user state while driving; therefore, we mainly explore sensor-based assessments to sample driver states. This approach provides quantifiable, fine-grained data in real-time. The feasibility of this approach has been validated in simulated driving experimentation (e.g., [12], [21]), and we apply this to naturalistic driving.

**Interruptions**

As the physical world becomes increasingly connected with our information spaces, so too is the likelihood that information will be pushed to people during the performance of real-world tasks. At best, those interruptions may alert users to important warnings or messages, inquire about people’s status (e.g., affective states or health conditions for health care), or deliver information that can benefit task performance. Despite this potential value, dealing with these interruptions through peripheral interaction (interaction not directly related to the primary task) demands cognitive attention that can negatively and variably impact user experience. An improved understanding of user availability or interruptibility is necessary for mediating this impact.

Task interruptions result in a time lag before users can resume their primary task, and thus decrease primary task performance [19, 27]. Appropriate timings of interruptions can reduce the impact on users. For example, in the context of desktop computing, interruptions delivered at points of lower mental workload reduced resumption lags and minimized disruption in primary task performance compared to interruptions at points of higher mental workload [1, 6]. In an experiment in which participants were interrupted with emails about consumer products and prices [22], users who experienced deferrable interruptions during high cognitive workload tasks frequently disregarded the notifications until they reached low workload periods. The results of another experiment showed that when peripheral tasks involving reasoning, reading or arithmetic interrupt the execution of primary tasks, users require more time to complete the primary tasks, commit more errors, and become more annoyed and increasingly anxious than when those same peripheral tasks were presented at the boundary between primary tasks [2].

These studies offer important insights for designing human-centered interruptions; however, they have mostly explored static, on-screen tasks mediated with conventional computers or mobile devices (e.g., the impact of call notifications for smartphone users [3]). Little research has been conducted to replicate these findings or approaches for delivering interruptions in situations in which users cannot fully divert their attention from the primary task (e.g., driving cars) and in which interruption timings can critically impact the user experience. We attempt to address this in our work.

**Naturalistic driving**

To ensure driver safety, driving studies about dual task paradigms have mostly been conducted in simulated environments [e.g., 21]. When using sensors to track drivers' eye gazes or physiological responses in these simulations, the experimental design requires special attention to achieve valid and realistic data. Due to this issue of driver safety, existing naturalistic driving datasets mostly include audiovisual records, traffic and lane information from vision-based systems [e.g., 18], and researcher-estimated driver states using image processing techniques [e.g., 15] on videos from multiple cameras installed in cars.

Recent advances in wearable technologies have resulted in sensors that are less intrusive and more comfortable to wear;
nevertheless little research (e.g., [23]) has been performed that tracks driver body motion or physiological responses during naturalistic driving. Driver workload has been mostly evoked by imposing artificial dual-task demands (e.g., auditory stimuli at pre-structured interaction times and intervals [26]). On-board diagnostic (OBD) systems or accelerometers on the steering wheel have been used to assess driver aggressiveness, driving environment and vehicle states (e.g., [7, 11]). In our work, we leverage all of these sensing technologies to help us identify which sensor streams and features are most useful for tracking driver state and the performance of peripheral activities; our novelty lies in detecting driver intermittibility by continuously monitoring naturalistic driver / driving states without artificially manipulating cognitive demands.

**EXPERIMENTAL SETUP**
Here we describe our experimental method for collecting naturalistic driving data for the purpose of identifying situations when peripheral interactions occur.

**Participants:** We recruited 25 drivers (age $M=32.0$, $SD=14.3$, age range: 19 - 69, gender: 14 female and 11 male). Participants were asked to drive their own cars and were compensated with $20. The experiment took approximately two hours to complete, including about 1.25 hours of driving.

Participants had 14.5 years of driving experience on average ($SD=13.2$, range: 1 - 52 years). They self-reported that they drove on city streets 8.7 times per week ($SD=7.1$, range: 2 - 35 times) and on highways 4.5 times per week ($SD=4.7$, range: 1 - 20 times). They used familiar routes 8.8 times per week ($SD=9.4$, range: 2 - 50 times) and unfamiliar routes 1.2 times per week ($SD=1.0$, range: 0-3 times). The average driving trip duration was 30.4 minutes ($SD=19.2$, range: 10 - 100 minutes).

![Figure 1: Experimental setup – two smartphone cameras, five body-worn sensors, and one on-board diagnostics in the cars.](image)

**Test-bed:** As shown in Figure 1, we plugged an on-board diagnostics (OBD) device in study participants’ cars, and they were asked to wear five body-worn sensor devices: four accelerometer sensors for capturing body motion and one chest belt sensor for tracking physiological responses. We installed two smartphones in each car - one on the front windshield and the other on the head-rest of the passenger seat. The phone cameras recorded traffic in adjacent lanes and drivers’ activities in the cars. They also received information from the OBD device and body-worn sensors via Bluetooth and logged sensor data streams in real-time.

**Main task:** Study participants performed two sessions of naturalistic field driving. Driving routes in both sessions were selected to combine both city streets and highways. The route in the 1st session consisted mainly of highway driving to a shopping mall. Drivers were allowed to take any preferred or familiar route to get the destination (according to Google maps, the destination was 14.7 miles away). The route in the 2nd session consisted mainly of driving on city streets to a sports stadium in our city’s downtown area, along with shorter lengths of highway driving compared to the previous session. In this session, drivers were asked to drive using a provided Garmin GPS to reach the destination (Google maps: 4.8 miles).

Some participants performed both driving sessions on the same day while others performed them on different days. By fixing the destinations, and in the 2nd session, the route, drivers were exposed to similar configuration of roads in terms of the number of signal lights and signs to encounter, the length of highways and streets, road curvature, and hill gradient. By leaving flexibility about the route to take in the 1st session, we were able to collect data from driving on familiar and unfamiliar routes.

**Sessions:** There was a set of activities performed in each session: we installed the devices in drivers’ cars and helped them wear the physiological sensors, time-synced the devices and collected baseline data, and had drivers complete a pre-questionnaire, perform the driving task, and complete a post-task questionnaire and interview.

For syncing the devices, experimenters shook all the wearable sensors for a period of twenty seconds, once after device setup both prior to and after completing the driving tasks. Those time-windows were used for synchronizing timestamps across data streams from the worn sensors.

For collecting baseline measurements, drivers were asked to grip the steering wheel at the three and nine o’clock positions, while closing their eyes after taking a deep breath. We collected sensor data streams from the body-worn sensors for 60 seconds of this “at rest” behavior.

In the pre-questionnaire prior to 1st driving session, we collected demographic information of study participants and information about driving frequency and patterns when performing peripheral interactions in cars.

In the post-questionnaire after each driving session, drivers were asked to describe situations during their just-completed session, (e.g., number of passengers, familiarity to driving routes, traffics, sensor comfort, driving performance) and complete a NASA-TLX assessment.

In the post-interview, we sought to validate our hypothesis that the driver and driving contexts can help us determine when drivers can perform peripheral interactions. Specifically, we presented drivers with 17 images of drivers...
performing secondary tasks while driving (e.g., eating, car radio interaction), and then asked them to comment on all the activities that they conducted during their driving session, explain what factors led them to perform those peripheral interactions, including traffic, in-car and on-road situations, and then comment on whether the moments when they chose to initiate peripheral interactions while driving would also be good for being interrupted.

**Peripheral Interaction Moments Correspond to Driver Interruptibility**

In regular driving situations, there are many instances of drivers engaging in peripheral interactions: actions that take place in the car, but are not related to driving the car. We believe that these instances can be used as ground truth of moments when the driver believes he is able to enter a situation of divided attention [5], and, therefore can better deal with interruptions. This belief needs to be validated.

For example, if a driver eats or drinks while at a red light and she usually interacts with the car radio interface searching for a particular radio station while driving at a constant speed on a street with no noticeable grade or curve in the road, these situations represent conditions in which the driver feels she is able to divide her attention. Drivers may consciously or subconsciously know about the moments in which they can perform multiple tasks at the same time. If they are not actually engaged in any peripheral interaction during such a situation, this may imply that this moment can be used to interrupt the driver with external information (e.g., upcoming traffic information, text message or even advertisements), more safely than in other conditions where no peripheral interactions are ever performed (e.g., during acceleration or driving on a sharp curve). It is also important to separate out interactions with activities that are performed to directly support the driving task – for example, operating blinkers for changing lanes and wipers for better visibility when raining. Moments involving these types of tasks cannot necessarily be viewed as opportunities for performing peripheral interactions or interrupting the user.

**Validation:** In the post interviews, all drivers described moments in which they initiated peripheral interactions during the driving session they just completed. They stated that these moments were appropriate for being interrupted.

One participant said he would perform peripheral interactions “usually if I was at a red light or if I was stuck in traffic, also depends on the situation - if I don't have to make a turn and it's a straight road. I didn't do anything while I was making a turn or going through an intersection. I waited until it was a straight stretch of road with no places where cars would pull out and low traffic.”

Another said “In most cases, I chose when to have the interaction - when to turn on the wipers, when to change the radio station or adjust the volume... except when I needed more attention on driving. Wipers I turned on because there was too much stuff on the view. Radio, I changed channels when I thought it was a good time too. If I were stopped in traffic, then yes, I would think that that's also a good time to [receive interruptions].”

**TIME-SERIES DATA**

Based on these interview results validating that moments in which peripheral interactions are performed are times when drivers could be interrupted, we proceeded with the analysis of the captured sensor data to automatically identify these moments. We started this process with identifying ground truth about what drivers were doing during driving sessions.

**Videos and Annotation features**

**Camera 1:** One of the two smartphone cameras was used for examining drivers’ activities in cars. We manually labeled the captured videos for moments when peripheral interactions were performed and not performed. We used five labels. **DRIVING I** includes activities that are quite central to the primary driving task (e.g., changing grip positions, operating levers for blinkers or wipers, switches for opening side windows), whereas **PI** includes activities that are not directly central to the primary task (e.g., eating food, manipulating the air conditioner or car radio, talking on the phone). **ONE_HAND_DRIVE WITH NO_PI** was used to label moments when one of the driver’s hands was off the steering wheel but that hand was not performing any specific activity. **NO_HAND_DRIVE** was used for moments when both hands were off the steering wheel and not performing any peripheral activity. **STEERING ONLY** was used for moments when both hands were on the wheel. From our interviews, moments that are labeled with PI or NO_HAND_DRIVE indicate that drivers have higher interruptibility, compared with moments of DRIVING I or STEERING ONLY.

**Camera 2:** The other smartphone camera was used for assessing the traffic around the driver. We labeled the videos for the amount of traffic visible in front of the car, to the left and right of the car, and in the oncoming lanes. The traffic labels we used were **NO_TRAFFIC**, **IGNOREABLE**, **SOME, A LOT** based on the approximate distance of adjacent cars and the number of cars. If the traffic in the oncoming lane was occluded by cars in the left lane or a center guardrail on a highway, we labeled it as **OCCCLUDED**. We also had a label indicating the car’s movement, whether it was **STOPPED** or **MOVING**.

Three experimenters, each having more than 4-years experience in psychology or human computer interaction, reviewed videos from the smartphones and labeled the kind, start time, and end time of drivers’ peripheral activities, and traffic. Note that while we manually labeled the videos from both cameras, these are labels that could be automatically provided using computer vision (e.g., [4]).

**Six Sensors and Applied features**

An **OBD device** provided information about the status of the vehicle (sampled at 1Hz) including longitude, latitude,
altitude, car speed, engine RPM, throttle position, and fuel flow rate. The data were transmitted via Bluetooth to the smartphone responsible for recording traffic video.

Four YEI 3-Space devices were used, placed on each of the drivers’ wrists, on the front of the head, and top of the foot. The devices collected information about the drivers’ motions using a tri-axial gyroscope, a tri-axial accelerometer, and compass sensors, all at 4-5 Hz. The sensor data was transmitted via Bluetooth to the other smartphone. In addition, drivers wore a BioHarness (BH) chest belt that collected drivers’ physiological data including electrocardiogram, heart rate, respiration rate, body orientation and activity at 20Hz sampling rate.

After time syncing the sensors, all sensor data were aggregated, and their means (µ) and standard deviations (σ) were calculated for every 1-second segment as statistical features (referred to as ‘basic’ features in this paper). In addition, we derived a series of additional features based on the basic features. For example, from the OBD data we derived road curvedness by tracking variations of longitude and latitude coordinates (See Figure 2 – top-right), centrifugal forces by combining car speed and curvedness information, types and gradient of road slope by using the 1st derivative of altitude data. From the YEI data, we derived motion information on how much each monitored body part moved from the baseline position (See Figure 2 – bottom). From the BH data, we quantified levels and duration of breathing-in, breathing-out, and holding-breath states (See Figure 2 – top-left). These supplementary features are referred to as ‘derived’ features in this paper. In total, we had 100 basic features and 52 derived features (OBD: 72, YEI: 40, BH: 40) and 5 manually annotated features related to traffic from videos (one car driving state and four traffic states around the vehicle – front, right, left, and oncoming).

DATA COLLECTION STATUS
Of our 25 drivers, 17 drivers successfully completed both driving sessions without any loss in data. The 17 drivers drove for an average of 43m 59.2s in the first driving session (SD = 7m 23.5s, range: 28m 42s - 54m 14s) and 33m 18.3s in the second driving session (SD = 7m 42.4s, range: 25m 36s - 51m 32s) on average. For the remaining 8 drivers, most data loss came from Bluetooth disconnections either when OBD devices were accidentally dislodged by drivers’ knees or when the Android data logging applications crashed. Occasionally smartphones fell from their mounts or turned enough to capture irrelevant data.

When examining the sensor data, we further excluded 2 drivers’ data. Their heart rate data contained errors (higher than 200 beats per minutes during the entire driving session). This happened when the BioHarness sensor was worn too loose around the driver’s torso during the driving session. We did not include data from these 10 drivers in the sensor data analysis although we did include them in the analysis of the questionnaires and interviews.

STATISTICAL ANALYSIS METHOD
With the data from our 15 remaining drivers, we examined how driver and driving states differ across our five driving state classes: STEERING ONLY, ONE HAND DRIVE WITH NO PI, DRIVING 1, PI, and NO HAND DRIVE.

For analyzing the continuous measures (e.g., car speed), we conducted a univariate ANOVA by using a general linear model and then used either Tukey HSD or Games-Howell as post-hoc tests after checking the homogeneity of variances (i.e., Levene Statistic), where ηp2 was examined as effect size. For ordinal measures (e.g., human-annotated data or fixed level data), we conducted the Kruskal Wallis Test followed by Mann-Whitney U test as post-hoc, where r was examined as effect size. In the analysis of Likert-scale rating data, Friedman tests and a Wilcoxon Signed Rank post-hoc test were conducted.

RESULTS
Driver Activities: Frequency and Duration
In total, 53 different activities were detected from the 15 drivers. In the remaining analysis, we focus on 22 of these activities that had a total duration of at least 90 seconds summed across multiple instances. As shown in Table 1, activities include interacting with the air conditioner or car radio, operating blinkers or windshield wipers, eating food and drinking, driving with one hand, smoking, dancing, changing hand grip, resting by removing both hands from the steering wheel, and so forth.
One-handed driving most frequently occurred (57.1 times per subject, with 6.9 instances for each 10-minute driving segment; See Table 1). When total durations were considered, summed up for each participant and summed over a 10-minute segment, it was also first (15.4 and 117.0 seconds, respectively). However, when ranking these by the duration per instance, using one’s cellphone was the longest (21.7 seconds), followed by eating food or drinking, smoking, and one-handed driving.

<table>
<thead>
<tr>
<th>Driver activities</th>
<th>Occurrence / 10-min drive</th>
<th>Duration (sec) / 10-min drive</th>
<th>M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE_HAND_DRIVE</td>
<td>57.1</td>
<td>15.4</td>
<td>117.0</td>
</tr>
<tr>
<td>GEAR_SHIFT</td>
<td>37.1</td>
<td>4.5</td>
<td>19.0</td>
</tr>
<tr>
<td>TOUCHING_SELF (e.g., face)</td>
<td>24.5</td>
<td>3.0</td>
<td>7.6</td>
</tr>
<tr>
<td>REST_OFF_WHEEL</td>
<td>20.6</td>
<td>2.5</td>
<td>14.7</td>
</tr>
<tr>
<td>L_BLINKER</td>
<td>15.7</td>
<td>1.9</td>
<td>3.9</td>
</tr>
<tr>
<td>R_BLINKER</td>
<td>14.0</td>
<td>1.7</td>
<td>4.4</td>
</tr>
<tr>
<td>CAR_RADIO</td>
<td>9.2</td>
<td>1.1</td>
<td>5.9</td>
</tr>
<tr>
<td>R_GRIP_CHANGE</td>
<td>8</td>
<td>1.0</td>
<td>2.9</td>
</tr>
<tr>
<td>HEAD_TURNING_WAY_LEFT</td>
<td>7.3</td>
<td>0.9</td>
<td>4.2</td>
</tr>
<tr>
<td>CELLPHONE</td>
<td>6.7</td>
<td>0.8</td>
<td>13.5</td>
</tr>
<tr>
<td>L_GRIP_CHANGE</td>
<td>6.7</td>
<td>0.8</td>
<td>1.8</td>
</tr>
<tr>
<td>HEAD_TURNING_WAY_RIGHT</td>
<td>5.7</td>
<td>0.7</td>
<td>2.6</td>
</tr>
<tr>
<td>GPS_PORTABLE</td>
<td>4.5</td>
<td>0.5</td>
<td>14.3</td>
</tr>
<tr>
<td>AC_CONTROL_SWITCHES</td>
<td>3.1</td>
<td>0.4</td>
<td>1.8</td>
</tr>
<tr>
<td>BOTH_HANDS_OFF_DRIVE</td>
<td>2.3</td>
<td>0.3</td>
<td>2.2</td>
</tr>
<tr>
<td>FOOD_OR_DRINK</td>
<td>2.6</td>
<td>0.3</td>
<td>11.6</td>
</tr>
<tr>
<td>OTHERS_MISCELLANEOUS</td>
<td>2.7</td>
<td>0.3</td>
<td>6.5</td>
</tr>
<tr>
<td>SMOKING</td>
<td>1.9</td>
<td>0.2</td>
<td>1.2</td>
</tr>
<tr>
<td>WIPER_WASHER</td>
<td>1.7</td>
<td>0.2</td>
<td>2.9</td>
</tr>
<tr>
<td>AC_AIR_VENTS</td>
<td>1.1</td>
<td>0.1</td>
<td>3.5</td>
</tr>
<tr>
<td>OTHERS_DANCE</td>
<td>0.1</td>
<td>0.0</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Total 235.328

Table 1. Frequency and duration of 22 top-ranked driver activities sorted by a descending order based on occurrence per unit, in which top-five measures and activities are in bold.

For most of the driving time, drivers used both their hands (48.0%) or one hand (22.8%) for controlling the vehicle without any peripheral interaction (See STEERING ONLY and ONE_HAND_DRIVE WITH NO PI in Figure 3).

The time spent in performing a task is indicative of user workload (e.g., [10]). Similarly, duration of a particular peripheral activity during driving may indicate how drivers are consciously or subconsciously estimating interaction workload for that activity. In other words, the longer the time spent on a peripheral interaction, the greater the availability for being interrupted. In this study, we found that the average duration was significantly different among the five interaction states (F4,4546=116.2, p =.000). The pairwise difference between each state was significant in the post-hoc test with the exception of the durations between ONE_HAND_DRIVE WITH NO PI and NO_HAND_DRIVE (See Figure 3 – two rightmost bars). This implies that although drivers did not actually execute any peripheral activity, one-handed driving is a distinct state from using both hands for steering. Similarly, there was no discernible distinction in durations between one handed and no-handed driving. We will discuss this finding later in the paper.

Driver activities that corresponded to the driving task (e.g., turning on blinkers) were completed in 2.4 seconds on average, while peripheral interactions lasted for 10.6 seconds (DRIVING I and PI in Figure 3).

Driving States: Car Speed and Road Conditions

The OBD data streams revealed that drivers significantly regulated car speed while performing peripheral activities or taking both hands off the steering wheel (See Figure 4 – the left-most bar in each group). Compared to the STEERING ONLY state, car speed was reduced to 60% while peripheral interactions were performed (PI state in Figure 4), and almost reduced completely (slower than 30 km/h) while driving without either hand on the steering wheel.

When drivers were in either of these two slower states, engine RPM was significantly lower and the proportion of time spent on flatter roads (i.e., low grades) was significantly higher, compared to the STEERING ONLY state. In particular, as indicated by * for the right-most two bars of Figure 4, drivers were on flatter and less curvy roads while driving with both hands off the steering wheel (statistically significant when compared to each of the other states). Interestingly, one-handed driving was performed at considerably high speeds, without necessarily being on flat...

Figure 3: Average time duration of drivers' interaction states. *p<.001 versus STEERING ONLY states; **p<.001 versus all other states.

Figure 4: Driving states and road conditions. *p<.001 versus STEERING ONLY states; **p<.001 versus all other states.
or straight roads.

These findings imply that driving status for the PI or NO_HAND_DRIVE states are most similar. In addition, they represent the moments when the driver is likely to be more interruptible. However, we do not include one-handed driving in this group. The driving state and road conditions were different for ONE_HAND_DRIVE_WITH_NO_PI.

In addition, when examining how drivers accelerate or decelerate across the five states, we see that drivers tend to drive at more consistent speeds during PI (47.8%) or NO_HAND_DRIVE (73.7%) states than in the ONE_HAND_DRIVE_WITH_NO_PI state (24.6%) (Figure 5).

**Figure 5: Time proportion of acceleration, deceleration, and constant speed across whole driving.**

**Driver State: Body Motion and Physiological States**

The deviation of gravity measures from the STEERING ONLY body pose were significantly different for every pair of the five states across every body element (Figure 6). In particular, the right hand movement was distinct when performing driving or peripheral activities (2nd and 3rd bars from the top). Head movement was larger when driving while taking both hands or one hand off the steering wheel without performing any peripheral activities (bottom two bars in Figure 6).

Drivers’ breathing states and heart rate also significantly varied across the five interaction states (\(F_{5,34693}=6.80, p=.000\) and \(F_{3,3665}=215.0, p=.000\), respectively; Figure 7). Interestingly, for NO_HAND_DRIVE, drivers’ heart rate was similarly low as the resting heart rate captured in the baseline session where no driving was performed (bottom 2 bars from Figure 7). Also, the heart rate during ONE_HAND_DRIVE_WITH_NO_PI state was similar to PI, but not to NO_HAND_DRIVE or baseline. The duration of drivers holding their breath was not statistically significant (\(F_{4,446}=1.99, p=.075\)). Rather, durations for inhaling and exhaling were significantly longer for the PI or NO_HAND_DRIVE states, and similar to when resting in the baseline session.

Overall, driver physiological states for the PI and NO_HAND_DRIVE states are most similar to the resting states. We therefore confirm that from our analysis of physiological states, driving states and road conditions, drivers are more interruptible than in the remaining states. However, interruptibility while one-handed driving without any peripheral interactions is still unclear since driver physiological states are similar to the two interruptible states and the baseline, but the driving states and road conditions differed significantly.

**Figure 6: Driver body motion revealed in sensor data streams.** *p<.01 in every pair of the five interaction states.

**Figure 7: Drivers’ physiological states revealed in sensor data streams.** *p<.01 versus STEERING ONLY states.

**Identifying Opportunities for Interruptions**

Given the statistical differences across the five interaction states, and particularly between the more interruptible and less interruptible states, we now seek to automatically identify opportune moments for driver interruption. Specifically, we want to build a classifier that accurately detects or predicts interruptibility every second.

To build this classifier, we combined the PI and NO_HAND_DRIVE states into a single class representing instances when drivers can be interrupted (INTERRUPTIBLE). We correspondingly combined data from the remaining classes into the LESS_INTERRUPTIBLE class; note that we do not call this “Uninterruptible” since we cannot know with absolute certainty that drivers are not
interruptible during these interaction states, but just that they are likely to be less so.

For this binary classification problem, we used a random forest classifier, which runs efficiently on large databases such as our sensor data. We also handled unbalanced classes using stratified 10-fold cross validation where each fold contains approximately the same percentage of samples from each target class as the complete set, and then applied different sample weights based on the ratio of samples that belong to each class. We automated this procedure for each driver and for each fold.

Our population-based classifier had an average classification accuracy of 94.3% classification accuracy across the data from all 15 drivers (SD=0.3%, average precision and recall: 0.94 and 0.89, respectively). From our initial list of 143 features, car speed (OBD, km/h), engine RPM, GPS bearing, centrifugal features, and the movement of the right hand deviating from the baseline position were rank-ordered as the most important features, along with two human-annotations about the car movement states (top-ranked feature) and the front traffic (7th-ranked). While we believe that these human annotations could be automated with today’s image processing technologies, we built a second classifier that did not use any human annotations. The classification accuracy dropped by only 2.2% (92.1%, SD=0.2). This means that the binary states representing driver interruptibility can be discriminated accurately every second using sensor data only.

Next, we examined the performance of the individual models for each of our fifteen drivers, to investigate individual difference in which features were most helpful and whether such models need to be personalized.

Promisingly, the average classification accuracy of the individual models of our drivers was 94.9% on average (SD=2.6%, range: 90.2% from 98.2%), almost the same as the population model (94.3%) Accuracy for every driver was greater than 90% (See Figure 8).

However, the recall rate (i.e., the fraction of relevant instances that are retrieved = the percentage of interruptible moments correctly classified as being interruptible) had some individual differences. Five of our drivers (driver 05, 10, 11, 21, and 23) had recall rates lower than 80%, while the precision rates (i.e., the fraction of retrieved instances that are relevant = the percentage of moments identified as interruptible that were actually interruptible) were fairly consistent along with the classification accuracy across every driver.

We then focused on the top fifteen features for each driver that contributed to classification accuracy, to identify which features were commonly used in the individual models. Of the 143 features, 46 features were chosen in at least one driver’s list of top fifteen features (23 from OBD device, 21 from the body-worn sensors, and 2 from the human-annotated data). Similar to the population model, engine RPM, car speed (OBD, km/h), and car movement annotation features were identified again as important features for more than ten of our drivers. However, unlike in our group model, heart rate, foot and body trunk/torso motion were also important for five drivers, centrifugal force for seven drivers and the front traffic annotation data for only six drivers.

**Discussion**

In this section, we discuss our finding on driving states, the features that contributed to our model, and our ability to detect moments of interruptibility. We use these to point out opportunities for future research that can build on the contributions of this paper.

**Driver and Driving states**

Our sensor data revealed that drivers perform peripheral activities when experiencing lower workload. Similar to driving without holding the steering wheel, drivers tend to initiate peripheral interactions at a near constant car speed on fairly flat and straight roads. Further, drivers’ physiological signals were similar to those during the resting states in the baseline session. This implies that drivers may defer processing peripheral tasks until they are experiencing lower mental workload, confirming findings reported in prior interruption studies (e.g., [22]).

An interesting situation is when drivers use only one hand but do not perform any peripheral activity with the idle hand. In general, drivers’ physiological responses during this state implied lower driver workload, closer to that of performing peripheral interaction. However, the driving states for one-handed driving were much more dynamic than when performing peripheral interactions. Drivers either accelerated or decelerated without avoiding curved roads, and more often even than when using both hands to steer. This provides an interesting hypothesis that drivers may have extra cognitive capacity to deal with peripheral interactions. It is unclear whether this “extra” capacity is actually in use for the driving task, for enjoying an idle moment without initiating peripheral activities. In our future work, we will investigate this issue.
Indicative sensor data features
The derived features used to create our classifiers had more discriminative power than the simple statistical (basic) features. For example, the centrifugal feature created with a combination of car speed and road curvedness (more precisely, the square of car velocity divided by the road curvedness feature not considering driver weight) was ranked as the 5th most important in the population model and was in the top-15 features for 7 of the drivers’ individual models. If we incorporated the gravity variation coming from the road slope or the driver’s seating pose, this new feature could be even more powerful.

Similarly, the derived physiological responses (e.g., breathing rates) and road attributes (e.g., categorized slope levels) had statistical differences between driver’s interaction states. However, they did not explicitly contribute to building high performance classification models. Only two basic features, heart rate (in top-15 list for five drivers) and electrocardiography amplitude (ranked 12th in the group model) highly contributed to the classification accuracy.

To explore how to limit the obtrusiveness of our sensing system, we examined the relative values of the different sensors. For building either the group or individual models for predicting driving interruptibility, the most important features come from the cameras and vehicle movement sensors, except for the right hand motion (which could be detected via camera rather than worn sensor). For individual models, though, data from the physiological sensors was also valuable. It is possible that this body part motion could contribute to the group model as well, particularly if combined across body parts or with the car motion. We leave these refinements to future work.

Detecting Driving Interruptibility
In this study, we presented promising evidence that driving interruptibility can be accurately assessed in real-time using sensor data streams, independent of human-annotated information. Further, our results also show that accurate personalized models can be constructed from only 2 driving sessions.

However, individual models for a few drivers (i.e., driver 05, 10, 11, 21, and 23) provide insufficient performance in correctly detecting driving interruptibility. Their recall was 75.3% (SD=6%) on average and the precision was 89.0% (SD=1.9%, indicating 11% false discovery rate). This (roughly) implies that one in four known opportunities for interruptions was missed. It also implies that one in ten times the driver was predicted to be interruptible, she was not. This could lead to risky driving situations, although arguably no worse than our current driving situation in which interruptions arrive unmediated to drivers. Given our finding that in a 10-minute driving window, approximately 7 and 3 minutes belong to the LESS_INTERRUPTIBLE and INTERRUPTIBLE states, respectively, this equates to drivers being interrupted when not interruptible for 42 seconds and missing opportunities for interrupting the driver for 45, over each 10 minute segment of driving. The performance of these individual models should be improved to be used in practice, but our results are already very promising. We may need longer-term data collections from these individuals or more personalized features (e.g., route familiarity or habitual behaviors in the car) to improve the individual model classification accuracy.

Overall, our models performed quite well in identifying moments when the driver is in the NO_HAND_DRIVE or the PI states. While we equated this to the driver being INTERRUPTIBLE, it could be the case that drivers in one of the other states (particularly one-handed driving) could have additional opportunities for being interrupted. In the future, we would like to more thoroughly investigate driver’s interruptibility in situations when they are not performing peripheral activities and not driving no-handed.

CONCLUSION
Nowadays, technology is giving us the ability to interact with information anywhere and anytime, even in the car while driving. Whether it is appropriate to interact with information depends a lot on the current driving situation and the driver state. In this paper, we sought to understand these situations and states particularly when drivers engage in peripheral interactions, actions not related to the primary task of driving the car, as they indicate opportunities to interact with information. In particular, we collected sensor and human-annotated driving data from 15 drivers and demonstrated that we could build a machine learning classifier that determined when they were interruptible with 94% accuracy. We discussed the features that contributed to this high accuracy and suggested directions for identifying features to improve on our classifier.

In continuing this work, we have a number of goals. First, we will collect additional data to confirm that our results generalize across a wider driving population and work to build models and apply features that more consistently perform across drivers and for the nuances of specific drivers. Second, with this accurate classifier, and relatively easy-to-deploy system, we can now investigate how to mediate interruptions to drivers. We plan to conduct another field study where we push interruptions at different timings to assess the real-world impact of being able to detect interruptible moments while driving. These interruptions will differ in terms of their temporal urgency, relevance to the driving tasks, and overall importance. Third, we plan to develop generalized guidelines for designing intelligent interruption systems for drivers. We will use these improved models and understanding to identify the attributes of opportune moments detected (e.g., expected duration, expected level of driver engagement). For example, a relevant local advertisement should impose a short interruption during an opportune moment; if a driver has extended interruptibility, interruptions such as an alternate route/highway exit to take can be delivered.
Fourth, we plan to apply our technology to identify breakpoint moments for prompting drivers to safely participate in experience sampling while driving, which will support others doing driving-related research in naturalistic driving situations.

REFERENCES


