Driving Behavior Analysis with Smartphones: Insights from a Controlled Field Study

ABSTRACT

We evaluate a mobile application that assesses driving behavior based on in-vehicle acceleration measurements and gives corresponding feedback to drivers. In the insurance business, such applications have recently gained traction as a viable alternative to the monitoring of drivers via “black boxes” installed in vehicles, which lacks interaction opportunities and is perceived as privacy intrusive by policyholders. However, pose uncertainty and other noise-inducing factors make smartphones potentially less reliable as sensor platforms. We therefore compare critical driving events generated by a smartphone with reference measurements from a vehicle-fixed IMU in a controlled field study. The study was designed to capture driver variability under real-world conditions, while minimizing the influence of external factors. We find that the mobile measurements tend to overestimate critical driving events, possibly due to deviation from the calibrated initial device pose. While weather and daytime do not appear to influence event counts, road type is a significant factor that is not considered in most current state-of-the-art implementations.

Categories and Subject Descriptors

I.5.2 [**Computing Methodologies**]: Pattern Recognition—*Feature Evaluation and Selection*; J.4 [**Computer Applications**]: Social and Behavioral Sciences

General Terms

Measurement, Design, Reliability, Experimentation, Human Factors.

Keywords

Mobile Sensing; Smartphone; IMU Data; Driving Behavior; Insurance.

1. INTRODUCTION

The advent of mobile sensing platforms facilitates the cost-effective capture and processing of fine-grained data from the physical world, thus increasing the information base of business processes and decision making [1]. In the insurance sector, such data can be used to improve the assessment, communication, and mitigation of insured risk, thereby creating value for insurers and policyholders alike. A particular instance of this proposition is motor insurance. Vehicular sensor data is indicative of accident risk and may be utilized to incentivize risk-minimizing behaviors among drivers, thus also contributing to overall traffic safety [2], [3]. The premise of this approach is that by providing feedback of recorded driving actions to drivers, they are encouraged to change their behavior and reduce their individual accident risk. Large-scale field studies have reported an average estimated accident reduction of some 20% as the result of such interventions [4]. However, as with all sensing technology, the required installation of sensor and data transmission technology in vehicles raises severe privacy concerns among potential users who perceive the continuous monitoring by an insurer as intrusive [5]. In most insurance markets, consumers have thus rejected so-called Pay-As-You-Drive or Pay-Per-Risk policies. Furthermore, installation and operation of the typical on-board units devised for this purpose incurs additional costs to insurers and consumers.

An alternative approach is the assessment of driving behavior through smartphone applications that are operated at the users’ discretion. With increasing market penetration of state-of-the-art smartphones, consumers carry with them a connected sensing platform which allows for interaction with their data via intuitive visual interfaces. If insurers can leverage these platforms, no investments in additional hardware are required, and connectivity is often available at no additional cost. Furthermore, smartphone-based driving assessment emphasizes the importance of user consent and is more of a driving support tool than a “black box” monitoring device. It is thus potentially more effective in persuasion towards improved driving behavior, and therefore in insured risk. Moreover, such mobile applications offer new touch points for the low-involvement product insurance, where conventionally interaction with consumers takes place only at the point of sale and in the case of an insurance claim.

Considering their objective of inducing behavioral change among drivers – e.g., in an insurance context – it is an imperative requirement for such applications to deliver reliable and well-understood measurements. While vehicle-fixed systems have been thoroughly tested in practice, the data quality provided by smartphones as a sensing platform is presumably inferior in comparison. New degrees of freedom are added to the system due to the generally unrestricted movement of a smartphone relative to the vehicle. While this fact is unchallenged in the literature, there is hardly any evidence on the degree of deterioration of data quality when smartphones are used, and there is little published evidence on the performance of such systems in the field. The paper at hand aims to address this issue by evaluating a specific application for the assessment of driving behavior in a controlled field study. We developed and implemented a reasonably complex driver rating application which is described in more detail in the

MUM’12, December 3–6, 2012, Ulm, Germany.

following Section. In the study, 78 participants drove a vehicle on a test course of approx. 45 minutes duration while the application was running. In order to evaluate the functional performance of the application, a dedicated, off-the-shelf sensing system was installed in the vehicle which recorded reference data. From a comparison of these two systems, we are able to assess the performance of the mobile sensing platform.

2. RELATED WORK

Our research builds on a broad body of literature on dedicated on-board units that are mounted in a vehicle and compute risk indices from recorded sensor data [6], [7]. Sensor data in these solutions comprise position, acceleration, odometer readings or visual information; risk indices derived from these range from basic variables like vehicle velocity to sophisticated maneuver recognition algorithms that capture lane changes, U-turns, or distance to lateral road marks and preceding vehicles. Acceleration is particularly rich source of risk-related information as it allows for the detection of extreme braking and vehicle acceleration, sharp cornering, and sudden lateral movements [8]. Coupled with wireless data transmission technologies, specific on-board unit solutions for insurance applications have been developed [2], and are currently available from several insurers such as Progressive in the US and SARA in Europe.

Several related mobile applications have been released by academics, insurers, and other interest groups in the past. Such applications have focused on braking events and road bumps [9], the sensing of travel mode and trip purpose [10], and the identification of critical driving event from sensor fusion including a smartphone's CMOS sensor [11]. It has also been proposed to extend a smartphone's sensor capabilities by connecting it to a vehicle's OBD-II diagnostics interface [12], [13] and retrieving external web-based data sources, e.g., weather conditions. Two exemplary applications are depicted in Figure 1. Insurance-themed driver-rating applications: Driver Feedback (iOS, left) and MotorMate (Android, right).

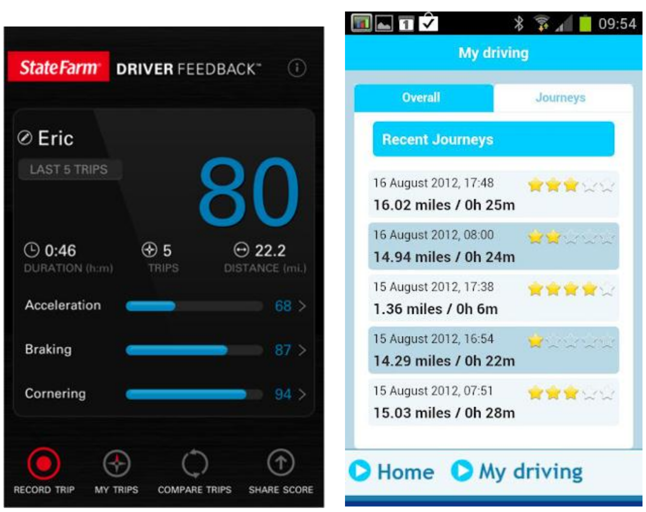

Figure 1. Insurance-themed driver-rating applications: Driver Feedback (iOS, left) and MotorMate (Android, right).

Arguably the most established method of assessing driving behavior among these approaches is to analyze the frequency of critical driving events [10]. Driving events are generated by different means of sensor fusion. They can be aggregated by summation and normalized over driven distance, and are thus suitable metrics of driving behavior. For the scope of this paper, we consider critical driving events as violations of certain thresholds imposed on vehicle acceleration. Acceleration is typically measured by inertial measurement units (IMU).

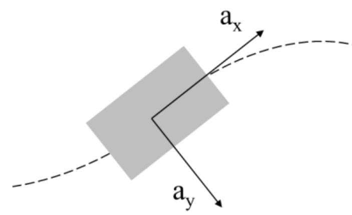

Figure 2. 2D-IMU measurements on a moving vehicle.

An abstract model of a vehicle following a driving path is depicted in Figure 2. In the plane of vehicle movement, let the vehicle-fixed x-axis be tangential to the path of the vehicle at any point in time, and the y-axis perpendicular. Then, four distinct maneuvers can be detected by thresholds on these measurements. Along the x-axis, forward acceleration corresponds to the throttle use of the driver, where abrupt peaks indicate aggressive increases of velocity. Deceleration is indicative of harsh braking, and therefore indirectly of not retaining a minimum distance to the vehicle ahead. Along the y-axis, high lateral acceleration points towards excessive velocity in left or right turns, and may result in the vehicle loosing traction. These four event types are concluded in Table 1 below, with exemplary thresholds.

Table 1. Event thresholds.

Event Type	Threshold Sensitivity*
Acceleration	$a_x > 0.1$ g
Braking	$a_x < -0.1$ g
Turn (Left)	$a_y > 0.2$ g
Turn (Right)	$a_y < -0.2$ g

* $g = 9.81$ m/s^2

This set of events therefore captures some essential characteristics of individual driving behavior. Through the definition of events in the form of geometrical properties and physical units, it is furthermore comparable and reproducible across different sensor platforms. A major part of smartphone-based applications for the assessment of driving behavior follow this approach, which is why we adapted it for our application as further elaborated below.

3. MOBILE SENSING APPLICATON

An overview of the application used in our study is depicted in Figure 3. On the hardware level, the application collects data from accelerometer, gyroscope, and GPS sensors in a smartphone, which for the prototype implementation is an iOS device. The sensor data forms the input to two functional components, calibration and trip recording. The calibration procedure is designed to determine the three-dimensional orientation of the device in the vehicle, and its performance contributes substantially to the reliability and accuracy of measurements. After it is completed, calibrated parameters are stored in the application data and trip recording may commence at the users discretion. During recording, sensor data is processed in a fast sampling loop which outputs aggregated trip profiles. Via the trip management component, users can access their data and receive driving feedback. A social network link enables users to share their performance data via a registered profile.

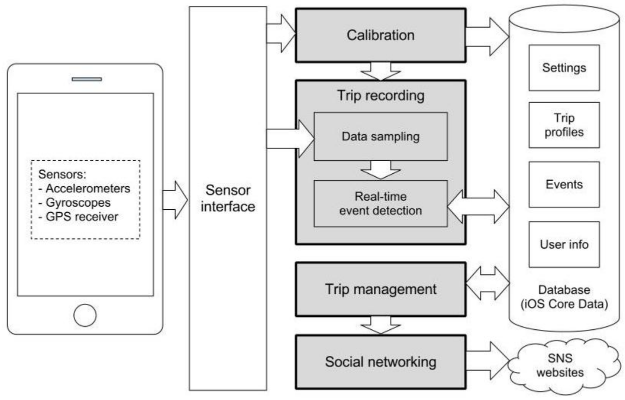

Figure 3. Application overview.

3.1 Calibration

There are three accelerometers and three gyroscopes available on the iPhone-integrated IMU for the measurement of lateral and angular accelerations along three fixed axes. These axes form the coordinate frame of the device. The *x*-*y* plane of the accelerometer sensor, for instance, is parallel to the touch screen of a device, with positive directions of *x* and *y* axes pointing to the right and top of the device, respectively. Ideally, the device is mounted in a vehicle in an orientation where all three axes are aligned with the relevant axes of vehicle manipulation, i.e., forwards and sideward. However, users are expected to place the device at an arbitrary rotated position. Next to the alignment problem which inhibits a direct interpretation of acceleration and rotation measurements, this also results in disturbed acceleration measurements due to the earth's gravitational field. In order to simplify data processing, we establish a virtual vehicle coordinate frame as a vector space that uses the principal directions of movement as a basis. In this coordinate frame, all gravitational accelerations should be measured along the *z*-vertical to the plane of vehicle movement. This premise can serve as a starting point to determine the device orientation and the necessary corrective transformations.

Our proposed calibration procedure consists of two steps and uses fundamental linear algebra to derive rotation matrices that change the iPhones coordinate system to the desired one. A rotation matrix R transforms the coordinates of an object, in this case an (lateral or angular) acceleration measurement vector *x*, to a new coordinate systems that is obtained by rotation relative to fixed axes [14]. An instance for a rotation matrix in three dimensions that corresponds to a turn of the coordinate frame around the *x* axis (1st vector component) by α degree is given below:

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}. \quad (1)$$

In order to align the virtual vehicle coordinate frame with the gravity vector, two rotational degrees of freedom have to be fixed, i.e., two sequential rotations are required. The third parameter is determined from turning the resulting coordinate frame around the vertical axis to align one of the horizontal axes with the principal direction of vehicle movement.

3.1.1 *Horizontal Plane*

In order to detect the horizontal plane, we assume that the application has been started and the device is placed and secured in the vehicle at an arbitrary orientation, while the vehicle is parked on a horizontal ground. Therefore, all accelerometer measurements originate solely from gravity. Let V_{dev} be the coordinate frame of the device and let V_{zv} be the z-vertical frame as illustrated in Figure 4 on the left. Given that the gravity vector in V_{zv} ought to take the values $g_{\mathrm{zv}} = (0, 0, -1)$, we can establish a constraint on the rotation matrix R1 as

$$g_{\mathrm{zv}} = R_1 g_{\mathrm{dev}}. \quad (2)$$

In order to obtain R1 that satisfies Equation (2) on an iPhone, we employ the ready-to-use-methods provided by the Core Motion Framework of iOS, which add the additional constraint of aligning one horizontal axis with compass north to achieve a fully determined system of equations. However, the corresponding function yields the inverse of R_1, i.e., the mapping from V_{zv} to V_{dev}. As rotation matrices have the useful quality of equivalence of inversion and transposition, we can take the transposed output to arrive at R_1. We acknowledge that magnetometer readings, and hence the direction of compass north, are possibly affected by the metal frame of a vehicle. However, in our approach it suffices to determine the rotation matrix for an arbitrary horizontal axes orientation, as the north alignment will be replaced by the second calibration step. We found that the iPhone's accelerometer is quite sensitive to vibrations induced by the vehicle engine even if the car is parked. Therefore, we average rotation matrices over a series of measurements comprising 50 individual acceleration samples. Testing proved the resulting rotation matrices to be stable.

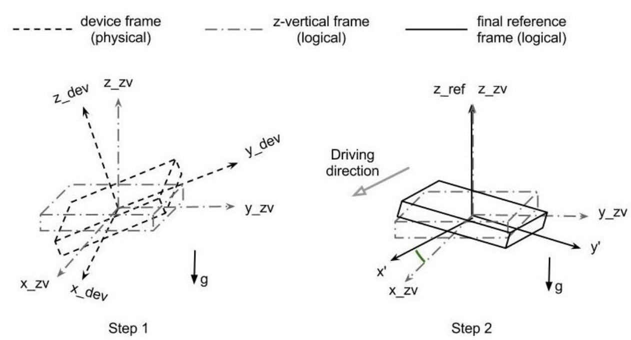

Figure 4. Two-step rotation to vehicle-aligned coordinate frame

3.1.2 *Principal Direction of Vehicle Movement*

The detection of the principal direction of vehicle movement, that is the "forwards" axis, is only possible after some indication of vehicle movement is available. If the vehicle accelerates straight ahead, the accelerometer should exhibit measurements a_{zv} that point in a somewhat constant direction when projected onto the horizontal *x*-*y* plane of V_{zv}. Based on the angle between this projection and the x axis of V_{zv}, we can than determine a rotation matrix R2 into the vehicle-aligned reference coordinate system V_{ref} . This second rotation is visualized in Figure 4 on the right. Our calibration algorithm for the inference of R_2 is based on [13] and aimed at detecting the first instance of straight ahead acceleration from which a valid direction can be inferred. We therefore use a triple threshold that comprises a minimum absolute acceleration value, a minimum time period during which this value is exceeded, and angular sector bound. These three thresholds are applied after subtraction of the gravity vector. Users may also choose to set vehicle orientation in the horizontal plane manually (Figure 5). We implemented this alternative for the second calibration step as an arrow visualization that can be adjusted by the user via the device's touch interface.

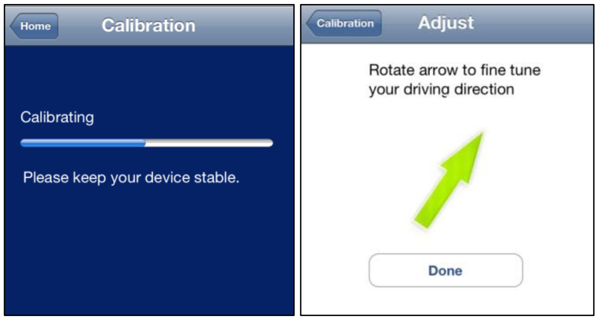

Figure 5. Automated and manual calibration of principal direction of vehicle movement

3.2 Trip Recording and Data Management

After calibration, the user can start a new trip to be recorded by the application. During operation, it collects three different types of raw data from which higher order constructs are computed:

- Trip start time, end time and duration,
- GPS latitude and longitude in regular intervals, from which velocity and driven distance are inferred, and
- Calibrated acceleration and gyroscope measurements sampled at 20 and 3⅓ Hz, respectively.

In order to record critical driving events, acceleration thresholds are introduced as elaborated in Section 2, subject to hysteresis: After a value has been exceeded for a defined amount of time, the application generates an event object together with GPS position and timestamp. Preliminary testing revealed that automatic detection of the principal driving direction did not work reliably and thus the manual angle setting was used for the field study. While a more sophisticated algorithm may solve this problem in the future, the manual configuration poses no relevant usability barrier in our opinion and gives an instant feedback of the applications internal parameters. Other than suggested by [15], we do not incorporate a dynamic re-adjustment of calibration parameters in R1 and R2 that arise due to (1) deviations in the horizontal plane if the vehicle is climbing or (2) movement of the device relative to the vehicle. Vehicles on regular road networks very seldom face inclinations of more than 10 degree, so that the resulting errors in acceleration measurement are negligible. With respect to user interaction with the smartphone that introduce pose alterations, we argue that European safety guidelines forbid phone usage while driving.

To limit memory usage, sensor data is discarded immediately after it is processed by the event detection algorithm. The application only stores trip objects that include time and distance, together with all events generated on the corresponding trip. The data structure of our application including user management objects is depicted in Figure 3. Users can access recorded trips in a list and access the registered events for each trip, which can also be displayed in a Google Maps mash-up based on the location where they were generated. This allows drivers to retrospectively understand where threshold violations occurred, and possibly identify locations of frequent inappropriate driving. Furthermore, users can post completed trips on their social network profile. Next to the driving score, the posting can contain trip location, time, and driven distance. Through frequent postings of on social networks, we anticipate that the application well receive greater awareness among potential users.

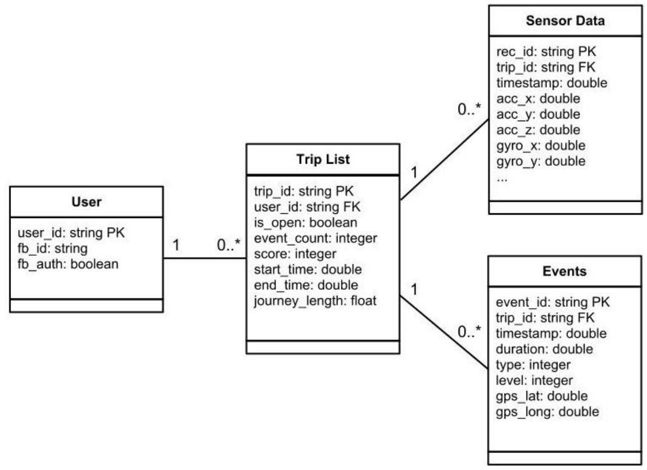

Figure 6. Data structure of application

4. EVALUATION

4.1 Objectives

As stipulated in the introduction, our objective is to assess the performance of a mobile sensing system by comparing its critical driving event counts with ground truth data obtained from vehicle-fixed reference sensor unit. Specifically, we expected the positioning and movement of a mobile sensing system in a vehicle to affect these event counts. Furthermore, next to these disturbances, event counts determined from a mobile sensing unit are also affected by external driving conditions. As scores determined by our application should only reflect individual driving behavior, it is important to identify and control for both these effects across different – i.e., "low" and "high" score – driving behaviors, and across the different types of driving events discussed in Section 2.

To subsume, our evaluation should take into account

- The specific type of critical driving events,
- Disturbances due to different configurations of the mobile sensing unit in a vehicle, and
- Disturbances due to driving conditions across a representative range of driving behaviors.

Previous studies known to the authors evaluated the performance of similar mobile applications in a laboratory setup and with questionable external validity. Therefore, our approach was to use a controlled field study setup, where test runs were obtained from a representative sample of drivers for constant external conditions in terms of vehicle type, driven distance, and road features.

4.2 Experimental Setup

4.2.1 Vehicle with Reference Sensor Unit

For the entire field study, one test vehicle was used. This decision was made based on the consideration that each additional test vehicle would require an increase in sample size to capture additional variance. As the large variety of vehicle models, motorizations and vehicle interfaces would ultimately render such an approach infeasible, we did not include vehicle type in our analysis. The test vehicle was a Renault Megane with manual shift, as a reasonable representation of mid-sized cars that are a major part of the overall vehicle population in Europe.

The car was equipped with a commercial sensor unit (MHub 837) that, next to a pre-calibrated IMU included GPS localization. The device was mounted directly on the vehicles OBD-II diagnostic port, from which it received its power supply and additional vehicle data such as speedometer readings. As such, the device was fixed in its orientation to the car, and in preliminary testing able to discriminate between lateral and longitudinal acceleration vectors. Its IMU measurements were therefore considered suitable as reference data for the mobile sensing application. The fixed sensor unit transmitted data wirelessly over GPRS to a backend system, from which aggregated information could be pulled using a standard fleet management solution. Threshold sensitivities on the unit could be preconfigured over a serial interface, so that no continuous transmission of acceleration data, which potentially strains data volume limitations, was required.

4.2.2 Test Route

In order to keep driving conditions constant between participants, we used a pre-defined route with a total length of 28.55 km, consisting of sections in city environment (11.85 km), on country roads (6.80 km) and highways (9.90 km). The route included inclined terrain with a maximum height difference of approximately 200m. Through the identical test route in every evaluation run, it was ensured that each participant driver would face the identical number of turns, traffic lights, ramps, intersections, and other infrastructure features.

Figure 7. Test route through different terrains and infrastructure features, approx. 45 minutes driving duration

4.2.3 Study Procedure

Recruitment of potential participants was primarily conducted among students at our university. Every participant was compensated by a gift voucher and participated in an additional prize raffle. Interested candidates were asked to fill out a short online questionnaire containing questions on their suitability. For ethical and insurance reasons, only persons with at least one year of driving experience, at most one self-induced accident and no past revocation of their driver's license were accepted. Interested and suitable subjects fixed an appointment with the test supervisor and were invited to join a testing session.

Introducing participants to the goal and contents of the study, the testing session started with a short questionnaire containing a declaration of liability and some items on car access and general driving behavior (e.g. "In average, how many trips do you take per month?"). Then, participants got an introduction to the use of the test car. Before starting, the test supervisor used the app to calibrate direction of vehicle movement and start trip recording. Preventing any potential influence by a co-driver, participants undertook the test drive on their own. Therefore, the pre-defined route was stored in the car-internal navigation system and participants were asked to follow navigation system's commands while driving. Participants who failed to follow the exact route were excluded from data evaluation.

Participants were randomly assigned to one of three subgroups, for each of which a different position of the mobile sensing device in the vehicle was used. After an initial pre-study with colleagues, it was determined that the common positions for a smartphone where (a) face-up on the co-driver's seat, (b) the middle console between driver's and co-driver's seat, usually made from hard plastic material, and (c) a designated smartphone holder attached to the front windshield, typically associated with mobile navigation systems.

5. RESULTS

Event count data was aggregated from the smartphone after each trip by accessing the application memory via the interface described in Section 3. Similarly, event counts from the reference IMU were exported from the commercial fleet management software in which transmitted data was stored. Left and right turns were summarized to general turn events, and events were aggregated for each participant. For the evaluation of collected data, IBM SPSS 19 was used.

5.1 Sample description

In total, 78 people participated in the study. Due to traffic jams or deviation from the test route, six subjects had a substantial rise in travel time or length and were thus excluded from further analysis.

The remaining 72 participants' mean age was 23.65 years ($SD = 4.77$; Range from 19 to 51). 83.3% were male, and 75.0% of them mainly used the car of their parents or other family members (only 22.2% owned a car by themselves). Average time since getting the driver's license was 5.46 years ($SD = 4.74$; $Mode = 2$) and participants reported to take 9.05 trips per month ($SD = 8.79$) with an average distance of 28.10 km per trip ($SD = 24.06$). The average time needed for the test drive was 46.37 minutes ($SD = 5.06$).

5.2 Control variables

A standardized procedure including an identical test route for all candidates was chosen to minimize possible side effects of varying traffic environments. However, external factors such as daytime or weather conditions might jeopardize standardization by resulting in different traffic densities or unexpected incidents such as road closure. Thus, resulting event data was controlled for the impact of daytime and weather.

To assess daytime, start time was noted at the beginning of every test drive. The earliest test drive was undertaken starting at 8:47 AM, the latest at 4:58 PM. As only low correlations were found between daytime and IMU data (Reference IMU: $r = -.02$, $p = .85$; mobile IMU: $r = -.13$, $p = .27$), significant influence of daytime on measured events could be excluded.

Weather conditions were assessed using a basic rating procedure. At the beginning of every test drive, the test supervisor rated actual weather conditions on a 5-point scale ranging from 1 (= rainstorm) to 5 (= sunshine). A calculated mean of 3.92 ($SD = 0.80$) over all test drives suggested mainly bright or sunny conditions. Actually, no test drive was undertaken under heavy rain. In connection to IMU data, no significant correlations were found (Reference IMU: $r = .23$, $p = .06$; mobile IMU: $r = -.05$,

$p = .65$). Thus, a systematic influence of weather conditions on measured data could be excluded as well.

5.3 Event Count Statistics

5.3.1 Reference IMU

In average, the on-board unit registered 43.40 events ($SD = 10.71$). With regard to the standardized route of 28.55km, this implies participants created about 1.5 events per km. Results of a Kolmogorov-Smirnov-Test indicated data to be normally distributed ($Z = 0.56$, $p = .92$) with two cases highly deviating from mean (see Figure 8). However, those outliers were not excluded from further analysis as there was no evidence for IMU malfunction or anomalies concerning standardization (e.g. traffic jams).

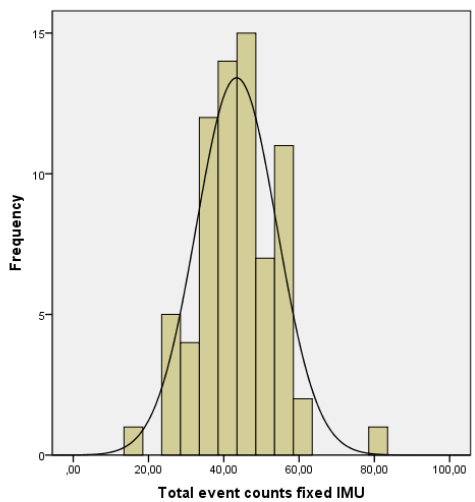

Figure 8. Distribution of event counts for fixed IMU

With regard to particular event types, turns and braking were registered more than twice as often as accelerating events (see Table 2). These findings suggest the need for braking and heavy turning to occur more frequently in real traffic situations. Additionally, road type manifested as important impact factor: As illustrated in Table 2, more than twice as many events were registered in city environment compared to country roads. In contrast, event counts on highways were surprisingly low, indicating the most defensive and constant driving style on freeways, despite high average speed.

Table 2. Means and standard deviations for event counts registered by fixed IMU

Events (road/type)		*Acc.*	*Braking*	*Turns*	*Total*
City	*M* *SD*	7.11 3.23	13.54 4.15	8.32 3.36	28.97 7.97
Country	*M* *SD*	1.24 1.24	3.22 1.53	8.93 3.54	13.38 4.78
Highway	*M* *SD*	0.24 0.43	0.36 0.68	0.44 0.95	1.04 1.48
Total	*M* *SD*	8.58 3.73	17.13 4.92	17.69 5.37	43.40 10.71

Quantified for each participant separately, events referring to a specific event or road type can be seen as repeated assessments of driving style. Thus, repeated-measure ANOVAs were conducted to test for differences in event counts based on event and road type. For both analyses, the assumption of sphericity was violated. Thus, the Greenhouse-Geiser approach was used to correct for degrees of freedom. Significant differences were found both for event ($F\ (1.84, 130.77) = 130.58$, $p < .001$) and road type ($F\ (1.54, 109.16) = 560.28$, $p < .001$). Post-Hoc Tests using the Bonferroni correction revealed differences in braking compared to acceleration ($t\ (71) = 14.11$, $p < .001$) and turn events ($t\ (71) = 16.16$, $p < .001$), and differences in city environments compared to country roads ($t\ (70) = 16.58$, $p < .01$) and highways ($t\ (70) = 28.81$, $p < .01$), respectively. Not surprisingly, city traffic requires more acceleration, braking and sharper turns. However, while braking and acceleration occur less frequently on country roads, turns seem equally important, probably due to increased rotary traffic outside the city.

5.3.2 Mobile IMU

In average, the iPhone prototype registered 48.57 events ($SD = 58.32$), signifying 1.7 events per driven km. However, histograms indicated a skewed distribution (see Figure 9). A Kolmogorov-Smirnov-Test confirmed this hypothesis, revealing data to be non-normally distributed ($Z = 2.26$, $p < .01$).

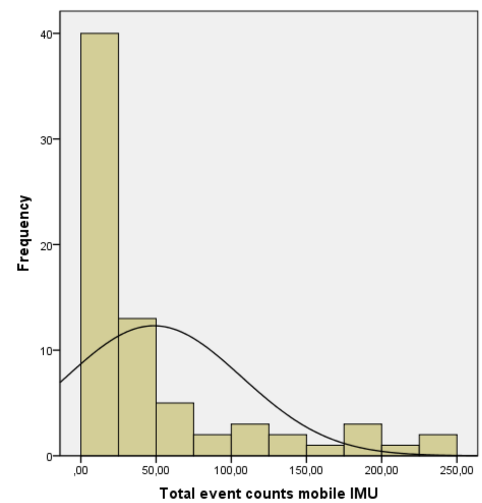

Figure 9. Distribution of event counts for mobile IMU

Based on a pre-study, the position of the smartphone in the vehicle had been systematically varied between test drives. Data revealed substantial differences both for different positions and event types (see Table 3). While braking events were the most prevalent event type occurring under co-driver`s seat condition, acceleration was the most noticeable event type when the smartphone resided in the car holder. Additionally, high standard deviations of acceleration and braking events pointed towards large variability in event counts between test drives, especially when the smartphone was positioned on the co-driver`s seat. In contrast, number of turning events did merely vary between conditions, and low standard deviations indicated homogenous measurement between test drives.

With respect to the violation of the assumption of normality, most common methods like ANOVA could not be used for further data analysis. Instead, non-parametric procedures were considered. Differences between mobile positions were assessed using Kruskal-Wallis tests, a non-parametrical equivalent to ANOVA [16]. Results revealed significant differences for all event types. (Acceleration: $\chi^2\ (2, N = 72) = 19.80$, $p < .01$; Braking: $\chi^2\ (2, N = 72) = 9.04$, $p < .01$; Turns: $\chi^2\ (2, N = 72) = 18.93$, $p < .01$; Total: $\chi^2\ (2, N = 72) = 11.72$, $p < .01$.). Post-Hoc tests using the Bonferroni approach indicated that these differences were mainly due to higher event counts on co-driver’s seat compared to the two other positions. Thus, smartphone attachment and shifting was a main source of variability in event counts.

Table 3. Means and standard deviations for event counts registered by mobile IMU

Events (pos./type)		*Acc.*	*Braking*	*Turns*	*Total*
Co-driver`s seat	*M*	9.97	60.11	4.49	74.57
	SD	24.73	72.75	2.22	71.82
Middle console	*M*	3.90	5.75	6.95	16.60
	SD	6.01	4.90	3.07	7.18
Windshield holder	*M*	15.67	3.60	7.80	27.07
	SD	10.51	2.92	2.76	13.17
Total	*M*	9.47	33.24	5.86	48.57
	SD	18.94	58.88	2.94	58.32

5.4 Comparison of Mobile Measurements with Reference IMU

Directly connected to the car, the reference IMU is considered to reliably measure acceleration, braking and turn events. In order to assess measurement quality of the mobile IMU, event registered by the smartphone app were systematically correlated with those measured by the fixed IMU. Using correlations to estimate measurement validity is a well-known procedure, especially in behavioral sciences [17]. As Pearson`s correlation expects data to be normally distributed, Spearman´s ρ as non-parametrical equivalent was used instead.

As explained above, a more differentiated view was needed due to differences in event types and smartphone position. Thus, correlations were systematically calculated for each combination of event type and position (see Table 4). Significant correlations for acceleration and braking events indicated that the smartphone app was able to capture driving events with similar accuracy as the reference IMU in general. Turn events captured by mobile phone, however, deviated from those measured by the reference IMU particularly for the smartphone positioned on the co-driver´s seat. With regard to the smartphone position, highest correlations were attained using a car holder. Still, a substantial correlation could be registered with the smartphone positioned on the co-driver´s seat.

In order to further investigate the cause diverging measurements, we analyzed scatter plots that compare mobile vs. reference IMU events of each study participant. Figure 10 displays such a plot for the total number of event counts. From these, it became evident that event counts generated by a smartphone are in specific instances much higher, while in the remaining majority of cases (clustered on the left side of the plot) measurements appear exhibit higher correlation.

Table 4. Correlations of iPhone data and OBD-II data for different event types and iPhone positions

ρ (iPhone, fixed IMU)	Event Counts			
	Acc.	*Braking*	*Turns*	*Total*
Co-driver’s seat	.19	.42**	.13	.42**
Middle console	.61**	-.11	.40^	.56**
Windshield holder	.68**	.32	.49^	.74**
Total	.45**	.27*	.06	.56**

^ p < .10; * p < .05; ** p < .01

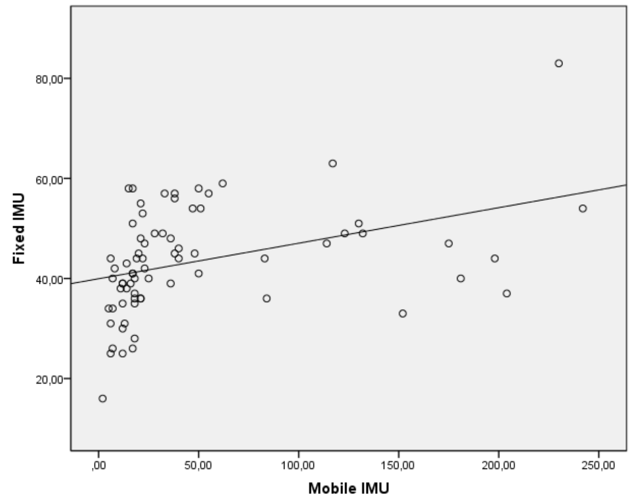

Figure 10. Scatter Plot of mobile vs. fixed IMU with linear approximation function

6. DISCUSSION

Data Analysis revealed significant correlations between mobile and reference IMU event counts, indicating sufficiently reliable event detection by the smartphone application. However, correlations differed depending on event type and smartphone position. In addition, for the smartphone lying on the co-driver’s seat, means and variances in registered event counts were higher for the smartphone compared to the on-board unit. These findings suggest a high error rate for smartphone-based event counts for specific smartphone positions and event types, respectively.

From Figure 10, it is evident that smartphone measurements appear to be exaggerated in certain instances, a fact that may be attributed to faulty calibration in these cases: If the mobile device reference frame deviates too strongly from the vehicle frame, thresholds are effectively altered. This effect was more noticeable for lesser-constrained positions of the smartphone in the vehicle.

With regard to the influence of driving conditions, significant differences of registered event counts due to different road types were found. This observation is in accordance with expectation, as city traffic requires more acceleration, braking and sharper turns. However, it also suggests that event counts for a particular driver reflect to a significant extent the road profile on which a vehicle was moved. Therefore, event counts should not only be normalized by driven distance, but also with respect to the average event count for the associated road type. Surprisingly, none of the currently available applications for the assessment of driving behavior takes this issue into account.

6.1 Limitations

Several limitations restrict the generalizability of our results. Firstly, though the controlled field study approach for evaluation was concerned with a broad range of variables, not all potentially relevant factors could be considered. Prominently, only one specific vehicle model was used. Others may differ in attainable accelerations or the locations in the vehicle where a smartphone running the application can be placed. Other conceivable factors are the traffic situation which was specific to Zurich, or unique characteristics of the chosen test route. However, we feel that the results of our study should not change substantially in a setting were these are altered.

Secondly, threshold values were deliberately chosen very low to achieve high resolution measurements. An alternative approach would use higher thresholds on acceleration, thus emphasizing “extreme” driving events. These would occur much less frequently and may not have delivered sufficient variability over the limited duration of the test run.

Thirdly, our setup did not consider smartphone usage by the driver during trip recording. This would have without doubt affected event counts. The handling of a smartphone during a call or the execution of other applications would have introduced additional acceleration noise; furthermore, after usage the device may not have been returned to its original place and orientation. Such disturbances make a strong case for adaptive, live calibration algorithms to be incorporated into our design.

6.2 Further Research

Our findings point towards research challenges arising in the domain of driving behavior applications to be met by the mobile sensing community in the future. A more sophisticated, online calibration algorithm is potentially capable of detecting changes in orientation of the smartphone with respect to the vehicle. Such an algorithm can alter its configuration accordingly and thus reduce measurement biases arising from a misalignment. Furthermore, it may be extended by plausibility checks that compare acceleration-derived events with gyroscope or GPS data and eliminate invalid or inconsistent events. However, we consider such algorithms to be inherently complex and difficult to verify, and thus point at another design option: If event thresholds in varying directions can be chosen identical, the calibration of driving direction is no longer necessary. In this scenario, only the horizontal plane alignment would have to be taken into account.

Another issue to be addressed by future research is the effect which the feedback of event count scores has on actual driving behavior. Opposed to conventional on-board units that monitor driving behavior, mobile applications allow for instant feedback, visual or acoustical, during driving or after completion of a trip. Instead of the rather technical and data-centric design of current driver rating applications, more advanced designs that build upon concepts such as Social Norms or Gamification may prove more successful in achieving lasting improvements of driving behavior.

7. CONCLUSION

This paper has evaluated a mobile application for the assessment of driving behavior based on critical driving events in a controlled field study. We found that significant correlations with reference measurements from a vehicle-fixed unit exist, albeit conditional on event type, position of the smartphone in a car, and external influence factors. We pointed out specific points as to how the performance of such systems may be further increased in the future. While there remains room for improvement, insurers are thus well-advised to consider smartphones as an interesting alternative to conventional “black boxes” for the monitoring of driving behavior. Thereby, they may leverage the broadly available base of smartphones with advanced sensing capabilities, reduce device and data transfer costs, and reduce the perceived intrusion into policyholders’ privacy.

The functional performance of such systems as considered in this paper is only a foundation – both researchers and practitioners should aim for innovative interfaces that enable drivers to make use of the collected information and improve their actual behavior. Therein lays an opportunity not only to reduce claims costs and reward insurance customers, but also to increase overall traffic safety and potentially safe human lives.

8. ACKNOWLEDGEMENTS

(omitted for review)

9. REFERENCES

[1] E. Fleisch and F. Thiesse, “On the management implications of ubiquitous computing: an IS perspective,” in Proceedings of the 15th European Conference on Information Systems, St. Gallen, Switzerland, 2007, pp. 1929–40.

[2] V. Coroama, “The smart tachograph - individual accounting of traffic costs and its implications,” in Proceedings of the 4th International Conference on Pervasive Computing, Dublin, Ireland: Lecture Notes in Computer Science, Vol. 3968, 2006, pp. 135–152.

[3] J. Paefgen, F. Michahelles, and T. Staake, “GPS trajectory feature extraction for driver risk profiling,” International Workshop on Trajectory Data Mining and Analysis, 13th Conference on Ubiquitous Computing, Beijing, China, 2011.

[4] I. J. Wouters and J. M. Bos, “Traffic accident reduction by monitoring driver behaviour with in-car data recorders.,” Accident Analysis & Prevention, vol. 32, no. 5, pp. 643-50, Sep. 2000.

[5] C. Troncoso, G. Danezis, and E. Kosta, “Pripayd: privacy friendly pay-as-you-drive insurance,” ACM workshop on Privacy, 2007.

[6] T. Toledo, O. Musicant, and T. Lotan, “In-vehicle data recorders for monitoring and feedback on drivers’ behavior,” Transportation Research Part C: Emerging Technologies, vol. 16, no. 3, pp. 320-331, Jun. 2008.

[7] A. Pérez, M. I. Garcia, M. Nieto, J. L. Pedraza, S. Rodríguez, and J. Zamorano, “Argos: An Advanced In-Vehicle Data Recorder on a Massively Sensorized Vehicle for Car Driver Behavior Experimentation,” IEEE Transactions on Intelligent Transportation Systems, vol. 11, no. 2, pp. 463-473, Jun. 2010.

[8] O. Musicant, H. Bar-gera, and E. Schechtman, “Electronic records of undesirable driving events,” Transportation Research Part F: Psychology and Behaviour, vol. 13, no. 2, pp. 71-79, 2010.

[9] R. Bhoraskar, N. Vankadhara, B. Raman, and P. Kulkarni, “Wolverine: Traffic and road condition estimation using smartphone sensors,” 2012 Fourth International Conference on Communication Systems and Networks (COMSNETS 2012), pp. 1-6, Jan. 2012.

[10] M. Li, J. Dai, S. Sahu, and M. Naphade, “Trip analyzer through smartphone apps,” Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS ’11, p. 537, 2011.

[11] D. a. Johnson and M. M. Trivedi, “Driving style recognition using a smartphone as a sensor platform,” 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1609-1615, Oct. 2011.

[12] A. K. Shaout and A. E. Bodenmiller, “A Mobile Application for Monitoring Inefficient and Unsafe Driving Behaviour,” 1996.

[13] M. Ruta, F. Scioscia, F. Gramegna, E. D. Sciascio, and P. Bari, “A Mobile Knowledge-based System for On-Board Diagnostics and Car Driving Assistance,” no. c, pp. 91-96, 2010.

[14] R. C. Hibbeler, Engineering Mechanics: Dynamics, 12th ed. Upper Saddle River, NJ: Prentice Hall, 2009.

[15] K. Li, M. Lu, F. Lu, Q. Lv, and L. Shang, “Personalized Driving Behavior Monitoring and Analysis for Emerging Hybrid Vehicles,” Pervasive 2012.

[16] A. Field, “Discovering Statistics using SPSS”, London: Sage Publications, pp. 559-571, 2009.

[17] C.J. Goodwin, “Research in Psychology: Methods and Design”, New York: Wiley, p. 132., 2009.