KTH Electrical Engineering

Insurance telematics: opportunities and challenges with the smartphone solution

PETER HÄNDEL, ISAAC SKOG, JOHAN WAHLSTRÖM, FARID BONAWIEDE, RICHARD WELCH, JENS OHLSSON, AND MARTIN OHLSSON

Stockholm 2014

Signal Processing
School of Electrical Engineering
Kungliga Tekniska Högskolan

Insurance telematics: opportunities and challenges with the smartphone solution

Peter Händel, *Senior Member, IEEE*, Isaac Skog *Member, IEEE*, Johan Wahlström, Farid Bonawiede, Richard Welch, Jens Ohlsson, and Martin Ohlsson

***Abstract*—Smartphone-based insurance telematics or usage based insurance is a disruptive technology which relies on insurance premiums that reflect the risk profile of the driver; measured via smartphones with appropriate installed software. A survey of smartphone-based insurance telematics is presented, including definitions; Figure-of-Merits (FoMs), describing the behavior of the driver and the characteristics of the trip; and risk profiling of the driver based on different sets of FoMs. The data quality provided by the smartphone is characterized in terms of Accuracy, Integrity, Availability, and Continuity of Service. The quality of the smartphone data is further compared with the quality of data from traditional in-car mounted devices for insurance telematics, revealing the obstacles that have to be combated for a successful smartphone-based installation, which are the poor integrity and low availability. Simply speaking, the reliability is lacking considering the smartphone measurements. Integrity enhancement of smartphone data is illustrated by both second-by-second low-level signal processing to combat outliers and perform integrity monitoring, and by trip-based map-matching for robustification of the recorded trip data. A plurality of FoMs are described, analyzed and categorized, including events and properties like harsh braking, speeding, and location. The categorization of the FoMs in terms of Observability, Stationarity, Driver influence, and Actuarial relevance are tools for robust risk profiling of the driver and the trip. Proper driver feedback is briefly discussed, and rule-of-thumbs for feedback design are included. The work is supported by experimental validation, statistical analysis, and experiences from a recent insurance telematics pilot run in Sweden.**

I. Introduction

A. What is insurance telematics?

Insurance telematics is a flavor of telematics [1] which relies on an insurance premium that is based not only on static measures like the drivers age, occupation or place of residence, car model and configuration, or expected mileage over the policy period, but also on dynamic measures like actual mileage, time spent on the road or the time of day when the trip is being made, location, and the driver's actual style of driving. These insurance schemes are often labeled as pay-as-you-drive (PAYD), pay-how-you drive (PHYD), manage-how-you-drive (MHYD), and the like. Accordingly, the premium is based on information gathered from car trips utilizing different types of measurement probes, spanning from original equipment manufacturers (OEM) installed black-boxes[1], dongles plugged into the vehicle's on-board diagnostics (OBD) outlet, to contemporary smartphones; two UBI measurement probes (sensor platforms) are shown in Fig. 1.

Fig. 1. Examples of measurement probes for insurance telematics: Progressive Insurance Snapshot measurement probe for the OBD outlet (left), (Courtesy: Progressive Insurance); Screenshot of Movelo Smartphone-UBI software RUBI (right) displaying alarm indicators for speeding, non-smooth drives, harsh accelerations, swerving, harsh braking, and heavy cornering (Courtesy: Movelo AB).

B. Trends in insurance telematics

The UBI programs are expected to take off during the coming years. The market forecasts predict that up to 30% of the vehicles in the United States and 60% of the vehicles in the United Kingdom will be insured trough some type of insurance telematics program by 2020 [2]. The number of policy holders corresponds to some 60 million in the United States alone. Programs using fixed vehicle installations are based on mature technologies, e.g., *Progressive Casualty Insurance Company* in the US has currently around 1.5 million end-users who have tried out their UBI program utilizing information extracted using the measurement probe displayed in Fig. 1. However, the cost for hardware, including the indirect cost for installation, maintenance, and logistics, limits the scalability of the UBI-

Manuscript received

P. Händel, I. Skog, J. Wahlström, F. Bonawiede, and M. Ohlsson are with the ACCESS Linnaeus Center, Dept. of Signal Processing, KTH Royal Institute of Technology, Stockholm, Sweden. (e-mail: ph@kth.se).

R. Welch is with REW Insurance Consulting Services, Holden, MA, USA.

J. Ohlsson is with the Department of Computer and System Sciences, Stockholm University, Kista, Sweden. He is also with Movelo AB, Stockholm, Sweden.

[1]Within the insurance telematics area, the term black-box refers to a for the purpose tailored measurement probe that is fixedly installed inside the vehicle. It may be equipped with its own sensors or piggy-backed onto the vehicle's internal sensors via the CAN-bus.

Fig. 2. Typical wind-shield mounting of the smartphone running an insurance telematics application (Courtesy: If P&C). The mounting supports smartphone based vision systems for smartphone-based looking in and looking out of a vehicle using the back and front mounted cameras, respectively; refer to e.g., [6].

programs and reduces their deployment. Insurance providers are therefore seeking scalable solutions.

The cellular phone and its use during car trips has early on been identified as a valuable tool for intelligent transportation systems [3], [4]. The use of (typically) wind-shield mounted contemporary smartphones, as illustrated in Fig. 2, has been identified as a promising option, thanks to the high penetration of smartphones among end-users, the development talent within the telecom industry, and the ease of deployment of smartphone functionality via distribution of downloadable applications. In fact, the smartphone is a technology driver. For example, considering receivers for the different global satellite navigation systems (GNSS), the first global positioning system (GPS) receiver equipped smartphones entered the market in 2007; handsets with receivers that work with the signals from both the GPS and the Russian GLONASS entered around 2011; and handsets that, in addition to the GPS and GLONASS signals, can also receive and utilize the signals from the European Galileo and the Chinese BeiDou systems are expected to appear during the end of 2014. In addition, the large manufacturing volumes have turned smartphones into a driving force in the development of low-cost microelectromechanical system (MEMS) sensors such as accelerometers, gyroscopes, magnetometers, etc. Accordingly, the price–performance metric of the measurement capabilities of the smartphone is continuously improving over time; refer to [5] for a survey on the sensing capabilities of a contemporary smartphone. Further, the smartphone also provides audio-visual means for user interaction prior to (eg. for trip planning), during (cf. the extended dashboard feedback displayed in Fig. 1), and after a trip (eg. trip analysis and travel report). Finally, the wireless connectivity of the smartphone provides an efficient means for data transfer, avoiding an additional data plan and SIM card, which are typically required by black-boxes and OBD dongles.

Noteworthy, is that automobile manufacturers are increasingly equipping new vehicles with telematics capabilities that do not require additional hardware. This technology, however, will take many years to penetrate the market to a level where it would represent a primary means of collecting UBI data. As a result, the reliance on fixed installed devices, OBD-dongles, or smartphones is likely to remain dominant for the next several years.

Despite the favorable properties of the smartphone, smartphone-based insurance telematics have not yet succeeded to enter the market on a large scale. During recent years, some pilot trials have been launched, but no full-blown commercial programs are currently available. The project *If SafeDrive*, launched by the Swedish insurer *If P&C* in 2013, is a reported example of a recent commercial pilot, where the smartphone was used as an advanced measurement probe [7]. During the campaign, some 1,000 signed-up drivers piloted the program which led to both a transformation of existing customers to the UBI program, as well as recruitment of new customers. One may note that these pilots typically are of the same size as the Berkeley Mobile Millennium Project [8], CA, USA, in terms of smartphone-collected road traffic information; e.g., the *If SafeDrive* campaign collected some 4,500 driving hours / 250,000 km of road vehicle traffic data [7], which is believed to have an interest in its own right.

C. *Smartphone data quality*

A major reason for the slow deployment of smartphone-based UBI programs is identified as the issues with data quality, and in particular the reliability of measurement data. Simply speaking, the smartphone does not provide data of required quality so that the state-of-the-art algorithms implemented in UBI tailored hardware measurement probes can be directly transferred to a smartphone application. An illustrative example is detection of harsh braking events, where thresholding the calculated change of speed typically gives a large number of false detections due to the occurrence of outlier data and variations in the data acquisition rate. We will come back to this example in Sec. III.

A *naïve* explanation for the poor data quality is, of course, that the smartphone *per se* is not designed to be a high-end measurement probe for capturing movements with high dynamics, such as harsh braking during a car trip. To take the discussion forward, the data quality is in Tab. I categorized according to the quality measures commonly employed within the (in-car) navigation system research field [9].

In traditional insurance telematics where the speed is captured directly from the vehicle, e.g., via the OBD outlet, the captured speed is quite accurate (it is, however, subject to some quantization and offset). Thanks to the proprioceptive sensing of the speed, the integrity, availability, and continuity of service are all high. In order to use the smartphone as the information source, on the other hand, the vehicle's speed has to be captured by some exteroceptive means, where second-

TABLE I
QUALITY MEASURES OF UBI SENSOR INFORMATION [9].

Measure	Description
Accuracy	The conformity between sensor information and actual value.
Integrity	The reliability (trust) that can be put in the sensor information and the systems quality (accuracy) indicators.
Availability	The geographical coverage for which the sensor information is available.
Continuity of service	The availability of the service over time without non-scheduled interruptions during the intended working period.

TABLE II
QUALITY OF THE SPEED DATA ORIGINATING FROM THE OBD-OUTLET AND THE SMARTPHONE GNSS-RECEIVER, AS WELL AS THE QUALITY OF THE GNSS-RECEIVER SPEED DATA AFTER THE INTEGRITY HAS BEEN ENHANCED THROUGH DIGITAL SIGNAL PROCESSING.

Sensor	Accuracy	Integrity	Availability	Continuity of Service
OBD	Medium	High	High	High
GNSS	High	Low	Low	High
Enhanced GNSS	High	High	Low	High

by-second speed data from the GNSS-receiver[2] typically is the method of choice. This speed data is accurate, but subject to quite frequent occurrences of undetected outliers as well as irregularities in the data acquisition rate, i.e., the data has low integrity. The availability is inferior compared with the speed data provided by the OBD outlet, because of the dependency on line-of-sight conditions to the navigation satellites [10]. In Tab. II, the vehicle's speed, captured by the proprioceptive OBD-dongle, and the exteroceptive smartphone GNSS-receiver is compared with respect to the four performance measures introduced in Tab. I.

D. Challenges in Smartphone-UBI

As displayed in Tab. II, the first engineering challenge for success of smartphone-based insurance telematics is how to handle the lack of integrity of the GNSS-receiver speed data. Both the occurrence of outliers as well as the irregularities in the data acquisition rate are, however, detectable and combatable by employing low-level (that is, close to the sensor) digital signal processing on the data streams. The result is a data sequence with improved quality - in Tab. II denoted enhanced GNSS. Such enhancement of data quality is one important aspect that has to be considered in the design of smartphone based UBI systems, which will be discussed within this article.

A second engineering challenge in the design of a smartphone based UBI system is to provide proper methodologies to ensure robust driver scoring based on data with low availability and low integrity (but, possibly enhanced). Clearly, lost measurements are not recreatable, but the sought after UBI measures for policy determinations may still be extractable. For example, a harsh braking cannot be detected if the data covering the seconds of the events are lost, whereas a measure like smoothness of the trip can still be calculated with some predetermined accuracy despite severe loss of data. This is another main topic to discuss within this article.

A third challenge is the validity of the scoring, i.e., the correlation between the measured figure of merits with the corresponding scoring and the actual risk profile of the driver; which has to be investigated using database information of claim statistics. Verifying the validity of scoring includes studying such (disputed) claims that drivers who sign up for UBI programs *per se* are safer drivers, than those who do not. Studies of the validity are of utmost importance, but also a challenge to be handled by the actuaries at the insurance companies, and not by the engineers designing and operating the systems. Accordingly, it is beyond the scope of this paper.

E. Contributions and outline

For the sake of the discussion, a simple yet descriptive definition of UBI or insurance telematics is given by:

Insurance telematics defines the process of using sensor measurements to extract relevant figure of merits (FoMs) of a car trip driven by a (human) driver. The FoMs are later used to calculate a driver safety profile (a score) based on several trips, where the score will influence the driver's insurance premium.

Based on the above definition, some remarks are in order. First, the general purpose of this work is to discuss the technology aspects of insurance telematics, and the particular purpose is to highlight the implications of using smartphones as measurement probes. In other words, the process of gathering sensor data and extracting trip-based FoMs, as studied in Sec. II. The technology aspects include characterization of the FoM in terms of: how important they are for the scoring in the underwriting process; to what degree they can be influenced by the driver to reduce the insurance premium; and their observability (i.e., the correlation between actual sensor measurements and the FoM) and stationarity (i.e., the time length of the associated events).

Secondly, the enhancement of information integrity is considered in Sec. III, where digital signal processing on different levels is discussed; spanning from sensor-near model-based enhancement of second-by-second data, to trip-based post processing using additional information sources such as digital maps.

Thirdly, we discuss how the gathered FoMs and their corresponding characterization in terms of importance, influence, observability, and stationarity can be utilized for the underwriting or scoring process. In particular, we discuss in Sec. IV the robustness issue, i.e., the robustness of the driver score with respect to the properties of the FoM. Not only the robustness of the scoring and reliability of data are of importance for a successful UBI program, but also a key concern is the feedback that should be provided to the individual drivers.

[2]The operation systems on some smartphones do not allow data to be read specifically from the GNSS-receiver, but only from the so called Location Service that uses a combination of Cellular, Wi-Fi, Bluetooth, and GNSS-receiver data to locate the phone. We will throughout the paper make no distinction between these data sources and only refer to it as GNSS-receiver data.

TABLE III
QUALITY MEASURES OF UBI FOMS.

Measure	Description
FoM observability	The correlation between the sensor measurements and the FoM.
Event stationarity	The time length during which the events which make up the FoM normally are registered.
Actuarial relevance	The importance of the FoM for the risk assessment of the driver.
Driver influence	The extent to which the driver influences the FoM.

Here, real-time feedback typically concerns the use of the audio visual capabilities of the smartphone, as exemplified by the extended dashboard screenshot in Fig. 1. Accordingly, not only the robustness of the feedback is of importance, but also the perception of it, which is discussed in terms of some examples in Sec. IV.

The conclusions are drawn in Sec. V. Before we continue, we should emphasize that insurance telematics by no means excludes the use of the traditional static measures for the risk calculation, such as drivers' age and the model of the car. They are, however, out of relevance for the discussions performed in this article, and accordingly not further discussed.

II. FIGURE OF MERITS

While there exists many different FoMs that may be used for insurance telematics, we have listed some of the most common in Tab. IV. We emphasize that the list by no means claims to be a full list of FoMs, as it excludes, e.g., FoMs reflecting the road conditions and traffic intensity, which may be measured via accelerometer based pothole detection and by analysis of the sound recorded with the microphone of the smartphone [11], [12]. The characteristics of an FoM depend on the properties of the driver's behavior that the FoM is to reflect and the quality of the sensor measurements that are used to calculate the FoM. We will in this section describe the characteristics of the FoMs listed in Tab. IV and how the quality of the sensor measurements affect them.

A. Characterization of FoMs

The FoMs are by different means calculated from the sensor measurements characterized by their corresponding accuracy, integrity, availability, and continuity of service. The effect of the data quality on the FoMs is a non-trivial task to determine, because it depends on the correlation between the sensor measurements and the FoM (that is, the observability) and the time length during which the events which make up the FoM normally are registered (that is, the stationarity of the FoM). Clearly, it is an easier task to measure an FoM that is based on events that have a long duration, e.g., the smoothness over the trip, as opposed to an FoM which consists of transient events like the number of harsh braking events. In a similar vein, it is an easier task to measure an FoM with high observability such as the elapsed time of a trip, as opposed to measuring the amount of swerving using GNSS-receiver data. The higher the observability or stationarity measures, the less sensitive the FoM will be with respect to the integrity and the availability of data.

From an underwriting perspective, the mentioned measures are not enough. The actuarial relevance of the FoM for the scoring also plays an essential role. Traditionally, FoMs such as the amount of speeding and the number of harsh brakings are considered as relevant; however, the choice of suitable FoMs for insurance telematics is a hot topic of discussion, and is typically a choice for the individual underwriters designing the UBI programs.

Concerning the chosen feedback given to the driver, it is more relevant to what extent the driver can improve his driving with respect to the given feedback, than how large of an impact the associated FoM has on the insurance premium. A proper driver feedback should be based on FoMs with high associated influence. For example, the driver can typically influence the smoothness of a ride, whereas the destination is set for a trip with a given purpose. The measures we have introduced to characterize the different FoMs are summarized in Tab. III.

B. FoM survey

In this section, the FoMs summarized in Tab IV are described together with their characteristics.

1) Acceleration and Braking: The acceleration FoM and braking FoM are commonly defined as the number of, per unit distance or unit time, hard acceleration events and harsh braking events, respectively. The latter FoM is by several insurance companies viewed as one of the best indicators of a driver's risk profile, as it reflects how aggressive and observant a driver is, and how well the driver plans the driving with respect to the other road users and his surrounding. The observability of the acceleration and braking FoMs is considered as medium since the acceleration and deceleration of the vehicle are not directly measured by the GNSS-receiver, but rather calculated by differentiating the speed measurements. Robust calculations of the accelerations and decelerations from the speed measurements are a non-trivial task as shown in Sec. III; since the speed measurements frequently contains outliers and the acquisition rate of the speed measurement on some smartphones may vary with time. The time duration of an acceleration and braking event is typically quite short, and thus the stationarity of the FoMs are classified as low. In detail, the acceleration and deceleration events used to calculate the FoMs have durations of a few seconds and the GNSS-receiver inside the smartphones nominally has an update rate of once per second, which means that there will be only a handful of measurements holding information about the event; data loss during an event can thus not be compensated for. The driver's influence on the acceleration and braking FoMs are classified as high and medium, respectively. The driver can mostly avoid hard accelerations, but there may be unforeseen and uncontrollable events that may force the driver to do a harsh braking.

2) Speeding: Speeding is a typical FoM of relevance. Lowering the average speed in road traffic has shown to be a valid means to reduce the number of road fatalities and thus,

TABLE IV
CHARACTERIZATION OF FOMS IN INSURANCE TELEMATICS CALCULATED USING GNSS-DATA IN TERMS OF FOM OBSERVABILITY, EVENT STATIONARITY, DRIVER INFLUENCE, AND ACTUARIAL RELEVANCE. CHARACTERIZATION PERFORMED AS "LOW", "MEDIUM", OR "HIGH".

FoM	Description	FoM observability	Event stationarity	Driver influence	Actuarial relevance
Acceleration	Number of rapid acceleration events and their harshness	Medium	Low	High	Medium
Braking	Number of harsh braking events and their harshness	Medium	Low	Medium	High
Speeding (absolute)	Amount of absolute speeding	High	High	High	Medium
Speeding (relative)	Amount of speeding relative a location dependent limit	Medium*	High	High	High‡
Smoothness	Long-term speed variations around a nominal speed	High	High	Medium	Low
Swerving	Number of abrupt steering maneuvers and their harshness	Low†	Low	Medium	Low
Cornering	Number of events when turning at too high speed and their harshness	Medium	Medium	High	Medium
Eco-ness	Instantaneous or trip-based energy consumption or carbon footprint	Low	Medium	High	Low
Elapsed time	Time duration of the trip	High	High	Low	Low
Elapsed distance	Distance of the trip	High	High	Low	High
Time of day	Actual time of day when making the trip	High	High	Low	High
Location	Geographical location of the trip	High*	High	Low	Medium

† Not observable using only GNSS-receiver data. Fusion with inertial measurements is required.
‡ Given that the database with speed limits is sufficiently accurate.
* Digital map or database required.

the driver's speeding is likely to correlate with a higher risk of road accidents. Speeding can be defined and measured in several ways. In addition to the more general definition: "*The act of driving a vehicle faster than is allowed by law*", one could argue that it is more relevant from a risk perspective that the speed either should not exceed a certain absolute fixed threshold, a legal or recommended speed for a road segment, or a certain percentile of the speed distribution on a specific road segment. The deviation from the reference speed can be measured both as a percentage and an absolute number.

The speed of the vehicle is directly measured by the GNSS-receiver, and the observability of the absolute speeding FoM is thus considered high. However, the relative speed FoM requires the position of the vehicle to be mapped to a road-segment, via some map-matching algorithm, in order to get the current reference speed limit [10]. The observability of the relative speeding FoM is, due to the need of additional information, considered to be medium. The stationarity of speeding is considered high since the time window that must elapse before an act is considered speeding can almost surely be set sufficiently long. Since a driver can fully control the speed of the car, its influence-value is regarded as high.

3) Smoothness: The FoM describing the smoothness of the trip is a measure of the driver's skills in driving softly, anticipating the traffic, and keeping a constant speed. The FoM is typically calculated as the variance of the measured speed over a predefined time window; a time window chosen so that accelerations and decelerations between road-segments with different speed limits only marginally affect the FoM. Since the speed is directly measured by the GNSS-receiver and the smoothness FoM is based on the speed measurements from the whole trip, its observability and stationarity are considered to be high, and the sensitivity to data of poor quality is low because of the inherent averaging. Since the driver cannot fully predict the behavior of the other road users and the occurrence of traffic queues, etc., the influence is considered as medium.

4) Swerving: Swerving is defined as an abrupt steering maneuver in either direction which changes the deflection of the wheels at a high rate. The event can be detected when a driver attempts to steer away from an obstacle in the road or suddenly detects the risk of an involuntary lane change. This is often a sign of insufficient awareness of the road on the driver's part, which can be related to drowsiness or other distracting factors. However, there are of course also unforeseen situations which can require an immediate and resolute steering response from the driver, in order to avoid collision or other danger. Therefore, the driver's influence on this FoM is considered as medium.

The standard update rate of once per second of the smartphone GNSS-receiver data is generally too low for the detection of swerving events, due to their short duration. Instead, one must rely on data from accelerometers and gyroscopes. The gyroscopes are used to capture high-frequency steering events, while the accelerometer fixes the orientation of the gyroscope, which is needed since we are only interested in rotations around the yaw axis (perpendicular to the road surface). The observability of swerving events is constrained by the limited accuracy in both the angular rates given by the gyroscope and the estimation of the yaw axis given by the accelerometer data. To further complicate the solution is the requirement to reject sensor data when the smartphone is picked up by the driver or is not rigidly mounted in the car. While detections of impaired driving would be important in any driver assessment, the low observability presently limits the effect that swerving can have on the final score.

5) Cornering: Many insurers have requested methods for the detection of risky driving during cornering. Sharp turning maneuvers at high speeds are associated with several dangerous driving actions, and can lead to both car rollovers and skidding events, in which the driver completely loses control of the vehicle. Although several measures have been taken to improve the stability of historically rollover inclined vehicles such as SUVs, the influence and overrepresentation (as compared to all car crashes) of rollovers in car crash fatality statistics remains. Moreover, sharp turns are more often conducted in urban or other areas with a high population density, than on e.g., isolated freeways. In these areas, many pedestrians and other drivers can be assumed to be in the vicinity, and the speed should be kept low enough to give the driver sufficient reaction time.

Whether or not a cornering event is near to cause a rollover or skidding, is determined by comparing the magnitude of the horizontal forces acting on the vehicle, with the vehicle's stability coefficient and the tires friction coefficient, respectively. These coefficients are uniquely given by the characteristics of the vehicle, the tires, and the road's surface. The horizontal forces can be estimated using several information sources. Since the cornering events are relatively short (a single event does not last longer than a typical turning maneuver, and the stationarity is therefore classified as medium), it is close at hand to locally approximate the driving trajectory during an event with a circle arc. It is then possible to estimate the horizontal forces by applying Newtonian mechanics to circular motions, and the cornering event detection becomes a problem of estimating the vehicle velocity, acceleration (in the longitudinal direction) and angular velocity (or equivalently, the radius of the driving trajectory). While these quantities can be estimated using only GNSS-receiver data, the update rate will, just as in the case of swerving, limit how short cornering events that can be captured.

The driver-influence on cornering is classified as high, since there are few events in which high horizontal forces improve the driving safety or are necessary from a practical point of view.

6) Eco-ness: The eco-ness FoM is a measure of how energy efficient a driver is driving. Even though the eco-ness FoM may be of low importance in the calculation of the insurance premium, it may be an attractive feature to include in the feedback to the user. Since the smartphone can not directly measure the fuel consumption, eco-ness FoM calculations using non-intrusive instantaneous fuel consumption monitored by smartphone GNSS data is a good example of advanced model based signal processing, where a parameterized model of a combustion engine in combination with a vehicle energy balance model is utilized to predict the instantaneous energy consumption. The observability of the FoM is clearly low because of the heavy involvement of approximate models used to transform the speed information to a measure of the energy consumption. Typically, the predictions of relevance are on a minute scale, which through averaging eliminates the influence of random variations due to data imperfections and modeling imperfections. Accordingly, the stationarity can in most applications be considered to be on a medium level. A detailed presentation including an extensive evaluation using experimental data can be found in [13].

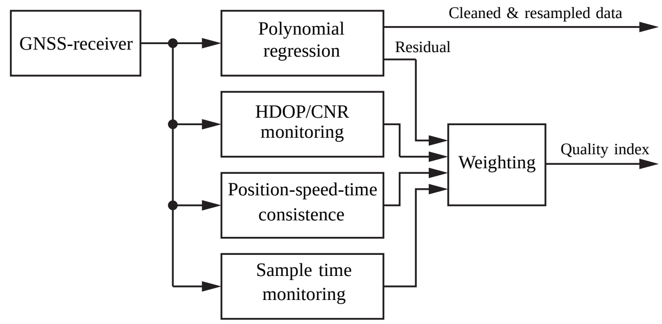

Fig. 3. Block diagram of a low-level signal processing algorithm that enhances the integrity of the GNSS-receiver data. It handles the possibly non-uniformly sampled GNSS-receiver data and provides uniformly resampled data as well as an estimate of the accuracy and reliability of the data, i.e., a quality index.

7) Elapsed time, distance, and time of day: The FoMs elapsed time, distance, and time of the day are defined as the total time the vehicle has been in motion, the total distance the vehicle has traveled, and the time of the day when the vehicle has been used, respectively. Since all of these FoMs are based on quantities that can be directly measured by the smartphone and the duration of the events are long, both the observability and the stationarity are classified as high. However, as the driver's daily-life needs dictate where and when the driver has to drive, the influence the driver has on these FoMs is considered to be low.

8) Location: The FoM location is defined as the location of the vehicle with respect to some of the insurer predefined geographical regions. For example, some areas of a city or an intersection may be associated with a high frequency of accidents, and the risk associated with driving in these areas is thus considered high. Even though the calculations of the FoM require the GNSS-receiver data to be mapped to a database of geographical regions, the regions are generally large and the mapping is easy. Further, the time spent in a region is often relatively long. The observability and stationarity are thus considered high. However, as the driver's daily-life needs dictate where and when the driver has to drive, the influence the driver has on these FoMs is considered low.

III. Integrity monitoring and enhancement

Inherent to all GNSS-receivers are their sensitivity to distortions and disturbances of the received satellite signals. If these distortions and disturbances, which may be due to, e.g., signal multipath, go undetected by the GNSS-receiver, the true accuracy of the calculated navigation solution and accuracy indicated by the GNSS-receiver may be in mismatch, causing overconfidence in the accuracy of the navigation data and a loss of data integrity. The loss of data integrity, which in data from smartphone GNSS-receivers often appears as undetected outliers, makes it difficult to reliably calculate some of the FoMs. Thus, to guarantee robust calculations of the different FoMs, additional data cleansing and integrity monitoring are needed. Such additional cleansing and integrity monitoring rely on additional information and constraints about the possible motion and behavior of the GNSS-receiver, and we will next describe two such methods. One that works on the second-by-second data, and one that works with the data from a whole trip.

A. Second-by-second data cleansing and integrity monitoring

The acceleration and deceleration of the car are used to calculate a plurality of the discussed FoMs, e.g., the harsh braking FoM. A direct differentiation of the GNSS-receivers speed data, i.e.,

$$\widehat{a}_k = \frac{s_k - s_{k-1}}{t_k - t_{k-1}} \tag{1}$$

is a typical example of a, so called, mimic hardware method for calculating the acceleration by the same method as used in black-boxes or OBD dongles. In (1), s_k is the speed at time instant t_k, and $\widehat{a}_k$ denotes the calculated acceleration. The differentiation method in (1) amplifies high frequency errors like noise and outliers, which in turn may cause false detections in a detector that signals the occurrence of a harsh braking event if the calculated deceleration falls below a threshold. In addition, the time difference $t_k - t_{k-1}$ is typically considered constant, since most GNSS-receivers calculate a navigation solution at a regular rate. (The update rate is typically once per second.) However, for some smartphone platforms there are significant variations in the update rate and in accordance the assumption of a fixed update rate will be another source of errors, or false detections.

In Fig. 3, a block diagram of a low-level signal processing data cleanser designed to overcome the problems with data outliers and a time varying data rate, is depicted. The main block of the cleanser is the polynomial regression block, in which the speed measurements are fitted to a polynomial model that locally describes the dynamics of the vehicle. From the fitted polynomial model, interpolated estimates of the vehicle's speed and accelerations can be calculated and by monitoring, the residual of the fitted data, outliers can be detected. To further check the data for erroneous data points, the consistence between the position, speed, and time stamps of the data; the variations in the data rate; and the dilution-of-precision (DOP) and carrier-to-noise (CNR) ratio of the GNSS-receiver data, is monitored. The monitored measures are weighted together into a new data quality index.

The performance of a harsh braking detector that uses (1) to calculate the speed derivative and a detection threshold of $-2\ m/s^2$ is illustrated in Tab. V; refer to [14] for details. Also illustrated is the performance of the detector that uses the data cleanser in Fig. 3 to estimate the speed derivative and remove outlier data. The results are based on the GNSS-receiver data recorded from seven, in the windshield cradled mounted iOS and Android smartphones, during a 90 minute drive. The GNSS coverage for the different phones spanned the interval 60.0-99.7%. As a gold standard reference, the harsh braking events detected using the speed read out from the vehicle's OBD-outlet, together with the differentiator in (1) were used. Because of the known quantization error of the OBD signal, the harsh braking threshold of $-2\ m/s^2$ was altered to include not only the nominal detector, but also the zero-missed-detection (a slightly higher threshold) and zero-false-alarm (a slightly lower threshold).

TABLE V
HARSH BRAKING DETECTION PERFORMANCE USING (1) AND THE DATA CLEANSER IN FIG. 3 TO ESTIMATE THE DECCELERATION OF THE VEHICLE. THE SMARTPHONE STATISTICS ARE BASED ON THE RESULTS OF 7 SMARTPHONES USED FOR THE TEST, WITH A GNSS-COVERAGE IN THE INTERVAL $60.0 - 99.7\%$. THE INDIVIDUAL DETECTION RESULTS HAVE BEEN SCALED WITH THE GNSS COVERAGE TO COMPENSATE FOR THE LOSS OF DATA. THE ESTIMATED STANDARD DEVIATION σ OF THE TOTAL NUMBER OF DETECTIONS IS GIVEN WITHIN PARENTHESES.

Gold standard reference				
Detector	Correct	Indeterminable	False	Total
Zero-false-alarm	3	0	0	3
Nominal	3	3	0	6
Zero-missed-detection	3	15	0	18
Average over 7 smartphones running in parallel				
Data type	Correct	Indeterminable	False	Total (σ)
Raw data	3.1	10.0	30.2	43.3 (11.7)
Cleansed data*	2.2	3.6	0.8	6.6 (1.7)

* Data cleansed using the cleanser in Fig. 3.

From Tab. V, one can conclude that the direct differentiation of the smartphone-data provides a large number of false detections of the harsh braking events. A proper low-level data cleansing is shown to remove the false detections and provide detection results close to the nominal detector that is using OBD speed data. In particular, the detection results using enhanced data are well centered in the interval determined by the number of detections provided by the OBD-based zero-false-alarm and zero-missed-detection detectors, respectively. As a guideline for the requirement on outlier detection of the data cleansing, one may note that harsh braking events typically occur once per 100 km. Thus, the harsh braking detector must have a false-alarm rate which is a magnitude less frequent.

B. Trip-based data cleansing and integrity monitoring

To calculate the relative speeding FoM, the GNSS-receiver measured speed must be compared to the speed limit of the road-segment on which the vehicle currently is traveling on. Hence, the data from the GNSS-receiver must be mapped to a database of the road-segments in the road-network, a process referred to as map-matching. A survey on map-matching technologies can be found in [15].

Clearly, the accuracy of the calculated relative speeding FoM will depend both on the reliability of the map-matching process and the accuracy of the speed limits stored in the database. The accuracy of the speed limits stored in the database is mainly dependent on how frequently the map vendor updates their map database, but the reliability of the map-matching can be enhanced, as will be illustrated next through trip-based data processing.

To illustrate how the reliability of the relative speeding FoM can be increased through trip-based data processing, the

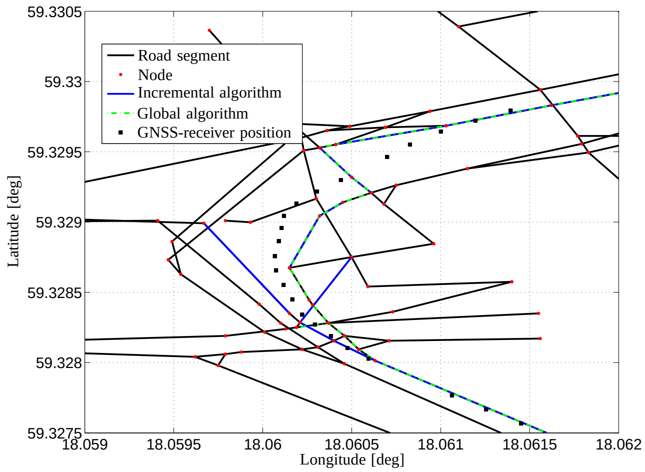

Fig. 4. Illustration of how trip-based map-matching can be used to improve the robustness of the relative speeding FoM calculations. The global map-matching algorithm, which uses the data from the whole trip, maps only one road-segment erroneously, whereas the incremental algorithm maps five road-segments erroneously.

number of correctly matched road-segments using a global and incremental map-matching algorithm is compared in Fig. 4. The global algorithm does the map-matching off-line and uses the data from the whole trajectory, whereas the incremental algorithm does the map-matching in real-time and thus only uses past data to determine the current location of the vehicle in the road-network; the algorithms use the same geometrical and topological information. From Fig. 4 it is clear that by using a global (off-line) map-matching algorithm, the robustness of the map-matching, and the relative speeding FoM calculations, can be be greatly enhanced.

IV. Scoring and its robustness

FoMs for insurance telematics are typically calculated over a trip, eventually differentiated into categories like speeding in urban areas, speeding on highways, daytime speeding, etc. Combining, possibly categorized, FoMs continuously collected over the policy period or collected during a qualification period into a measure of the actual risk for different kinds of insurance claims is clearly an intricate problem. By necessity, the description below is simplified and adopted to the purpose of the paper.

A. Introduction to scoring

The scoring is defined as the process

$$\mathcal{F}:\ \{f_1,\dots,f_N\}\rightarrow\mathcal{S}, \tag{2}$$

where f_n denotes the n:th FoM, and $\mathcal{S}$ defines the resulting score. For the sake of the discussion, we will consider a scalar score $\mathcal{S}\in[0,\,1]$, where a higher value of $\mathcal{S}$ indicates a safer driver than a lower value.

Often asked for is an explicit rule $\mathcal{F}$. However, it is clear that no such unique rule exists, but rather, it is dependent upon the considered customer segment, traffic situation, or particular risk models employed by the insurer. The way ahead to determine a suitable $\mathcal{F}$ is based on doing by learning, which requires pilots such as the *If SafeDrive* in [7].

B. Scoring based on number of harsh braking events

Let f denote the number of detected harsh brakings per unit time, that is, the actual number of detected harsh brakings over a trip divided by the elapsed time of the trip. Further, consider the scoring

$$\mathcal{S}=\frac{1}{1+\alpha f} \tag{3}$$

where α is a constant to be determined. The scalar α could for instance be determined by letting the average number of harsh brakings per unit time, f_A, for some given customer segment and regional location, result in some predetermined score, $\mathcal{S}_A$ (say $\mathcal{S}_A=1/2$). This means that we set $\alpha=(1-\mathcal{S}_A)/(\mathcal{S}_A f_A)$. Note, that no triggered events ($f=0$) give a score of $\mathcal{S}=1$, while $\mathcal{S}$ approaches zero as the number of detections per unit time grows.

Now consider a driver in the above mentioned population. As we are primarily interested in deviations in the scoring due to imperfect measurements, and not deviations due to varying driving or trip characteristics, we assume that the driver's actual number of braking events per unit time is exactly μ for each trip. Since the scoring function $\mathcal{F}$ is strictly convex for all relevant parameter values, a loose application of Jensen's inequality shows that the scoring is biased in the sense that the expectation of the obtained scoring $\mathcal{S}_{\text{Obtained}}$ exceeds, due to the random fluctuations inherent in the measurements, the actual scoring $\mathcal{S}_{\text{Actual}}$ as

$$\mathcal{S}_{\text{Actual}}=\frac{1}{1+\alpha\mu}<E\left[\frac{1}{1+\alpha f}\right]=\mathcal{S}_{\text{Obtained}}, \tag{4}$$

where $E[\cdot]$ denotes statistical expectation operator.

Now assume that we would like to classify each trip as either *good*, *acceptable* or *poor* (in terms of safety, or low risk) by quantizing $\mathcal{S}$. For example, the limits can be chosen so that the outcomes of $\mathcal{S}$ are divided into three equally large intervals. We then arrive at the classification system

$$\mathcal{C}(\mathcal{S})=\begin{cases}\textit{good}, & \mathcal{S}>2/3\\ \textit{acceptable}, & 1/3<\mathcal{S}\leq 2/3\\ \textit{poor}, & \mathcal{S}\leq 1/3.\end{cases} \tag{5}$$

Given some distribution for f, it is at this point easy to calculate the probability of different classifications, or the risk of an erroneous classification, e.g., that a driver with $\mathcal{S}_{\text{Actual}}<1/3$ has a trip erroneously classified as *acceptable* or *good*. We will exemplify this by studying Tab. V, which shows that after proper applied low level digital signal processing on the smartphone data, the number of detected harsh brakings in a typical trip has only a negligible bias and a standard deviation $\sigma\approx 0.25\mu$, where μ is the actual number of harsh brakings. By assuming that this $(\mu,\,\sigma)$-relation holds in general, we can in a first attempt model f as a Gaussian stochastic variable, i.e. $f\sim\mathcal{N}(\mu,\,(0.25\mu)^2)$. Now, assuming that $f_A=6$, $\mathcal{S}_A=1/2$ and the scalar $\alpha=(1-\mathcal{S}_A)/(\mathcal{S}_A f_A)=1/6$, a good driver with half the number of harsh brakings compared to the average of the considered population has $\mu=f_A/2=3$ (which means that the driver is exactly on the limit between *good* and *acceptable* in the sense that $1/(1+\alpha\mu)=2/3$) and will be classified as *good* or *acceptable* with a probability of

$1/2$ each. Similarly, a poor driver on with $\mu = 2f_A = 12$ will to a very good precision be classified as *acceptable* or *poor* with a probability of $1/2$ each.

C. Design of scoring

In the previous example, the scoring was calculated using a single FoM. In practice, at least a handful of different FoMs is combined into a score $\mathcal{S}$. For a *valid* scoring, the actuarial importance of the FoMs has to be taken into account, i.e., a scoring should result in a risk profile that correlates with the actual risk of the driver. The detailed characterization of the FoMs in terms of actuarial importance is of the discretion of the underwriters at the insurance companies, based on analysis of their databases with claim statistics.

Given a set of sufficiently important FOMs, the scoring process (2) has to be designed, now taking into account the effects of the reliability (that is the combination of integrity and availability) of the gathered sensor data and its influence onto the observability of the FoMs, as well as the physical meaning of the FoMs as given by their time duration, i.e., their stationarity. Sensor data integrity should in general be enhanced to reliably calculate FOMs which are transient in nature and rely on measurements with weak correlation to the measurand. Additional uncertainty has to be taken into account as well, like uncertainty originating from fusion of data from a plurality of sensors; or originating because of the use of external databases like digital maps. It is an intricate problem where the robustness of the obtained $\mathcal{F}$ has to be studied with respect to the uncertainties introduced in the chain from sensor to score.

In summary, the process of designing *reliable* scoring is an engineering trade-off between actuarial importance of the FoMs and their corresponding properties in terms of FoM observability and event stationarity. The *validity* of the scoring is then an actuarial exercise where success relies on reliable data.

D. Robustness and perception of driver feedback

We end this section with a discussion on the robustness and perception of the feedback provided to the driver.

An example (taken from the *If SafeDrive* campaign reported in [7]) of after-trip feedback is to provide the driver with a map indicating the locations of harsh braking events as a means to provide constructive feedback of how their driving performance can be improved. As discussed, a harsh braking event is an event of short time duration; medium correlation between sensor data and an event to be detected; and is also a quite rare event as such. A typical trip usually contains none, or a very small number of actual harsh braking events. Because of the lacking availability of data (recall the span of GNSS-coverage reported in Tab. V), actual events may be missed with a quite high probability. In addition, the lack of reliability (relative the requirements) may introduce some false indications. The effect is that the driver most probably reacts in a negative manner if the provided feedback does not comply with the driver's actual driving pattern during the trip. A harsh braking is typically an event that you remember as a driver, both as an event as well as the location of it. In summary, the feedback is not robust enough with respect to false or missed detections and should be avoided because it reduces the overall trust in the UBI program.

Now, consider the following scoring

$$\mathcal{S} = \frac{1}{3}\left(\frac{1}{1+f_1} + f_2 + f_3\right), \tag{6}$$

where f_1 denotes the number of harsh braking events; f_2 the time of day, with $f_2 = 1$ during daytime and $f_2 = 0$ during nighttime; and f_3 the location, with $f_3 = 1$ for very safe locations and $f_3 = 0$ for roads with historically a very high number of accidents, and that f_3 is available with some kind of granularity in the interval $[0, 1]$. From the previous discussions, it is clear that the scoring (6) reflects the driver's risk and is a suitable score for calculating an insurance premium.

From a perception of scoring, however, it may be considered counterintuitive. Consider an extremely rough ride ($f_1 \to \infty$) during daytime ($f_2 = 1$), and also a smooth ride ($f_1 = 0$) during nighttime ($f_2 = 0$). Both trips result in the score

$$\mathcal{S} = \frac{1+f_3}{3}, \tag{7}$$

where, for these trips, $\mathcal{S} \in [1/3, 2/3]$. Now, considering the classification system in (5) (where the region classified as *acceptable* is slightly shrunk, to be mathematically stringent), that is with transition levels $1/3+\delta$ and $2/3-\delta$ for some positive δ close to zero, respectively. We note that both trips most likely are classified as *acceptable*, which as such may be counterintuitive for the driver – expecting a classification of *poor* for the former ride, and *good* for the latter. Clearly, the obtained scoring is not only dependent on the number of harsh braking events, but also due to the time of day. Even worse, taking the location f_3 into account, a smooth (and perceived safe) ride may in fact be classified as *poor* as well as a rough ride may be classified as *good* – a result that clearly by the driver may be perceived as counterintuitive.

Results like the ones obtained in the example above are typically considered to be counterintuitive by the driver, and thus will influence the driver's opinion about the reliability of the UBI program. Here, as classified in Tab. IV, the time of day f_2 and location f_3 are FoMs with a low influence factor. Typically, the driver's opinion on the level of risk correlates more with the driver's own behavior, rather than on parameters that typically cannot be influenced, such as the time of day and location of the trip.

V. Conclusions

Usage based insurance using smartphones as measurement probes is a good example of an application where the information and communication technologies (ICT) enable new applications for intelligent transportation systems (ITS). The ubiquitous smartphone enables commercial insurance telematics here and now, and bridges the gap to future UBI programs based on OEM installed devices following (yet to be agreed upon) standards. The smartphone offers a scalable solution where the driver downloads a UBI-application, after which the driver may sign-up with a program, obtaining an insurance fee that is determined by dynamic measures describing when, where and how the trip has been performed. This technology has the capability to disrupt the contemporary and existing business models in the insurance industry. As outlined in this paper, it also affects how the process of risk assessment is conducted. Other core business processes in car insurance, such as the sales process can be radically innovated with this technology, getting more points of interaction with the customer, and enabling new relations with the customers.

The paper has discussed the technical challenges and highlighted the main obstacles for successful UBI programs in terms of sensor information quality with respect to integrity and availability; figure of merits describing driver behavior; and the forming of a scoring procedure describing the risk profile based on the figure of merits of actuarial relevance. The challenges on the higher system levels when providing sensing as a service, and challenges related to the provided incentives to the insurance customers, and related to the insurer's business model are relevant, but are also beyond the scope of this paper. Discussion around these topics can be found in [7].

In the paper, we have tried to highlight and discuss the challenges involved in transforming a UBI program from a program that uses tailored hardware electrically connected to the vehicle, to a program based on a stand-alone smartphone as a measurement probe and communication device. With the introduction of the smartphone as the in-vehicle device, the possibilities to feed back information to the driver increase tremendously, thanks to the high resolution screen, audio-visual means, but also the easy access to remotely located database information. We have exemplified the risk of providing the driver with information that reflects the risk profile, because it may be perceived in the wrong way. Accordingly, the design of driver feedback has to be considered seriously in the design of any smartphone-based UBI system.

References

[1] Y. Zhao, "Telematics: safe and fun driving," *IEEE Intelligent Systems*, vol. 17, pp. 10–14, Jan. 2002.

[2] "Insurance telematics report 2013." Telematics update, 2013.

[3] Y. Zhao, "Mobile phone location determination and its impact on intelligent transportation systems," *IEEE Trans. Intell. Transp. Sys.,*, vol. 1, pp. 55–64, Mar. 2000.

[4] F. Calabrese, M. Colonna, P. Lovisolo, D. Parata, and C. Ratti, "Real-time urban monitoring using cell phones: A case study in rome," *IEEE Trans. Intell. Transp. Sys.,*, vol. 12, pp. 141–151, Mar. 2011.

[5] N. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. Campbell, "A survey of mobile phone sensing," *IEEE Communications Magazine*, vol. 48, pp. 140–150, Sept. 2010.

[6] M. Trivedi, T. Gandhi, and J. McCall, "Looking-in and looking-out of a vehicle: Computer-vision-based enhanced vehicle safety," *IEEE Trans. Intell. Transp. Sys.,*, vol. 8, pp. 108–120, Mar. 2007.

[7] P. Händel, J. Ohlsson, M. Ohlsson, I. Skog, and E. Nygren, "Smartphone-based measurement systems for road vehicle traffic monitoring and usage-based insurance," *IEEE Systems Journal*, Mar. 2014.

[8] K. Jasper, S. Miller, C. Armstrong, and G. Golembiewski, "National Evaluation of the SafeTrip-21 Initiative: California Connected Traveler Test Bed Final Evaluation Report: Mobile Millennium," tech. rep., United States Department of Transportation, ITS Joint Program Office, January 2011.

[9] I. Skog and P. Händel, "In-car positioning and navigation technologies – A survey," *IEEE Trans. Intell. Transp. Sys.,*, vol. 10, pp. 4–21, Mar. 2009.

[10] A. E. Eskandarian, *Handbook of Intelligent Vehicles*. Springer Reference, 1:st ed., 2012.

[11] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: rich monitoring of road and traffic conditions using mobile smartphones," in *Proc. of the 6th ACM conference on Embedded network sensor systems*, SenSys '08, (New York, NY, USA), pp. 323–336, 2008.

[12] M. Fazeen, B. Gozick, R. Dantu, M. Bhukhiya, and M. Gonzalez, "Safe driving using mobile phones," *IEEE Trans. Intell. Transp. Sys.,*, vol. 13, pp. 1462–1468, Sept. 2012.

[13] I. Skog and P. Händel, "Indirect instantaneous car-fuel consumption measurements," *IEEE Trans. Instrum. Meas*, 2014, accepted.

[14] I. Skog, P. Händel, M. Ohlsson, and J. Ohlsson, "Challenges in smartphone-driven usage based insurance," in *Proce. of IEEE Global Conference on Signal and Information Processing*, (Austin, TX, USA), Dec. 2013.

[15] M. A. Quddus, W. Y. Ochieng, and R. B. Noland, "Current map-matching algorithms for transport applications: State-of-the art and future research directions," *Transportation Research Part C: Emerging Technologies*, vol. 15, pp. 312 – 328, Oct. 2007.