Personalized Driving Behavior Monitoring and Analysis for Emerging Hybrid Vehicles

Kun Li[1], Man Lu[1], Fenglong Lu[1],
Qin Lv[2], Li Shang[1], and Dragan Maksimovic[1]

[1] Department of Electrical, Computer, and Energy Engineering,
[2] Department of Computer Science
University of Colorado Boulder, CO 80309 USA
{kun.li,man.lu,fenglong.lu,qin.lv,li.shang,maksimov}@colorado.edu

Abstract. Emerging electric-drive vehicles, such as hybrid electric vehicles (HEVs) and plug-in HEVs (PHEVs), hold the potential for substantial reduction of fuel consumption and greenhouse gas emissions. *User driving behavior*, which varies from person to person, can significantly affect (P)HEV operation and the corresponding energy and environmental impacts. Although some studies exist that investigate vehicle performance under different driving behaviors, either directed by vehicle manufacturers or via on-board diagnostic (OBD) devices, they are typically vehicle-specific and require extra device/effort. Moreover, there is no or very limited feedback to an individual driver regarding how his/her personalized driving behavior affects (P)HEV performance.
This paper presents a personalized driving behavior monitoring and analysis system for emerging hybrid vehicles. Our design is *fully automated and non-intrusive*. We propose *phone-based multi-modality sensing* that captures precise driver–vehicle information through de-noise, calibration, synchronization, and disorientation compensation. We also provide *quantitative driver-specific (P)HEV analysis* through operation mode classification, energy use and fuel use modeling. The proposed system has been deployed and evaluated with real-world user studies. System evaluation demonstrates highly-accurate (0.88-0.996 correlation and low error) driving behavior sensing, mode classification, energy use and fuel use modeling.

1 Introduction

Energy use for transportation represents a pressing challenge, due to the heavy and growing reliance on petroleum and the environmental impacts of emissions from fossil fuel combustion. Recent studies have shown that *transportation electrification*, such as (P)HEVs, holds the potential to significantly reduce greenhouse gas emissions and the ever-growing dependence on oil [22]. (P)HEVs integrate an internal combustion engine (ICE) powered by gasoline with an electric motor powered by the battery system. Active market penetration of HEVs, such as Toyota Prius and Ford Escape, has been observed in recent years. Automobile manufacturers are poised to introduce new models of PHEVs and their market penetration is expected to increase rapidly in the coming years.

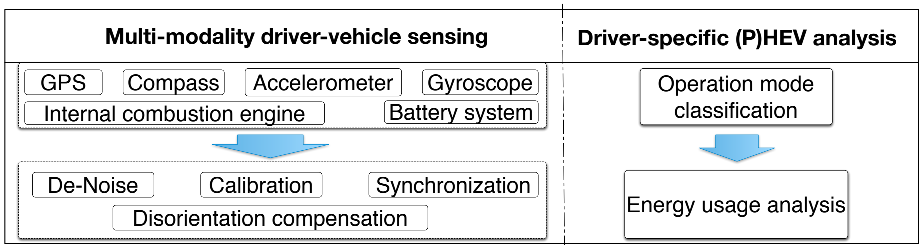

Fig. 1. Personalized driving behavior monitoring and analysis for emerging (P)HEVs.

Advancing (P)HEV technology and promoting its market adoption require *comprehensive and quantitative investigations of how user-specific driving behavior affects run-time (P)HEV operation, which in turn determines the energy and environmental impacts of (P)HEVs*. To date, little is known regarding the relationships between user-specific driving behavior and (P)HEV performance. Previous studies on vehicle fuel economy and emissions were typically tested or modeled over a limited number of driving cycles designed to represent "typical" driving profiles [3]. These standard cycles ignore the diverse driving behaviors among users, which strongly influence vehicle fuel consumption and emissions [17]. Furthermore, previous studies have focused on conventional gasoline vehicles, which work very differently than emerging (P)HEVs [7].

Automobile manufacturers have recently started to consider user-specific driving behavior. Toyota, for example, conducted a "world's first large scale" user study with around 100 PHEVs in September 2007, which aimed to understand how the new technology performs under real-world user driving behaviors [2]. While such manufacturer conducted user studies can be well organized and detailed data can be easily collected, they are limited in that: (1) Little feedback is provided to individual drivers, even with dashboard display. MPG (miles per gallon) alone can be misleading for (P)HEVs. It is not clear how specific driving behavior translates to (P)HEV energy and environmental impacts. (2) Building personalized analysis software directly into the vehicles would require support and adoption by automobile manufacturers, and such systems would be vehicle-specific and significantly limited in ubiquity.

Feedback of vehicle performance is also possible through the on-board diagnostic (OBD-II) interface. Still, the format of vehicular sensor readings differ by manufacturer and vehicle, and the readings do not translate directly to (P)HEV energy and environmental impacts. In addition, this approach requires special hardware like OBD scanner and extra effort from user to connect and pair the device each time, which limit its applicability to the general public.

In this work, our goal is to bridge this information gap through *a personalized driving behavior monitoring and analysis system* (Figure 1). The system is designed to be *fully automated and non-intrusive*. It is based on smartphones, which are widely available nowadays. A user only needs to download the phone application and have it running during a trip, and personalized quantitative reports will be generated given his/her specific driving behavior. Such simplicity and functionality make it attractive for most drivers. Even drivers who do not currently own a (P)HEV could use this application to predict the energy and environmental impacts if they switch to (P)HEVs. Moreover, the aggregated data from a large number of drivers can be invaluable for vehicle manufacturers and researchers as they keep improving the design of (P)HEVs.

A number of sensing capabilities are available in modern smartphones, making it possible for personalized driving behavior monitoring and analysis. However, signals obtained from individual sensors suffer from a number of intrinsic and contextual issues: (1) *noise*, *drift*, and *mis-synchronization* of multi-modality driver–vehicle data; (2) substantial *disorientation* of vehicle movement sensing at the start of and during each driving trip; and (3) complex (P)HEV *operation mechanisms* and driver-dependent vehicle performance.

To address these issues, we propose innovative techniques that "fuse" together multiple types of sensor readings to enhance signal quality and driver–vehicle sensing, as well as techniques that classify driver-specific operation modes and analyze the corresponding energy and environmental impacts. Our work makes the following contributions:

(1) **Multi-modality driver–vehicle sensing.** We propose fully-automated, phone-based sensing techniques that effectively correct the noise, drift, mis-synchronization in multiple types of sensor data, as well as compensating initial/dynamic phone–vehicle disorientation using wavelet-based analysis.

(2) **Driver-specific (P)HEV analysis.** We propose operation mode classification, run-time energy use and fuel–CO_2 emission modeling, to map specific driving behavior to (P)HEV energy and environmental impacts; and

(3) **Real-world deployment and user driving studies.** System evaluation demonstrates high correlation for vehicle sensing (0.88-0.96), energy use and fuel use modeling (0.918, 0.996), operation mode classification (89.9%, 87.8% accuracy).

The rest of the paper is organized as follows. Section 2 presents the problem formulation and design overview. Section 3 and Section 4 describe in detail multi-modality driver-vehicle sensing and driver-specific (P)HEV analysis. Section 5 describes system deployment and user studies. Section 6 evaluates the proposed system. Section 7 discusses related work. Section 8 concludes.

2 Problem Formulation and Design Overview

HEVs and PHEVs are the emerging solutions for transportation electrification. HEVs feature a gasoline internal combustion engine (ICE) and an electric motor equipped with a battery system for harnessing and storing run-time braking energy. PHEVs have an additional electrical plug to directly recharge the battery system from the electrical grid. The energy and environmental impacts of (P)HEVs are primarily determined by (P)HEV operation, which in turn is heavily affected by user-specific run-time driving behavior, such as speed, acceleration, and road condition, etc. Figure 2 shows the fuel use, CO_2 emissions, and battery system long-term capacity degradation based on eight different users'

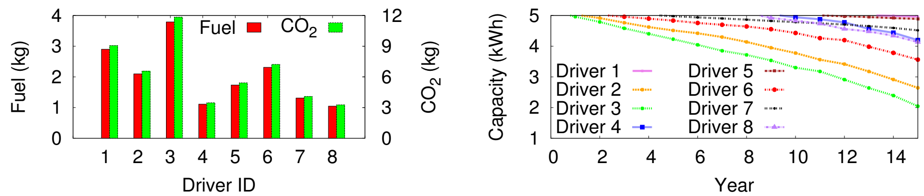

Fig. 2. Heterogeneous fuel use, CO_2 emissions, and battery system long-term capacity degradation based on eight different users' daily commute driving profiles.

daily commute driving profiles. Among the eight drivers, over 3× variation is observed in terms of fuel use and CO_2 emissions for battery system, and based on the system-level battery model developed by Li et al. [16], over 9× long-term capacity variation can be expected.

A comprehensive and quantitative sensing and analysis system is thus essential for advancing (P)HEV technology and promoting its market adoption by individual drivers. We propose a personalized multi-modality sensing and analysis system that effectively captures and fuses the following signals: (1) **user-specific driving behavior**, including speed, acceleration, road and traffic conditions; and (2) **(P)HEV operation profile**, including fuel use, battery system charge/discharge current and voltage.

Accurate characterization and quantification of the relationships between user-specific driving behavior and (P)HEV energy and environmental impacts require fine-grained, time-stamped, robust sensor readings during users' driving trips, as well as accurate modeling of (P)HEV internal operation mechanisms. We propose a two-staged process, as illustrated in Figure 1.

1. Multi-modality driver–vehicle sensing: The first stage captures and enhances the quality of multiple types of sensor data using novel de-noise, calibration, and synchronization techniques. It then automatically identifies potential phone–vehicle disorientation and compensates the corresponding sensor readings at run-time for accurate vehicle movement sensing.

2. Driver-specific (P)HEV analysis: In the second stage, we propose to first map users' driving behaviors to the corresponding (P)HEV operation modes via an effective mode classifier. Then, leveraging battery system modeling and fuel–CO_2 emission modeling, we quantitatively analyze the energy and environmental impacts of (P)HEVs under specific user driving behavior.

3 Multi-modality Driver–Vehicle Sensing

Using multiple types of sensors that are readily available on mobile phones, we propose techniques to obtain high-quality sensor data for both user driving behavior and vehicle movement. The challenge is to achieve personalized, accurate, and run-time data acquisition with minimum inconvenience and obstruction to individual drivers. Specifically, we allow the phone to be unrestricted in the vehicle (i.e., not mounted in a fixed position), yet still effectively identify and remove noises and inconsistencies in multiple types of sensor data, as well as compensating for potential phone-vehicle disorientation.

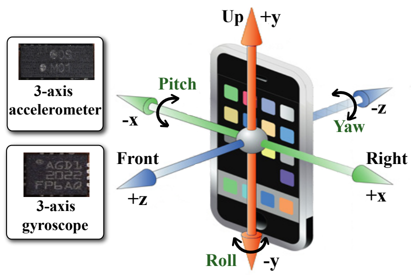

Fig. 3. Linear & angular acceleration sensing using accelerometer and gyroscope.

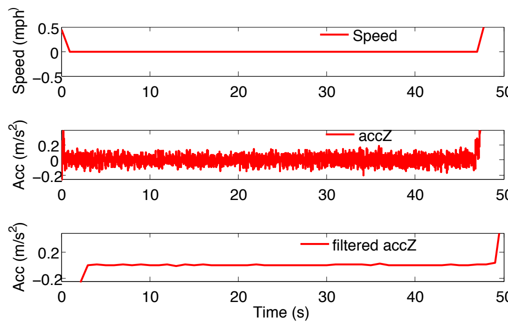

Fig. 4. Acceleration noise when vehicle is stationary; de-noise via low-pass filtering.

3.1 User Driving Behavior Sensing

A user's specific driving behavior can be represented by his/her driving trips with regard to speed, acceleration, slope, and turning at individual time points. Specifically, **speed** is directly reported by GPS; **slope** of the road can be calculated from altitude reported by GPS; **acceleration** is reported by accelerometer but requires further compensation by gyroscope; and **vehicle turning** information is derived from gyroscope readings and calibrated by digital compass. As illustrated in Figure 3, accelerometer and gyroscope can be used to sense linear and angular acceleration. By combining the readings of both sensors, detailed information about the device's six-axis movement in space can be derived.

Noise of sensor is a major technical barrier to precise sensing. There are two primary kinds of noise sources. One is *intrinsic* high frequency noise due to the combined effects of thermally dependent electrical and mechanical noise [21]. The other is *contextual* noise caused by vehicle vibration during a trip, whose frequency usually peaks at around 3-5Hz. For example (Figure 4), acceleration readings ranging from -0.2 to $0.2m/s^2$ are reported even when the vehicle is stationary (speed = 0). A low-pass filter with 2Hz cutoff frequency is applied to improve the signal to noise ratio. The noise characteristics of gyroscope are more complex, which we address later in this section.

Another source of error in motion sensing is the *drift* of sensor. In particular, calculating the angular position using gyroscope requires integration of noisy angle change rate readings, which accumulates over time and results in large drift. Such drift can be potentially compensated by periodically resetting the gyroscope to the known directional source: gravity, which is collected whenever the phone is determined to be stationary, e.g., when the vehicle is stopped for traffic light. The acceleration vector of gravity is parallel to the Yaw-axis of the reference coordinate system, which requires special calibration with the digital compass. Figure 5 shows an experiment in which the phone was returned to the original position after 14 minutes of driving. We can see that digital com-

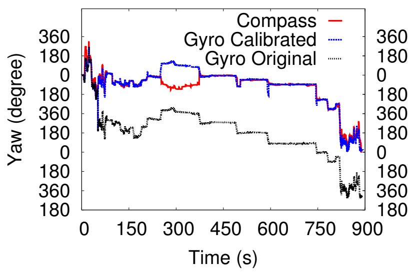

Fig. 5. Degree of drift: Digital compass vs. gyroscope (original and calibrated).

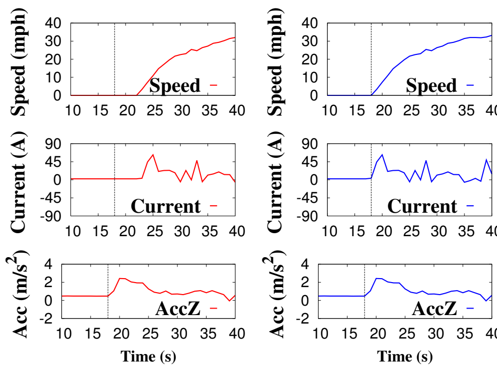

Fig. 6. Unsynchronized (left) vs. synchronized (right) signals.

pass has much less drift than gyroscope. In our system, the gyroscope value is calibrated whenever the compass reading is steady for a period of time.

Using OBD devices, information regarding the (P)HEV operations can also be collected, including speed, steering, battery system charge/discharge and fuel use. Such information is used as ground truth in our evaluations. One major challenge lies in appropriate *synchronization* of multi-modality data from three different data sources: mobile phone, internal combustion engine, and battery system. As demonstrated in Figure 6, when the vehicle speeds up from stationary status, both current and acceleration should change to non-zero values at the same time, but they did not in the raw monitored data. We propose to synchronize the multi-modality data streams by maximizing the correlation between each data stream and the reference data stream:

$$\mathrm{corr}(A, B) = \frac{E[(A - \mu_A)(B - \mu_B)]}{\sigma_A \sigma_B} \tag{1}$$

We select acceleration as the reference stream since its noise can be effectively removed and is more precise than others. We also exploit stops (based on speed and acceleration to segment each trip into several sub-trips and apply synchronization to each sub-trip.

3.2 Vehicle Movement Sensing

Although it is possible to obtain vehicle movement information by mounting a phone in a fixed position before each trip, it is cumbersome and inconvenient for users. Our solution allows the phone to be unrestricted in the vehicle. This leads to potential phone–vehicle disorientation. Let (X, Y, Z) be the vehicle's Cartesian frame of reference, and (x, y, z) be the phone's frame of reference. As illustrated in Figure 7, these two frames of reference should be well-oriented in the ideal case. However, the phone may be placed anywhere in the vehicle (e.g., driver's pocket) at trip start (*initial disorientation*), and may shift around during a driving trip (*dynamic disorientation*). Such disorientation can be substantial and highly dynamic, as shown in Figure 8.

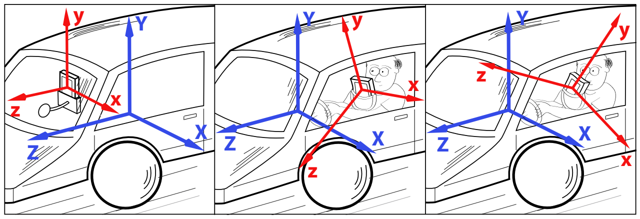

Fig. 7. Frame of reference orientation of vehicle (blue) and phone (red). Left: oriented; Middle: initial disorientation; Right: dynamic disorientation.

Initial Disorientation Compensation. The phone's (x, y, z) axes are disoriented with respect to the vehicle's (X, Y, Z) axes. Leveraging the automatic attitude initialization method originally proposed by Mohan et al. [20], the rotation matrix needs to be calculated and applied to x, y, and z in sequence in order to transform arbitrary orientation of (x, y, z) to (X, Y, Z). According to the definition of Euler Angle [1], any orientation of the accelerometer can be represented by a pre-rotation ϕ of Y, followed by a tilt θ of Z, and then a post-rotation α of Y. Thus, the rotation matrices associated with these three rotation angles are:

$$R_\phi = \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix}, R_\theta = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, R_\alpha = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix} \tag{2}$$

ϕ and θ can be calculated by applying the rotation matrix to the accelerometer reading $[a_x, a_y, a_z]^T$ when the vehicle is stationary, i.e., gravity $[0, 1, 0]^T$:

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = R_\theta * R_\phi * \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix} * \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \tag{3}$$

Thus, $\theta = \cos^{-1}(a_y), \phi = \tan^{-1}(a_z/z_x)$.

Next, we need to estimate the post-rotation α of Y. When a trip starts, the vehicle goes from stationary to acceleration, producing a force in a known direction. Suppose that after rotation of ϕ and θ, the acceleration vector $[a_x, a_y, a_z]^T$ changes to $[a'_z, a'_y, a'_z]^T$, reflecting the force produced. While in the vehicle's coordinate system, only a_Z has a significant value and a_X should be 0. So the summation vector of a_x and a_z is exactly a_Z, and we get $\alpha = \tan^{-1}(a'_x/a'_z)$. In the end, we get a rotation matrix $R = R_\alpha * R_\theta * R_\phi$, which can transform accelerometer readings of the mobile phone to the vehicle's true accelerations. Note that this approach also works when car starts on a non-flat surface.

Dynamic Disorientation Compensation. Dynamic disorientation is highly unpredictable and may occur at anytime during a trip. If not detected and corrected in real time, it introduces significant error to the sensing value, as illustrated in Figure 8. Thus, it needs to be compensated before the measured acceleration values can be used. We propose a wavelet-based technique, which analyzes the rotation information received from the gyroscope in order to separate true vehicle movement from contextual noise. This technique is based on multi-resolution analysis [26], by which a given time-series signal can be decomposed into multiple wavelet levels, each corresponding to a specific frequency range.

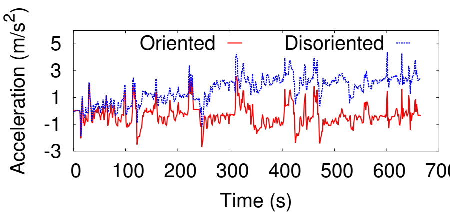

Fig. 8. Error of dynamic disorientation.

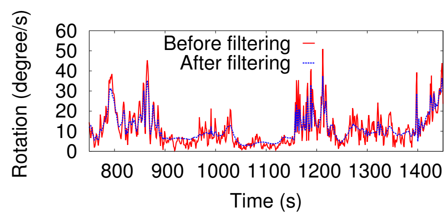

Fig. 9. Gyroscope de-noise using wavelets.

One important observation is that *vehicle movements and phone movements have different characteristics and tend to occur in different time–frequency domains*, e.g., vehicle changing lane vs. phone moving in driver's pocket. Wavelets are particularly suitable for such joint time–frequency analysis.

Wavelet-based gyroscope noise filtering. The gyroscope signal has unique noise characteristics. True rotation can appear in both high frequency and low frequency, and signal noise cannot be removed via simple low/high-pass filtering. We leverage a wavelet-based de-noise method [23] to improve the signal quality from gyroscope. Let $s(n) = f(n) + \sigma e(n)$ be the raw gyroscope signal, where n is the index of equally-spaced time points, $f(n)$ is the true signal and $e(n)$ models the noise. We assume $e(n)$ is a Gaussian white noise and the noise level σ equals to 1. We make the following observations: (i) Signal from gyroscope should be mostly smooth, with a few abrupt changes caused by vehicle turning or phone's sudden movement. Therefore, it should have only a few non-zero wavelet coefficients. (ii) A white noise signal is reflected by the coefficients at all levels. Also, a Gaussian noise after orthogonal wavelet transform still preserves the Gaussian property. Thus, noise can be estimated by removing the correlated signal at each level. Thus, for each level of decomposed signal, we select a threshold (0.2506 based on our experiments) and ignore high-frequency coefficients whose absolute values are lower than the threshold. We then reconstruct the signal using all remaining coefficients. As we can see in Figure 9, the gyroscope signal becomes much smoother after wavelet-based de-noise.

Wavelet-based movement analysis. Figure 10 schematically depicts the wavelet-based procedure that is applied to the rotation rate signal acquired from gyroscope, which is the combination of three-dimensional signals $s_{total} = \sqrt{s_{pitch}^2 + s_{roll}^2 + s_{yaw}^2}$. After DWT decomposition using the db6 wavelet function, the energy content of each level is calculated and plotted against the observed movements. As shown in Figure 10, attitude change of the phone caused by vehicle movement and phone movement are well-separated after the decomposition – the former by higher-level wavelet coefficients (low frequency domain) and the latter by lower-level wavelet coefficients (high frequency domain). In the left figure, level 4 and 5 coefficients are selected as they correlated well with true phone movements relative to vehicle (blue vertical lines). In the right figure, level 1 and 2 coefficients are selected as they correlated well with true vehicle turning events.

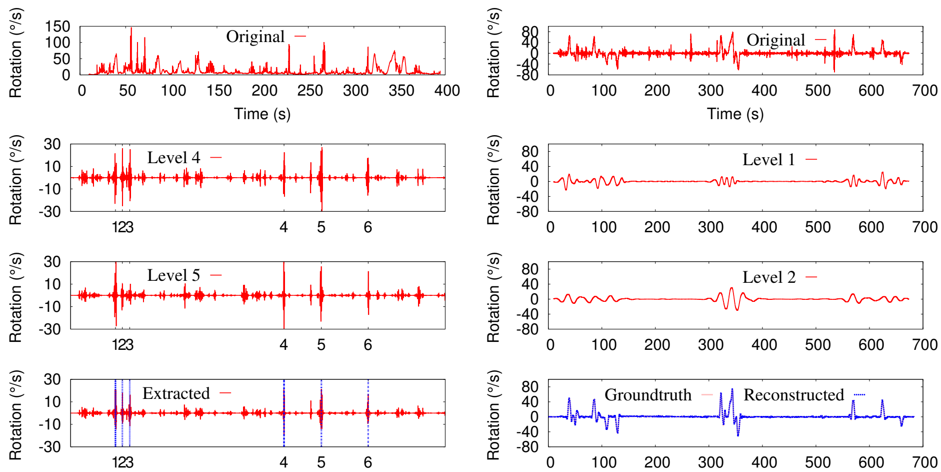

Fig. 10. Wavelet-based detection of phone movement (Left) and vehicle turning (Right).

4 Driver-specific (P)HEV Analysis

Given the multi-modality driver-vehicle data that our sensing system collects, our goal is to analyze, quantitatively, how user-specific driving behavior affects (P)HEV operation, which in turn results in different energy use and environmental impacts. This has been a challenging problem due to the high complexity of (P)HEVs. Leveraging the multi-modality sensing data, we propose an analysis approach that consists of three key components: (1) **Operation mode classification** characterizes the key (P)HEV operation modes and maps user-specific driving behavior to corresponding operation mode; (2) **Energy profile analysis** identifies and quantifies the underlying relationship between (P)HEV electricity and fuel use and different operation modes; and (3) **Fuel–CO_2 emission analysis** characterizes the relationship between fuel use and greenhouse gas emissions.

4.1 (P)HEV Operation Mode Classification

First, we investigate the relationship between user driving behavior and (P)HEV operation. Modeling (P)HEV energy use is a difficult task. On one hand, users' driving behaviors are diverse and vary by road and traffic conditions. On the other hand, under different conditions, a (P)HEV may be powered by either battery or fuel, or both, each of which is a complex process. To address these issues, we have identified five operation modes for the Toyota Prius control system, based on the operation animation displayed on dashboard. The modes are illustrated in Figure 11. **Mode 1**: Only the battery system is used to drive the vehicle. It occurs primarily during city driving with low speed. **Mode 2**: Extra energy (e.g., breaking) is harnessed by the vehicle to charge the battery system. It happens when the vehicle is decelerating. **Mode 3**: Both the ICE and battery system are providing energy to drive the vehicle. It happens when the

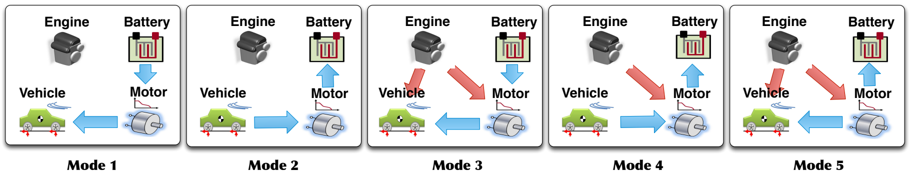

Fig. 11. Categorization of (P)HEV operation modes under different driving scenarios.

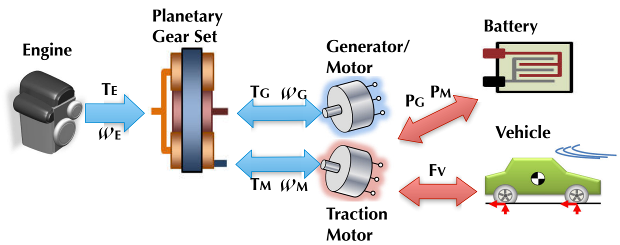

Fig. 12. (P)HEV energy profile analysis. Engine and battery system work together to balance the power demand of the vehicle.

vehicle is driving at high speed and needs massive amount of energy to sustain the movement. **Mode 4**: ICE is charging the battery system. It occurs when the energy generated by ICE exceeds the need to maintain the vehicle movement. **Mode 5**: ICE is both charging the battery system and powering the vehicle. It typically occurs when the battery's SOC (state of charge) is very low while the vehicle has high energy demand.

Based on our analysis of real-world user driving data, we choose to use speed, acceleration, acceleration change, and altitude as the input for (P)HEV operation mode classification. We then construct a decision tree using the CART method [5]. The output of the model is the working status of the engine (on or off) and the battery system (charging or discharging).

4.2 (P)HEV Energy Profile Analysis

Based on the different operation modes, we develop an analytical (P)HEV model specifically for Toyota Prius in MATLAB. Prius adopts series-parallel structure in its drive-train, where the engine, generator/motor and traction motor are coupled together through the planetary gear set to provide the demanded traction power. The driving power requirement is distributed between the Motor and ICE depending on the different modes and user-specific driving behavior. This procedure is illustrated in Figure 12. Mathematically, **battery energy use** can be expressed as [13]:

$$-\Delta SOC = \frac{T_G \omega_G \Delta t}{\eta_G(T_G, \omega_G)} + \frac{T_M \omega_M \Delta t}{\eta_M(T_M, \omega_M)} \tag{4}$$

where ΔSOC is the change of SOC, which is the integration of the exchanged energy between motors and battery system. η_G and η_M are the efficiency of generator/motor and traction motor respectively. They are functions of (T_G, ω_G) and (T_M, ω_M) respectively, where T represents output torque and ω represents angular speed. And **fuel consumption** can be calculated by:

$$fuel = \frac{T_E \times \omega_E \times \Delta t}{m \times \eta_E(T_E, \omega_E)} \tag{5}$$

Fig. 13. Sensing devices (OBD and personal mobile phones) deployed in (P)HEVs for real-time monitoring. OBD data are only used as ground truth in system evaluation.

Table 1. Types of Data Collected in Multi-modality Driver–Vehicle Sensing

Category		Types of Data	Frequency (Hz)	Category		Types of Data	Frequency (Hz)
Vehicle	battery system	temperature	10	User	trip information		
		current	125			elevation	1
		voltage	125			GPS location	1
		state-of-charge/health	10				
	gasoline engine	air intake	3-4		driving behavior	speed	1
		engine RPM	3-4			acceleration	50
	other	speed	3-4			rotation rate	50
		steering	3-4			heading	1

where m is the unit energy generated by ICE, η_E is the efficiency of the ICE and is determined by T_E and ω_E.

4.3 Fuel–CO_2 Emission Analysis

To determine the amount of CO_2 emissions from fuel consumption, we consider the chemical reaction model that assumes predominantly complete combustion of the gasoline to carbon dioxide. With current motor vehicle emission control technology, this assumption has no more than 1% error. The chemical reaction and stoichiometry is shown below([12]):

$$C_8H_{18} + 12.5O_2 + 46.4N_2 \rightarrow 8CO_2 + 9H_20 + 46.4N_2 \tag{6}$$

Using this model, we determine the ratio of the mass of CO_2 emissions to the mass of air input: 0.208 gram of CO_2/gram of air. Since the vehicle run-time fuel consumption can be accurately measured by monitoring its air intake rate (gram/s). The proposed model provides efficient and accurate estimation of vehicle CO_2 run-time emissions.

5 System Deployment and User Studies

The proposed personalized driving behavior monitoring and analysis system has been implemented and deployed for real-world user driving studies. In this section, we describe the details of system deployment and user studies. The data collected in our user studies are then used for system evaluation and driver-specific (P)HEV analysis in Section 6.

5.1 System Deployment

Our sensing system is designed as an application on personal smartphones. It collects, stores, and transmits personalized vehicle and user driving data, and supports user interaction and presentation of analysis results. As shown in Table 1, our sensing system supports real-time collection of comprehensive vehicle and user driving data that are relevant to energy use and environmental impact. For ground truth collection, the OBD cables we use are OBDLink Scan Tool, which supports the OBD-II specification and WiFi communication (with Baud rates 9600 up to 2M). For the personal mobile devices, we use two iPhone 4 smartphones for each trip, one mounted and one unmounted. The smartphones communicate with the OBD devices and computer server via WiFi. Figure 13 shows the OBD and personal mobile phone sensing devices deployed in (P)HEVs for real-time monitoring.

In total, we have deployed the sensing system on ten different vehicles. Two of them are PHEVs, one Ford Escape and one Toyota Prius, both using customized plug-in battery system developed by a clean-energy transportation company. The PHEV battery system consists of over 1000 Li-ion battery units, organized into over 100 modules, with an overall 5.2 KWh energy storage capacity. The other eight cars are owned by the participants. Four of the vehicles are regular HEVs (Toyota Prius), and the other four are conventional gasoline-powered vehicles.

5.2 User Studies

Over a period of one year, we have conducted a series of user studies using the deployed sensing systems. Our goal is to: (1) demonstrate the feasibility and correctness of the sensing system we have developed; (2) collect real-world user driving data and (P)HEV operation information at run-time; and (3) analyze quantitatively how user-specific driving behavior affects (P)HEV energy and environmental impacts. Besides user studies for prototyping and debugging purposes, we have conducted the following user studies:

(1) *Macro-driving studies*: In this set of studies, 8 different participants have been asked to drive their own vehicles (with the deployed sensing systems) according to their daily driving needs. Sensing data are collected and stored locally on the mobile devices whenever a user drives, and the data are uploaded to the server machine when WiFi connection is available. Over 150 trips have been recorded in this set of studies. Table 2 summarizes the key characteristics of the eight participants' daily driving trips, which vary in both time and distance. Figure 14 shows the routes taken by our eight drivers in their regular driving activities. These routes vary significantly from driver to driver. For instance, driver 4 traveled more on freeway, while drive 5 spent most of his driving time on city roads. These trips cover diverse weather and road conditions, such as temperatures from -7°F to 71°F, city vs. highway driving, sunny vs. rainy vs. snowy weather, and driving time throughout the day. The participants also drive differently and have different impact on vehicle, which can be seen in Figure 2 (e.g., driver 3 vs. driver 5).

(2) *Micro-driving studies*: This set of controlled user studies are designed to collect all required ground truth data for evaluation of proposed models and investigate specific factors of users' driving behavior and vehicle operation. 4

Table 2. Comparison of Different Participants' Driving Trips

Driver ID	1	2	3	4	5	6	7	8
Total time (s)	43,176	26,218	22,188	37,498	44,615	103,608	4,702	10,057
Total distance (mile)	366.4	177.8	166.5	322.4	185.6	537.1	40.5	112.0
Total days	14	9	5	27	11	25	3	10
Total trips	26	22	18	37	43	35	5	10
Time per day (s)	3,084	2,913	4,438	1,388	4,056	4,144	1,567	1,006
Distance per day (mile)	26.2	19.8	33.3	11.9	16.9	21.5	13.5	11.2

different participants have been asked to drive the PHEVs under different road and traffic conditions. Over 20 trips have been conducted in this set of studies.

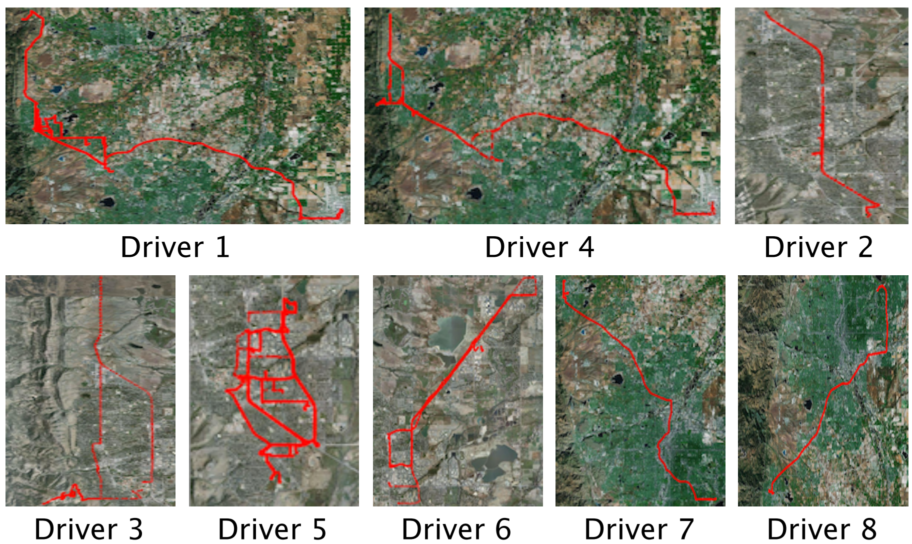

Fig. 14. Heterogeneous routes driven by the eight participants in the macro-driving user studies.

Our real-world system deployment and user studies have demonstrated the feasibility and effectiveness of the multi-modality driving–vehicle sensing system we have developed. The comprehensive user driving and vehicle operation data we have gathered are then used in the following section for detailed system evaluation and driver-specific (P)HEV analysis. The data provide valuable insights into users' driving behaviors, as run-time acceleration, speed, and slope are critical factors in determining the power demands of the hybrid vehicle components.

6 System Evaluations and Analysis Results

In this section, we evaluate the proposed personalized driving behavior monitoring and analysis system for emerging hybrid vehicles. Using the real-world

Position	Seat	Pocket Upper	Pocket Lower	Bag	Mounted (reference)
MAE-before (m/s^2)	5.19	9.99	9.57	8.10	0.14
Corr-before	0.08	0.12	0.16	0.28	0.97
MAE-after (m/s^2)	0.26	0.46	0.22	0.17	0.14
Corr-after	0.88	0.80	0.92	0.96	0.97

Table 3. Accuracy of Acceleration Sensing with Different Phone Positions. Results shown are MAE and correlation values before and after sensor data correction.

user-driving data we have gathered from the deployed systems, we would like to answer the following questions:

1. Does our system achieve high accuracy for driver–vehicle sensing?
2. Does our system achieve high accuracy with regard to operation mode classification, (P)HEV modeling of battery system use and fuel use?

6.1 Sensing System Validation

We have demonstrated in Section 3 that individual sensor readings can be effectively processed to enhance quality. Here, we focus on evaluating the "end results", i.e., vehicle movement in terms of acceleration and speed, whose accuracy are crucial in (P)HEV analysis. In particular, obtaining the correct vehicle acceleration information requires complex processing of accelerometer, gyroscope, and digital compass readings, as well as disorientation compensation.

In our experiments, we consider four different ways of positioning a phone in the vehicle: (1) *Seat*: Phone lays flat on the passenger seat. (2) *Pocket Upper*: Phone is in the driver's jacket pocket. (3) *Pocket Lower*: Phone is in the driver's pant pocket. (4) *Bag*: Phone is in a bag on the passenger seat. We use the *mounted* position as our reference, where the phone is securely mounted to the front of the windshield. We use **mean absolute error (MAE)** and **cross-correlation coefficient** (Equation 1) to measure the accuracy of the corrected sensor readings. Given two series A and B each with n values a_i and b_i ($i = 1, \ldots, n$), we define

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|a_i - b_i| \quad (7)$$

Table 3 shows the accuracy of acceleration readings before and after correction, using the mounted phone reading as reference. Note that even under exactly the same context, there are still differences between samples from different sensors. The acceleration readings between two mounted phones have an MAE of $0.14 m/s^2$ and correlation of 0.97. We can also see that our corrected

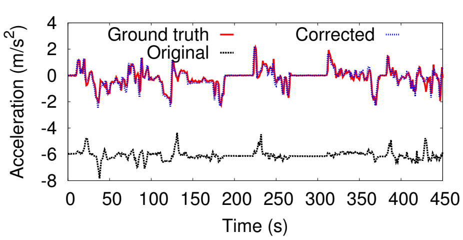

Fig. 15. Accuracy of acceleration sensing.

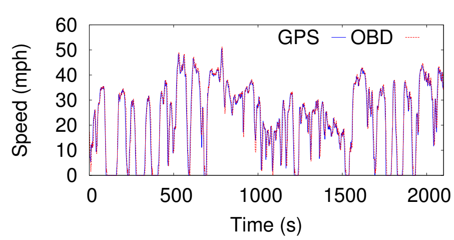

Fig. 16. Accuracy of speed sensing.

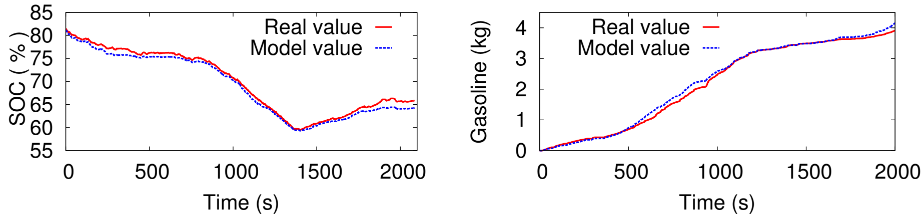

Fig. 17. Accuracy of (P)HEV run-time energy use and fuel use modeling.

acceleration readings effectively improve the vehicle acceleration sensing accuracy over the raw readings, under all four typical positions, e.g., 0.28 vs. 0.96 correlation and 8.10 vs. 0.17 MAE, before and after correction, when phone is placed in a bag on the passenger seat. As shown in Figure 15, the corrected acceleration readings match well with the ground truth.

Figure 16 also shows that the GPS speed readings match very well with the OBD (ground truth) speed readings (0.99 correlation).

6.2 Analysis Model Validation

We first evaluate the accuracy of the operation mode classifier we have developed, which maps user driving behavior to the corresponding (P)HEV operations mode. The ground truth is generated from the engine RPM and electrical current reported by OBD device. We use k rounds of 0.632 bootstrapping [6], a widely used approach, to generate k pairs of training and testing data sets. For each pair, about 63.2% of the instances are in the training set and the remaining instances are in the testing set. Let M_i be the classifier obtained using the i-th round training set, $\zeta(M_i)_{test}$ and $\zeta(M_i)_{train}$ be the accuracy of M_i on the i-th round testing set and training set, respectively, the expected accuracy of our operation mode classifier is

$$\zeta(M) = \sum_{i=1}^{k}(0.632 * \zeta(M_i)_{test} + 0.368 * \zeta(M_i)_{train}) \tag{8}$$

Using $k = 10$ rounds and real-world user driving data, the expected classification accuracy is 89.9% for engine status and 87.8% for battery status.

Figure 17 shows the modeling results for run-time battery use (left) and fuel use (right), respectively. Battery energy use is reported as SOC (%) over time, and fuel use is based on gasoline (kg). Both figures show that the values predicted by the models match well with the real (measured) values. The battery use model achieved 0.41% MAE and 0.918 correlation; while the fuel use model also achieved a very low MAE of 0.07 kg and a high correlation of 0.996.

6.3 Discussions

Vehicle systems are highly complex, and a large number of factors come into play, with varying impacts on fuel efficiency, battery system run-time performance and long-term aging. Although it is difficult to model all factors in a complex vehicle system, our solution does provide good approximation of vehicle performance in real-world driving scenarios, thus bridging the information gap between driver and vehicle. Here, we further discuss other related factors and limitations of the proposed solution.

For fuel consumption, the ground truth is derived from air flow mass reported by the OBD device, using formula proposed in (http://www.lightner.net/lightner/bruce/Lightner-183.pdf). This approach works very well in a modern automobile, which is accurate to within a few percent. Although some factors are not explicitly modeled in our system, our evaluation data set does contain diverse weather and road conditions, and the high accuracy of our system would be generalizable in many scenarios. As our future work, we plan to conduct larger-scale user driving studies and incorporate explicit modeling of other related factors such as temperature, humidity, road conditions, etc.

Although the speed information reported by GPS is accurate (i.e., highly consistent with the speed reported by OBD, as validated in our studies), the altitude information from GPS may suffer from significant noises and can impact our results to some extent. Our evaluation data set contains substantial altitude variations, and our system achieves high accuracy. We are therefore confident that the proposed system can be deployed in almost all physical settings. Nevertheless, places with large altitude variations may benefit from more accurate GPS reporting and better algorithm to reduce the noise in altitude sensing data.

We have evaluated the proposed system with four different phone positions in a vehicle (seat, pocket upper, pocket lower, and bag), which should capture most driving scenarios. Still, other phone positions may be possible and may affect the accuracy of our system. In addition, due to the lower quality sensors used in mobile phones, other types of noises may arise and may not be effectively removed by our current solution.

Our battery system model is constructed based on the type of PHEVs developed by Toyota. The specific vehicle design is orthogonal to our solution and adaptation to different PHEVs can be achieved through vehicle-specific driving studies and data-based parameter tuning. We expect the overall trend of fuel consumption and battery system use to stay the same for the same type of driving behavior.

7 Related Work

Our work focuses on personalized driving behavior monitoring and analysis for emerge hybrid vehicles. It draws upon research works in several related areas, which we summarize in this section.

Mobile Sensing System for Vehicles. Sensing systems that utilize mobile platform and other devices such as GPS, accelerometer, and OBD device, have been developed for monitoring road and traffic conditions [24], commute time, WiFi deployment, and automotive diagnostics [11], or finding the optimal route in terms of lowest fuel consumption [9]. These systems require purchasing extra devices, which limits user adoption. Privacy of driving trips has also been studied [14]. Several techniques have been proposed to classify a wide variety of human movements and activities using mobile device [15]. However, previous works typically require stable sensor placement for data collection, e.g., mounting the phone in a fixed place. Such requirement is intrusive and inconvenient for users. During a driving trip, the phone may move relative to the vehicle (e.g., sliding or user picking up a phone call). Such abnormal movements make statistical models inadequate and affect the quality of monitored data [19]. This may be addressed by taking extraneous activities into consideration [18]. However, these methods are not suitable for detecting transient movements of phone relative to vehicle, nor can they dynamically compensate for the error. Our work differs from these systems in that we focus on the energy and environmental impacts of user-specific driving behavior on emerging electric-drive vehicles. Moreover, our proposed sensing techniques and analysis system make it possible to separate transient phone/vehicle movements at run-time and use the mobile platform to automatically fulfill all functionalities, thus eliminating the inconvenience of user intervention and extra OBD devices.

Driving Cycle Analysis. Driving cycle analysis for performance assessment of conventional vehicles has drawn significant attention in the past [25]. Due to the challenge of collecting real-world trip data, standard driving cycles have been developed to describe various driving modes in different regions, such as UDDS, HFEDS, US06, etc. These driving cycles are used in previous research to simulate real-world conditions for vehicle performance analysis [8]. Recently, driving cycle analysis studies began to consider hybrid vehicle technologies [10]. For instance, standard driving cycles are used to investigate different power train configurations [8]. A few works have used speed information in real-world trips to investigate (P)HEV battery performance [4]. We take a systematic approach to identify the corresponding operation mode under specific driving behavior, model (P)HEV battery system use and fuel–CO_2 emissions, therefore enabling personalized, comprehensive, and quantitative analysis of (P)HEV performance and environmental impacts.

8 Conclusions

We have developed a personalized driving behavior monitoring and analysis system for emerging hybrid vehicles. We propose techniques for multi-modality drive–vehicle sensing, including de-noise, drift calibration, synchronization, and wavelet-based disorientation compensation for precise vehicle movement sensing. In addition, via operation mode classification, (P)HEV battery system modeling and fuel-CO_2 emissions modeling, our system enables comprehensive and quantitative analysis of (P)HEV energy environmental impacts under user-specific driving behavior. Real-world system deployment and user driving studies demonstrate the feasibility and high accuracy of the proposed system, with 0.88–0.996 correlation values.

ACKNOWLEDGEMENTS. This work was supported in part by the National Science Foundation award CNS–0910995. We thank our shepherd, Hao-Hua Chu, and the anonymous reviewers for constructive comments.

References

1. Euler angles, http://en.wikipedia.org/wiki/Euler_angles
2. Toyota edf, http://www.edfenergy.com/media-centre/press-news/edf-energy-and-toyota-launch-uk-trials-of-plug-in-hybrid-vehicle.shtml
3. Development of speed correction cycles. Tech. rep., US EPA Assessment and Modeling Division, NVFEL (1997), EPA report no. M6.5PD.001
4. Adornato, B., Patil, R., Filipi, Z., Baraket, Z., Gordon, T.: Characterizing naturalistic driving patterns for plug-in hybrid electric vehicle analysis. In: IEEE VPPC (2009)
5. Breiman, L., Friedman, J., Olshen, R., Stone, C.: Classification and Regression Trees. CRC Press (1984)
6. Davison, A.C., Hinkley, D.: Bootstrap Methods and their Application. Cambridge (2006)
7. Ehsani, M., Gao, Y., Emadi, A.: Modern Electric, Hybrid Electric, and Fuel Cell Vehicles: Fundamentals, Theory, and Design. CRC Press. (2009)
8. Ganji, B., Kouzani, A.Z., Trinh, H.: Drive cycle analysis of the performance of hybrid electric vehicles. LSMS/ICSEE '10, Part I pp. 434–444 (2010)
9. Ganti, R.K., Pham, N., Ahmadi, H., Nangia, S., Abdelzaher, T.F.: GreenGPS: a participatory sensing fuel-efficient maps application. In: MobiSys '10 (2010)
10. Huang, X., Tan, Y., He, X.: An intelligent multi-feature statistical approach for discrimination of driving conditions of hybrid electric vehicle. In: IJCNN '09
11. Hull, B., Bychkovsky, V., Zhang, Y., Chen, K., Goraczko, M., Miu, A., Shih, E., Balakrishnan, H., Madden, S.: CarTel: a distributed mobile sensor computing system. In: SenSys '06. pp. 125–138 (2006)
12. J, H.: Consider a Spherical Cow. University Science Books (1988)
13. Jonasson, K.: Analysing hybrid drive system topologies. Licentiate thesis (2002)
14. Krumm, J.: Realistic driving trips for location privacy. In: Pervasive '09. pp. 25–41
15. Lane, N., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., Campbell: A survey of mobile phone sensing. Communications Magazine, IEEE 48(9), 140 – 150 (2010)
16. Li, K., Wu, J., Jiang, Y., Hassan, Z., Lv, Q., Shang, L., Maksimovic, D.: Large-scale battery system modeling and analysis for emerging electric-drive vehicles. In: ISLPED '10. pp. 277–282 (2010)
17. Lin, J., Niemeier, D.A.: Regional driving characteristics, regional driving cycles. Transportation Research Part D 8, 361–381 (2003)
18. Lu, H., Yang, J., Liu, Z., Lane, N.D., Choudhury, T., Campbell, A.T.: The jigsaw continuous sensing engine for mobile phone applications. In: SenSys'10 (2010)
19. Miluzzo, E., Papandrea, M., Lane, N., Lu, H., Campbell, A.T.: Pocket, bag, hand, etc. - automatically detecting phone context through discovery. In: PhoneSense'10 (2010)
20. Mohan, P., Padmanabhan, V.N., Ramjee, R.: Nericell: rich monitoring of road and traffic conditions using mobile smartphones. In: SenSys '08. pp. 323–336 (2008)
21. Mohd-Yasin, F., Nagel, D.J., Ong, D.S., Korman, C.E., Chuah, H.T.: Low frequency noise measurement and analysis of capacitive micro-accelerometers: Temperature effect. Japanese Journal of Applied Physics 47(6), 5270–5273 (2008)
22. Samaras, C., Meisterling, K.: Life cycle assessment of greenhouse gas emissions from plug-in hybrid vehicles: Implications for policy. Environ. Sci. Technol. 42 (2008)
23. Sardy, S., Tseng, P., Bruce, A.: Robust wavelet denoising. Signal Processing, IEEE Transactions on 49(6), 1146 –1152 (2001)
24. Thiagarajan, A., Ravindranath, L., LaCurts, K., Madden, S., Balakrishnan, H., Toledo, S., Eriksson, J.: VTrack: accurate, energy-aware road traffic delay estimation using mobile phones. In: SenSys '09. pp. 85–98 (2009)
25. Tong, H.Y., Hung, W.T., Chun-shun, C.: On-road motor vehicle emissions and fuel consumption in urban driving conditions. Joural of the Air and Waste Management Association 50, 543–554 (Apr 2000)
26. Torrence, C., Compo, G.P.: A practical guide to wavelet analysis. Bulletin of the American Meteorological Society 79, 61–78 (1998)