The Car as an Ambient Sensing Platform

BY EMANUELE MASSARO, CHAEWON AHN, AND CARLO RATTI
MIT Senseable City Lab, Cambridge, MA 02139 USA

PAOLO SANTI
MIT Senseable City Lab Cambridge, MA 02139 USA and Istituto di Informatica e Telematica del CNR, Pisa 56124, Italy

RAINER STAHLMANN AND ANDREAS LAMPRECHT
AUDI AG, Ingolstadt D-85045, Germany

MARTIN ROEHDER AND MARKUS HUBER
Volkswagen Group of America Electronics Research Lab Belmont, MA 02139 USA

Fig. 1. ***The vision of cars used as pervasive sensing platforms: CAN data are communicated to remote servers through wireless communication, possibly after onboard processing and aggregation. GPS signal is fundamental to build time- and spatially-resolved databases from the multiple data streams collected by a single vehicle, as well as from data generated by different vehicles.***

In recent years, cars have evolved from purely mechanical to veritable cyber-physical systems that generate large amounts of real-time data. These data are instrumental to the proper working of the vehicle itself, but make them amenable to a multitude of other uses. For instance, GPS information has recently been used for a large number of mobility studies in the academic community [1], [5], as well as to feed traffic apps such as Google Traffic and Waze. This use of vehicle data is already having a profound impact in science, industry, economy, and society at large. Now, imagine that instead of accessing one single source of vehicle-generated data (GPS), one can access the entire wealth of data exchanged on the controller area network (CAN) bus in near real time amounting to over 4000 signals sampled at high frequency, corresponding to a few gigabytes of data per hour. What would be the implications, opportunities, and challenges sparked by this transition?

This transition is now being made possible by the so-called connected car paradigm, which allows vehicle CAN bus data to be recorded and wirelessly transmitted to central servers for analysis. Thus, the car sensing dimension, which can be informally understood as the number of different signals that a vehicle records and makes available for data analysis, is increasing from 1 (or a few) to 1000 or above.

In the rest of the article, we discuss the groundbreaking effects this transition will have on our ability of sensing roads and the urban living environment. Guided by the preliminary

Digital Object Identifier: 10.1109/JPROC.2016.2634938

exploration of an immense CAN sensory data set, we suggest possible research directions that the connected vehicle of the future will open, as well as discuss the implied research, technological, and societal challenges.

I. EVOLUTION OF CAR SENSING TECHNOLOGY

The idea of using the car as a data collection device is not new: starting several decades ago, scholars, urban planners, and policy makers have looked at the car not only as a means of transport, but also as a data collector probe [6]. This especially occurred with the first applications of Global Positioning System (GPS) and, immediately after, with its integration with cellular communication. GPS data sets have been used to study, analyze, predict, and understand a large variety of problems: traffic congestion, vehicle energy consumption and emissions [7], urban mobility [5], [8], human reactions to emergent technologies [9], and many others.

However, car data collection capacity is now well beyond GPS, with a large number of components that exchange data through a technology invented in 1986 by Robert Bosch GmbH: the so-called CAN [10], a serial broadcast bus that allows near-real-time management of most sensors and electronic devices embedded in the car.

Today, silicon-rich vehicles include many sophisticated systems such as ABS, ESP, and adaptive cruise control, each of which requires a number of sensors for its functioning. These are highly integrated sensors, which can actually be considered as a single sensor system. For instance, the Ethernet rear-view camera helps the driver in safely handling the vehicle during the parking operations [11]. On top of sensing, these devices also have embedded computational resources for image and graphics processing.

A second driver of the increasing siliconization of cars is in-vehicle entertainment: features such as managing multimedia information, communication, and personalized services are increasingly being demanded by customers. To support this evolution, heterogeneous multicore processors and different sensors must be available onboard. An example of this is integrity multivisor technology developed by Green Hills, which provides an operating environment capable of effectively managing both comfort and safety [13].

Another milestone in the evolution of the car is digital communication between cars with a vehicle-to-vehicle (V2V) network, and with the roadside infrastructure (toll booths, traffic lights, etc.) with a vehicle-to-infrastructure (V2I) communication system. The data generated by the internal networks of the car can then be transmitted automatically, allowing near-real-time communication of in-vehicle sensing information. For instance, Google has recently filed a patent for a system that is able to detect and share pothole information online, using a combination of onboard sensors, GPS signal, and cloud technology [14].

Finally, we mention the convergence of vehicle software platforms with mobile operative systems: both Apple and Google are working on integrating their mobile platforms into vehicles with the CarPlay and Android Auto technologies. Car manufacturers are also partnering with phone manufacturers to enable seamless interactions with existing mobile platforms [15], [16].

In this article, we posit that the car has the potential to make a breakthrough transition from single sensing probe, to a veritable mobile sensing platform. Fig. 1 graphically synthesizes this concept: by accessing data from the vehicles networking architecture, in-vehicle data can be collected, possibly aggregated, and reported to a remote server through wireless communication. Data collected from different cars can then be space and time aligned across different vehicles thanks to GPS [17]. The time and spatial accuracy that can potentially be achieved by this new sensing paradigm (in the order of a second and a few meters, respectively) are unprecedented, opening up the way to a multitude of novel applications, as well as computational, communication, and privacy challenges.

II. THE CAR AS AN URBAN SENSING PLATFORM

Since most vehicles travel in urban environments, a wide class of applications enabled by the car sensing platform will likely focus on urban sensing. In this section, we instantiate this vision through the preliminary analysis of a pseudonymous data set composed of the recordings of over 900 in-vehicle signals. Contractors drove ten test vehicles equipped with data loggers that captured CAN bus signals as part of a research project. All signals were collected with a frequency of at least 1 Hz. Vehicles were driven by 64 different users over 86 days in the surroundings of a German city, for a total of over 1900 driving session with an average duration of 70 min. The spatial coverage of these sessions is reported in Fig. 2(a).

In the following, we focus our attention on specific examples dealing with:

- weather and environmental sensing;
- road safety;
- driver behavior analysis.

A. Weather and Environmental Sensing

The global trend toward urbanization explains the growing interest in the study of urban climate, and of its change due to human intervention. Phenomena such as heat island on a local scale, and global warming at the planet scale, are well known, yet not fully understood mostly due to lack of fine-grained data. Scholars have recently pointed out the importance of developing accurate local microclimate urban emission models to face climate change at the global scale [12]. A recent experimental study, for instance, suggests the use of moving cars to measure rainfall [18] through analysis of wipers speed.

Notwithstanding advances in sensing and communication technology, our current ability to monitor weather and environmental conditions is still severely limited in both time and space. Considering weather, even pervasive,

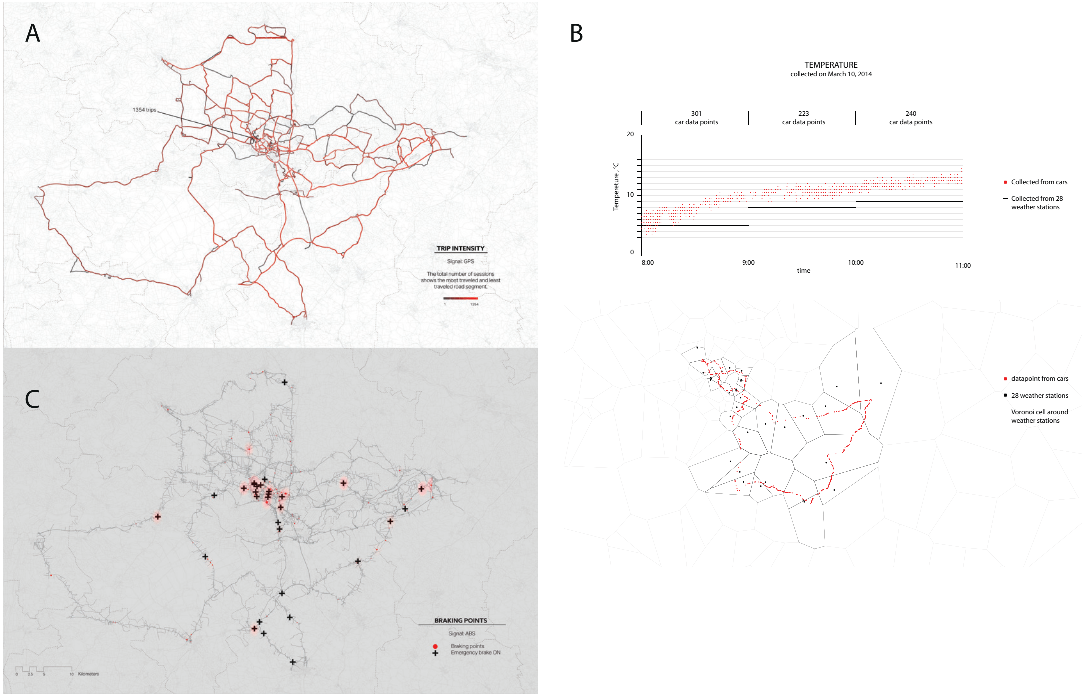

Fig. 2. ***(a) Visualization of urban, sub-urban, and highway coverage through the different sessions performed by the users. (b) Weather data resolution improvement with car sensing: the sample trajectory intersects the area covered by 28 weather stations. Both temporal and spatial resolutions are improved by at least one order of magnitude with respect to state-of-the-art monitoring capability. (c) ABS and emergency braking activations indicate potentially dangerous points in the road network.***

crowdsourced weather services such as those provided by Weather Underground still collect data with spatial granularity in the order of several square kilometers, and time resolution in the order of 1 h. Conversely, the car sensing platform allows collecting weather information with a typical 1-Hz frequency, and accurate spatial resolution thanks to GPS.

Some weather data are directly sensed in most vehicles (e.g., outside temperature, rain intensity if sensors are present, etc.), while other can be easily obtained indirectly (e.g., activation of wipers indicating the presence of rain, of fog lights to detect fog, etc.). By collecting data from a multitude of cars, it will be possible to significantly improve the accuracy of weather and environmental sensing in regions covered by vehicle traffic. The map reported in Fig. 2(b) shows the location of Weather Underground stations in the region where our vehicular data were collected: these amount to 224 stations in an area of approximately 104.6 km × 66.9 km ≃ 6998 km^2, indicating that each station covers about 31 km^2 on average.

The map also depicts the data points recorded by the ten vehicles that collected CAN data. As seen, for a single temperature reading performed by a few weather stations in a fixed location, there are hundreds of readings performed by the vehicles in the same coverage area, and these readings are taken at different times and in different positions. Considering that our data set was produced by only ten vehicles, and that the sensing resolution is proportional to the number of cars driving in the territory, it is easy to see that the number of available readings can scale up to several thousands. This increased data granularity opens the way to the development of much more accurate microclimate models.

B. Road Safety

While most of the accidents are related to human driver mishandlings, road conditions often are the root cause of, or at least contribute to, those mishandlings. As such, a prompt identification of potentially dangerous road conditions is a key to improve road safety and reducing accidents. This requires fast and accurate sensing of road conditions, as well as implementation of low-latency communication with roadside infrastructure and surrounding vehicles to prevent or reduce the impact of accidents. Low-latency V2I and V2V communication has attracted significant interest in the research and industry community in recent years, including standardization activities [19]. As a matter of fact, the development of applications [20], [23] based on sensing of road conditions goes hand in hand with standardization activities for wireless communication (e.g., ETSI TC ITS) [24], [25].

CAN collected data can potentially be useful not only for near-real-time road condition alerts, but also to systematically identify dangerous portions of the road network, as resulting, e.g., from recurrent activations of ABS or emergency braking in specific road s egments. An example of this application is reported in Fig. 2(c), where ABS activation points are reported. A number of potentially dangerous road segments can be identified. Since ABS is activated also in presence of sudden accelerations, we have overlaid ABS data with those generated by the emergency braking system: this is an onboard ADAS that assists the drivers in case of sudden vehicle deceleration. As seen from Fig. 2(c), in most cases, ABS and emergency braking activation do occur in the same road segments, indicating potentially dangerous portions of the road network. By crosschecking this information with accident databases, it will be possible to correlate ABS/emergency braking maps with likelihood of accident occurrence, opening up the way to preventive measures to increase the road safety.

C. Driver Behavior Characterization

Characterizing driver behavior is important since, as outlined above, a large fraction of car accidents are caused by human mishandlings [26]. By characterizing driver behavior, potentially dangerous habits such as aggressive or distracted driving can be identified, and corrective actions undertaken.

However, current technology is able to characterize only coarse-grained driving features, which are mostly limited to travel distance, covered driving areas, and occurrence of sudden braking events. Indeed, accident risk is only mildly related to some of these features. Consider, for instance, total traveled miles: a large number of traveled miles might indeed be an indication of a very experienced driver, which is relatively less likely to mishandle potentially dangerous situations than an inexperienced driver who drives only a few miles daily. A more accurate characterization of driver behavior would require analyzing several parameters related to driving, such as speed profile, intensity of braking/acceleration, distance to surrounding vehicles, etc. Most of this information is currently collected through the vehicles CAN bus. Such fine-grained classification of driver behavior could be useful not only to reduce accident risk, but also to help drivers improve their driving styles in terms of, e.g., reducing fuel consumption and emissions.

In scientific literature only few works have been presented that targeted driver identity inference from sensor data [27]. Most work on driver behavior characterization is based on simulator data [28], and only few recent works use CAN generated data [29], [30]. Also the latter works, however, analyze driver behavior in a scenario where drivers were: 1) aware of the fact that a driving experiment was ongoing; and 2) instructed to perform specific maneuvers and/or trips. However, to provide a breakthrough in driver behavior characterization, a methodology for near-real-time classification of driver behavior in uncontrolled environments should be defined. Ample availability of CAN-collected data will give the opportunity to improve models such as [31] and characterize driver behavior in an open, uncontrolled environment.

III. CHALLENGES

The vision of the car as an ambient sensing platform, while opening the way to exciting applications and opportunities, poses also a number of challenges to the research and engineering community, as well as to the society at large.

By looking at the sheer amount of generated datain the order of a few gigabytes per hour per vehicleit is clear that simply dumping CAN bus data to a remote server, if desirable at all, is not a feasible option. So the dynamic, context-aware selection of which sensor signals to discard, store onboard, or send to a remote server is an important research question to be addressed. This problem is relevant also for the design of responsive driver dashboards, where information to be displayed to the user is dynamically selected based on context and information exchanged with surrounding vehicles.

How and where to preprocess and aggregate CAN-generated data is also an important question to be addressed. Different architectural choices, on vehicle, edge computing, cloud, etc., as well as their combination, will be discussed, evaluated, and tested in terms of both data processing performance, as well as induced communication overhead.

Another formidable challenge relates to security and privacy aspects. The connected car paradigm extends the realm of potential security attacks to vehicles. Thus, secure data exchange between vehicles and infrastructure is a fundamental building block of our proposed vision. For what concerns privacy, it has recently been shown that driver fingerprinting is possible based on access to a limited number of car sensor information [30]. While this could be useful in identifying potentially dangerous drivers or vehicle misuse, on the other hand, it opens substantial privacy issues. Similarly to other realms where big data approaches find application, privacy issues should be critically discussed at societal level, and compared with the benefits that advanced urban sensingas made possible by the car sensing platformwould imply. We stress that giving information and transparency about usage of data and the purpose for its collection to the vehicle owner is very important. This needs to be done not only to comply with data privacy laws and regulations, but also to support customers awareness and self-determination, especially in cases where the realization of an application requires providing personal data to third parties. It is the decision of the customer based on a declaration of consent, if personal data may be collected and for which purpose it may be used.

IV. CONCLUSION

Fully exploiting the opportunities enabled by a scenario in which CAN-collected data can be recorded, communicated, and analyzed in near-real-time requires tightly knit contributions from a wide range of disciplines. This ranges from research fields within the realm of engineering, such as sensor and actuator design, signal processing, onboard processor design, data communication, to fields in the related computer science discipline, machine and deep learning, computer graphics, human-computer interaction, data security and privacy, middleware and mobile computing, to encompass other disciplines in science and humanities such as physics, environmental science, and urban design and planning. More broadly speaking, the privacy and security issues implied by the transition of the car into a connected sensing platform will be discussed not only within the engineering community, but also with industry, government, and society at large to steer this evolution of the car in a direction that is socially desirable. ■

REFERENCES

[1] C. de Fabritiis, R. Ragona, and G. Valenti, "Traffic estimation and prediction based on real time floating car data," in *Proc. 11th Int. IEEE Conf. Intell. Transp. Syst.*, Oct. 2008, pp. 197–203.

[2] G. Taylor, G. Blewitt, D. Steup, S. Corbett, and A. Car, "Road reduction filtering for GPS-GIS navigation," *Trans. GIS*, vol. 5, no. 3, pp. 193207, Jun. 2001.

[3] S. T. S. Thong, C. T. Han, and T. A. Rahman, "Intelligent fleet management system with concurrent GPS & GSM real-time positioning technology," in *Proc. 7th Int. Conf. ITS Telecommun.*, 2007, pp. 1–6.

[4] J. H. Rillings and R. J. Betsold, "Advanced driver information systems," *IEEE Trans. Veh. Technol.*, vol. 40, no. 1, pp. 31–40, Feb. 1991.

[5] L. Vincent, "Taking online maps down to street level," *Computer*, vol. 40, no. 12, pp. 118–120, Dec. 2007.

[6] J. P. Womack, D. T. Jones, and D. Roos, *The Machine That Changed the World: The Story of Lean ProductionToyotas Secret Weapon in the Global Car Wars That is Now Revolutionizing World Industry*, 1st ed., New York, NY, USA: Free Press, 1990.

[7] K. Ahn and H. Rakha, "The effects of route choice decisions on vehicle energy consumption and emissions," *Transp. Res. D, Transp. Environ.*, vol. 13, no. 3, pp. 151–167, May 2008.

[8] A. Bazzani, B. Giorgini, S. Rambaldi, R. Gallotti, and L. Giovannini, "Statistical laws in urban mobility from microscopic GPS data in the area of Florence," *J. Stat. Mech., Theory Exp.*, vol. 2010, no. 5, p. P05001, May 2010.

[9] G. Leshed, T. Velden, O. Rieger, B. Kot, and P. Sengers, "In-car GPS navigation: Engagement with and disengagement from the environment," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2008, pp. 1675–1684.

[10] U. Kiencke, S. Dais, and M. Litschel, "Automotive serial controller area network," SAE Tech. Paper 860391, 1986.

[11] R. Ramanath, J. Bochenek, and G. Rottner, "Vehicle rear view camera system and method," U.S. Patent 8994825, Mar. 31, 2015.

[12] K. R. Gurney *et al.*, "Climate change: Track urban emissions on a human scale," *Nature*, vol. 525, no. 7568, pp. 179–181, Sep. 2015.

[13] K. Sandstrm, A. Vulgarakis, M. Lindgren, and T. Nolte, "Virtualization technologies in embedded real-time systems," in *Proc. IEEE 18th Conf. Emerg. Technol. Factory Autom.*, Sep. 2013, pp. 1–8.

[14] D. K. Jackson, "Systems and Methods for monitoring and reporting road quality," U.S. Patent 9108640, Aug. 18, 2015.

[15] L. D. Burns, "Sustainable mobility: A vision of our transport future," *Nature*, vol. 497, no. 7448, pp. 181–182, May 2013.

[16] J. White, C. Thompson, H. Turner, B. Dougherty, and D. C. Schmidt, "WreckWatch: Automatic traffic accident detection and notification with smartphones," *Mobile Netw. Appl.*, vol. 16, no. 3, pp. 285–303, Jun. 2011.

[17] J. Ryu and J. C. Gerdes, "Integrating inertial sensors with global positioning system (GPS) for vehicle dynamics control," *J. Dyn. Syst., Meas., Control*, vol. 126, no. 2, pp. 243–254, 2004.

[18] E. Rabiei, U. Haberlandt, M. Sester, and D. Fitzner, "Rainfall estimation using moving cars as rain gauges—Laboratory experiments," *Hydrol. Earth Syst. Sci.*, vol. 17, no. 11, pp. 4701–4712, Nov. 2013.

[19] J. B. Kenney, "Dedicated short-range communications (DSRC) standards in the United States," *Proc. IEEE*, vol. 99, no. 7, pp. 1162–1182, Jul. 2011.

[20] T. Kosch, I. Kulp, M. Bechler, M. Strassberger, B. Weyl, and R. Lasowski, "Communication architecture for cooperative systems in Europe," *IEEE Commun. Mag.*, vol. 47, no. 5, pp. 116–125, May 2009.

[21] A. Festag, L. Le, and M. Goleva, "Field operational tests for cooperative systems: A tussle between research, standardization and deployment," in *Proc. 8th ACM Int. Workshop Veh. Inter-Netw.*, 2011, pp. 73–78.

[22] H. Stbing *et al.*, "simTD: A car-to-X system architecture for field operational tests [topics in automotive networking]," *IEEE Commun. Mag.*, vol. 48, no. 5, pp. 148–154, May 2010.

[23] R. Stahlmann, A. Festag, A. Tomatis, I. Radusch, and F. Fischer, "Starting European field tests for CAR-2-X communication: The DRIVE C2X framework," in *Proc. 18th ITS World Congr.*, Orlando, FL, USA, 2011, pp. 1–9.

[24] Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Definitions, document ETSI TR 102 638 V1.1.1, ETSI TC ITS, 2009.

[25] Intelligent Transport Systems (ITS); European Profile Standard on the Physical and Medium Access Layer of 5 GHz ITSs, document Draft ETSI ES 202 663 V0.0.6, ETSI TC ITS, 2009.

[26] European Commission, "Mobility and transport: Road safety, mobility and transport: Road safety," Oct. 12, 2015. [Online]. Available: http://ec.europa.eu/transport/road_safety/index_en.htm

[27] W. Wang, J. Xi, and H. Chen, "Modeling and recognizing driver behavior based on driving data: A survey," *Math. Problems Eng.*, vol. 2014, pp. 1–20, Feb. 2014.

[28] C. Miyajima *et al.*, "Driver modeling based on driving behavior and its evaluation in driver identification," *Proc. IEEE*, vol. 95, no. 2, pp. 427–437, Feb. 2007.

[29] J. Carmona, F. Garca, D. Martn, A. de la Escalera, and J. M. Armingol, "Data fusion for driver behaviour analysis," *Sensors*, vol. 15, no. 10, pp. 25968–25991, Oct. 2015.

[30] M. Enev, A. Takakuwa, K. Koscher, and T. Kohno, "Automobile driver fingerprinting," *Proc. Privacy Enhancing Technol.*, vol. 2016, no. 1, pp. 34–50, Jan. 2016.

[31] A. Pentland and A. Liu, Modeling and prediction of human behavior, *Neural Comput.*, vol. 11, no. 1, pp. 229–242, Jan. 1999.