Collaborative Vehicle Data Collection Through Mobile and Web Technologies

Giuseppe Loseto*, Filippo Gramegna*, Saverio Ieva*, Agnese Pinto*, Michele Ruta*, Floriano Scioscia*
* Polytechnic University of Bari – via E. Orabona 4, Bari (I-70125), Italy
{name.surname}@poliba.it

***Abstract*—Analyzing On-Board Diagnostics 2 (OBD-II) vehicle status data is highly useful for Smart Mobility research, but collection costs and competitiveness in the automotive sector currently prevent large-scale OBD dataset availability. Inspired by worldwide crowdsourcing efforts like OpenStreetMap (OSM) for cartography, we propose a framework based on open standards to enable collaborative collection of OBD data. Basic technological elements are: a fork of an open source OBD-II logger for Android, exploiting an extension of GPX data format for recording route logs enriched with OBD-II parameters; a Web platform enriching OSM to upload, manage and search routes with embedded OBD data. A full prototype has been developed to be exploited in Smart Mobility use cases.**

I. Introduction and motivation

Current Intelligent Transportation Systems (ITS) research and development increasingly rely on real-time processing of (i) detailed cartographic data, (ii) vehicle-to-everything (V2X) communication [1] inputs, (iii) high-throughput environmental sensors such as radar, lidar or cameras, and (iv) data streams generated by internal vehicle sensors and embedded devices. State-of-the-art machine learning methods are increasingly effective, but model training requires huge amounts of data. This implies the need, among others, for large vehicle status corpuses. On-Board Diagnostics, version 2 (OBD-II) [2] is the standard protocol for real-time access to diagnostic trouble codes and vehicle status parameters.

OBD datasets published by industry players refer to a very limited set of parameters and vehicle models. Research projects have also produced corpuses, but collection costs have strongly limited their size so far and collection has rarely occurred in real driving conditions. The availability of larger and more diverse OBD-II datasets is curbed by the effort required to produce them as well as by the highly competitive nature of the automotive sector, as companies value data corpuses as critical industrial assets and do not share or release publicly them. Holding back data collection, in turn, limits the potential for open research. Conversely, large collections of OBD-II vehicle data would support the creation and improvement of several kinds of Smart Mobility services with positive societal impact, including vehicle maintenance and remote assistance, platooning, fleet management (also for public services such as ambulances and police or fire departments), pollution reduction, and road traffic management.

Crowdsourcing a.k.a. *peer production* is defined as the Internet-mediated decentralized collaboration of large volunteer communities [3]. This paper proposes a framework to apply this paradigm for generating real-world OBD-II data collections. It consists of three elements:

- the extension of the *GPX* [4] open data format for embedding OBD-II parameters into points composing GNSS (Global Navigation Satellite System) route logs;
- the extension of an open source Android mobile client for logging GNSS traces enriched with OBD-II data and uploading them to the Web;
- a Web-based platform, enriching the open source *OpenStreetMap* (OSM: https://www.openstreetmap.org/) software stack, for managing route traces providing collection, search and editing functionalities.

The proposed framework and tools enable car drivers to participate in the creation and evolution of a shared corpus of vehicle status data gathered in fully real driving conditions. This implies the potential to collect OBD-II data with unprecedented scale and variety, in turn enabling a wide range of novel ITS research projects, applications and services, analogously to how location-based services have been leveraging the worldwide OSM cartography.

A working prototype of the overall framework has already been developed and scalability tests have been carried out to evaluate its computational sustainability [5].

II. Crowdsourcing framework

A Web platform has been developed to store and share vehicle data collected through the OBD-II protocol. As shown in Figure 1, the system monitors vehicles equipped with an OBD-compliant scan tool able to expose raw data through a Wi-Fi or Bluetooth connection. Particularly, *PLX Kiwi 2+* (http://www.plxkiwi.com/kiwiwifi/hardware.html) wireless adapter has been exploited in our tests. When turned on, different driving information (*e.g.*, speed, acceleration, fuel consumption, driving style, status of roads and traffic) can be acquired by means of a mobile application and forwarded to the Web platform. The mobile application provides further data related to embedded smartphone micro-devices, such as GNSS location and accelerometer values. Data are gathered within short observation intervals and annotated according to a specific GPX [4] format extension introduced to embed OBD-II and environmental parameters in route descriptions. More details about technologies and tools exploited to define the proposed framework are in the following paragraphs.

Mobile application. *AndrOBD* (https://github.com/fr3ts0n/AndrOBD/wiki) open source Android application has been

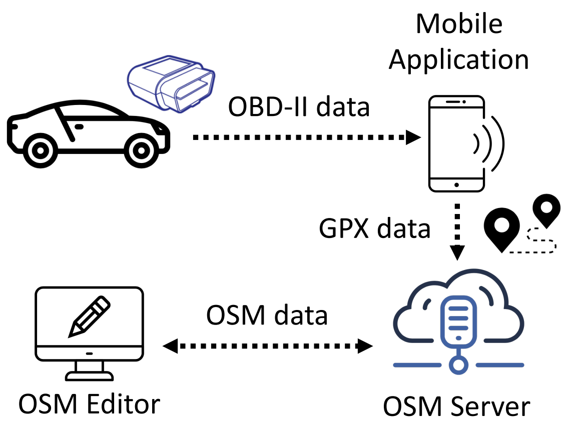

Fig. 1. Framework architecture

extended to record GPS data and collect OBD parameters. Basically, it allows to: (i) show information stored in the vehicle Electronic Control Unit (ECU); (ii) monitor OBD-II parameters in real-time; (iii) export the recorded data in Comma-Separated Values (CSV) files. The original application has been extended to introduce the following functionalities: (a) selection of a variable logging interval from 1 to 5 seconds; (b) tracking of GNSS coordinates, in order to draw a route on a map; (c) compliance with the adopted GPX file format to export logged data; (d) direct upload of recorded routes on the Web platform at the end of a logging session. In particular, through the Graphical User Interface (GUI), the user can select the reference scan tool from a list of nearby Bluetooth devices, start a gathering session and collect vehicle data. OBD parameters supported by the vehicle are listed along with a short description, interpreted value (repeatedly updated) and related measurement unit. The application begins recording data, sampled as specified in the settings, until the user clicks on the *Save* button. Then collected information is exported as a GPX file and uploaded to the Web platform. The user can also define a description for the logged route, visibility criteria and any associated tags.

Web platform. The proposed Web application is based on the OpenStreetMap open cartographic platform, aiming to create and share freely editable data on a worldwide basis. According to the crowdsourcing paradigm, OSM provides a collective effort scheme where users record routes to update and enrich a map. Basically, an OSM map is a data model consisting of the following elements: *nodes*, cartographic points with geographic coordinates; *ways*, sequences of nodes, forming a polyline or polygon; *relations*, groups of nodes, ways and other relations to which specific properties can be assigned; *tags*, key-value labels applicable to nodes, ways or relations. Any change to OSM data must be saved within a *changeset*, representing a group of modifications. OSM has been also enhanced to manage GPX traces containing vehicle data. A user should be able to load a route, show it and make changes without modifying the entire map, which is simply used as cartographic reference. For this purpose the *OSM Database* schema has been extended with: (i) a new `gpx_id` attribute added to a `way` entry for connecting each uploaded GPX file to the reference `way` object; (ii) a new `aux_info` table containing additional data related to each node, such as the values for OBD parameters. Users can set new types of map tags without restriction to accommodate previously unforeseen cartography usage. Nevertheless, following OSM community guidelines is highly recommended. Users can also query or modify stored information exploiting the OSM editor making any authorized contributor able to enrich maps with additional elements. In particular, the Web application user interface provides a direct access to several functionalities: (a) show a simple map and an explanation section listing all the changesets and uploaded routes within the currently displayed area; (b) filter data to select only paths, in the area of interest, including the specified OBD parameters; (c) manually edit a path through a modified version of the OSM *iD* editor (https://github.com/openstreetmap/iD), if needed; (d) export in GPX format all the paths within a selected zone or returned by a search.

III. Conclusion and Future Work

This paper has introduced a framework based on mobile and Web technologies for collaborative OBD-II vehicle status data collection in real driving conditions. It should ground a widespread initiative to facilitate research and development in the Smart Mobility sector. A complete prototype of the framework has been developed by extending an open source Android data logger and the OpenStreetMap Web platform.

Future work aims at optimizing performance, security and privacy aspects before publishing the framework online as a service. Ongoing research concerns the integration of a blockchain platform in the framework to support fleet management with verifiable storage of digests of routes enriched with vehicle and payload status information.

Acknowledgments

This work was supported by Italian Ministry of Economic Development R&D project *BARIUM5G* (Blockchain and ARtificial Intelligence for Ubiquitous coMputing via 5G).

References

[1] K. Abboud, H. A. Omar, and W. Zhuang, “Interworking of DSRC and Cellular Network Technologies for V2X Communications: A Survey,” *IEEE transactions on vehicular technology*, vol. 65, no. 12, pp. 9457–9470, 2016.
[2] K. McCord, *Automotive Diagnostic Systems: Understanding OBD I and OBD II*. CarTech Inc, 2011.
[3] Y. Benkler, “Peer production and cooperation,” in *Handbook on the Economics of the Internet*. Edward Elgar Publishing, 2016.
[4] D. Foster, “GPX: the GPS Exchange Format,” https://www.topografix.com/gpx.asp, 2004, accessed: 2021-06-24.
[5] G. Loseto, F. Gramegna, A. Pinto, M. Ruta, and F. Scioscia, “A Mobile and Web platform for crowdsourcing OBD-II vehicle data,” *Open Journal of Internet Of Things (OJIOT)*, 2021, in press.