Nisrine Osman, Francesco Bellotti, Ricardo Berta, Alessandro De Gloria

DITEN, University of Genoa, Genoa, Italy

L3Pilot Driving Automation

SUMMER SCHOOL 2020

Data Segmentation on Driving Scenarios for L3 Pilot

Building a Toolchain for Data Segmentation on Driving scenarios and Gather for L3 Automation Pilot Test Analysis

Introduction

- L3Pilot is the first comprehensive test of SAE level3 automated driving functions (ADFs) with hands-off the wheel on public roads across Europe [1]. In this project, we contributed in the development of the data tool chain [2], which processes vehicular information converted to a common data format (CDF), which has been designed to seamlessly deal with data from proprietary sources [3]. This data are segmented in driving scenarios on different road types and experimental conditions, then aggregated and posted to a set of local databases and one global consolidated Database (CDB) for information shared among all the vehicle owners [4]. The goal is to answer research questions about the impact of ADFs on road traffic [5] which are generated through the top-down approach recommended by the FESTA Handbook [6]. The driving scenarios segmentation process has been made through algorithms based on expert knowledge. This poster shows the software architecture and workflow we have developed in order to support data processing from the source HDF5 file to the Json files to be posted to the CDB.

The Overall System Workflow

- To unlock the challenge of the high-volume data collected from autonomous vehicle [7-8] testing fields and to enable the seamless and secure sharing of this data, a whole chain of data processing and management took place in the system Architecture. This data is segmented on different driving scenarios and experimental conditions to be then shared into the CDB among analyzers and research partners who participate in the setting of the research questions for the technical & traffic evaluation, user & acceptance evaluation, impact assessment, and socio-economic impact assessment. At a first step, the data is logged on the pilot test sites where different signals from different sensors, lidars, cameras and radars are collected. At a second step, this time series signals data is converted to an HDF5 common format then is post processed and enriched with different derived measures and performance indicators, driving scenarios and external additional information in the form of meta data in the HDF5 file [2-4]. Likewise, the subjective data is collected from different participants at the various pilot sites through pre and post driving questionnaires [9]. These were implemented using the online tool LimeSurvey [10] then translated in each pilot site language. A parallel workflow is supported for the subjective data to convert it to a common format before being transferred to the CDB. The sensitive data like the driver and trip ID was pseudonymized through a simple procedure, based on a SHA-256 hashing to be ready to the final step of data sharing in the CDB. The Figure 1 shows an overall schema of the overall system architecture.

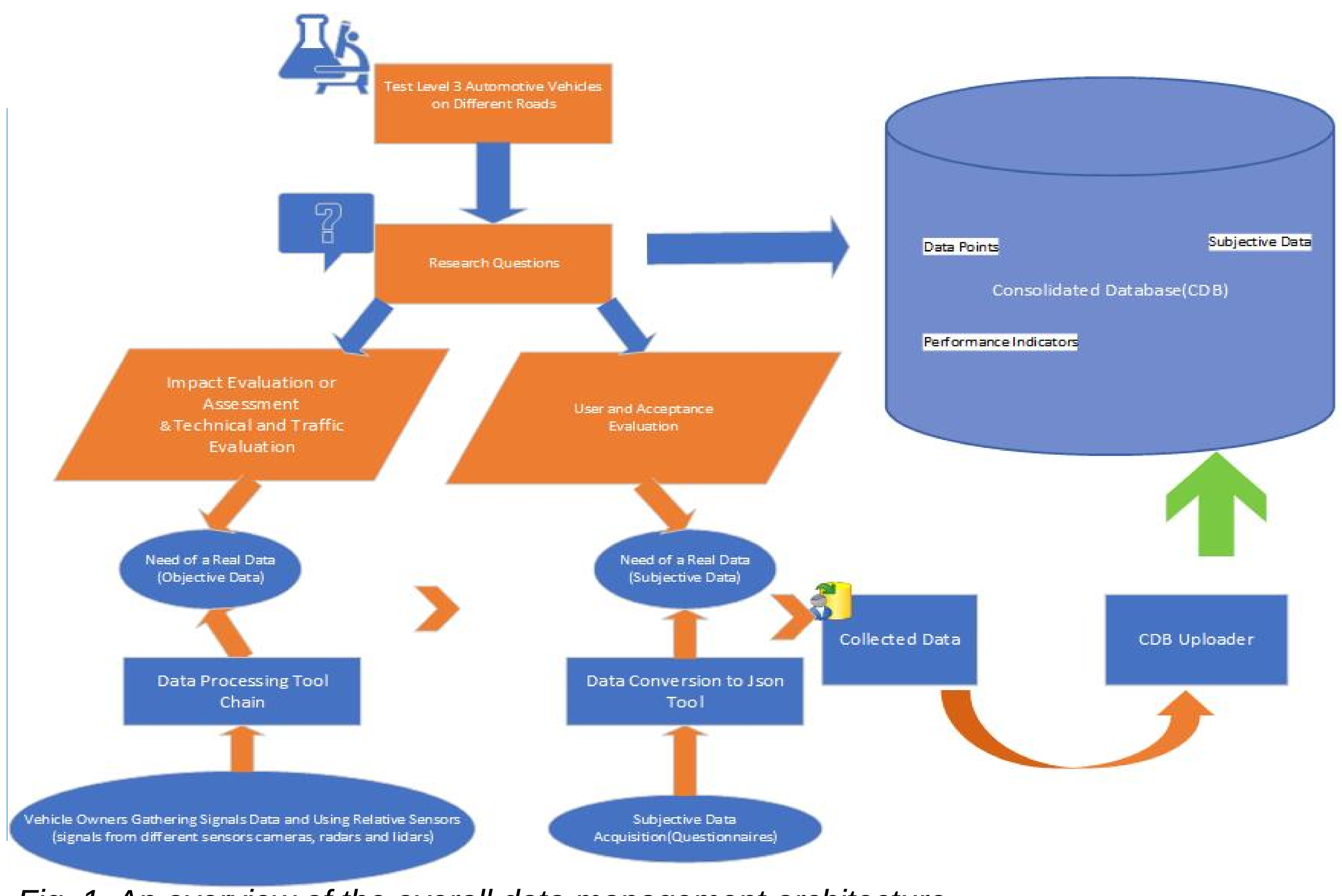

Fig. 1. An overview of the overall data management architecture

The Consolidated Database Toolchain

- The consolidated Database in a transitional step to the impact and socio-economic impact analysis, performs an integration of pseudonymized data collected from the various pilot sites within L3Pilot. It consists of three components. The first is a MongoDB non-rational database that is accessed through a Node.js application programming interface (API). This Database is designed to support the storage and the retrieval of big data for self-driving cars based on the open source Atmosphere framework [11]. It exposes a set of RESTFuL APIs for inserting and retrieving data [12]. The data is screened through a web Graphical User Interface (GUI) [4] and is uploaded through the Uploader, the third component of the toolchain. Three types of user roles are defined to access the consortium-wide shared database that are: An admin who can configures the DB and has the right to access(read, write and cancel) a data from the DB. A provider who can only access his own data and an analyst whose role is limited to view all the data shared on the DB. A MATLAB framework is the second component that supports the computation and segmentation of enriched and processed data provided in the HDF5 common format. Thus, it splits the data on different driving scenarios, road types (parking, urban, motorway), experimental conditions (baseline, treatment, ADF available, active) and specific transitions between treatment conditions during driving e.g. a transition from an ADF active status to and ADF not available. This approach results in high-level data types as following:
  - Performance indicators (PIs): These are statistical descriptors that compute frequencies, min, max, avg and std dev of relevant signals, scenarios or events. E.g. the percentage of the time ADF is available along a whole trip. They are computed either on trip level or on each conditional driving scenario data segment.
  - Data Points per driving scenario instances: These are particular types of data used for the impact assessment by comparing different simulations enabled by the collected data. Some of this data are significant values at the beginning of the scenario instance like the relative speed of the involved vehicles, the positions of the relative objects etc. Others are PIs for the considered scenario instance (e.g., mean speed of the ego vehicle, min and mean longitudinal acceleration).

  Finally, The MATLAB tool converts the post processed data structures to a .json format. It supports also the subjective data that is collected among driving test participants and converts it from excel to a .json format. The Uploader mentioned above is the third component of the toolchain and it consists of Node Js scripts that check and upload the converted Json structured data to the database.

• References

[1] L3Pilot, https://www.l3pilot.eu/, accessed 14 June 2020
[2] Nagy, B.; Hiller, J.; Osman, N.; Koskinen, S.; Svanberg, E.; Bellotti, F.; Berta, R.; Kobeissi, A.; De Gloria, A. Building a Data Management Toolchain for a Level 3 Vehicle Automation Pilot'. In Proceedings of the 26th ITS World Congress, Singapore, 21–25 October 2019
[3] Hiller, J., Svanberg, E., Koskinen, S., Bellotti, F., Osman, N.,: 'The L3Pilot Common Data Format – Enabling Efficient Autonomous Driving Data Analysis', to appear in Proceedings of NHTSA ESV 2019, Eindhoven
[4] Hiller, J., Koskinen, S., Berta, R., Osman, N., Nagy, B., Bellotti, F., Rahman, A., Svanberg, E., Weber, H., Arnold, E.H. and Dianati, M., 2020. The L3Pilot Data Management Toolchain for a Level 3 Vehicle Automation Pilot. *Electronics*, *9*(5), p.809.
[5] Hibberd, D., Louw, T., Aittoniemi, E., Brouwer, R., Dotzauer, M., Fahrenkrog, F., Innamaa, S., Kuisma, A., Merat, N., Metz, B., Neila, N., Penttinen, M., Puente Guillen, P., Rösener, C., Silla, A., Streubel, T., Tango, F., van den Boom, B., Weber, H., Woerle, J., & Zerbe, A. (2018). Deliverable D3.1 From research questions to logging requirements. Deliverable D3.1 of L3Pilot project.
[6] FOT-Net & CARTRE (2018). FESTA Handbook. Version 7. December 2018. Retrieved from https://connectedautomateddriving.eu/wp-content/uploads/2019/01/FESTA-Handbook-Version-7.pdf
[7] Cheng, N., Lyu, F., Chen, J., Xu, W., Zhou, H., Zhang, S. and Shen, X., 2018. Big data driven vehicular networks. *IEEE Network*, *32*(6), pp.160-167
[8] Xu, W., Zhou, H., Cheng, N., Lyu, F., Shi, W., Chen, J. and Shen, X., 2017. Internet of vehicles in big data era. *IEEE/CAA Journal of Automatica Sinica*, *5*(1), pp.19-35.
[9] Metz, B., Rösener, C., Louw, T., Aittoniemi, E., Bjorvatn, A., Wörle, J., Weber, H., Torrao, G., Silla, A., Innamaa, S., Fahrenkrog, F., Heum, P., Pedersen, K., Merat, N., Nordhoff, S., Beuster, A., Dotzauer, M., Streubel, T. (2019). Deliverable D3.3 Evaluation methods. Deliverable D3.3 of L3Pilot project.
[10] LimeSurvey. Available online: https://www.limesurvey.org (accessed on 14 June 2020).
[11] Atmosphere. Availableonline: https://github.com/Atmosphere-IoT-Framework/api org (accessed on 14 June 2020).
[12] Steiner, T.; Algermissen, J. Fulfilling the hypermedia constraint via http options, the http vocabulary in rdf, andlinkheaders. InProceedingsoftheSecondInternationalWorkshoponRESTfulDesignACM,Eindhoven, The Netherlands, 10–13 June 2011; pp. 11–14.

www.L3Pilot.eu Twitter@_L3Pilot_ LinkedInL3Pilot

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 723051.

Supported by the European Council for Automotive R&D.

eucar