An integrated sensing-based urban freight data collection framework: methodology and pilot projects in Singapore

Vittorio Marzano\textsuperscript{a}, Carlos Azevedo\textsuperscript{b}, Lynette Cheah\textsuperscript{c}, Fang Zhao\textsuperscript{d}, Jorge Santos\textsuperscript{a}, Yin Jin Lee\textsuperscript{d}, Moshe E. Ben-Akiva\textsuperscript{a}

\textsuperscript{a} University of Napoli “Federico II” – Department of civil, architectural and environmental engineering (vmarzano@unina.it)
\textsuperscript{b} Singapore-MIT alliance for research and technology (SMART-MIT) – Singapore (cam@mit.edu)
\textsuperscript{c} Singapore University of Technology and Design (SUTD) – Engineering Systems and Design – Singapore (lynette@sutd.edu.sg)
\textsuperscript{d} Singapore-MIT alliance for research and technology (SMART-MIT) – Singapore (fang.zhao@smart.mit.edu)
\textsuperscript{e} Singapore-MIT alliance for research and technology (SMART-MIT) – Singapore (jassant@mit.edu)
\textsuperscript{f} Massachusetts Institute of Technology (MIT) – Cambridge (USA) (yinjin@mit.edu)
\textsuperscript{g} Massachusetts Institute of Technology (MIT) – Cambridge (USA) (mba@mit.edu)

Extended abstract

Objectives and motivation
The paper illustrates the design and preliminary results of an integrated framework for urban freight data collection, under development in the pilot test-site of Singapore, based on next-generation sensing/surveying capabilities to enable future modelling and policy-making in urban freight systems. Specifically, primary objective of the research is to develop proof-of-concept of a coherent, scalable and holistic collection of all freight data (production, logistics, transport), with the aim of tracking vehicles and shipments and for surveying relevant freight agents, leveraging state-of-the-art sensing technologies and approaches, obtaining unprecedented urban freight data. The research is motivated by the strong simplifying assumptions and limited behavioural foundations of current modelling and policy-making for urban freight: the explanatory power of existing freight models is limited by lack of data, implying many feasibility studies and subsequent policies and investments to be based on biased forecasts.

Various research groups are active worldwide in urban freight research. In Europe, the BESTUFS project (http://www.bestufs.net) is one of the best known projects on urban freight, including data harmonization and a published Good Practice Guide on Urban Freight. The Volvo Research and Educational Foundations Center of Excellence on Sustainable Urban Freight (https://www.coef-sufs.org), and the related MetroFreight consortium (http://www.metrans.org/metrofreight) were established in 2013. Their primary objective is urban freight research, education and outreach, and a state-of-the-art data collection effort was carried out in Paris, encompassing an establishment survey, a delivery truck driver survey and major transport companies survey. Other similar data collection methods are being designed for Los Angeles, New York, and Seoul (Seo and Lee, 2014), along with freight-specific policy design case-studies. While some technologies are deployed in this effort, such as GPS for tracking a sample of vehicles, this data collection effort is still traditional with no consistent and integrated modelling effort as yet. In terms of solutions and policy-making, it is worth mentioning the leading-edge pilot case studies for consolidation centres and off-peak deliveries in New York, by the group led by Prof. Holguin-Veras. In Japan, the Tokyo Metropolitan Freight Survey (TMFS) in 2003 involved a questionnaire survey to logistics establishments, to large truck drivers and a local delivery survey of loading/unloading activities in the Central Business District (CBD). Finally, it is worth mentioning the modelling activity by the research group at Kyoto University under the supervision of Prof. E. Taniguchi (Hyodo et al, 2007). However, in spite of ambitious and ground-breaking objectives undertaken by these groups, data collection is still acknowledged as a major limitation in understanding and modelling urban freight (DLR workshop, 2014), along with the lack of proper behavioural models to support micro-simulation agent-based models.

On one side, traditional, manual surveys to a sample of all relevant agents are very costly and time consuming, with questionable quality and validity of the answers, for many reasons: low response rates, non-representative samples based on convenience and intercept sampling, short respondent attention span and limited ability to accurately recall information, inability to capture a variety of important information (e.g. routing of vehicles, shipment paths). On the other side, the “big data” paradigm enables collection of huge amount of passive observations (e.g. truck movements through GPS traces), but underlying logistics and transport choices remain unrevealed, as does key information on explanatory variables (e.g. attributes of the decision-maker, perception of alternatives, decision patterns) needed for behavioural modelling. As a result, current modelling of urban freight is based on strong simplifying assumptions and limited behavioural foundation, thus leading to limited knowledge and biased forecasts. In fact, current microscopic transportation simulation platforms have been restricted to the simulation of very specific logistic models, such as individual shipper’s and carrier’s operational decisions regarding vehicle assignment and routing (Barcelo et al, 2005). Only a few advanced modelling and simulation frameworks focused on multi-modal passenger travelling decisions, such as SimMobility (Lu et al, 2015), have recently taken the first steps towards full freight and logistic simulation integration, but have been struggling with the lack of data to build the necessary models to represent the complex interactions.

General description
The proposed next-generation freight data collection effort leverages pervasive GPS loggers, advanced sensing and communication technologies and machine learning architecture to deliver previously unobtainable data that reflect observed rather than stated information on the decisions of shippers and carriers. The general structure of the proposed data collection framework embeds: (a) sensing devices for tracking shipments, vehicles and driver behavior; (b) a backend server database which collects and processes collected data (c) web-based and tablet-based reporting and surveying tools. The backend system has already been developed for passenger transport at SMART-MIT (Zhao et. al., 2015), with the capability of processing and analysing all types of information (from sensors, surveys, GIS tools), with a scalable architecture. Specifically, raw data from vehicles and shipments are cleaned and map
matched, and coupled with GIS-related information, e.g. point of interests (POI). This enables using machine learning algorithms that infer the user’s stops and activities based on the collected data as well as contextual information such as POIs, user history, weather, incidents etc. The output of these algorithms is displayed in an easy-to-understand web-interface for the user to validate. The user validation is then fed back to the backend to refine the machine learning algorithms. Validated data allow for further personalized visualizations and elaborations (maps, fuel use and carbon footprint analyses) to be shared with survey respondents. The same system can also process all web-based and tablet-based surveys of all relevant freight agents (shippers, logistics operators, carriers, end-consumers), transmitting information from GIS sources to the surveys and collecting back GIS-located survey data.

Importantly, acknowledged that an urban commodity flow survey is needed to understand commodity generation (as a function of establishment characteristics), as well as the distribution and efficiency of the urban logistics systems when coupled with freight mode choices and capacity utilization, the key value of the proposed research is also to provide a holistic synthesis of all data collection activities, e.g. the commodity flow survey will be coupled with passive observations of shipments, and truck drivers’ surveys will be coupled with GPS/OBD-based vehicle observations, forming an innovative framework that we call commodity flow sensing (CFS).

Along with the CFS, i.e. the core of the proposed integrated data collection framework, additional efforts are considered, either for data scalability or for getting insight on specific features of the logistic system with localized high resolution data. The fusion between the CFS data and these additional efforts has been targeted to:

- collect all needed data/information for the generation of the entire population of freight agents (establishments, retailers, end-consumer, transport providers, …) for statistical and modelling purposes;
- network-wide commercial vehicle flow measurements, where our team has carried out a feasibility study to derive traffic counts and vehicle classification from traffic camera images or videos. Such data will allow us to determine partial o-d observation through automatic identification of vehicles at different locations and the calibration of estimated vehicle flows;
- detailed operational high-resolution data, with a feasibility study for camera and sensor-based observations and driver interviews at loading/unloading bays;
- implementing innovative truck drivers’ surveys using tablets on-board the vehicles, coupled with GPS and On-Board Diagnostics (OBD) technologies plus vehicle telematics for real-time analysis;
- detailed driving cycle and vehicle performance data including instantaneous fuel consumption, engine RPM, air intake temperature, etc, collected within the logger-based truck drivers survey by introducing OBD devices. This additional information enables one to gain insight on fuel use, emissions, and truck idling.

Results and conclusions
The project is in its implementation stage, and some of the proposed approaches are under testing. Promising results have been already achieved, with specific reference to the capability of unobtrusively collect a wealth of valuable information, leading to better quality and quantity of data. Specifically, completed activities include:

- observations at loading/unloading bays in Bugis and Orchard, two central shopping and business districts in Singapore;
- a pilot of a combination of tablet-based and traditional surveys, investigating both within-mall and outside-mall retailers, with 210 tablet-based pilot surveys in Spring 2014 in commercial districts and 555 surveys in Summer 2014 island-wide;
- an urban truck route choices study, aiming to track a variety of commercial vehicles with more frequent stops and more diverse activities.

Focusing on the Singapore context, the ongoing research designs and demonstrates scalable collection of high-resolution data on freight agents’ behaviours/choices and freight movements, not captured through traditional surveys, leading to integrated knowledge of logistics and transport characteristics of shipments. It sets the foundation for continued freight data collection to better understand long-term trends in city-scale commodity flows and urban goods movements.

References

Keywords: urban freight data collection; GPS and OBD based surveys; urban commodity flow survey.