Improving public transport decision making, planning and operations by using Big Data
Cases from Sweden and the Netherlands

dr. Niels van Oort
Transport and Planning
Delft University of Technology
Delft, the Netherlands
N.vanOort@TUDelft.nl

dr. Oded Cats
Transport and Planning
Delft University of Technology
Delft, the Netherlands
O.Cats@tudelft.nl

ABSTRACT— **New big data (sources) in the public transport industry enable to deal with major challenges such as elevating efficiency, increasing passenger ridership and satisfaction and facilitate the information flow between service providers and service users. This paper presents two actual cases from the Netherlands and Sweden in which automated data sources were utilized to support the planning and operational processes. The cases illustrate the benefits of using smartcard and vehicle positioning data. Due to the data (processing), valuable insights were gained helping to make the right choices and improve the public transport system.**

Keywords—Public transport, Big data, Smartcard data, Vehicle positioning

I. INTRODUCTION

The public transport industry is facing several challenges. It is a conservative industry in the age of rapid change and information. Passengers require higher quality of the services and at the same time there is more attention to cost efficiency and subsidies allocation. Concerning quality improvements, robustness and service reliability gain more attention [1,2]. Public transport systems are increasingly equipped with automated data collection systems that can be instrumental in addressing these challenges. For instance, the design and refinement of cost-effective measures can be supported by measuring, computing and projecting the expected impacts on costs and passengers. Furthermore, the availability of data sources such as mobile phone data [3] can provide insights into passenger demand of all modes, thereby enabling the identification of potential public transport connections. Automated Vehicle Location (AVL) data, has already been available for a long time (e.g. [4,5]) and recently much more passenger data (Automated Passenger Counting (APC) data) has become available as well [6]. These data sources support public transport design and decision making, in addition to efficient and high quality operations.

Data could be used to gain better understanding of passenger needs and behavior, system performance and real-time conditions. Moreover, data enable planners to investigate and quantify the costs of service deficiencies and the potential benefits of alternative solutions, for instance the extension of a turning facility [7] or timetable synchronization [8]. These costs and benefits are relevant for decision making and may be incorporated in cost-benefit analyses [9].

The aim of this paper is to provide an overview of the potential public transport data sources and illustrate their value through Dutch and Swedish cases. The following section starts with a short introduction of data sources. Section III describes the role of data in the planning process and in Section IV our cases are presented. Finally, conclusions and further research opportunities are given in Section V.

II. BIG DATA IN PUBLIC TRANSPORT

A. Data sources

A single bus equipped with AVL transmits its positioning every 15 seconds, amounting to approximately 3000 positioning records per day. The daily number of smartcard transactions are in the millions in large metropolitan areas. Not to mention social media data, such as user data of Twitter, Facebook and Flickr, which may yield new insights on public transport usage [10]. Furthermore, video cameras (e.g. surveillance systems in stations and on-board vehicles), Wi-Fi and Bluetooth trackers may provide knowledge of pedestrian flows in stations, at platforms and in-vehicle [11]. Sensors connected to different types of assets, signals and switches for instances, enable optimization of maintenance schemes. In this paper, we focus on the traditional and basic data sources in public transport, namely passengers (APC) and vehicles (AVL) data. These data sources are most prevalent. Nevertheless, their applications are typically limited to performance measurement. Recent developments created a huge leap in the availability and applicability of this kind of data, as will be demonstrated in Sections III and IV of this paper, following the description of these data sources.

B. Automated vehicle location systems

Automated Vehicle Location Systems (AVL) were originally installed for monitoring safety in railway operations. However, AVL emerged as the primary vehicular-data and are often used for analyzing public transport performance in terms of commercial speed and service reliability (e.g. [12,13]). AVL systems are either time-based or event-based – implying that vehicles either transmit information on their positioning with fixed time intervals or when a change occurs (e.g. driving, stopping, doors open), respectively. Management report programs (several tools are available, see for instance [14]), transfer the board computer data into information (e.g. graphs and tables), showing for example schedule adherence and speed. Although, traditionally, this type of monitoring is performed off-line, recent developments also enable real-time loops for monitoring services. Accurate real-time vehicle location data has become available for public transport operators with the wide availability of GPS and mobile phone devices. Control center operations facilitate real-time interventions based on instantaneous information flow and real-time predictions.

C. *Automated passenger counts and fare validation*

For analysis, design and optimization of public transport, actual and future demand are essential. The number of passenger (-kilometers) in the network, per line and per stop are crucial. In addition to traditional counting, smart cards can (partly) provide a richer data source albeit they require path inference. In recent years, counting have become available for different service, temporal and user group segments.

The major advantages of smart card data for transport service providers were identified by [15] as

- Large volumes of personal travel data.
- Market analysis of travel patterns of individual card holders and user groups
- Having access to continuous trip data covering longer periods of time.

Depending on the exact characteristics of the system, more insights may be gained. The number of areas where smart cards are applied and analyzed increases rapidly. Prominent examples are London (Oyster card) and Hong Kong (Octopus card), but many more examples are presented in literature (e.g. Seoul [16], Beijing [17], Santiago de Chile [18], Shenzen [19] and Brisbane [20]). An overview by [6] describes a range of smart card data applications, varying from strategic and tactical planning optimization to operational improvements. Most applications aim at assessing OD-patterns [18,21], route choice behavior [22] and transfer analysis [23].

III. FROM DATA TO APPLICATIONS

Big data can foster innovation in both planning and operation of public transport services. The abundance and diversity of data that is collected in the daily operations of public transport systems calls for a data utilization framework in order to support its utilization. The public transport planning process consists of strategic, tactical, operations and control decisions (see Figure 1 for the planning process including feedback [24]). Big data is instrumental throughout this process. The following sections describe the role of big data in off-line (long term) and real-time (short term) applications.

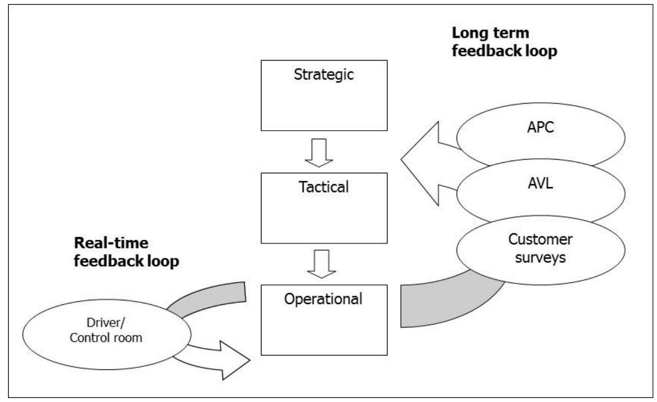

Fig. 1. Planning and operations of public transport [24]

A. *Offline applications: Strategic and tactical planning*

Strategic planning is concerned with the long-term planning of overall network and service design, including stop positioning, line topology and the design of respective capacities. Thereafter, tactical planning specifies the mid-term service frequencies, timetable development, and vehicle and crew scheduling. These offline applications require large amounts of data which are conventionally obtained from large and costly travel habit and stated preference surveys. The availability of large amounts of APC and AFC data facilitates the development of methods and algorithms to directly construct, instead of estimate, travel demand based on observed travel patterns and trip inference algorithms [17,18, 25,26].

The growing availability of data sources paves the way to unravelling travel patterns based on automatically collected mobility data. For example, smartcard data was used for analyzing temporal and spatial variations of destination choices in London and Shenzhen [19, 27] and identifying transfer and activity locations in Brisbane [20]. Moreover, several recent studies used public transport flows at the urban area level to shed light on the underlying urban structure [e.g. 28].

The analysis of big data can also support continuous evaluation of service performance, such as the analysis of seasonal effects and system-wide demand variations [29]. Similarly, analyzing data from passenger satisfaction survey can reveal how satisfaction and its determinants evolve over the study period, as demonstrated by [30] which analyzed more than half a million records that were collected between 2001-2013 in Sweden.

The tactical planning of timetables and resource allocation require large-scale and detailed datasets. A report by the Transit Cooperative Research Program in the US [4] provides a review of offline applications of AVL and APC data for public transport service management. The analysis of within-day and day-to-day variations is essential for planning a robust timetable. Data clustering techniques can support the design of the minimal number of distinctive timetables required for operations [31].

B. *Online applications: Operations and control*

Operations and control involve the deployment of real-time strategies to improve service performance. Real-time strategies include the dissemination of travel information, control (e.g. holding, stop skipping), fleet management (e.g. short turning, deadheading) and priority measures (e.g. traffic signal priority, dynamic lanes). The deployment of such strategies require instantaneous access to large amounts of online data and tools to assess and implement alternative measures in real-time. These applications often involve generating predictions on future system conditions.

Predicting the progress of public transport vehicles requires the collection, integration and processing of instantaneous and historical data, which are then used as an input to prediction algorithms [32]. Information concerning current traffic conditions could be generated by analyzing traffic data (such as traffic camera counts, loop detectors, plate recognition or floating car data [33]), public transport vehicular data or integrating these two sources. Various machine learning techniques have been applied, due to their capability to utilize large amounts of data, to reveal complex patterns and to address noise in data streams.

Real-time operations and control decisions are especially critical in case of service disruptions. Offline analysis of smartcard data can be used to analyze passenger behavior during severe disruptions [34]. Such a model can be used for predicting the consequences of information provision or supply adjustments to mitigate the impact of disruptions within a short decision horizon.

Fleet management and real-time control strategies can vary greatly in the amount of data that needs to be processed as part of the decision making process. Methods that are more technology-enabled and are more data-hungry typically perform better (e.g. [35]). Ultimately, prediction schemes and tools to evaluate the performance under alternative scenarios will be integrated in a decision support system which evaluates and implements monitoring, real-time scheduling, disruption management and information provision schemes, as was for example developed in [36] for a bus rapid transit corridor.

IV. APPLICATIONS

The previous section described the wide range and ample potential of big data in planning and managing public transport systems. Previous studies demonstrated the benefits of specific applications. In the following, we present two series of integrative applications that were implemented in the planning practices of public transport authorities and operators.

A. *Passenger behaviour estimation and ridership prediction*

In the Netherlands, public transport operators started to develop a smart card system in 2001. The system was introduced in Rotterdam only in 2005 and in 2012 the full country was equipped [37]. The Dutch smart card uses NFC (near filed communication)-chip technology and passengers have to check in and check out when using different vehicles and operators. All public transport, including national train services, is accessible with the same smart card. Thus, valuable information is measured about origin-destination patterns (on station/stop level) of all public transport users. In the Netherlands, card validation devices are either located on the platform (for trains and metros) or located inside the vehicle (for buses and trams). The most detailed information is available in the latter case, where each trip leg in a journey - a journey may consist of multiple trip legs separated by interchanges - is tracked, whereas when the smart card devices are located on the platforms, information is only available of the first and the last station, making route search through the public transport network necessary for the analyst (e.g. [22]). The complete passenger journey can be therefore traceable.

Our first case of applying data focuses on predicting ridership by applying smartcard of HTM, the tram operator in The Hague (about 500,000 inhabitants, 3rd largest city of the Netherlands). The city of The Hague has 12 tram/light rail lines with a total network length of about 335km. Checking-in and out is done in the vehicle. We applied the smartcard data in two ways:

- Inferring passenger behavior (revealed preference)
- Constructing a reference network load as a basis for what-if predictions

The first example consists of determining the optimal set of parameters used in public transport forecast models. Transit assignment models distribute passenger flows over the public transport network by superimposing passengers' route choice decisions. Route choice models are conventionally estimated based on stated preference surveys. By using revealed travel data, we investigated which values of different model parameters are most adequate for predicting passenger loads and lead to best model calibration results. Moreover, passengers' preferences may differ in case of disruptions from normal operations.

Route choice model estimation were applied for investigating the consequences of a planned disturbance (e.g. due to construction works). For large maintenance works on a tram line in The Hague, we examined the effect on demand and route choice using smartcard data by comparing the number of trips made on the HTM tram network as a whole and for each line separately between the undisrupted and disrupted scenarios. After correcting for structural differences between these periods which cannot be attributed to the maintenance works, we determined the empirical mode and route choice effects which could be attributed to the specific disruption. We tested a variety of parameter sets to obtain model predictions as close as possible to the realized travel patterns.

Results showed that a 25% higher value of the elasticity parameter (more negative) was required for a better fit between realized and predicted mode and route choice in case of the disturbed situation compared to structural changes. This shows that passengers are more sensitive to changes in supplied quality (expressed in generalized travel costs) during planned disruptions, compared to the structural undisrupted situation. This might be explained because of lack of knowledge or additional disutility associated with temporary services.

The second part involved the construction of a simple prediction model based on smart card data and an existing modelling software, Omnitrans. This model enables for analysis of what-if analyses by using transport planning software. Detailed insights into this approach are presented in [38]. The smart card data is converted to passengers per line and origin-destination matrix between stops. This matrix is assigned to the network (in the existing transport model) to reproduce the measured passenger flows (see Figure 2). Once the assignment can reproduce the passenger flows simple what-if analysis can be examined. With the introduction of an elasticity method on the demand matrix, simple modal-split calculations are possible.

Figure 3 shows the outcomes (in terms of change in passenger load) of an example of a specific timetable adjustment: a frequency increase of two public transport lines. The main contribution of this method is that we can calculate the network impacts in detail. In addition to the affected line, we can also see how other lines are affected. Figure 3 shows the expected ridership growth after this adjustment on these two lines (green) as well as a decrease in ridership on a parallel line (red).

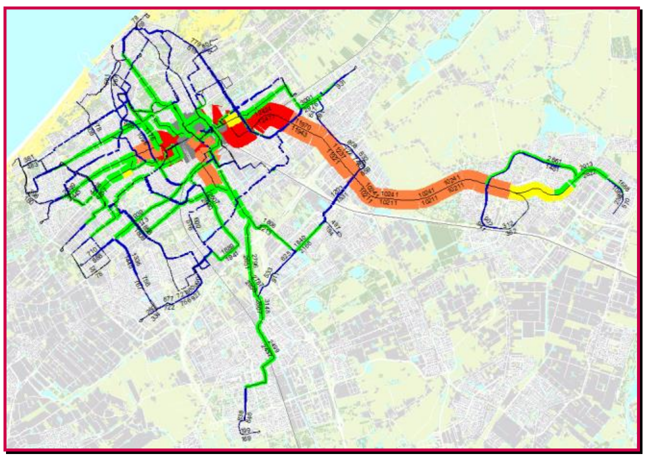

Fig. 2. Example of passenger flows in The Hague based on smart card data

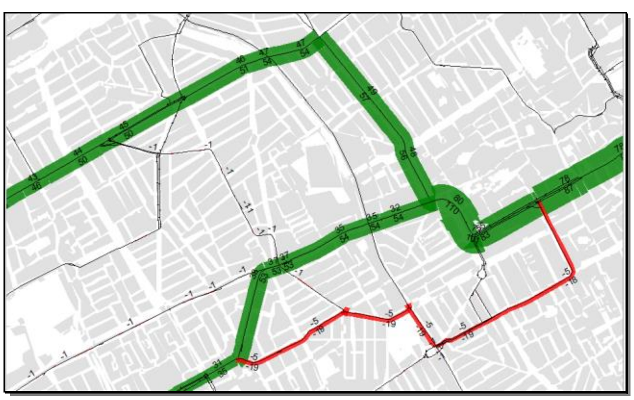

Fig. 3. Example of whatif results: the impact of a frequency increase of two PT lines on ridership (green = increase; red= decrease)

The tool turned out to be very valuable for the operator to gain insights into small changes and calculate both the costs and benefits of such a measure at the network level. However, the approach has some limitations and shortcomings. First of all, the elasticity method is only valid for short-term predictions and only unimodal (public transport) results are provided. We recommend further research on region-specific elasticity values. With the availability of smart cards, valuable revealed preference research is possible by performing before-after analysis following service changes.

B. Real-time information provision and control strategies

The entire bus fleet in Stockholm is equipped with an AVL system. The system is used for several purposes including radio communication, real-time monitoring and control of vehicles, fleet management strategies and the generation of real-time passenger information. Moreover, the real-time vehicle location probes are automatically processed for computing service performance indicators such as punctuality and efficiency. In addition, 10-15% of the bus fleet is equipped with APC devices which are circulated to sample from all services. The introduction of a new fare collection system based on an entry-only smartcard validation, enables the spatial and temporal analysis of travel demand generation.

The backbone of the bus network in Stockholm inner-city consists of four trunk lines. A series of big data applications were performed for these lines with the objective of improving service reliability by

- Developing and evaluating operational and control strategies
- Developing and evaluating algorithms for generating real-time passenger information

Service reliability was first evaluated at the stop, line and system levels using detailed AVL data. In particular, the analysis of service regularity – variation in headways between consecutive bus arrivals – require data from each individual vehicle. The AVL data was integrated with APC data in order to investigate the relation between regularity and crowding. As illustrated in Figure 4, passenger loads vary greatly as a result of irregular headways, leading to a poor capacity utilization and on-board congestion.

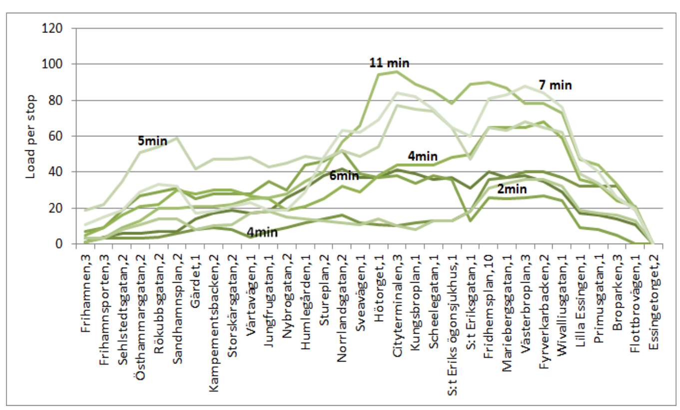

Fig. 4. Passenger loads on consecutive bus trips on the westbound direction of trunk line 1 in Stockholm and their respective headways

Communication and monitoring systems enable the design and implementation of real-time control strategies which are based on actual conditions as well as anticipated downstream conditions. A strategy based on the headways between consecutive vehicles was tested and refined in a simulation model based on local empirical data [39] and was then implemented in a series of field experiments in Stockholm.

Each of the field experiments was evaluated by considering its impact on total passenger travel time. The computation of waiting and in-vehicle times require detailed and comprehensive data on vehicle arrival times and passenger flows. The AVL and APC data were processed in order to estimate travel time and headway distributions, and origin-destination matrixes, respectively. The overall headway distribution is presented in Figure 5 for all observed headway throughout the line for the entire day time (7:00-19:00). The range of acceptable headways – up to 50% difference from the planned headway of 5 minutes – is marked. Headway variability decreased significantly and the service became much more regular in 2014 where headway-based control was in place. Headway distribution became narrower during the field experiment period with a large decrease in cases of extremely short or extremely long headways. The number of headways close to the planned headway increased.

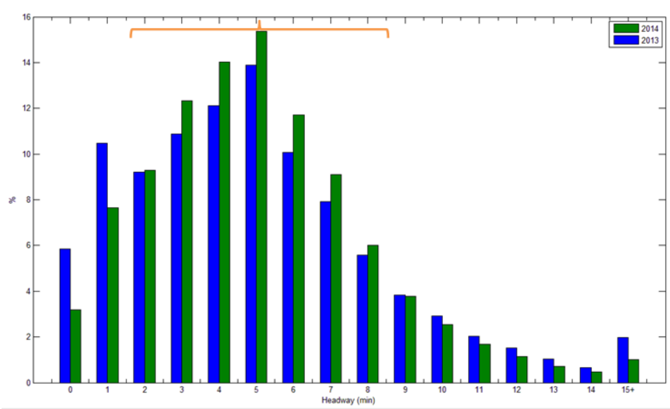

Fig. 5. Headway distribution on line 4 in Stockholm, 7:00-19:00

The origin-destination matrix was estimated for each line using the iterative proportional fitting method. The total passenger travel times savings were obtained by comparing for each origin-destination the waiting time and in-vehicle times in the before and after periods and assessing their monetary values. The full-scale implementation of the new real-time control scheme was supported by the introduction of new incentive schemes that rely on AVL data analytics [40].

Real-time control schemes require online predictions on downstream service conditions. Furthermore, real-time passenger information rely on such predictions. Passengers' perception of service reliability depends not only on service provision but also on information provision reliability. The accuracy and reliability of the current system was evaluated by analyzing big data concerning vehicle trajectories and the corresponding predictions that were generated by the prediction scheme [41]. A cross-network sampling of bus arrival information was compared with vehicle positioning data and enabled the analysis of the added-value of real-time information provision as compared with static information and identify its shortcomings.

The analysis of big data facilitates the development of more elaborate prediction schemes which takes into consideration the current traffic, fleet and travel demand conditions. By mining historical data and integrating information from downstream traffic conditions and the schedule, a hybrid scheme was developed and yielded better predictions [42]. It should be stressed that in addition to passenger information provision, better online predictions are also instrumental in selecting and deploying real-time operations and control strategies.

V. Conclusions

The abundance of data in the public transport industry facilitates addressing the major challenges such as enhancing efficiency, increasing passenger ridership and satisfaction. This paper presents two actual cases from the Netherlands and Sweden in which data sources were successfully applied throughout the planning and operational processes. The cases showed the benefits of both smartcard and vehicle data (APC and AVL). Big data analytics resulted with valuable insights and supported improvements in the public transport system. AVL and APC data were designed to support fleet monitoring and revenue distribution, which could hinder gaining new information and knowledge from Big data analytics.

In addition to the value of using individual data sources as demonstrated by the applications, the combination of data sources might provide synergic advantages. For instance, combining passenger and vehicle data enables gaining insights into passenger reliability. New data sources like mobile phones and Wi-Fi/Bluetooth tracking can also shed light on individual travel patterns. Mobile phone data (showing traveler flows of all modes) may reveal modal shares thereby illustrating weak spots in the public transport network. Wi-Fi/Bluetooth data shows passenger flows in a micro level, for example within an interchange hub. All these data sources may be combined in an urban mobility lab. The authors are currently constructing such a lab in Amsterdam with the objective to collect and process multimodal data aiming at understanding and modelling traveler behavior.

Acknowledgment

The authors are thankful for the data and tooling provided by HTM The Hague, SLL Stockholm and Goudappel Coffeng.

References

[1] N. van Oort, N., T. Brands, E. de Romph, J.A. Flores, Unreliability effects in public transport modelling, International Journal of Transportation Vol.3, No.1, 2015, pp.113-130.

[2] O. Cats, M.D. Yap, N. van Oort, Exposing the role of exposure: identifying and evaluating critical links in public transport networks, INSTR conference, Nara, Japan, 2015.

[3] Calabrese, Francesco et al.. Estimating origin-destination flows using mobile phone location data. In: Pervasive Computing, IEEE. 2011.

[4] P.G. Furth, B. Hemily,, T.H.J. Muller, J.G. Strathman TCRP Report 113: Using Archived AVL-APC Data to Improve Transit Performance and Management. Washington, D.C. 2006.

[5] M. Hickman, Evaluating the Benefits of Bus Automatic Vehicle Location (AVL) Systems, in: D. Levinson and D. Gillen (eds.), Assessing the Benefits and Costs of Intelligent Transportation Systems, Chapter 5, Kluwer, Boston, 2004.

[6] M. Pelletier, M. Trepanier and C.Morency. Smart card data use in public transit: A literature review. Transportation Research Part C: Emerging Technologies, Vol. 19, No. 4, 2011, pp. 557-568.

[7] N. van Oort, N. and R. van Nes, The impact of rail terminal design on transit service reliability, Transportation Research Record, No. 2146, 2010, pp. 109-118.

[8] A. Lee, N. van Oort, R. van Nes, Service reliability in a network context, Transportation Research Record, No. 2417, 2014, pp. 18-26

[9] N. van Oort, Quantifying benefits of enhanced service reliability in public transport. Presented at the 12th Conference on Advanced Systems for Public Transport (CASPT), Chile, 2012.

[10] S. Bregman, Uses of Social Media in Public Transportation. Transit Cooperative Research Program (TCRP) Synthesis 99. Transportation Research Board, Washington, 2012.

[11] J. van den Heuvel, J., A. Voskamp, W. Daamen, S.P. Hoogendoorn, Using bluetooth to estimate the impact of congestion on pedestrian route choice at train stations, *Traffic and Granular flow '13*, M.Charaibi et. al (eds), Switzerland, 2015.

[12] T.H.J. Muller, P.G. Furth, Trip Time Analyzers: Key to Transit Service Quality, *Transportation Research Record*, No. 1760, 2001, pp. 10-19.

[13] J. Strathman, T. Kimple, K. Dueker, R. Gerhart and S. Callas, Evaluation of transit operations: data applications of Tri-Met's automated bus dispatching system, *Transportation 29*, 2002, pp. 321-345.

[14] N. van Oort, D. Sparing, T. Brands, R. Goverde. Optimizing Public Transport Planning and Operations Using Automatic Vehicle Location Data: The Dutch Example. Presented at the 3rd International Conference on Models and Technology for ITS, Dresden, Germany, 2013.

[15] M. Bagchi, P. White. The Potential of Public Transport Smart Card Data. *Transport Policy*, Vol. 12, No. 5, 2005, pp 464-474.

[16] J. Park, D.J. Kim, D.J., Y. Lim. Use of Smart Card Data to Define Public Transit Use in Seoul, South Korea. Transportation Research Record, Journal of the Transportation Research Board, No. 2063, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp.3-9.

[17] X. Ma, Y.J. Wu, Y. Wang, F. Chen, J. Liu. Mining smart card data for transit riders' travel patterns. Transportation Research Part C: Emerging Technologies, Vol. 36, 2013, pp. 1-12.

[18] M. Munizaga, C. Palma. Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smart card data from Santiago, Chile. *Transportation Research C*, Vol. 24, 2012, pp. 9-18.

[19] S. Hasan, C.M. Schneider, S.V. Ukkusuri and M.C. Gonzalez, "Spatiotemporal patterns of urban human mobility", J. of Statistical Physics, vol. 151 (1-2), 2013, pp. 304-318.

[20] N. Neema, Hickman M. and Z-L. Ma, "Activity detection and transfer idenitifiaction for public transit fare card data", Transportation, Vol. 42, No. 4, 2015, pp. 683-705.

[21] M. Trépanier, N. Tranchant, R. Chapleau. Individual trip destination estimation in a transit smart card automated fare collection system. *Journal of Intelligent Transportation Systems*, Vol. 11, 2007, pp. 1-14.

[22] J.D. Schmöcker, H. Shimamoto, F. Kurauchi. Generation and calibration of transit hyperpaths. *Transportation Research Part C: Emerging Technologies*, Vol. 36, 2013, pp. 406-418.

[23] C. Seaborn, J. Attanucci, N.H.M. Wilson. Analyzing Multimodal Public Transport Journeys in London with Smart Card Fare Payment Data. *Transportation Research Record, Journal of the Transportation Research Board, No. 2121*, Washington, D.C., 2009, pp. 55–62.

[24] N. van Oort, Service Reliability and Urban Public Transport Design, T2011/2, TRAIL PhD Thesis Series, Delft, 2011.

[25] J. Zhao, A. Rahbee and N.H.M. Wilson, "Estimating rail passenger trip origin-destination matrix using automatic data collection systems", Computer Aided Civil and Infra. Eng., Vol. 22, pp. 376-387, 2007.

[26] Y. Ji, Q. You, S. Jiang and H.M. Zhang, "Statistical inference on transit route-level origin-destination flows using automatic passenger counter data", J. of Advanced Trans., DOI: 10.1002/atr.1301, 2015.

[27] L. Liu, A. Hou, A. Biderman, C. Ratti and J. Chen, "Understanding individual and collective mobility patterns from smart card records", 12th International IEEE Conference on Intelligent Transport Systems, pp. 1-6, 2009.

[28] C. Roth, S.M. Kang, M. Batty and M. Barthelemy, "Structure of urban movements: Polycentric activity and entangled hierarchical flows", PLoS ONE, Vol. 6(1), 2011.

[29] M. Berkow, A.M. El-Geneidy, Bertini R.L. and D. Crout, "Beyond generating transit performance measures", Trans. Res. Rec., vol. 2111, pp. 158-168, 2009.

[30] O. Cats, R.F. Abenoza, C. Liu and Y.O. Susilo, "Identifying priority areas based on a thirteen years evolution of satisfaction with public transport and its determinants", 94th Annual Meeting of the Trans. Res. Board, 2015.

[31] L. Matias. J. Gama, J. Mendes-Moreira and J. Freire de Sousa, "Validation of both number and coverage of bus schedules using AVL data", 13th International IEEE Conference on Intelligent Transport Systems, pp. 131-136, 2010.

[32] F.W. Cathey, Dailey D.J, "A prescription for transit arrival/departure prediction using automatic vehicle location data", Trans. Res. Part C, vol. 11, pp. 241-264, 2003.

[33] M. Rahmani, Koutsopoulos H.N. and Ranganathan A., "Requirements and potential of GPS-based floating car data for traffic management: Stockholm case study", 13th International IEEE Conference on Intelligent Transport Systems, pp. 730-735, 2010.

[34] D. Yokoyama, Itoh M., Toyoda M., Tomitta Y., Kawamura S. and Kitsuregawa M., "A framework for large-scale train trip record analysis and its application to passengers' flow prediction after train accidents", In Advanceds in Knowledge Discovery and Data Mining, Lecutre Notes in Computer Science, vol. 8443, 2014, pp. 533-544.

[35] M. Dessouky, Hall R., Zhang L. and Singh A., "Real-time control of buses for schedule corrdiantion at a terminal", Trans. Res. Part A, Vol. 37, No. 2, 2003, pp. 145-164.

[36] G. Xiong, Dong X., Fan D. and Zhu F., "Parallel bus rapid transit (BRT) operation management system based on ACP approach", 9th International IEEE Conference on Networking, Sensing and Control, 2012, pp. 22-27.

[37] F. Cheung, Implementation of Nationwide Public Transport Smart Card in the Netherlands: Cost-Benefit Analysis. *Transportation Research Record, Journal of the Transportation Research Board, No. 1971*, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 127-132.

[38] N. van Oort, T. Brands, E. de Romph , Short term ridership prediction in public transport by processing smart card data, Transportation Research Record, 2015 (in press)

[39] O. Cats O., Larijani A.N., Ólafsdóttir A., Burghout W., Andreasson I. and Koutsopoulos H.N.. Holding Control Strategies: A Simulation-Based Evaluation and Guidelines for Implementation. Transportation Research Record, 2274, 2012, pp.100-108**.**

[40] O. Cats, Regularity-Driven Bus Operations: Principles, Implementation and Business Models, Transport Policy, vol. 36, pp. 223-230, 2014.

[41] O. Cats, Loutos G. "Real-Time Bus Arrival Information System – An Empirical Evaluation", Journal of Intelligent Transportation Systems, In press, DOI: 10.1080/15472450.2015.1011638 , 2015.

[42] M. Fadaei Oshyani and O. Cats, "Real-time bus departure time predictions: Vehicle trajectory and countdown display analysis", 17th International IEEE Conference on Intelligent Transport Systems, 2014, pp. 2556-2561.