SHORT TERM RIDERSHIP PREDICTION IN PUBLIC TRANSPORT BY PROCESSING SMART CARD DATA

Prepared for the 94th Annual Meeting of the Transportation Research Board 2015

Dr.ir. Niels van Oort
Delft University of Technology / Goudappel Coffeng
Faculty of Civil Engineering and Geosciences
Transport & Planning
P.O. Box 5048
2600 GA Delft, The Netherlands
Telephone: +31.6.15908644
E-mail: N.vanOort@TUDelft.nl*(corresponding author)*

Ir. Ties Brands
University of Twente / Goudappel Coffeng
Faculty of Engineering Technology
Centre for Transport Studies
P.O. Box 217
7500 AE Enschede, The Netherlands
Telephone: +31 53 489 4704
E-mail: tbrands@goudappel.nl

Prof.dr. ir. Erik de Romph
Delft University of Technology / DAT.Mobility
Faculty of Civil Engineering and Geosciences
Transport & Planning
P.O. Box 5048
2600 GA Delft, The Netherlands
Telephone: +31.6.20604452
E-mail: E.deRomph@TUDelft.nl

July 2014

Word count:
Abstract (226) + Text (5055) + Figures/Tables (7*250) = 7031

ABSTRACT

Public transport operators are exposed to massive data collection from their smart card systems. In the Netherlands, every passenger needs to check in and to check out, resulting in detailed information on the demand pattern. In buses and trams, checking in and checking out takes place in the vehicle, providing good information on route choice. This paper explores options for using this smart card data for analysis and performing what-if analyses by using transport planning software. This new generation of transport demand models, based on big data, is an addition to the existing range of transport demand models and approaches. The intention is to provide public transport operators with a simple (easy to build) model to perform these what-if analyses. The data is converted to passengers per line and an OD-matrix between stops. This matrix is assigned to the network to reproduce the measured passenger flows. After this step, what-if analysis becomes possible. With fixed demand, line changes can be investigated. With the introduction of an elastic demand model, changes in level of service realistically affect passenger numbers. This method was applied on a case study in The Hague. We imported the smart card data into a transport model and connected the data with the network. The tool turned out to be very valuable for the operator to gain insights into the effect of small changes.

Keywords: public transport, smart card data, demand modeling, forecasting

1. INTRODUCTION

Recently, many cities and regions introduced a smart card system for their public transport systems (e.g. *1, 2, 3, 4* and *5*). In addition to ticket handling, being an alternative for individual regional or urban tickets, these systems also provide valuable data. Without these systems, information of origins and destinations, number of passengers, trip lengths, etc. can only be made available by time and budget consuming surveys. These surveys often only provide limited data due to time and budget restrictions. Smart card systems have the potential of providing more and better insights of revealed passenger behavior. These insights are helpful when dealing with the main current challenges in the public transport industry.

Within the public transport industry, we see several challenges. Due to the increased focus on cost savings, there is more attention to measures that increase cost efficiency of public transport. In the meantime, passengers require higher quality of the services. Although both developments seem to be a contradiction, measures do exist that serve both objectives. Improving operational speed and service reliability, for instance, will lead to higher quality and lower costs at the same time, as shown by *(6)*. However, to find and optimize cost-effective measures, data is required. Fortunately, the amount of data is increasing rapidly. Automated Vehicle Location (AVL) data, has already been available for a long time (e.g. *7, 8* and *9*) and recently much more passenger data (Automated Passenger Counting (APC) data) has become available as well (*1*)). These data support public transport design and decision making, since they enable planners to illustrate the costs of certain problems and the benefits (and additional costs) of potential solutions, for instance the transformation of a regular bus line into a high quality Bus Rapid Transport (BRT) system or into a tram line (*10*) or optimized synchronization between tram and train (*11*). These costs and benefits are relevant for decision making and may be incorporated in cost-benefit analyses.

This paper deals with the Dutch smart card system, the so-called *OV-Chipkaart*, and illustrates potential application of the data. Our objective is to process the data in such a way that it supports optimization of the level of service by improved network and timetable design. The outline of this paper is as follows. Chapter 2 will elaborate on smart card systems in general and the Dutch smart card data specifically. The next section describes our approach to apply the data to predict future ridership, which is applied in a case study, presented by Section 4. The conclusion and reflection on the approach are provided in Section 5.

2. SMART CARDS AND DATA

2.1 Smart Card Data Applications

In (*12*) the following major advantages of smart card data for transport service providers are distinguished:

- Large volumes of personal travel data.
- Being able to link those data to the individual card and/or traveler.
- Having access to continuous trip data covering longer periods of time.
- Knowing who the most frequent customers are.

Depending on the exact characteristics of the system, more insights may be gained. The number of areas where smart cards are applied increases rapidly. Famous examples are London (Oyster card) and Hong Kong (Octopus card). Depending on the technology used, limitations in applications arise. On the London buses, for instance, passengers only tap in and techniques are required to determine the destination stop (*5*). In (*2*), an example of Beijing is presented where no location information is connected to the smart card data and several data sources have to be connected to get that information. In (*12*), it is mentioned that travel purpose is hard to obtain from the data and in (*13)*, a method is presented to assess the number of boardings, since the smart card data does not provide this.

Canadian researchers (*1*) present a broad overview of applications of smart card data, varying from strategic and tactical planning optimization to operational improvements. Most applications aim at assessing OD-patterns (e.g. *14*, *15*) and route choice behavior (e.g. *16*) and transfer analysis (e.g. *17*). Surprisingly, improved forecasting based on historical data is only mentioned once.

2.2 The Dutch Smart Card System: *OV-Chipkaart*

The Dutch smart card, the *OV-Chipkaart*, replaced the former payment system, which was called *Strippenkaart* (*18*). The latter system was introduced in 1980 and replaced all individual urban and regional systems by one nationwide paying system. The *Strippenkaart* was valid in the whole country and the travel costs depended on the number of zones through which one travelled. The size of these zones differed per region and so did the total price. The advantage of this system was that everybody could travel with one ticket in bus, tram and metro throughout the country. The major disadvantage was that no information was available to the operators and authorities where people travelled. It was known where the tickets were sold (shops and counters) but not where they were devaluated. Expensive surveys were required to determine the distribution key over all operators. Every year, total revenues had to be distributed by this key. To solve this, the 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.

The Dutch smart card uses nfc-chip technology and passengers have to check in and to check out. All public transport (including train services) is accessible with this smart card. Thus, valuable information is measured about origin-destination patterns (on station/stop level) of all public transport users. In the Netherlands, the check-in and check-out 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 in a journey (a journey may consist of multiple trips, with a transfer in between) is tracked. The complete route through the public transport network is therefore traceable. When the smart card devices are located on the platforms, only information is available of the first and the last station, making route search through the public transport network necessary for the analyst (e.g. *16*). In the remainder of this paper we describe the situation where a check-in and a check-out are available in the vehicle. This is the case in a vast majority of the regional and urban public transport lines in the Netherlands (all bus and tram lines).

Until March 2014, 19 million smart cards were produced and every week, about 2.8 million people travel using their smart card, producing about 42 million transactions per week, consisting of checking in and out, but also e.g. checking account *(19)*.

2.3 Dutch Smart Card Data

An example of the raw data format, provided by the Dutch smart card transactions, is given in Table 1. Every record contains a trip, with a check-in station, check-in time, check-out station and check-out time. The anonymous smart card ID can be used to combine multiple trips to a journey, hence identifying the transfers. Furthermore, a public transport line number is given, so that the trip may be matched to a specific service in case multiple public transport lines run in parallel. Potentially, the vehicle number and /or run number are also given, enabling detailed analysis of distribution among individual services, for example to provide solutions for capacity problems. Furthermore, some information may be provided on smart card type / ticket type to predict trip purpose, for example an annual season ticket (usually used for commuting), a student card (usually used for education) or a special offer tickets (usually used for recreational purposes).

TABLE 1 A sample of fictitious smart card data: every record represents a trip in a public transport vehicle. For simplicity, all data in this example is for a specific date

Chip ID	Check in stop	Check out stop	Check in time	Check out time	Line number	(vehicle number)	(ticket type)
1	35	488	10:27	10:52	9	..	Regular single
2	23	86	8:01	8:09	1	..	Student
2	86	90	8:17	8:55	3	..	Student
3	73	94	7:20	7:53	4	..	Annual ticket
3	94	73	16:55	17:27	4	..	Annual ticket

When looking at Table 1, the first trip is the only trip conducted on this day by chip ID 1. This may be a trip for visiting family (including an overnight stay) or the return trip may be made by car (as a passenger). The second and third records are from the same chip ID. Furthermore, the trips are very close to each other in time, so we may assume these trips are a part of the same journey, which includes a transfer. In this example we can even see that the alighting stop for the first trip and boarding stop for the second stop of the second chip ID are the same, but this is not necessary to form a transfer: a short walking leg may be in between (for example in the situation of a large station with various tracks or in the situation of several on street stops). Finally, the last two records are of the same chip ID as well, but these trips are apart from each other in time, so these trips should not be seen as one journey. We can observe that, very likely, this is a typical commuting pattern: in the morning the traveler goes to work and returns home in the evening. The ticket type 'annual ticket' is another indication for commuting.

For the train and metro system the line number is omitted because the check-in and check-out takes place at the station instead of in the vehicle. Mostly the line is easily derived from the check-in and check-out station using path-finding.

The technical system of the Dutch smart card system contains several components ((*19*); see Figure 1):

- Level 0: the smart card, which may contain personal products, for example to get discounts or unlimited travel.
- Level 1: Devices that have direct contact with the smart cards: check-in and check-out devices and ticket vending machines.
- Level 2: Local systems at public transport operators that collect data from level 1 devices and temporarily store it (for example located at a bus garage).
- Level 3: Central system for each public transport company, where all data of a company is available and where the data is prepared to be send to the national (public transport smart card) data collecting agency (i.e. TLS).
- Level 4: The database of the national data collecting agency. Here the smart card transactions are verified and the financial consequences of the transactions are determined (the actual payment takes place). Another function is providing the personal transaction history to the smart card user.

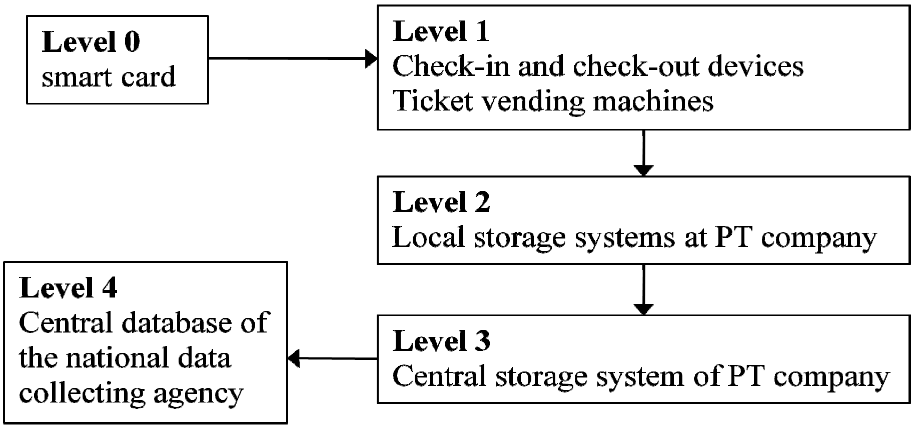

FIGURE 1 Dataflow through the Dutch smart card system, from the smart card in level 0 to the national database in level 4. A distinction may be made between components of the national agency (left) and components of the PT companies (right).

So technically speaking, the data is available at the individual level, giving large possibilities for detailed analysis. However, there are some concerns about the availability of the data and privacy agreements that must be obtained. Privacy is the most important issue, since individual data is used. It may not be possible to identify a unique profile from the data of one day, but if the data of, say, a month is analyzed, it may become possible to derive weekly patterns for individual users, which, combined with other data like home address and work location, might result in a match between an (anonymous) smart card ID and a name. From that moment, this individual can be followed on all his / her movements in the Dutch public transport network. Therefore, Dutch privacy law states that these individual data cannot be kept more than 18 months.

Another concern is the availability for analysis. The data is owned by public transport operators and most of them see it as confidential company information (mainly relevant due to the tendering system of the Dutch public transport concessions). Data of only one public transport operator now regularly becomes available for analysis, since this confidentiality can be regulated in contracts. But combining data from more operators (for example to analyze transfer movements on a train station) is still complicated due to this issue. At this moment, both the National and regional governments are developing methods to solve these issues in cooperation with the operators.

3. PREDICTING RIDERSHIP BY SMART CARD DATA

3.1 Introduction

Making predictions for public transport can be done in several ways, ranging from multimodal activity based models to simple rules using spread sheets. In the Netherlands, a hierarchy of traffic forecast models exists. The national model is a disaggregated model mainly focused on road travel. Four more detailed regional models exist using the same principles as the national model, but with more detailed networks. Public transport is modeled during the distribution phase of the model, but the level of service matrices are mainly exogenous.

On the urban level, many cities in The Netherlands have their own models. In most of these models public transport is modeled in more detail on the network level. However, the models are generally simpler. Most of them are multimodal gravity models for estimating the demand. Recently, the importance of the bicycle as access mode to public transport was recognized, resulting in slightly more sophisticated models using a nested logit structure that distinguishes between different access and egress modes in public transport (*20*).

Most of the public transport operators in The Netherlands do not use transport models for predicting ridership or changes in demand. Mostly spread sheets are used with relatively simple rules. However, the transport models could provide valuable insights for most public transport operators. Enriching these models with smart card data improves the potential of these models. Many regions, however, do not have a multimodal model, or the level of detail of these models does not match the level of operation within the public transport company.

With the introduction of the smart card system several public transport operators wonder what they can do with this massive data collection. Mainly because of the continuous nature of this data, systems and ideas have emerged to use this for gaining more insight into current use, but also to produce predictions. Patterns over time could provide valuable information, if suitable software to deal with these existed. Modeling software, capable of analyzing massive model output, seems to be a logical candidate to perform this task. Apart from visualization and analysis of the measured data, it seems an attractive option to perform what-if analysis using the same software.

Several options exist to allow what-if analysis using smart card data. A station to station matrix can be derived directly from the data. This matrix needs to be converted to a zone-to-zone matrix. This can be done using the Zenith assignment model to estimate the stop usage per zone (20). In the case described in this paper this step is skipped. The resulting public transport-matrix could replace the estimated matrix from the model. Once such a matrix is derived the assignment needs to reproduce the passenger occupancies per line from the smart card data. This requires calibrating the model route choice parameters.

If a multimodal model exists the growth factors calculated in the model could be used to produce forecasts. When no model exists a unimodal approach could be followed using elasticities to estimate demand shifts. When no elasticities are available what-if analysis are possible with fixed demand. It would only show the effects on route choice.

Which approach seems most feasible depends on the availability of existing models and the time horizon for the decision making. Table 2 gives an overview of the possibilities.

TABLE 2 Possibilities public transport modeling

	Multimodal model	Elasticity model	Quick-Scan model
Modes	Car, public transport, bike	Public transport	Public transport
Scale	National, regional, urban	Regional, Urban	Urban
Time horizon	10-20 years	< 10 years	< 5 years
Project type	Strategic, policies, infrastructure changes	Tactical, changing lines, frequencies, stops	Tactical, changing lines, frequencies
Usage	Modal split, cost-benefit analysis	Network effect	Route choice effects

In this paper we will focus on the elasticity approach. These kind of models are relatively simple to make and can make good use of the available data.

3.2 Elasticity model

Given an OD matrix that is observed in smart card data, the step to short and medium time prediction can be made, for example to assess the network effects of changing the frequency of lines, changing routes of lines, introducing new routes and increasing the speed of a line. These measures may be temporary or permanent. In this paper we present a method that is based on demand elasticity: the relative change in costs per OD pair have an effect on transportation demand on that OD pair. The costs of a trip are the generalized costs, comprised of in-vehicle time, waiting time, number of transfers (penalties) and fare. All attributes of the trip are expressed in monetary values by the coefficients α. A Value of Time for the Dutch situation of 6 Euros per hour is used for in-vehicle time *(21)*. For waiting time, a factor is used that is one and a half times as high as the factor for in-vehicle time (i.e. 9 Euros per hour) (*22*). A transfer penalty of 5 minutes is used. Equation 1 shows the calculation of generalized costs for OD pair *i,j*. Note that the coefficient of fare α_4 is equal to 1, because costs are expressed in monetary values.

$$C_{ij} = \alpha_1 T_{ij} + \alpha_2 WT_{ij} + \alpha_3 NT_{ij} + \alpha_4 F_{ij} \qquad (1)$$

With:

C_{ij}	Generalized costs on OD pair *i,j*
α_1,α_2,α_3,α_4	Weight coefficients in generalized costs calculation
T_{ij}	In-vehicle travel time on OD pair *i,j*
WT_{ij}	Waiting time on OD pair *i,j*
NT_{ij}	Number of transfers on OD pair *i,j*
F_{ij}	Fare to be paid by the traveler on OD pair *i,j*

Figure 2 shows the steps in our elastic demand calculation. First, using a public transport route choice algorithm (e.g. *20*), generalized cost matrices are calculated for the base situation and the situation that includes a network scenario. Note that this requires successful calibration of the route choice parameters: we here assume that the route choice algorithm is able to reproduce the line loads in the base situation. Comparing the cost matrices results in relative cost changes per OD pair. Using the OD matrix for the base situation (from smart card data) and an elasticity value (e.g. *23, 24* and *25*), the relative changes in OD flows are calculated, resulting in an OD matrix for the network scenario. The final step is to assign this OD demand to the public transport network, again using a public transport route choice algorithm. This process is also captured in Equation 2 that calculates the new OD demand (in the situation with the network scenario) from the base demand, the costs in both situations and the elasticity values. The subtraction and later addition of 1 in the equation is to convert from a growth factor to relative growth or vice versa. Note that in this definition the value for elasticity should be negative to be realistic, since an increase in costs then leads to a decrease in demand. Consequently, the demand change is directly calculated from generalized costs. This is different from using, for example, travel time elasticity or fare elasticity, since those values only include specific components of the generalized costs. The value of generalized costs elasticity is chosen in a way that the effect of a travel time or fare change roughly corresponds with the changes that would occur when using travel time or fare elasticity.

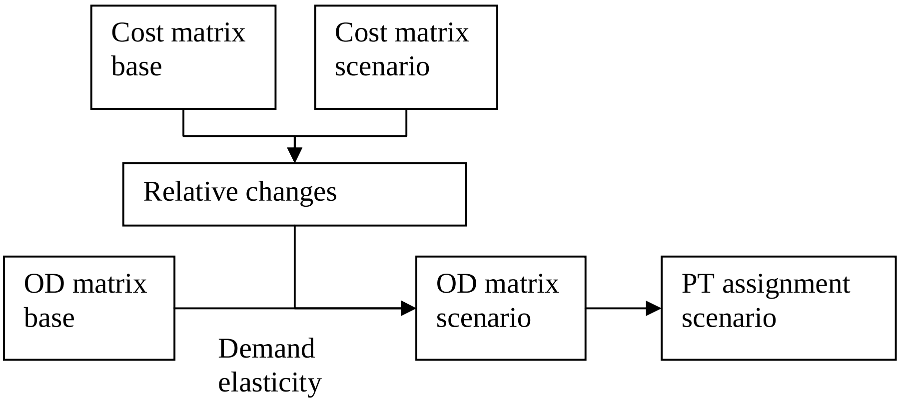

FIGURE 2 Schematic representation of the demand prediction model.

$$D_{ij}^{1} = \left(E\left(\frac{C_{ij}^{1}}{C_{ij}^{0}} - 1\right) + 1\right) * D_{ij}^{0} \qquad (2)$$

With:

D_{ij}^{1} Demand on OD pair i,j in the scenario
E Elasticity
C_{ij}^{1} Generalized costs in the scenario
C_{ij}^{0} Generalized costs in the base situation
D_{ij}^{0} Demand on OD pair i,j in the base situation

Extensions of this model can be made when new housing or job developments take place in the region at study. The relative growth of housing or jobs around public transport stops may be converted into growth factors to be applied to rows or columns of the OD matrix. Then the assumption is made that the distribution of trips among destinations or origins does not change from the observed distribution (based on smart card data) in the base situation. When both rows and columns are adjusted, a balancing method should be applied, for example the Furness method.

4. APPLICATION ON THE TRAM NETWORK OF THE HAGUE

4.1 Introduction

We applied the approach of connecting data to a transport model in a case study. In this case study, we tested whether our approach presented in the previous section would work with actual data. We connected smart card data of HTM, the tram operator in The Hague (about 500,000 inhabitants, 3rd largest city of the Netherlands) to a transport model built in OmniTRANS. The city of The Hague has 12 tram/light rail lines with a total length of about 335km. In addition to these tram lines, the public transport in and around the city consists of urban and regional bus lines and railway lines. In this case, we have only investigated the tram lines. This made the calibration of the route choice model relatively simple because in most cases only one route choice option existed. In future research we will add the bus and train services to the model. This will give us more possibilities to calibrate the route choice model.

4.2 Evaluation

A first step in supporting public transport planners and designers is visualizing historical data. In (*6* and *26*), examples of AVL data visualization are provided. In addition, illustrating smart card data on a geographical layer is beneficial as well.

To combine the smart card data with geographical information, we imported the public transport network into the software environment OmniTRANS using timetable data which is publicly available in GTFS (General Transit Feed System) format. This format was introduced by Google to allow public transport operators to feed their timetables to Google Maps. This data contains the lines, positions of stops and the departure and arrival times of each run at each stop. It is translated into frequencies and travel times per line per time period (AM peak, PM peak, off-peak day period and evening). The information of the lines (including the locations of stops) is mapped geographically on the underlying infrastructure, in this case the tram rail network of The Hague. The resulting network can be seen in Figure 3.

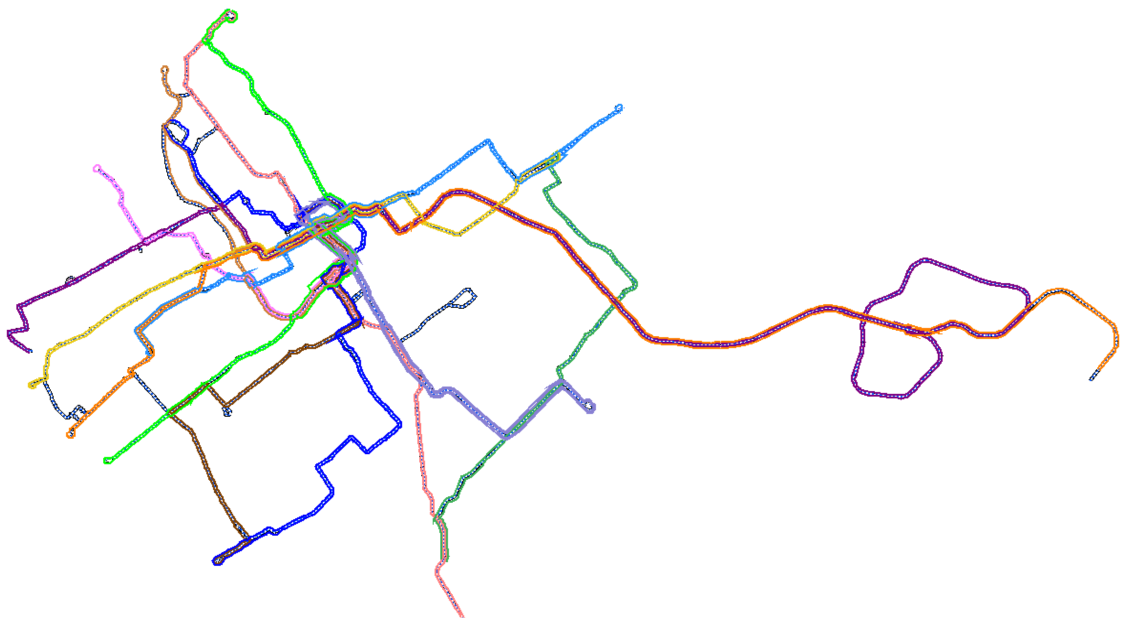

FIGURE 3 All Lines in the tram network of The Hague (each line is plotted in a separate color).

In this case the decision was made to put the zones directly at the stops. In the anticipated extension with bus lines we will use modeling zones instead.

The combination of geographical data of stops and lines and the smart card data is used to visualize passenger flows on the network. To this end, the smart card data (in the format of Table 1) is first preprocessed: invalid records are removed (for example records with the same stop for check-in and for check-out) and trips are combined to journeys by identifying transfers, based on smart card ID and check-out / check-in time. After that, the journeys are loaded onto the network, following the check-in and check-out stop and public transport line number in the data. When the network data (from GTFS) and smart card data (from the public transport company) of the same date are used, these two data sources fit very well: almost all records from the smart card data can be directly imported.

The resulting geographical visualization can be shown over time, since the check-in and check-out times are known. Given an assumption which time stamp determines the time block of the trip (check-in time, check-out time or an average between the two), the data can be visualized per aggregated time period, for example per one-hour period. Figure 4 shows the AM peak for a one-day sample: it can clearly be observed that before and after the peak period, the flows are much lower than during the peak period (see the presented loads in the added circle for instance). This time-dependent data may as well be visualized in an animation. The visualization helps to understand the past: identifying high or low flows, identifying important (transfer) stops and understanding the difference among time periods over the day.

Since the capacity of transit vehicles is known per transit line, as well as the frequency of the transit lines (stored in the network data), these figures are easily extended to intensity-capacity ratios and plots, showing the interaction between transit supply and demand. This enables detecting capacity problems, dependent on the hour of the day.

FIGURE 4 An example of time dependent visualization of smart card data: upper left: 6AM-7AM; upper right: 7AM-8AM; lower left: 8AM-9AM; lower right: 9AM-10AM.

4.3 Predicting

In addition to showing historical data, by connecting network and smart card data in the transport model, we also tested our elasticity method on the actual data simulating several measures. We investigated frequency changes, fare adjustments and rerouting of a line. The elasticities we used were based on literature (*25*) and also on rules of thumb of HTM. For instance, we used an elasticity value of -0,5 for travel time changes. This means that an increase of 10% in travel time will lead to 5% less travelers. The rules of thumb of HTM were audited and proven to be valid by independent research (*27*).

We adjusted the original skim matrix to the measures and calculated the new passenger OD-matrix accordingly. We assigned this matrix, using the Zenith-algorithm (*20*). Similar visualizations as shown in Figure 4 may be generated showing the new link loads. Figure 5 shows the outcomes (in terms of change in passenger load) of two examples of specific network scenarios: a frequency increase and a route change in a public transport line. The main contribution of this method is that we clearly see the network impacts. Figure 5 (left) shows a frequency increase on two lines, with expected ridership growth on these two lines (green), but also a decrease in a nearby line (red). In Figure 5 (right), we see the impacts of a route change (a link was blocked and trams had to be diverted). Due to higher travel costs on the new route, the total number of passenger decreased (increase on the diverted line route (green) is smaller than decrease on the original route (red)).

FIGURE 5 The effect of two example network scenarios in the public transport network: a frequency increase (left) and a change in route (right).

We did a validity check on the results, which were in line with the existing methods (traditional models). However, the next step would be detailed research on revealed behavior after changes to find updated elasticity values, specifically focusing on this area and the types of passengers.

4.4 Reflection

In this research we have chosen an initially practical approach by choosing a zonal system corresponding to the actual stops. A more desirable option is to choose a zonal system as used in the model system for this region. This would allow direct usage of modal split factors from this model while having a much more accurate matrix for the current situation.

Because our case was limited to tram lines only, route choice in this network did not play a significant role. This meant that calibrating the route choice model using smart card data was not very challenging as mostly just one route was feasible. In an anticipated extension of this study we intend to increase the network with all the bus lines of The Hague, resulting in a network with significant route choice options. A public transport route choice algorithm needs parameters, for example logit parameters in stop choice and line choice models, or the weight factors for cost components in the generalized costs function (as are also defined in this paper). The detailed smart card data presented in this paper could be used to calibrate these parameters, since the actual routes chosen by travelers are observed. If multiple routes are available, a distribution among the routes can be derived from the data that should be estimated by the route choice model and its parameter settings. Furthermore, it can be tested whether these parameters are approximately equal for different situations (i.e. short trips vs. long trips).

5. CONCLUSIONS

Public transport operators are exposed to massive data collection from their smart card systems. In the Netherlands, every passenger needs to check in and check out, resulting in detailed information on the demand pattern. In buses and trams, check in and checkout takes place in the vehicle, providing also information on route choice. This paper explores options for using this smart card data for analysis and performing simple what-if analyses by using transport planning software. The intention is to provide public transport operators with relatively simple (easy to build) models to perform these what-if analyses.

The data is relatively easily converted to passengers per line and matrices between stops. This matrix is assigned to the network to reproduce the measured passenger flows. Once the assignment can reproduce the passenger flows simple what-if analysis becomes possible. With fixed demand, line changes can be investigated. With the introduction of an elasticity method on the demand matrix, simple modal-split calculations are possible.

The method described above was applied on a case study in The Hague. The tool turned out to be very valuable for the operator to gain insights into small changes. 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 elasticities. With the availability of smart cards, valuable revealed preference research is possible after changes in level of service. Another anticipated improvement is related to the zonal system. In this case the zones are at the stops making what-if analysis on stop choice rather limited. In an anticipated extension the smart card data station-to-station matrix will be converted to a proper zone matrix.

ACKNOWLEDGEMENTS

This research is performed in cooperation with HTM, Delft University of Technology, Department of Transport & Planning, Goudappel Coffeng and DAT.Mobility.

REFERENCES

(1) Pelletier, M., 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.

(2) Ma, X., 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.

(3) Kurauchi, F., J.D. Schmöcker, H. Shimamoto, S.M. Hassan. Variability of commuters' bus line choice: an analysis of oyster card data. *Public Transport*, Vol. 6, No. 1-2, 2014, pp. 21-34.

(4) Wang, W., Attanucci, J.P., N.H.M. Wilson. Bus Passenger Origin-Destination Estimation and Related Analyses. *Journal of Public Transportation*, Vol. 14, No. 4, 2011, pp. 131-150.

(5) Park, J., D.J. Kim, D.J., Y. Lim. Use of Smart Card Data to Define Public Transit Use in Seoul, South Korea. *Transportation Research Record,* No. 2063, pp.3-9, 2008.

(6) Van Oort, N., 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.

(7) Furth, P. G., 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.

(8) Hickman, M. 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.

(9) Van Oort, N., T. Brands, E. de Romph, J. Aceves Flores. Incorporating unreliability of transit in transport models: theoretical and practical approach. Proceedings of the 93rd Annual Meeting of the Transportation Research Board, Washington D.C., 2014.

(10) Van Oort, N. . Quantifying benefits of enhanced service reliability in public transport. Presented at the 12th Conference on Advanced Systems for Public Transport (CASPT), Santiago, Chili, 2012.

(11) Lee, A., N. van Oort, R. van Nes. Service Reliability in a network context: impacts of synchronizing schedules in long headway services, *Transportation Research Record*, 2014. (Printing in progress)

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

(13) Morency, Catherine, M. Trepanier, B. Agard. Measuring transit use variability with smart-card data. *Transport Policy*, Vol. 14, No. 3, 2007, pp 193-203.

(14) Munizaga, M., 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, pp. 9-18, 2012.

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

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

(17) Seaborn, C., J. Attanucci, N.H.M. Wilson. Analyzing Multimodal Public Transport Journeys in London with Smart Card Fare Payment Data. *Transportation Research Record*, No. 2121.-1, 2009, pp. 55–62.

(18) Cheung, F. Implementation of Nationwide Public Transport Smart Card in the Netherlands: Cost-Benefit Analysis. *Transportation Research Record*, No. 1971, 2006, pp. 127-132.

(19) Translink systems. https://www.translink.nl/nl-NL/OV-chipkaart/Techniek-OV-chipkaart. retrieved on 17th of June, 2014.

(20) Brands, T., E. de Romph, T. Veitch, J. Cook. Modelling public transport route choice with multiple access and egress modes. *Transportation Research Procedia*, Vol 1, 2014, pp. 12-23.

(21) Significance, VU University, John Bates Services, TNO, NEA, TNS NIPO en PanelClix. Values of time and reliability in passenger and freight transport in The Netherlands. Report for the Ministry of Infrastructure and the Environment, Significance, The Hague, 2013.

(22) Wardman, M. Public transport values of time. *Transport Policy*, Vol. 11, No. 4, 2004, pp. 363-377.

(23) Wardman, M. Review and meta-analysis of U.K. time elasticities of travel demand. Transportation, Vol. 39, No. 3, 2012, pp. 465-490. DOI: 10.1007/s11116-011-9369-2

(*24)* Transportation Research Board. TCRP REPORT 95: Transit Scheduling and Frequency Traveler Response to Transportation System Changes. Chapter 9, 2004.

(25) Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., White, P. The demand for public transport: a practical guide. 2004.

(26) Van Oort, N., R. van Nes. Line length versus operational reliability: network design dilemma in urban public transportation. *Transportation research record*, No. 2112, 2009, pp.104-110.

(27) Oostra, R. Elasticiteitsonderzoek binnen het vervoergebied van HTM. TU Delft, 2004 (In Dutch)