Estimating Urban Road Traffic States Using Mobile Network Signaling Data

Thierry Derrmann, Raphael Frank, Francesco Viti, Thomas Engel
University of Luxembourg
E-Mail: {*firstname.lastname*}@uni.lu

***Abstract*—It is intuitive that there is a causal relationship between human mobility and signaling events in mobile phone networks. Among these events, not only the initiation of calls and data sessions can be used in analyses, but also handovers between different locations that reflect mobility. In this work, we investigate if handovers can be used as a proxy metric for flows in the underlying road network, especially in urban environments. More precisely, we show that characteristic profiles of handovers within and between clusters of mobile network cells exist. We base these profiles on models from road traffic flow theory, and show that they can be used for traffic state estimation using floating-car data as ground truth. The presented model can be beneficial in areas with good mobile network coverage but low road traffic counting infrastructure, e.g. in developing countries, but also serve as an additional predictor for existing traffic state monitoring systems.**

I. Introduction

Mobile data is becoming increasingly popular as a research topic, as mobile network operators (MNOs) are looking for novel applications of their data. With Internet-of-Things (IoT) devices, connected vehicles and innovations in radio access technologies, the data available to MNOs is getting richer. Its spatio-temporal resolution is growing as the locations of users are known more frequently because of increased mobile connectivity, and with higher precision due to the decrease in cell size of recent radio access technologies. The types of data that are readily available to MNOs which are commonly used for data analytics are Call Detail Records and Signaling Data.

Call Detail Records (CDRs) are the most popular type of mobile data used for analysis. CDR entries describe single transactions of a user equipment on the mobile network, e.g. the initiation of a call or data session, or the transmission of a text message (SMS). They are easily extracted from the MNO's billing systems and are routinely evaluated by MNOs. However, there are privacy concerns with CDR data, as individual users can be identified from a large set of users based on the frequented locations and spatio-temporal patterns present in the data, as demonstrated by De Montjoye et al. [1]. CDR data contains only a single, initial location information per activity (call/SMS/data transmission), which can be a limiting factor in some studies.

Signaling Data is a different kind of mobile data. It consists of meta-data from the mobile network infrastructure that is generated e.g. if a user initiates a call or moves from one antenna to another (a so-called *handover*). This kind of data is related to the operation of the mobile network. Unlike CDR, signaling data can include information on a user's precise movement trajectory within the mobile network during an active connection (in the so-called *active mode*). In the absence of an active connection, the location of a user is known to the network only at Location-Area (LA) level, i.e. a set of cells covering a large area. In aggregated form, data regarding handovers can serve as a privacy-neutral way of assessing mobility, in the form of counts or the distribution of time spent within a cell (*cell dwell time*).

In this vein, we want to make use of mobile network signaling data in order to estimate the road traffic state in an urban context. More precisely, we want to show the correlation between road traffic and mobile network handovers, to enable using the mobile network infrastructure as a distributed traffic sensor. For this purpose, it is necessary to deal with the restriction that mobile signaling data includes stationary and pedestrian users, while traditional traffic detectors (e.g. loop detectors and floating car data) only include vehicles. Hence, part of the contribution in this paper will lie in the vehicular density estimation from aggregate signaling data. This is a privacy-neutral and almost cost-neutral model, that can serve as an additional input to existing traffic state monitoring systems, or stand-alone for areas with sufficient mobile network coverage but little road traffic counting infrastructure – in particular developing countries.

II. Related Work

A. Mobile Data Analytics in Transportation

1) CDR-Based Analysis: There are several examples of making use of CDR data for estimating the demand in the underlying transportation networks. Iqbal et al. propose a method for estimating OD matrices from CDR data and traffic counts jointly, using trip patterns from the mobile phone data and traffic data as ground truth [2]. They scale the CDR-based demand to the loop detector data using MITSIMLab, a microscopic simulator, in order to obtain realistic OD-flows.

Di Donna et al. have shown in [3] that demand can be estimated from CDR data by means of a Markov chain of user movements between clusters of antennas obtained using k-Means clustering. They show that the change in transition matrices is slow and their results indicate strong spatial demand similarity across multiple days of studied data.

In [4], Gundlegård et al. propose a full methodology for going from CDR data to OD matrices, separating the demand and route choice model estimation. They propose algorithms for estimating the temporal distribution of demand, route choices as well as travel time estimation and suggest how to implement mode choice models.

2) Signaling Data Studies: Various studies have been performed on Location/Tracking Area Updates, which concern idle phones (in a disconnected state) and are useful at larger distances, e.g. on Interstate highways.

Janecek et al. [5] have demonstrated the utility of Location Area Codes (LACs) and handovers to detect congestion on a highway case study. In particular, they monitor the fastest cell-switches of users under the assumption that they represent vehicular mobility. They use the LAC updates from idle mobile phones to detect traffic slow-downs, and then use handovers of active, connected phones to more precisely locate the source of congestion. They also show that mobile signaling data can be among the fastest ways of detecting highway congestion.

Bar-Gera evaluated the predictive power of cell dwell times, using a system that logs handovers of phones that are in-call (i.e. in active mode) and the handover time stamps [6]. The case study focused on the Ayalon freeway in Israel, and the author shows the strong correlation between travel times and speeds estimated from the mobile phone dwell times and those measured using loop detectors.

Hui et al. have investigated inter-city travel volumes using signaling data, inferring the modal split (road and air travel) of users moving between Calgary and Edmonton [7]. This shows that for large distances, passive mode data on Location Areas can be a valuable source for mobility studies.

The aforementioned studies on LAC updates and handovers primarily concern highways and long distance travel. To the best of our knowledge, there are no studies on the relationship between handovers of mobile phones in active mode and traffic states in urban environments. In a previous simulation study, we have shown that characteristic profiles of flow and density exist in mobile cell clusters, and that their behavior correlates with the underlying road network [8]. In this work, we want to confirm these findings using real data. The resulting model can be beneficial for real-world applications. It can be applied in areas where there is little traffic counting infrastructure, but sufficient mobile network coverage, e.g. in developing countries. It can also be used for anomaly detection in the relative states of road and mobile networks. Since we want to establish a link between mobile data and traffic flow, we will now present some relevant related work from the domain of traffic flow theory.

B. *Traffic Flow Theory: Macroscopic Fundamental Diagram*

The Macroscopic Fundamental Diagram (MFD) describes the distinct relationship between density and border flows of homogeneous road network regions. In [9], Geroliminis and Daganzo show the existence of urban macroscopic fundamental diagrams on data from Yokohama, a very homogeneous road network with a high loop-detector coverage of 500 fixed detectors placed 100 m upstream of intersections. They show that the phenomenon emerges for areas greater than 10 km^2.

In order to observe MFD characteristics, it is necessary to adequately partition the road network. In [10], Ji et al. show that the normalized-cut algorithm can be used for initial partitioning of the network, but they also propose a method for further improving the network partitioning with respect to the homogeneity of the obtained clusters.

Buisson and Ladier explore the impact of homogeneity on the variance of MFDs in [11]. They show how the loop detector spacing and heterogeneity in road types impact both the shape and scatter of the resulting MFDs. They conclude that link similarity, regular data collection and comparable congestion patterns are the key attributes of zones that exhibit low-scatter MFDs. These results are relevant to our work as they help explain some of the behavior we will observe in the road network partitions in this study.

III. Datasets

A. *Mobile Network Signaling Dataset*

1) Description: The mobile dataset contains aggregate data from 1839 3G (UMTS) cells within the country of Luxembourg, 611 of which are located in and around its capital, Luxembourg City, the region relevant to this study. More specifically, the data consists of:

- the number of handovers between cell pairs per hour
- the number of calls initiated from each cell per hour

The motivation behind choosing these two metrics was to be able to approximate flows and density from them. The data was made available for a whole week at the end of September 2016. We organize the handovers in a *handover matrix*, i.e. the weighted, directional adjacency matrix of 3G cells in the study area summed over the data of Monday, in order to identify the amount of flows between cells across a single typical workday.

2) Clustering: We want to consider internal and exiting handovers for different partitions of the mobile network, and thus have to partition it into mobile cell clusters.

Ji et al. have shown in [10] that partitioning by normalized graph cuts is a valid starting point for finding homogeneous road network partitions. *Spectral clustering* is a relaxation of the normalized graph cuts algorithm, and has proven effective for clustering the mobile network in a previous study [8] that we performed on simulated handover data.

Thus, in this work, we apply spectral clustering to the handover matrix (i.e. the weighted adjacency matrix) of the mobile network cells. The weights in the matrix correspond to the number of handovers between cell pairs, and spectral clustering allows defining the desired number of clusters of this matrix.

Geroliminis et al. have shown the emergence of MFDs in areas greater than 10 km^2 [9]. Our study area of Luxembourg City and its highway ring consist of 82 km^2, so we opted for 8 clusters.

B. *Floating-Car Dataset*

1) Description: As ground truth data, we use Floating-Car Data (FCD) collected during the same study week. This is a set of time-stamped location updates and travel speeds which was collected in the area of Luxembourg City and its highway ring, and consists of 600 trips and 220000 location updates. In particular, we are interested in *traffic states*, i.e. the ratio between actually driven speeds and the speed limit ($\frac{v}{v_{max}}$). Thus, we performed map-matching on the FCD to obtain the values of v_{max} for every location update.

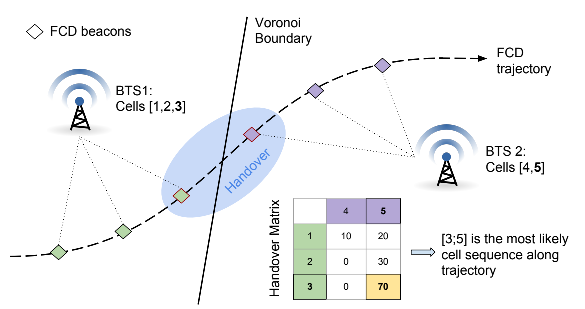

Figure 1: Floating-Car Data and Mobile Network Mapping

2) Mapping FCD to the Mobile Network: In order to enable the use of FCD for validation purposes, we need to map the most likely associated mobile network cell to each GPS location entry.

Fig. 1 shows an example of the method we used: First, we can easily find each location entry's nearest base station (BTS) by distance. In the example, that is BTS1 for the first three floating car beacons, and BTS2 for the next three. Usually, a BTS hosts a set of mobile network cells emitting into different directions, e.g. cells 1, 2 and 3 for the first part of the trajectory and cells 4 and 5 for the second part.

From an FCD trajectory, we can thus identify a sequence of these sets of potentially associated cells corresponding to the taken road path. Now, in order to identify the single, most likely visited cell sequence, we use the handover matrix. We choose the most frequent cell transition to be the likely cell pair visited, thus building a chain of visited cells over the entire trip. In the example above, the most likely cell transition (handover) is $3 \rightarrow 5$, because there are the most handovers between these two cells. Thus, we pick these two cells as the most likely occurred sequence.

Using this method, we get a single likely associated cell for each Floating-Car Data entry, i.e. the cell that the driver's phone was most likely connected to at their current location. This allows to compute road traffic statistics relative to the connected cell.

Finally, in order to compute the *traffic state* variable, i.e. the ratio between the actual observed link speeds and their respective speed limits, we perform map-matching of the Floating-Car Data entries to the OpenStreetMap road network (to obtain the speed limit at each entry).

C. Merging FCD and Mobile Datasets into Training and Validation Sets

1) Training Set: We build a training set with the road network traffic state (the response variable) as well as each cluster's hourly statistics and profile function parameters and derivative (as introduced in the following section). Using this training set, we can then estimate a global equation for all clusters linking the inputs to the traffic state. Thus, we construct the following training dataset using FCD and mobile network data from Monday through Wednesday of the study week:

For each FCD entry we have computed the current traffic state variable and most likely associated cell, mapping to a mobile network cluster as defined by the spectral clustering of the handover matrix. Thus, we have a mapping of each FCD entry to its respective mobile network cluster, i.e. we can associate the response variable (traffic state) to the predictors (mobile network statistics). Among these predictors, there are aggregate handover statistics between and within cell clusters. By the term *internal handovers*, we understand handovers between cells belonging to the same cluster, while *exiting handovers* describe those leaving the cluster.

To summarize, for each cluster and hour, we compute:

- entering, internal and exiting handovers
- aggregate amount of calls emitted from this cluster
- fit parameters of the profile function
- the derivative of the profile function
- the average traffic state ($\frac{v}{v_{max}}$)

Using these features, we can then evaluate our approach: We study whether the aggregate mobile network statistics and the previously learned profile functions (as described in the following section) can serve as sufficient predictors for the underlying road network traffic state.

2) Validation Set: The validation dataset consists of FCD data and mobile network data of Thursday and Friday. The values for the profile function coefficients are adopted as learned from the training dataset, representing the learned profiles from past observations.

IV. Correlation between Mobile and Road Networks

We evaluate the predictive power of mobile network signaling data using Floating-Car Data as ground truth. In particular, we estimate for each road network partition the *traffic state* variable, i.e. the ratio between the actual observed link speeds and their respective speed limits.

We want to identify if there are profile functions of the mobility inside and between mobile network cell clusters, analogous to the Macroscopic Fundamental Diagrams (MFD) in road networks. The MFD describes the relationship between outgoing flows and internal density of vehicles in a homogeneous partition of a road network.

Similarly, we want to make use of exiting and internal handovers of mobile cell clusters to build such profile functions for mobile cell clusters. Their relationship can be summarized in a profile function, which is analogous to an MFD in the road network, but based on mobile network signaling data. We then want to use these profile functions estimate the current degree of saturation of the underlying road network, if more or fewer handovers happen within clusters or across their boundaries.

Concretely, these profile functions are modelled as quadratic polynomials. They express the relationship between flows exiting a cluster and the density within the cluster, which we approximate using inner-cluster flows and the number of calls initiated within the cluster.

As with the MFD, velocity can then be expressed by the ratio of outflows to density, which are approximated from the mobile data (cf. following subsections). In further analogy to an MFD, the derivative of the profile function should be positively correlated to the velocity inside the cluster, as it assumes positive values in low density/high velocity situations, and decreases as density increases/velocity decreases.

A. Proxy Function for Vehicle Density from Mobile Data

In order to model the relationship between flows exiting a cluster and the velocity within the cluster, it is necessary to estimate the density inside the cluster. For this purpose, we estimate the parameters of a function connecting internal (within-cluster) flows and calls emitted from inside the cluster. The motivation behind this is that a higher density of vehicles will lead to an increase of the within-cluster handover count (q_{inner}). Also, more people and vehicles within an area lead to more call initiations. Hence, we chose to express our density proxy through a product of these two metrics weighted by the exponents θ_Q and θ_C, respectively:

$$\tilde{k}\ (q_{inner}, n_{calls}) := {q_{inner}}^{\theta_Q} \times {n_{calls}}^{\theta_C} \tag{1}$$

Further, we can then get a proxy function for velocity estimation according to Greenshields' linear speed-density relationship model [12]:

$$\tilde{v}\ (q_{exiting}, q_{inner}, n_{calls}) := \frac{q_{exiting}}{\tilde{k}\ (q_{inner}, n_{calls})} \tag{2}$$

We have performed optimization on this function with respect to the training dataset. More precisely, we ran a hill-climbing algorithm to maximize –for each cluster– the correlation between the true traffic states and the velocity proxy function. Further, we introduced another parameter by adding an exponent (θ_M) to the handover matrix that we use as an input for the spectral clustering. The goal is to obtain balanced clusters for which the correlations above hold particularly well. The parameter basically allows transforming the adjacency (handover) matrix such that spectral clustering considers more or fewer cuts as transitions are rendered less ($\theta_M > 1$) or more ($\theta_M < 1$) similar.

B. Profile Functions

Following the ideas above, we estimate the coefficients a, b, c of a quadratic fit to the relationship between $q_{exiting}$ and $\tilde{k}$ for each of the clusters:

$$p(\tilde{k}) = a\tilde{k}^2 + b\tilde{k} + c \tag{3}$$

Depending on the approximate density $\tilde{k}$ at a given moment, we can evaluate the derivative of the profile, which should correlate positively with the traffic state inside the cluster:

$$p' = \frac{dp}{d\tilde{k}} = 2a\tilde{k} + b \tag{4}$$

We have now introduced all the functions that are necessary to proceed with our analysis and to compute correlations between the true traffic states and $\tilde{v}$ and p', respectively.

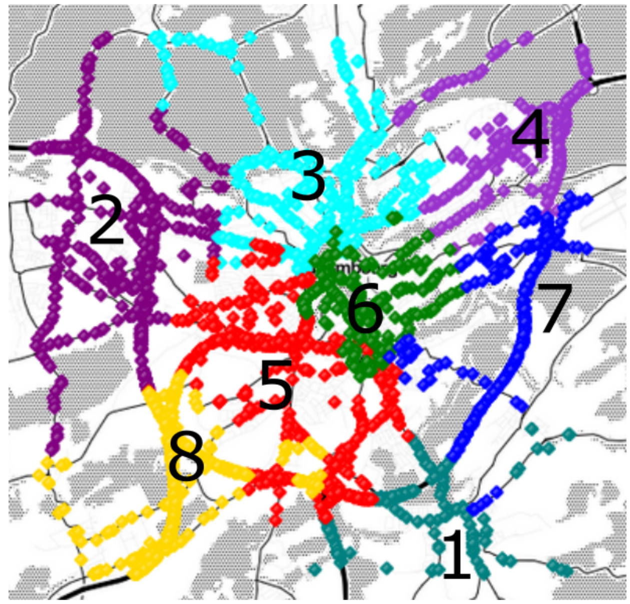

Figure 2: The road coverage of the 8 mobile network clusters as generated with spectral clustering for this study.

C. Parameter Optimization Results

The parameters resulting from the hill-climbing optimizer are $\theta_Q = 2.36$, $\theta_C = -0.36$ and $\theta_M = 0.059$. This means that the density approximation function ($\tilde{k}$) is primarily influenced by the within-cluster flows, but balanced by the amount of calls at roughly the cubic root. This likely reduces density overestimation in urban/business districts where a relatively large number of calls are initiated in comparison to suburban or highway environments. The low exponent for the handover matrix (θ_M) indicates that the matrix is flattened, allowing a larger set of cuts to be taken into consideration by the spectral clustering algorithm.

The optimum solution with $\theta_M = 0.059$ resulted in the clustering shown in Fig. 2. On this map, we plot Floating-Car data points colored by their associated mobile network cell cluster, according to the mapping procedure explained previously in Fig. 1. In terms of coverage areas, the clustering appears to be balanced, and the clusters are mostly cohesive. There are some small incohesive areas, which are due to the proximity-based approach taken during the mapping process.

D. Correlation Results

We computed the Pearson correlation coefficient between the traffic state with $\tilde{v}$ and p' for each cluster. Table I shows the resulting correlation values for the validation set. We observe that the resulting values are in the medium-high range for the majority of clusters. They are weak for clusters 3 and 4, which are situated in the uptown and business districts, which could be due to differences in the distribution of placed calls and within-cluster flows, as these clusters contain more stationary users during business hours. Hence, the single equation we have established for density estimation might not hold for these two clusters specifically. Also, the clusters might simply be too

Cluster	Area	r($\tilde{v}$)	r(p')
1	Hesperange, Ring: A3 LU ↔ FR	0.33	0.49
2	Ring: A6 LU ↔ BE	0.60	0.45
3	Limpertsberg, uptown	0.10	-0.04
4	Kirchberg	-0.08	-0.22
5	Cessange, Gasperich	0.51	0.58
6	Train station, downtown	0.52	0.29
7	Ring: A1 LU ↔ GER	0.65	0.54
8	Ring: A4 LU ↔ Esch/Alz.	0.59	0.7

Table I: Pearson correlation coefficients by cluster: Traffic state vs. velocity proxy and derivative of the profile function

small and/or heterogeneous to reliably estimate density using our method.

Given the relatively low resolution both in time and space, the error in FCD ↔ handover mapping and the very heterogeneous base station placement (in comparison to e.g. loop detectors), these results are very encouraging, in particular for the highway ring clusters. While these correlation values are insufficient as stand-alone predictors of traffic states, they can serve among others in a regression model, which we will discuss in the following section.

Fig. 3 shows results for three different clusters. The blue cluster (7) is a cluster that consists primarily of data from the A1 highway, a part of the ring around Luxembourg City. The red cluster (5) is a heterogeneous cluster that consists both of highway on-/offramp and urban data, specifically the Gasperich-Cessange area in the south of Luxembourg City and the A3 and A4 highway onramps and exits. Finally, the green cluster (6) corresponds to an urban cluster, more specifically the train station and old town areas.

The first line shows a set of plots corresponding to the profile functions computed, relating outflows to the approximated density. We observe low residuals against the quadratic regression line for the urban and highway clusters (green and blue), and larger variance on the mixed cluster (red), which is likely due to its mixed road types, matching the findings of Buisson et al. [11].

The second line shows plots of the distributions of the true traffic states (as measured with the Floating-Car Data) and the estimated proxy function of velocity ($\tilde{v}$). We can observe that the highway and mixed clusters correlate better than the urban cluster.

The plots show that the highway clusters are better represented by the model than urban clusters. Likely, the profile function captures the highway's fundamental diagram function, while the urban areas are quite heterogeneous and more difficult to estimate. Also, our approach of fitting a single density-estimation function to all clusters fails to capture the dynamics of all cluster types. Hence, we propose to go beyond this work in the future, by fitting different parameters to different cluster types.

V. Road Traffic State Regression

We want to find a single regression equation that uses the (previously trained) profile functions of mobile cell clusters, along with each cluster's handover and call counts, to estimate the road traffic states. If that is possible, then we have sufficiently characterized the mobility inside the clusters, as the same distinct regressive relation holds for all clusters.

A. Regression Results

In order to test the adequacy of mobile phone MFDs to represent the full spectrum of road traffic states, we fit a single linear regression model to all the clusters. This will yield good prediction accuracy across all clusters only if the information in the fitted profile functions can explain the traffic state variance sufficiently well.

As described above, we train our profile functions of each clusters on the data from Monday through Wednesday, and evaluate them on data of Thursday and Friday.

Fig. 4 shows the scatter between estimated and actual traffic states for the Thursday and Friday data. Individual points represent a clusters' mean traffic state during an hourly time slot. The green line is the identity line, while the blue line represents the regression trend between prediction and true values. The proximity of both lines indicates a good fit. The mean absolute percentage error (MAPE) amounts to 12.0%, which – given the constraints of our data sets – shows that mobile network data can be used for estimating road network traffic conditions. The low count of traffic state values < 0.5 has a negative impact on the prediction for low values, which can likely be remedied with more training data with sufficient congestion.

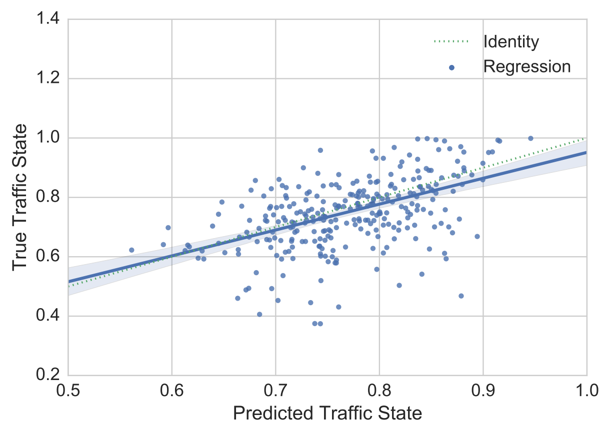

Figure 4: Regression results of traffic states on the validation set

B. Limitations and Possible Extensions

The main limitation of this work is the temporal aggregation of the mobile phone data set (1 hour granularity). However, we believe that the results above are sufficiently good to show that traffic state estimation is feasible using only mobile phone data. In particular, there is room for improvement of the predictions, using more fine-grained data, a longer training period and multiple radio technologies (2G, 3G and 4G).

Another limitation is the low degree of congestion in the data, which prevents us from directly comparing our profile functions to other models. Generally speaking, the functions follow the free-flow and *sweet-spot* parts of a Macroscopic

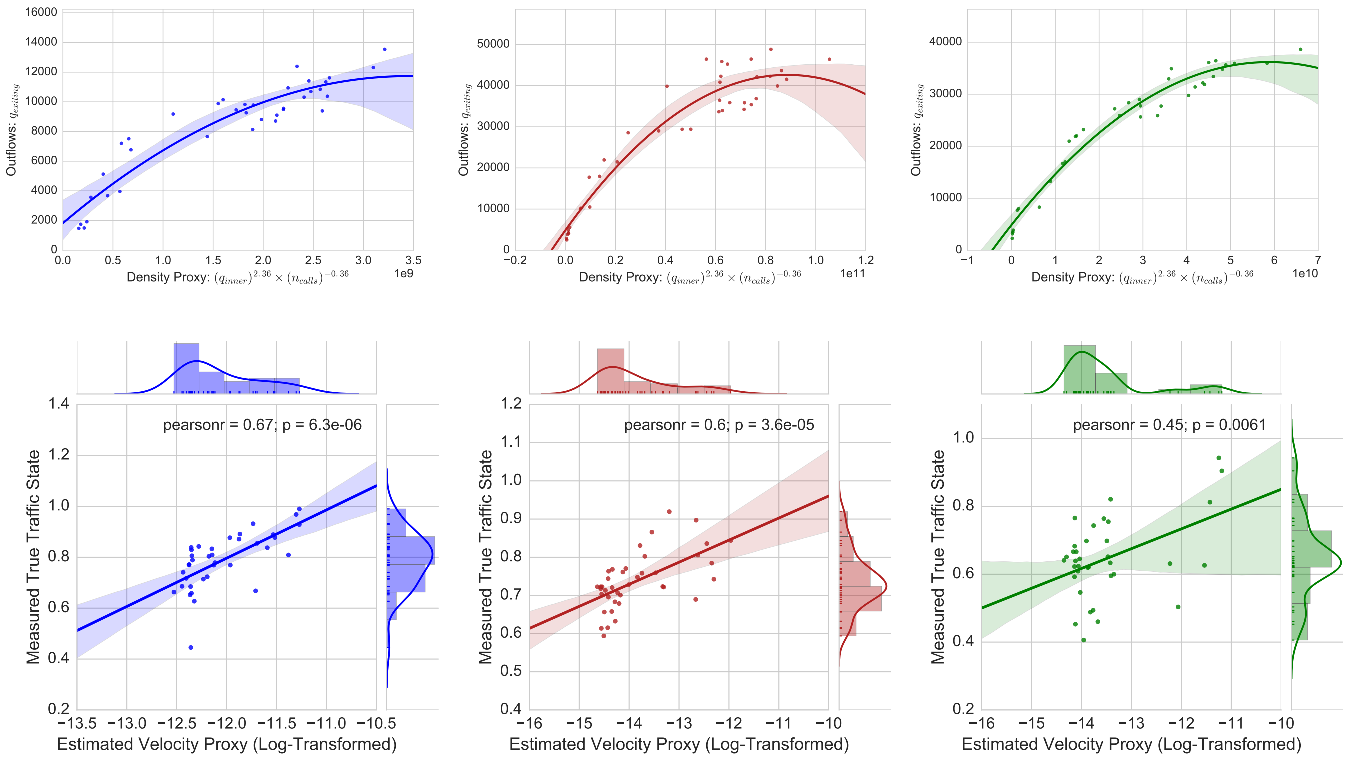

Figure 3: Clusters 7, 5 and 6: Profile functions, correlation between traffic states and velocity proxy (only validation dataset used). Colors match those in Fig. 2.

Fundamental Diagram (MFD), as known from traffic flow theory [9]. However, for a lack of very severe congestion in the road network, we do not observe the negative slope of the flow ratio characteristic of grid-lock and spill-back phenomena, but this is in line with studies on traffic data from other cities [11].

The results show that for the clusters containing a large percentage of mobile users, the presented data set is sufficient for traffic state prediction. The limitation of not knowing the proportion of stationary and pedestrian users is problematic in the business districts, and remains a drawback in comparison with more traditional traffic sensing technologies. This could be remedied with cell dwell time metrics. However, we are confident that the presented method works suitably well for the districts with higher vehicle/user ratios.

VI. Conclusion and Future Work

We have shown that profile functions of partitions of mobile networks exist, and that they exhibit predictive power for estimating the road network traffic state. Future work consists in evaluating whether or not these profile functions and Macroscopic Fundamental Diagrams (MFD) possess a theoretical link, and if that link holds for extreme congestion conditions, i.e. grid-lock. This could be studied further in a simulation setting. We made a first step in this direction with our work in [8], but a more extensive study and comparison to MFDs is necessary.

We have shown the feasibility of traffic state estimation using a common density proxy function along with a single regression model for highway and mixed clusters. While we showed that this works for the highway and transit clusters, the density function proved to be unfitting for the purely urban clusters. Therefore, we believe that fitting multiple density proxy function dependent on the cluster's road topology will improve density and velocity estimation. This will be the subject of a follow-up study.

As mentioned above, there are various extensions and directions for future work regarding the results we found. We believe that by using multiple radio technologies and finer-grained data can lead to better models of the underlying topology. Further improvements can be expected from improved clustering algorithms of the mobile network that lead to more homogeneous road network partitions.

The goal of this work was to get a single regression equation mapping to all the clusters, in order to show that mobile network clusters have characteristic profiles with predictive power. We have shown that this is indeed possible, and have reached a MAPE of 12.0% with respect to the true traffic states. It is possible to improve the resulting error by using fixed temporal effects for the typical daily road patterns, more radio access technologies and longer periods of data. Finally, comparing mobile network profile functions directly to the underlying road network fundamental diagrams would be a helpful next step, because this would tighten the theoretical link between both domains.

ACKNOWLEDGMENT

The authors would like to thank Post Luxembourg for providing the mobile network data and the National Research Fund of Luxembourg (FNR) for providing financial support through the CORE 2013 MAMBA project (C13/IS/5825301).

REFERENCES

[1] Y.-A. De Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel, "Unique in the crowd: The privacy bounds of human mobility," *Scientific reports*, vol. 3, p. 1376, 2013.

[2] M. S. Iqbal, C. F. Choudhury, P. Wang, and M. C. González, "Development of origin–destination matrices using mobile phone call data," *Transportation Research Part C: Emerging Technologies*, vol. 40, pp. 63–74, 2014.

[3] S. A. Di Donna, G. Cantelmo, and F. Viti, "A markov chain dynamic model for trip generation and distribution based on cdr," in *Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2015 International Conference on*. IEEE, 2015, pp. 243–250.

[4] D. Gundlegård, C. Rydergren, N. Breyer, and B. Rajna, "Travel demand estimation and network assignment based on cellular network data," *Computer Communications*, vol. 95, pp. 29–42, 2016.

[5] A. Janecek, K. A. Hummel, D. Valerio, F. Ricciato, and H. Hlavacs, "Cellular data meet vehicular traffic theory: Location area updates and cell transitions for travel time estimation," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp '12. New York, NY, USA: ACM, 2012, pp. 361–370. [Online]. Available: http://doi.acm.org/10.1145/2370216.2370272

[6] H. Bar-Gera, "Evaluation of a cellular phone-based system for measurements of traffic speeds and travel times: A case study from israel," *Transportation Research Part C: Emerging Technologies*, vol. 15, no. 6, pp. 380–391, 2007.

[7] K. Hui, C. Wang, and A. Kim, "Investigating the use of anonymous cellular phone data to determine intercity travel volumes and modes," 2017.

[8] T. Derrmann, R. Frank, and F. Viti, "Towards estimating urban macroscopic fundamental diagrams from mobile phone signaling data: A simulation study," 2017.

[9] N. Geroliminis and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings," *Transportation Research Part B: Methodological*, vol. 42, no. 9, pp. 759–770, 2008.

[10] Y. Ji and N. Geroliminis, "On the spatial partitioning of urban transportation networks," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1639–1656, 2012.

[11] C. Buisson and C. Ladier, "Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2124, pp. 127–136, 2009.

[12] B. Greenshields, W. Channing, H. Miller *et al.*, "A study of traffic capacity," in *Highway research board proceedings*, vol. 1935. National Research Council (USA), Highway Research Board, 1935.