Spatial and Temporal Analysis of Mismatch Between Planned Road Infrastructure and Traffic Demand in Large Cities

Jordan Ivanchev[1], Heiko Aydt[2], Alois Knoll[3]

Abstract—As large cities evolve in time, the traffic demand and the road network adapt to the mutually presented changes by each other. As a result of this process, previously planned roads and intersections that were designed according to some optimality criteria at the time, turn out to be suboptimal when traffic conditions change. This paper presents a method that can be used in order to identify intersections whose capacity is no longer in correspondence with the demand of vehicles on them and the choices agent make at those locations. Using real data from a survey describing the travel patterns of people in the city of Singapore we are able to model the routing choices of commuters and simulate the traffic demands on the road network. After calculating the turning probabilities on every intersection we are able to compare the traffic demand for every turn with the planned physical roads' capabilities. Furthermore we define a measure, which quantifies the deviation of the whole road network from the ideal demand-calculated values. We use these measures to evaluate the temporal and spatial profile of the mismatch between the roads and the demand for them. The measure is designed such that it is universal in nature and invariant to the absolute values of the traffic flows in the city. It can, therefore, be used to compare the proper utilisation of road networks among different cities.

I. Introduction

TRAFFIC conditions are mostly determined by two factors. The first one is the traffic demand profile of the population and the second one is the topology of the traffic network itself. Those two factors are strongly connected and a change in one of them will inevitable have an effect on the other. For example if the road network is changed by building a new highway or extending an already existing road, the commuters will adapt to this change and make use of the newly built road while reducing the traffic on others. Entrepreneurs might construct new business or industrial centres in proximity to the road, attracted by its high throughput. In a similar fashion, although at a relatively slower pace, if there is a much bigger demand in a certain area the corresponding officials will take action in order to somehow change the road network in order to satisfy those demands.

In order to formalise the description of the first factor, the traffic demand of the population, transportation researchers usually use the Origin Destination (OD) Matrix. It is used to structure information about the typical origins, destinations and timings of the commuters. The OD matrix answers the questions from where, to where and when does the city population want to travel. There are two main distinct ways to estimate the real origin destination profile of the population. The first one is by surveying a representative part of the population about their traffic habits and then scaling the results for the whole population. The second one makes use of information about the traffic flows collected from sensors on the road segments. Given the traffic flows there are methods to estimate the Origin Destination pairs and their intensities as in [1].

The second factor that mainly determines traffic is the road infrastructure, which is the medium that allows the traffic demands to be met. The topology of a network is a result of the evolution of the city. Nowadays, it is relatively easy to acquire a connected road network of almost every city. It is important to note that some vital attributes of a network are, however, quite hard to come upon such as the number of lanes for every road segment. This is a crucial trait of a road since it determines its capacity. It is rather uncommon to gain complete knowledge of a network including the capacity, which is one of the reasons why until now there is no full scale analysis of overall network capacity performance in big cities described in literature.

As a result of the constant interaction between the population and its road system, this very infrastructure is a result of an evolutionary process. As such, although it is fulfilling its functional requirements it may be suboptimal. It is possible that some already constructed roads that have been thought to be optimal at the time of planning turn out not to meet the demands any more or it is even possible that the demand is no longer existing. In other words, some roads may even become obsolete as a result of better alternatives appearing with time. This ever changing nature of network topology-traffic demand dynamics is the reason why there is a need of a measure that can evaluate if the traffic demands correspond to the potential the network has and vice versa. In this way the respective authorities can react to changes in demand in the most adequate way.

It is of great importance that the traffic demand for a road matches as closely as possible the capacity of the road. In case the demand is higher than the capacity there will be a congestion in the system that might spread to the whole network. This case is quite familiar and straight forward. The other extreme though is also highly undesirable. There are

[1]TUM CREATE, 1 CREATE Way, 10-02 CREATE Tower, Singapore 138602 `jordan.ivanchev@tum-create.edu.sg`
[2]TUM CREATE, 1 CREATE Way, 10-02 CREATE Tower, Singapore 138602 `heiko.aydt@tum-create.edu.sg`
[3]Technische Universität München (TUM), Institute for Informatics, Robotics and Embedded Systems, Germany `knoll@in.tum.de`

studies that suggest that if roads are widened thus increasing their capacity, the outcome may not be that intuitive. The increase of capacity improves the quality of travel and more people prefer to make trips using their own vehicles. In other words increasing the quality of the road network makes people travel more and even increase the level of congestion as a result of this as shown in [2].

In conclusion choosing just the right capacity of every road is the only way to ensure smoothly operating traffic. It is, therefore, desirable that the way traffic is spread onto the network is in sound agreement with the network capacity and its topology. Intersections are the places where cars make decisions to go either way. Therefore, the most fundamental approach to evaluating the mismatch between infrastructure and demand will be to take a closer look at what is happening at the intersections of the road network. Those are the locations where we should examine the choices that commuters make and see if they correspond to the actual capacities of the chosen road segments as shown in Fig. 1. We can then quantify the discrepancies between the two and observe their temporal and spatial distribution.

The main contributions of this work are:
- Measure of intersection capacity deviation from demanded capacities extracted from turning option probabilities
- Definition of an overall network proper utilisation factor
- Case study with real world data for the city of Singapore identifying problematic intersections and providing spatial and temporal analysis of road network utilisation.

II. LITERATURE REVIEW

In literature, intersections with high throughput of vehicles are referred to as "critical" [3] and [4]. Their performance can be modelled in various ways depending if they are unsignalised [5], [6] or use a control strategy such as traffic lights [7]. Modelling approaches include the use of Kirchhoff's law for traffic flows at nodes [8], queue dynamics as a function of demands [9] or turning probabilities [10], or modelling the routing choices [11].

If the number of passing commuters is higher than a certain threshold [12] it is recommended that the intersection is controlled using a traffic light. Furthermore, a statistical analysis of intersection performance has been done in [13] demonstrating that the performance significantly varies from site to site. It should be noted that all methods use data for and examine exclusively traffic conditions during rush hours, while neglecting the temporal nature of traffic dynamics throughout the day.

Traffic conditions are mainly determined by the flows of vehicles and the network structure [14]. Traffic analysis literature is mostly concerned with the traffic demands. In [15] for example the importance of a link is determined only by counting the number of commuters that pass through it. The latter factor, however, has also been analysed solely in topological studies in order to identify critical links [16] and vital links [17].

Graphical measures of whole networks are also studied and reviewed in [18] and information theoretic measures are
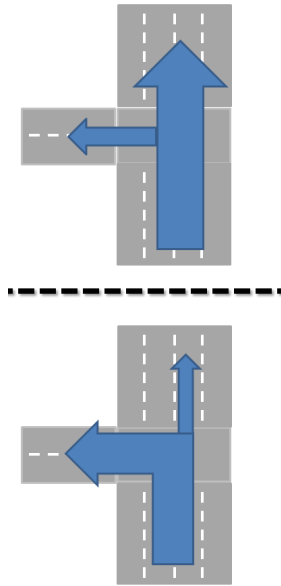

Fig. 1: A simplistic example showing a properly performing intersection in agreement with the traffic demand on top and an intersection where the infrastructure is significantly deviating from the traffic demand. In order to meet it, the roads corresponding to a left turn should have more lanes and the road corresponding to the vehicles going straight should have a reduced number of lanes. For the sake of simplicity of the visualisation we have assumed that there is no other incoming traffic for the intersection.

described in [19]. There are also efforts to combine the information from both the transportation network and traffic demand as in [20] where centrality measures are defined. The usage of the topological part, however, is not extensive.

Road networks are subject to evolution as a result of adaptation to changes as studied in Switzerland [21] and Milan [22]. As societies change and cities grow the traffic demands and the road network itself changes with a high degree of self-organization and spontaneous organization of hierarchies . The phenomenon is also observed in [23] where it is shown that it leads to imperfections of the once planned infrastructure.

III. THE MEASURE OF DEVIATION BETWEEN NETWORK CAPACITY AND TRAFFIC DEMAND AT INTERSECTIONS

In this section we introduce the measure of deviation of nodes capacity from traffic demand. We present and justify a mathematical formulation of the deviation of a node (intersections) and an overall deviation of a transportation network. Furthermore, we define an absolute measure of mismatch of an intersection that is measured in the number of lanes that need to be redistributed. Throughout the rest of the paper vehicles will be referred to as agents. Let us introduce some notation that will be used:

N_{ij} - number of cars that moves from node i to node j for the whole day

P_l - the path of the l-th agent

f_{ij}^l - function that is 1 if the sequence of nodes ij is in the path of agent l and 0 otherwise

A - a set containing all the agents

p_{ij}^t - probability that an agent that is at node i will continue on to node j during time period t

q_{ij}^t - turning probability that would be a perfect match for the infrastructure

S_i - set of nodes that are successors to node i

N_{ij}^t - number of cars that pass sequentially through node i and j during time period t

T - number of regions the day is split into

L - length of a time period

R - number of road segments in the network

w_{ik} - number of lanes on the road between nodes i and k

r_{ik} - ideal number of lanes between nodes i and k based on turning probabilities

m_i - absolute mismatch between number of lanes at an intersection i

Δ_{ij}^t - deviation measure of road from node i to node j during time period t

Δ_i^t - deviation measure of node i during time period t

Δ_{ij} - daily deviation measure of road from node i to node j

Δ_i - daily deviation measure of node i

$\hat{\Delta}_{ij}^t$ - corrected deviation measure of road from node i to node j during time period t

$\hat{\Delta}_i^t$ - corrected deviation measure of node i during time period t

$\hat{\Delta}_{ij}$ - corrected daily deviation measure of road from node i to node j

$\hat{\Delta}_i$ - corrected daily deviation measure of node i

$\hat{\Delta}_t$ - corrected overall deviation measure of whole network for time period t

$\hat{\Delta}$ - corrected overall deviation measure of whole network

C_{ij} - capacity of road from node i to road j per hour

G_{ij}^t - congestion factor for road from node i to node j for time period t

c_v - coefficient of variation

The following are the steps taken in order to calculate the deviation of a node and the degree of mismatch between the network and the traffic demand:

1) **Calculate turning probabilities:**
   Let N_{ij} be the number of cars that pass through the i-th node and after that through the j-th node and let P_l be the path of the l-th agent. Then let the function f_{ij}^l:

$$f_{ij}(P_l) = \begin{cases} 1 & \text{if nodes } ij \text{ are in } P_l \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

   Then:

$$N_{ij} = \sum_{l=1}^{|A|} f_{ij}^l(P_l) \quad (2)$$

   where $|A|$ is the number of agents.
   Let p_{ij}^t be the probability that an agent at node i continues to node j during time period t

Let S_i be the set of nodes that are successors of node i. Then we can define the turning probability as the ratio between the number of cars that pass through node i and then proceed to node j and the total number of cars that pass through node i :

$$p_{ij}^t = \frac{N_{ij}^t}{\sum_{k \in S_i} N_{ik}^t} \quad (3)$$

2) **Calculate intersection demand deviation value:**
   The goal of this measure is to estimate the degree to which the road infrastructure at the intersection is corresponding to the traffic demand. We have already defined turning probabilities at an intersection for every period of the day. We measure the degree of discrepancy between the demand and actual roads by comparing the ideal ratios between the roads' capacities extracted from the turning probabilities and the physical number of lanes of the respective roads.

   Let us first try to calculate what would be the best demand distribution such that the existing road width ratios are fulfilling the traffic needs perfectly. The ideal turning probability q_{ij} from node i onto a certain successor road ij can be calculated by dividing the width (number of lanes) of the successor road by the total width of all possible successors as shown in Equation 5. As a result of this, the optimal turning probabilities can only assume value of fractions of integers, since the number of lanes w_{ij} is an integer.

   The next step is to take the difference between the real turning probability and the ideal turning probability with respect to the already existing road infrastructure as shown in Equation 6. It is important to note that it is possible that the deviation value is non-zero even if the distribution of lanes on the successor roads is optimal due to the integer fractions that constitute the ideal turning probability, see Fig.2. Due to this fact, we also define the absolute mismatch measure m_i in Equation 4, which states how many lanes must be redistributed within an intersection in order to have an optimal performance. This measure constitutes of finding the difference between the actual number of lanes of all successors w_{ij} and the ideal number r_{ik}, where the total number of lanes coming out of the intersection is kept constant.

$$m_i = \frac{\sum_{k \in S_i} \|w_{ik} - r_{ik}\|}{2} \quad (4)$$

   Finally, in order to calculate the intersection demand deviation value Δ_i^t we take the average of all possible successors' deviations as shown in 7.

$$q_{ij}^t = \frac{w_{ij}}{\sum_{k \in S_i} w_{ik}} \quad (5)$$

$$\Delta_{ij}^t = \|p_{ij}^t - q_{ij}^t\| \quad (6)$$

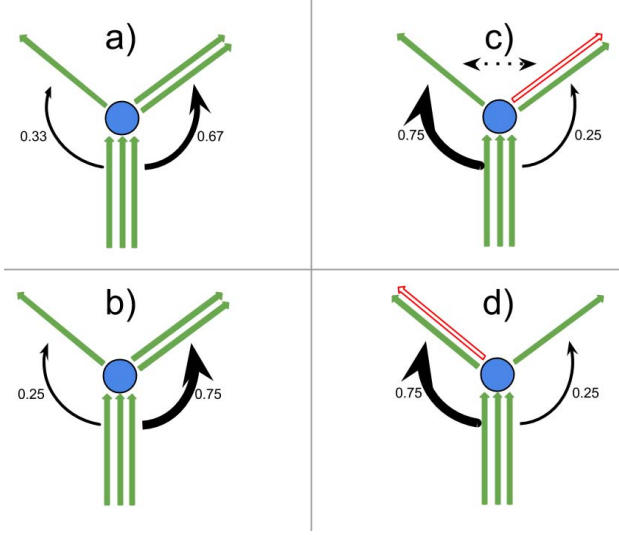$$\Delta_i^t = \frac{\sum_{k \in S_i} \Delta_{ik}^t}{\|S_i\|} \quad (7)$$

Fig. 2: Visualisation of the difference between the deviation measure and the absolute mismatch measure. In a) we can see a perfect agreement between the turning probabilities and the number of lanes, since both ratios are $2 : 1$. In b), however the ratio of the turning probabilities is $3 : 1$, while the lane ratio is still $2 : 1$. In this case the deviation measure Δ_i will be non-zero. The absolute mismatch measure m_i, however will still be 0 since given those turning probabilities and the total number of lanes to be distributed, which is 3, the optimal ratio of lanes is still $2 : 1$. In c) we can see a case where the absolute mismatch measure is non-zero. The turning probabilities are in ratio $1 : 3$, however, the lanes are in ratio $2 : 1$. In this case one lane should be moved from the successor road on the right to the successor road on the left as shown in d). The absolute mismatch measure is therefore $m_i = 1$.

3) **Weigh the deviation by the flow during the time periods:**

The time dependent deviation measures Δ_{ij}^t and Δ_i^t in the previous part can be used in order to evaluate the degree of change of the turning probabilities throughout the day and consequentially optimal intersection lane ratios. In order to make our analysis more complete we are, however, ultimately interested in a deviation value that represents the whole day rather than just one time period.

In order to achieve that, an averaging technique that would turn our time period observations into a representative all day value is needed. Naturally, the values during some periods are more important than others. We have chosen to measure this importance by the flow of vehicles through the respective turn ij relative to the overall flow for the whole day. As shown in Equation 8 the deviation of the turn for period t is just weighed by the ratio between the cars that passed through it during this time period and the whole flow throughout the day. In a similar fashion as in Equation 7 the total

daily deviation of the node is calculated by averaging the deviations of all possible turns as shown in Equation 9.

$$\Delta_{ij} = \left\langle \Delta_{ij}^t \frac{N_{ij}^t}{\sum\limits_{k=1}^{T} N_{ij}^k} \right\rangle_t \tag{8}$$

$$\Delta_i = \frac{\sum_{k \in S_i} \Delta_{ik}}{\|S_i\|} \tag{9}$$

4) **Extension for computation of the overall deviation of the network infrastructure from the traffic demand**

In order to extend this methodology to compute a measure that represents the mismatch between the whole network infrastructure and the traffic demand, our already defined measures should be normalised in order to enable comparison among distinct cities. The term that actually needs normalisation is the number of vehicles that pass through every road segment ij. In order to be able to compare one city to another a more global measure must be used. One that takes into consideration the performance of a road rather than the absolute value of its throughput. We propose the usage of the congestion factor measure that constitutes the flow through the node over its capacity as shown in equation 11. The capacity of a segment ij, C_{ij} is defined as the number of lanes multiplied by a standard number of cars that can pass per lane per hour and is usually set to 2000 [24].

$$C_{ij} = 2000 L w_{ij} \tag{10}$$

$$G_{ij}^t = \frac{N_{ij}^t}{C_{ij}} = \frac{N_{ij}^t}{2000 L w_{ij}} \tag{11}$$

In order to get a better picture of the overall road network deviation from the traffic demand we should weigh the individual intersection deviation values by the already computed congestion factor. In this way, if an intersection has a high deviation value and is congested it will receive a higher weight that an intersection that has the same degree of deviation but the traffic conditions on it are still on a satisfactory level. Therefore, we include the congestion factor as a correction in all the already computed node deviations expressions as shown in Equations 15 to 17

$$t_{max} = \max_t G_{ij}^t \tag{12}$$

$$G_{ij} = G_{ij}^{t_{max}} \tag{13}$$

$$\hat{\Delta}_{ij}^t = \Delta_{ij}^t G_{ij}^t \tag{14}$$

$$\hat{\Delta}_{ij} = \Delta_{ij} G_{ij} \tag{15}$$

$$\hat{\Delta}_i^t = \frac{\sum_{k \in S_i} \hat{\Delta}_{ik}^t}{\|S_i\|} \tag{16}$$

$$\hat{\Delta}_i = \frac{\sum_{k \in S_i} \hat{\Delta}_{ik}}{\|S_i\|} \tag{17}$$

After the measures are normalised we can proceed to calculate the overall deviations of the network. In order to do that we take the generalized mean with power

factor $\alpha = 2$ over all $\hat{\Delta}_i$ as in equation 19 and 18. The generalised or power mean is used in order to put an emphasis on the extreme values in the distributions that demonstrate higher deviations. We also calculate the whole city deviation value for every time period separately in order to observe if there are any peculiar dynamics of the deviation value throughout the day.

$$\hat{\Delta}^t = \left(\frac{1}{N} \sum_{i=1}^{N} \left(\hat{\Delta}_i^t \right)^\alpha \right)^{\frac{1}{\alpha}} \tag{18}$$

$$\hat{\Delta} = \left(\frac{1}{N} \sum_{i=1}^{N} \left(\hat{\Delta}_i \right)^\alpha \right)^{\frac{1}{\alpha}} \tag{19}$$

IV. MODEL AND CASE STUDY

In order to collect all the data needed by the methodologies described in the previous sections, the traffic in a specific realistic scenario should be modelled and simulated. We have chosen the city of Singapore with population 5.4 million people spread over 715 squared km of land area and around 1 million registered vehicles. It is an island city, which further simplifies the scenario since there is no significant flow of vehicles in or out of the city. There are 652 km of major arterial roads and 161 km of express ways in the city. The scenario for our experiment is as follows:

A. Agent Generation:

The data source used for agent generation is the Household Interview Travel Survey (HITS) of Singapore done in 2012. It contains information of a representative sample of people (about 0.7 % of the population) that state their daily travel patterns. For every person that participated we can extract the information about all trips that are made on a daily basis consisting of origin point, destination point and starting time of the trip. We use those "sample" points in order estimate an OD matrix for the city and its variation in time.

After that we use the OD matrix in order to generate agents with itineraries that consist of several trips throughout the day(depending on the agent) consisting of origin, destination and a start time. The sampling procedure basically consists of choosing one itinerary from the survey data at random and then adding noise to it in both space and time domains (sample random origins and destinations from a neighbourhood around the original ones and add mean 0 Gaussian noise to the starting time). The total number of agents that we generate is around $300,000$ with starting times of trips varying throughout the day. Most of the agents have two trips in their itineraries, one in the morning (between $6:30$ and $9:30$) and one in the evening (between $5:30$ and $7:30$), however there is traffic in the city throughout the whole day.

B. Route Calculation:

After the agent generation is completed, the origins and destinations in the itineraries of every agent are used to calculate the shortest routes that should be taken. Since we aim to represent reality as much as possible the routing of the generated agents is stochastic. Some people prefer the shortest path, some the fastest and some prefer comfort rather than speed or time. We therefore have 3 different ways to calculate our routes. We are able to realize the distinct routing variations by calculating the weights on our routing graph in a different way. After that we use a shortest path algorithm that minimizes the sum of the weights for a path between origin and destination. The three types of weights are:

1) $w_d =$ road length - minimising distance

2) $w_t = \dfrac{\text{road length}}{\text{road speed}}$ - minimising time

3) $w_c = \dfrac{\text{road length}}{\text{road speed} \times \text{number of lanes}}$ - maximising comfort

and the probabilities that an agent has a certain preference have been set to $\dfrac{1}{3}$ for each type of weights, which is in agreement with literature about route choices as [25].

After the generation of every agent one of the three preferences is chosen at random and the corresponding route is calculated. After the paths are computed, we have the sequence of links that every agent passes through, which is then used in order to calculate the turning probabilities. The day is split into 48 pieces consisting of 30 minute intervals. In order to calculate the dynamic factor of every node for the whole day, we use the starting time of the trip as a time-stamp of the whole route and in this way organize the passages through the links in time. The measure is calculated as described in section III.

C. Deviation Measures Calculation:

In order for the presented information to be easier to assimilate we have only examined nodes (intersections) with a throughput higher than 10000 cars per day. Even though, we account for the vehicle flow in the congestion factor part of our definitions, ultimately we are interested in the more central locations. First, we calculate the deviation factor (Δ_i) of every node that satisfies our constraints and examine its distribution as shown in Fig. 3a

On Fig. 3a it can be observed that the distribution resembles a log normal one with a peak at around 0.3. The maximum deviation of a node in the examined network is 0.8. It should be mentioned that the maximum possible value of Δ_i is 1. An example case of how the value can become 1 is one where all agents are systematically (throughout the whole day) turning onto a one lane road while no agents are turning onto a road with more than one lane coming from the same intersection. It can also be observed that the intersections that have a deviation value of 0 (perfect match) or lower than 0.1 are only around 100 while the ones with a value around the peak, which is at 0.3 sum up to almost 1500.

Although most of the intersections are not perfectly matching the demand, the lane ratios between the successors might be optimal. In other words, the distribution of the lanes leading
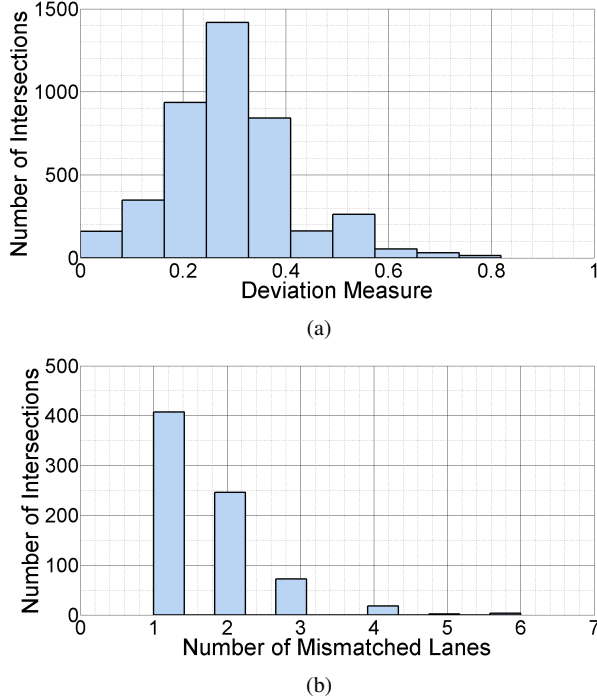
(a)

(b)

Fig. 3: Distributions of the measure Δ_i and the mismatch of number of lanes between road capacities and traffic demand. 3a presents the distribution of the deviation measure for every intersection in the city of Singapore that has a throughput higher than 10000 vehicles per day. The maximum value of the deviation is 1, which means that the demand of the commuters is inversely proportional to the capacities of the road system. The minimum is 0, which shows that the demand is perfectly met by the infrastructure. 3b is the distribution of the number of number of mismatched lanes, based on the deviation of the road infrastructure and the traffic demand, on intersections with throughput higher than 10000 in cases where this number is bigger than 0. A value of 1 means that one lane should be moved from one successor road of the node to another in order for the road structure to be in optimal agreement with the traffic demand.

to the different roads can still be optimal even if the deviation value is still not 0 as shown in Fig.2. It will be, therefore, interesting to examine the absolute mismatch measure m_i as well. It is looking at specific changes that need to be made in order to get an optimal match. The measure represents the number of lanes that should be transferred from one road to another in order to optimally satisfy the demand. We can see the distribution of this measure for the same set of intersections on Fig.3b, however we only include the intersections that need a redistribution of lanes.

Out of 4223 intersections 3476 do not need any readjustment in the number of lanes of the roads. We see that there are around 400 intersections that can benefit from one lane being moved from one road to the other and there are 3 intersections that can benefit from redistributing 6 lanes within

the intersection. The existence of places with such extreme deviations indicates that actually we are observing a fat tail in the distribution of the deviation between traffic demand and infrastructure capacity among the network. Such fat tail that cannot be observed from the distribution of the deviation measure since it has an upper bound of 1.

Next, let us examine the temporal nature of our measures. On Fig. 4a we see the evolution in time of the measure Δ^t, which summarizes the deviation of the whole network into one value for every time period. We can observe that the deviation has a stable nature with a coefficient of variation $c_v = \dfrac{\sigma}{\mu}$ of only 0.0019. This allows us to claim that the overall deviation of the network is time invariant.

On Fig.4b we can see the evolution in time of the measure Δ_i^t for the intersection that shows the highest overall deviation. We can still observe that there is no significant variation as the coefficient of variation c_v is 0.0470. We can observe, however, that the curve is relatively peaking throughout the morning $(7:30-9:30)$ and evening $(17:30-19:30)$ rush hours, which means that at this particular intersection, the mismatch between demand and road capacities is growing bigger with increasing traffic volumes.

One would expect the opposite to be true since the intersections should have been optimised to perform best during rush hour, since the deviations during them, are actually the ones that can lead to traffic jams. A high deviation value at midnight may not be of such a concern since even if the road capacities do not match the demands, the vehicle flows are not big enough for a feasible effect on the traffic conditions to be sensed. A significant mismatch at rush hour, however, will inevitably lead to high degrees of congestion and overall reduced network performance.

Interesting peculiarities can also be observed in the spatial distribution of mismatched intersections on Fig. 5. We can observe the biggest cluster of mismatches forming at the central business district (CBD), as it is the most dynamically changing location in the city. The high pace of emerging of new buildings and businesses that attract new employees or move old ones to new places inevitably results in mismatch of the traffic demand with the relatively slowly changing road structure.

We can also observe that the other most mismatched intersections (red and dark orange) are indeed key intersections that connect the down town area with the east, west and northern parts of the city. Those areas have been growing extensively in the past decade as a results of Singapore's government attempt to relax traffic demand in the down town area by creating self sufficient districts in different parts of the island. As a result of this fast growth, the road network is lagging in its development and thus we can observe such levels of mismatch between demand and infrastructure capabilities.

V. Conclusion

In the presented work we point out the possible mismatches between intersection capacities and traffic demand. We predict that this phenomenon will be observed due to the fact that commuter's origins and destination vary in the long term and

Fig. 5: Spatial distribution of mismatched intersections in the city of Singapore according to the m_i measure. Big and red dots represent high number of mismatching lanes (maximum is 6), while green and small dots represent intersections with lower number of mismatched lanes (minimum is 1).

the topology of the network may not be able to adapt to this process with the desired pace.

First, we define a set of measures that can help us identify intersections that are not meeting the demands on them by using the turning probabilities at every intersection as a central source of information. After that, we define measures for the overall deviation of the network's topology from the demands of the commuters. Thirdly, we make use of a congestion factor metric in order to evaluate the utilisation degree of roads rather than the absolute flow values. This approach makes the defined measure more universal in the sense that it can be used at a later stage of research in order to compare individual cities to each other invariant of city specific traits.

We have conducted a real life case study for the city of Singapore in order to test our hypothesis and methods. Following, we have modelled the routing choices of drivers using a stochastic routing approach and calibrated it using a statistical survey performed among 0.7% of the population. The defined measures were implemented and the spatial and temporal nature of the mismatches between road infrastructure and traffic demand were examined.

We have observed that the distribution of the deviation of the intersections is log normal with a fat tail. This agrees with our initial hypothesis that there are intersections that severely "disagree" with the traffic demands. Using our absolute mismatch measure of an intersection we have located intersections that need as many as 6 lanes redistributed in order to be in agreement with the turning probabilities.

The measure of the overall deviation of the city road net-

work topology seems to be time invariant with extremely small degree of variation. The evolution in time of the individual intersection deviation measure demonstrates weak correlations and relatively higher values throughout rush hours, which further stresses on the seriousness of the mismatch issue, since the degree of deviation is peaking at the worst possible time, when the demand on the roads is the highest.

The spatial distribution of the intersections with high degrees of mismatch from the traffic demand is also examined. We observe that the connections to and from regions with high degree of dynamics, such as business districts and fast growing sub-cities are experiencing the highest levels of mismatch between infrastructure capacity and traffic demand. This further strengthens our hypothesis that those deviations occur in cases where the traffic demands of the population change faster than the road topology can adapt.

It would be of great interest to perform the same study for other big cities with dynamic traffic conditions and to compare the overall city measure of deviation. Such studies can be used in order to validate our findings from Singapore that intersections with high degrees of deviation from traffic demand exist. As we compare the distributions of our measure we can also observe qualitative differences between the deviation profiles. We expect that bigger cities might experience an even "fatter" tail in the distribution of their deviation measures.

The authors of this paper would like to carry out a study where the traffic network is actually changed in order to minimise the deviation of the intersections' infrastructure from the traffic demands and to quantify the overall traffic
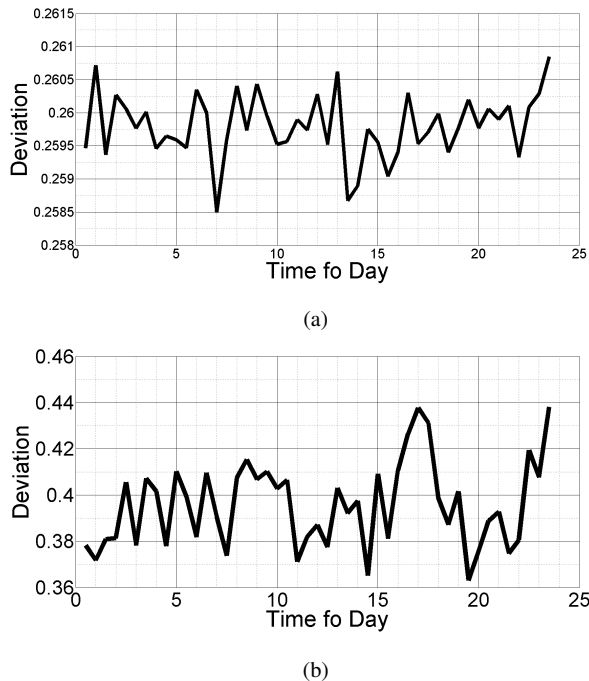
(a)

(b)

Fig. 4: Evolution of the deviation measures in time. Fig. 4a shows the measure Δ^t for different values of t. This is the deviation of the road infrastructure of the whole network from the traffic demands. On Fig.4b we can see the evolution in time of the measure Δ_i^t of the intersection with highest degree of deviation.

improvement of the system as a result of using our measure as an optimisation heuristic function.

REFERENCES

[1] P. Robillard, "Estimating the o-d matrix from observed link volumes," *Transportation Research*, vol. 9, no. 2, pp. 123 – 128, 1975. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0041164775900490

[2] G. Duranton and M. A. Turner, "The fundamental law of road congestion: Evidence from us cities," *The American Economic Review*, pp. 2616–2652, 2011.

[3] R. L. Gordon, "A technique for control of traffic at critical intersections," *Transportation Science*, vol. 3, no. 4, pp. 279–288, 1969. [Online]. Available: http://dx.doi.org/10.1287/trsc.3.4.279

[4] M. Aymerich and A. Novo, "Madrid critical intersections antiblocking strategies." in *Road Traffic Monitoring, 1992 (IEE Conf. Pub. 355)*, 1992, pp. 70–.

[5] S.-X. Yuan, Y.-S. An, X.-M. Zhao, and H.-W. Fan, "Modeling and simulation of dynamic traffic flows at non-signalized t-intersections," in *Power Electronics and Intelligent Transportation System (PEITS), 2009 2nd International Conference on*, vol. 2, Dec 2009, pp. 1–5.

[6] H. J. Ruskin and R. Wang, "Modeling traffic flow at an urban unsignalized intersection," in *Computational Science-ICCS 2002*. Springer, 2002, pp. 381–390.

[7] C. Gershenson and C. L. Apostel, "Self-organizing traffic lights," in *Complex Systems*. Citeseer, 2005.

[8] S. Lämmer and D. Helbing, "Self-control of traffic lights and vehicle flows in urban road networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 04, p. P04019, 2008.

[9] G. Stephanopoulos and P. Michalopoulos, "Real time control of critical intersections based on queue dynamics," in *Decision and Control including the 17th Symposium on Adaptive Processes, 1978 IEEE Conference on*, Jan 1978, pp. 1018–1024.

[10] J. Budiarto, S. Sulistyo, I. Mustika, and A. Infantono, "Road density prediction: Updated methods of turning probabilities and highway capacities manual for achieving the best route," in *Electrical Engineering and Computer Science (ICEECS), 2014 International Conference on*, Nov 2014, pp. 168–173.

[11] T. xing and X. Zhou, "Reformulation and solution algorithms for absolute and percentile robust shortest path problems," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 14, no. 2, pp. 943–954, June 2013.

[12] M. Dongfang, S. Xianmin, T. Pengfei, and W. Dianhai, "Critical traffic volume warrant of signal installing at equal weight intersections," in *Intelligent Computation Technology and Automation (ICICTA), 2011 International Conference on*, vol. 1, March 2011, pp. 388–391.

[13] L. Qian and X. Chen, "Peak volume characteristics of urban road intersections: Case in shanghai," in *Computer and Network Technology (ICCNT), 2010 Second International Conference on*, April 2010, pp. 397–401.

[14] E. Jenelius, "Network structure and travel patterns: explaining the geographical disparities of road network vulnerability," *Journal of Transport Geography*, vol. 17, no. 3, pp. 234 – 244, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0966692308000550

[15] M. Schreckenberg and R. Selten, *Human behaviour and traffic networks*. Springer, 2004.

[16] D. M. Scott, D. C. Novak, L. Aultman-Hall, and F. Guo, "Network robustness index: A new method for identifying critical links and evaluating the performance of transportation networks," *Journal of Transport Geography*, vol. 14, no. 3, pp. 215 – 227, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0966692305000694

[17] H. Corley and D. Y. Sha, "Most vital links and nodes in weighted networks," *Operations Research Letters*, vol. 1, no. 4, pp. 157 – 160, 1982. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0167637782900207

[18] F. Xie and D. Levinson, "Measuring the structure of road networks," *Geographical Analysis*, vol. 39, no. 3, pp. 336–356, 2007. [Online]. Available: http://dx.doi.org/10.1111/j.1538-4632.2007.00707.x

[19] M. Dehmer and A. Mowshowitz, "A history of graph entropy measures," *Information Sciences*, vol. 181, no. 1, pp. 57 – 78, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0020025510004147

[20] S. Porta, V. Latora, F. Wang, S. Rueda, E. Strano, S. Scellato, A. Cardillo, E. Belli, F. Cardenas, B. Cormenzana *et al.*, "Street centrality and the location of economic activities in barcelona," *Urban Studies*, vol. 49, no. 7, pp. 1471–1488, 2012.

[21] A. Erath, M. Löchl, and K. W. Axhausen, "Graph-theoretical analysis of the swiss road and railway networks over time," *Networks and Spatial Economics*, vol. 9, no. 3, pp. 379–400, 2009.

[22] E. Strano, V. Nicosia, V. Latora, S. Porta, and M. Barthélemy, "Elementary processes governing the evolution of road networks," *Scientific reports*, vol. 2, 2012.

[23] F. Xie and D. Levinson, "Topological evolution of surface transportation networks," *Computers, Environment and Urban Systems*, vol. 33, no. 3, pp. 211 – 223, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0198971508000653

[24] R. Roess and W. McShane, "Changing concepts of level of service in the 1985 highway capacity manual: some examples," *ITE Journal*, vol. 57, no. 5, 1987.

[25] S. Bekhor, M. E. Ben-Akiva, and M. S. Ramming, "Evaluation of choice set generation algorithms for route choice models," *Annals of Operations Research*, vol. 144, no. 1, pp. 235–247, 2006.