Analysis of Urban Traffic Network Vulnerability and Classification of Signalized Intersections

Georgios Sarlas and Anastasios Kouvelas

Institute for Transport Planning and Systems
Department of Civil, Environmental and Geomatic Engineering
ETH Zurich
Stefano-Franscini-Platz 5, 8093 Switzerland
{georgios.sarlas, anastasios.kouvelas}@ivt.baug.ethz.ch

Abstract—This work studies urban road traffic networks and discusses potential ways to assess the criticality of urban links and nodes. A series of indicators are defined that are useful to evaluate spatial and temporal features and facilitate the identification of congestion hot spots. One way to achieve this is by utilizing concepts from graph theory (e.g. connectivity, efficiency, betweenness). These indicators are based on the structure of the network (graph) and do not require any traffic data. On the other hand, one could utilize historical traffic data (e.g. flows, densities, average speeds) in order to evaluate the importance of an urban link (or node) to the overall performance of the network. Such an analysis could capture congestion dynamics, spill-backs, and queues propagation in the network. Here, we combine both approaches mentioned above to come up with a classification (or ordering) of the different links and nodes inside an urban zone. This classification can then be used for real-time traffic control purposes, e.g. a city can choose the intersections that need to be instrumented and thus reduce the operational cost of online traffic management. Finally, we discuss the correlations between the different indicators and how one could take this into account in the decision making process.

Index Terms—urban traffic congestion; graph theory; simulation; intersection classification; congestion hot spots.

I. INTRODUCTION

Urban road traffic networks constitute an important component of the infrastructure of modern metropolitan areas. Maintaining a good level of operation is crucial for the mobility of people and goods, and hence for the overall function of a city. The importance of traffic networks is highlighted further when various emergency and evacuation situations emerge. For all these reasons, traffic management has attracted a lot of attention in the recent literature. Real-time control can assist in increasing the capacity of the infrastructure, by identifying critical bottlenecks in the network and taking some actions in the correct direction. Recurrent and non-recurrent congestion is a daily phenomenon in most cities that has negative social, economic, and environmental impacts. Locations that can be characterized as congestion hot spots may arise due to problematic design of the road network, but also atypical congestion patterns. Therefore, both spatial topological analysis and temporal congestion propagation are key elements of defining the critical links or nodes for a traffic network.

A recent approach that tries to tackle the problem of real-time network-wide traffic control is the perimeter flow control (or gating) [1]. The basic concept of such an approach is to partition heterogeneous large-scale cities into a small number of homogeneous regions (zones), and apply perimeter control to the inter-regional flows along the boundaries between regions. It utilizes the concept of the urban Network or Macroscopic Fundamental Diagram (NFD or MFD), which comprises an aggregated relationship between zone-wide density and flow [2]. Previous research [3], has shown that the master-slave concept may arise, as some region can be “sacrificed” and led to congested states in favour of other regions that are more “important” for the total performance of the system. This perimeter control concept can be viewed as an upper-level control layer.

Subsequently, it has been proposed to add another lower-level control layer (see e.g. [4], [5] for two recent studies), in order to tackle local disruptive congestion phenomena. Local distributed and/or coordinated regulators (see e.g. [6]) can effectively deal with fast queue dynamics and short links, and prevent spill-backs and local close-to-gridlock conditions. As a result, by combining the upper NFD/MFD layer with the lower local layer in a hierarchical control framework one can achieve better performance for the network (see the results reported in [4], [5]). An interesting research question is how to choose the critical intersections inside each zone, so as to minimize the operational cost for the city authorities.

Furthermore, a plethora of approaches have emerged over the years aiming at tackling the criticality issue. Essentially, the use of such approaches is commonly referred to as vulnerability analysis. As defined by [7], “Vulnerability in the road transportation system is a susceptibility to incidents that can result in considerable reductions in road network serviceability”. As summarized by [8], the various suggested vulnerability analysis approaches have developed along two main lines; the first one attempts to identify the most critical elements by studying the demand-and-supply interaction mechanism (e.g. [9], [10]), while the second examines the issue in terms of topological properties (e.g. [11], [12]). An overview of the recent literature on the topic can be found in [8], [13].
For the case of urban networks, and in connection to the aforementioned identification approaches, a possible pathway is to study and analyze the historical data that are collected over the years in traffic management departments, usually through inductive loop detectors. There are rich databases available with measurements of flow, density, and speed for urban links. The most commonly used state variable in traffic control literature is the queue length, which is quite difficult to measure, thus is usually estimated from available raw measurements. A good proxy of queue (or density) are the time occupancy measurements that are collected by loop detectors; under certain conditions they can provide unbiased estimates of queue lengths. In the current work, we utilize simulated occupancy measurements to compute traffic indicators, i.e. queue length heterogeneity and level of congestion. The simulated inductive loop detectors replicate the measurements that are collected in real life (e.g. percentage of time that the detector is occupied). Section II-A provides a comprehensive description of the data analysis and the definition of the corresponding traffic-related indicators.

Another possible pathway is to focus on the network structure and exploit different graph theory concepts. Essentially, the underlying assumption in that case is that the topological properties of a network constitute a focal element of its performance. To that end, centrality indicators can quantify the importance of each network component by calculating the number of shortest paths passing through them; as a matter of fact, it has been found that they can serve as proxy variables for traffic flow (see e.g. [14]). In addition, identifying the impact of various network disruptions on the connectivity, in terms of network efficiency, can also be viewed as a reflective indicator of criticality.

This work integrates the indicators provided by graph theory with the ones related to traffic data, in order to obtain a generic metric that can capture the importance of each signalized intersections. Essentially, by using this metric, one can rank the nodes of a network (or the links) according to their criticality. The most critical intersections of an urban network (hot spots) can be chosen to be instrumented (with hardware and software)\(^1\) and apply control in real-time. In the next sections we describe in details all the metrics that have been used, we visualize the results for a case study, and we discuss the correlations of the different indicators along with the various rankings.

II. METHODOLOGICAL FRAMEWORK

A. Dynamic analysis using traffic data

In this section we use data generated by a microsimulation environment (Aimsun). Virtual detectors are installed in the middle of every arterial link, and the software provides the time-occupancy measurements for every control cycle (i.e. 90 sec). As mentioned earlier, these measurements are proxies of link densities (or queues) and can be utilized for control purposes. Nevertheless, here we investigate the impact that different locations of the network have to the overall performance of the system. For our analysis, we define two indicators that describe the state of each link (or node). More precisely, the following metrics are determined for all the incoming links of a signalized intersection:

- **Mean occupancy**: The average occupancy is computed over time and space; essentially this represents the space-time mean occupancy.
- **Mean SD occupancy**: The standard deviation (SD) of occupancy is computed over space (i.e. for all incoming links). Then, for all the simulation horizon, the average of all SDs over time is computed; essentially this represents the time mean of the space SD of occupancy measurements.

These two indicators are used to describe the traffic conditions, and represent the level of congestion (former) and the heterogeneity (latter) of each node. The higher their value the higher the probability of spill-backs in this intersection and local breakdown (failure) of the system. These two indicators are utilized to quantify the criticality of each node in terms of traffic loads and heterogeneity among links, and together with the ones described in the next section, provide a topological network-wide evaluation and result in a comprehensive classification described in Section II-C.

B. Static network analysis using graph theory

In addition to the indicators constructed on the basis of the dynamic analysis of traffic data, graph theory indicators are constructed as well. In contrast to traffic data indicators, their graph theory counterparts are lacking the ability to capture various dynamics of the network, since topologically the network remains unchanged over time. Nevertheless, two indicators are calculated to quantify the importance of links (or nodes), namely betweenness centrality (BC) and network efficiency change (NEC). In the case of the former, the indicator measures the number of shortest paths between all node pairs that pass through each link and it was introduced in [15] for the case of nodes (vertices) and in [16] for links (edges). Thereupon, the disruption of links/nodes with high centrality values deteriorates the network performance, as a high number of paths are diverted to longer paths. Note that the shortest paths are retrieved for a directed graph in terms of free-flow travel time.

Mathematically speaking, let \( n_{s,t}^i \) denote the number of shortest paths from \( s \) to \( t \) that pass through link \( i \), and let \( n_{s,t} \) be the total number of shortest paths from \( s \) to \( t \); then the betweenness centrality of link \( i \) is defined as:

\[
b_i = \sum_{s \neq t} \frac{n_{s,t}^i}{n_{s,t}}
\]

In the case of NEC, the indicator is constructed based on the average efficiency \( E(G) \) of the graph \( G \) that contains \( N \) nodes,
defined as:

\[ E(G) = \frac{1}{N(N-1)} \sum_{s \neq t} \frac{1}{d_{st}}, \]  

with \( d_{st} \) denoting the shortest path distance (in this case free-flow travel time). Subsequently, each link \( i \) of the network is removed iteratively and the average efficiency is re-calculated. The relative change on average efficiency is then defined as:

\[ \text{NEC}_i = \frac{E^i(G) - E(G)}{E(G)} \]  

It should be noted that both indicators are calculated on a link level in order to quantify the importance of the links. Given that the focus of traffic control lies on an intersection level, and in analogy to the previously introduced traffic indicators (Mean occupancy and Mean SD occupancy), we define the mean and the standard deviation of BC and NEC per node, respectively, by taking into account the values of all incoming links. Therefore, the following indicators are defined:

- **Mean BC**: The average betweenness centrality is computed over space.
- **SD BC**: The standard deviation (SD) of betweenness centrality is computed over space (i.e. for all incoming links).
- **Mean NEC**: The average network efficiency change is computed over space.
- **SD NEC**: The standard deviation (SD) of network efficiency change is computed over space (i.e. for all incoming links).

In general, the four indicators constructed above are utilized to quantify the criticality of each node, and together with the ones described in the previous section, provide a topological network-wide evaluation, and result in the comprehensive classification described in the next section.

**C. Discussion of indicators used for ranking and classification**

In the previous two sections six performance indicators have been introduced to assess the importance/criticality of an urban signalized intersection. Half of them correspond to a metric about the level of congestion of each node, and the other half to the spatial homogeneity of the measurements around this node. In this section we present a simple but elegant way to combine all the different indicators and derive a generic ranking of the nodes based on their criticality. The procedure that we follow is described below:

1) We compute each of the 6 indicators independently, and for each of them we make a ranking \( r \) of the nodes; this results in 6 different rankings.

2) For each node \( i \) we compute the following measure:

\[ R_i = \sum_{j=1}^{6} r_{ij}^2 \]  

where \( r_{ij} \) is the position in the ranking (e.g. 1, 2, \ldots, N) that each node \( i \) has for the different indicators \( j \in \{1, 2, \ldots, 6\} \). Then we sort the nodes based on the new measure \( R_i \) and the ones with the higher values are the most critical.

This process generates a comprehensive ranking for the nodes of the network and utilizes all the performance indicators presented here, both from graph theory and real traffic data. Obviously, this is not a universal methodology and more sophisticated techniques could be utilized to integrate the different metrics. In Section V we discuss the correlations between different indicators and some limitations of the proposed approach. Essentially, one could come up with many different measures (e.g. considering also the correlations between indicators), and there are many aspects that could be exploited further in future research.

**III. CASE STUDY**

**A. Urban network description**

The urban network of Barcelona, Spain is used as a test site to demonstrate the results (Figure 1(a)). The network covers an area of 12 km\(^2\) with about 600 intersections and 1500 links of various lengths, and it is modelled and calibrated in Aimsun microsimulation software. The number of lanes for through traffic varies from 2 to 5 and free flow speed is 45 km/h. Traffic lights at signalized intersections operate on multiphase fixed-time plans with constant (but not equal) cycle

![Fig. 1. Test case network: (a) map of Barcelona CBD (source: www.here.com maps); (b) partitioning into 4 homogeneous regions.](image-url)
lengths. For the simulation experiments, typical loop-detectors have been installed around the middle of each network link. The OD-based demand consists of 123 origin centroids and 132 destination centroids and provides a good replication of real life conditions, as it generates realistic traffic congestion patterns in the network.

The duration of the simulation is 5 hours, including 15 minutes of warm-up (beginning) and 45 minutes of cool-down (end) periods. A set of the real fixed-time plans that have been acquired by the city authorities are applied to the signalized intersections, and the network faces some serious congestion problems, with queues spilling back to upstream links. Note, that in our simulated scenario drivers adapt to the traffic conditions in real-time (3 minutes intervals) through a C-Logit route choice model which comprises a standard module of the simulator software; therefore, the distribution of demand into the network is more realistic. Note that previous works [17], [18] have shown (for different models and network configurations) that driver adaptivity increases the performance of large-scale networks and decreases hysteresis loops in the NFD/MFD, which is closer to real-life observations.

B. Network partitioning and NFDs/MFDs

Traffic congestion in the city of Barcelona is unevenly distributed, creating multiple pockets of congestion in different areas of the network. As NFD/MFD depends on the distribution of link densities (occupancies, speeds), partitioning heterogeneous loaded cities with uneven distribution of congestion into homogeneous regions is a possible solution to take advantage of well-defined NFDs/MFDs. In fact, the outflow of the network is a function of both the average and variance of link densities. By partitioning, we aim at grouping spatially-connected links with close density values within a cluster, which increases the network flow for the same average density. Spatial connectivity is a necessary condition that makes the application of perimeter control strategies feasible.

The partitioning algorithm used in this study is an optimization framework called “Snake” algorithm [19], which considers heterogeneity index as a main objective function, and contiguity is imposed explicitly by linear constraints. Note that this approach requires the desired number of clusters as a predefined input and obtains the optimal number of clusters by evaluating the heterogeneity metric for different cases. By applying this algorithm, the network of Barcelona is partitioned into 4 homogeneous regions that are shown in Figure 1(b). This partitioning simplifies the network dynamics as there is no need for routing information/decisions (i.e. due to the configuration of the regions there is only one choice to move from one region to another).

IV. CLASSIFICATION RESULTS

In this section we present the results of our analysis. We aim at classifying the signalized intersections into two groups, these of high importance (critical set of nodes) and the complement set. The important intersections are critical for the traffic conditions of the network and the propagation of congestion, thus they are selected to be instrumented and apply real-time adaptive control. This selection/classification can be done offline like in the present work, but essentially, it could also be performed in an online manner. In that case, the traffic operators would be able to choose the critical intersections in real-time from a set that includes all the instrumented junctions (with hardware, communications, etc.).

A. Graph theory analysis

Figure 2 presents a plot of mean betweenness (for all the incoming links) versus the SD of this indicator for all the intersections of the test network. Note that this can be computed only by utilizing the topological information of the network and the shortest paths (as described in Section II-B), and the units of the axes is number of shortest paths. Figure 3 demonstrates the locations of the most critical intersections according to mean and SD of betweenness and equation 4.

Similarly, in Figure 4 we report the same analysis for the efficiency of each intersection of the network. Again, the mean and SD are computed for all the incoming links, and we follow

![Fig. 2. All the intersections for the 4 regions and the critical ones according to BC criterion (red points).](image)

![Fig. 3. Network visualization for critical intersections according to BC criterion.](image)
equation 3 to compute the relative change in the efficiency of the graph. Finally, Figure 5 depicts the positions of these critical intersections on the Aimsun model of the Barcelona CBD.

B. Traffic data analysis

Figure 6 presents a plot of (spatio-temporal) average occupancy versus the SD of the occupancy for all the intersections of the test network. The data come from 5 hours simulation and the indicators are calculated by aggregating first in space (i.e. over the node’s incoming links) and then in time (i.e. over the control cycles of the simulation horizon). With red color we indicate the intersections that have the highest values for the criterion of equation 4, if we apply it for the two indicators (axes) of this plot. Figure 7 demonstrates the locations of these critical intersections on the Aimsun model of the city.

Note that depending on the indicator that we study the pattern of critical intersections on the map of the network can be substantially different. Graph theory is focusing on the vulnerability of the network, which is a structural component, while traffic data provide results that depend on demand patterns and the physical behavior of traffic flow dynamics. Next section presents a final classification that integrates both sources of information.

C. Final classification

All the aforementioned network performance indicators generate different rankings regarding the criticality of the signalized nodes and different topmost sets (compare the node configurations in Figures 3, 5, and 7). The question at hand is to come up with a unique classification of intersections to critical and non-critical. Here, we utilize equation 4 and consider all the rankings for the different indicators. This way, we integrate the information we get from graph theory (vulnerability) and traffic flow data (congestion propagation). Figure 8 presents the final result. It is worth noting that the final configuration that we obtain for the critical nodes of each region is quite reasonable, as it matches well with well-known locations that constitute congestion hot spots in reality (e.g. Avinguda Diagonal).
the different indicators and try to derive a more sophisticated model which takes into account the different rankings and their interactions (i.e. equation 4). Here, a basic methodology is proposed (in conjunction with other spatial configurations), in order to identify congestion hot spots in urban road traffic networks. Indicating the vulnerability of a graph network are combined with traffic measurements in order to come up with an integrated assessment. Fusing these two different sources of information can lead to more concise results as demonstrated with the case study. Our future plans include the application of max-pressure controller [6] to the critical set of nodes (in conjunction with other spatial configurations), in order to validate this classification in terms of traffic performance (e.g. total delays).

Another topic that deserves further investigation is the method used to combine the different performance indicators (i.e. equation 4). Here, a basic methodology is proposed which takes into account the different rankings and their variance. For example, one could consider the correlations of the different indicators and try to derive a more sophisticated model that could take this into account. In Figure 9 we present all the computed correlations among the different indicators. Clearly some of them are highly correlated, providing similar information. Equation 4 does not take this into account, and it would be interesting to investigate how the results reported here would change with the use of a more statistically rigorous model. Finally, the ultimate validity test for all the results would be to apply the controller to all different produced sets of critical intersections and evaluate the results.

REFERENCES