A Multiclass BCMP Queueing Modeling and Simulation-Based Road Traffic Flow Analysis

Jouhra Dad, Mohammed Ouali, and Yahia Lebbah

Abstract—Urban road network traffic has become one of the most studied research topics in the last decades. This is mainly due to the enlargement of the cities and the growing number of motor vehicles traveling in this road network. One of the most sensitive problems is to verify if the network is congestion-free. Another related problem is the automatic reconfiguration of the network without building new roads to alleviate congestions. These problems require an accurate model of the traffic to determine the steady state of the system. An alternative is to simulate the traffic to see if there are congestions and when and where they occur. One key issue is to find an adequate model for road intersections. Once the model established, either a large scale model is built or the intersection is represented by its performance measures and simulation for analysis. In both cases, it is important to seek the queueing model to represent the road intersection. In this paper, we propose to model the road intersection as a BCMP queueing network and we compare this analytical model against a simulation model for validation.

Keywords—Queueing theory, transportation systems, BCMP queueing network, performance measures, modeling, simulation

I. INTRODUCTION

The size and complexity of transport problems continue to increase with the growth of cities, road networks, and the number of motor vehicles. The main issue in urban transportations is road congestion. Road congestion is due to several factors such as the infrastructure, the ratio of number of vehicles with respect to the capacity of the road network, and the traffic signaling to name a few. Moreover, road traffic depends heavily on the time of the day: rush hours generally occur at the time people commute to and from work, 8am and 4pm, and around lunch time, 12pm. This pattern makes road traffic non ergodic. Despite this problem, a decent amount of research effort was devoted to traffic flow modeling and simulation. When dealing with road traffic analysis, both modeling and simulation are viable alternatives. However, depending on the nature of problems at hand, one alternative may overtake over the other. The demand in terms of road space continues to grow for the reasons mentioned above. If the status quo persists—no new roads are built or no structural nor organizational changes are made, congestions are unavoidable.

Their impacts are important and multiple. They result in economic, social, and environmental costs. It is thus necessary to limit, or at least, to manage road congestions. This can be done by limiting the request for traffic or by managing the flow of vehicles. We propose to consider tools to probe road traffic network performances. These tools may be viewed at several levels: a verification level, where the urban network is subject to probing to detect potential congestions, and a surveillance level, where the tool analyzes the traffic flow in real time to detect future congestions and their upstream repercussions on the road network. Vehicular traffic problems are usually treated in the literature, and much research has focused on methodologies for the optimization and evaluation of transportation systems [1]. Lozano et al. [2] present an algorithm for identifying levels of congestion in traffic problems. D’Ambrogio et al. [3] propose a model for an urban road network made up of traffic intersections. Other research presented an analytical queueing model that preserves finite capacity queues and uses parameters to investigate the correlation between the queues [4]. This model is validated later by mathematical methods and existing simulation results. Some studies measure the size of the queues of road intersections in order to find points of congestion in urban networks. In this paper, we propose to revisit the modeling and simulation of road intersections. This study is twofold. First, we want to find an adequate analysis model to represent an intersection at the analysis level. Second, we want to validate the analysis model by performing a traffic simulation and determining performance measures of the road intersection for comparison with the analysis model. For the time being, we will consider the simulation model as a verification means to validate the analysis model. Road intersections are divided into sectors. The analysis model is built as a queueing network, where each queue represents a sector. Ultimately, the whole intersection is modeled by a queue having a given performance measures. The analysis of a road network will be the natural extension of this model. The remaining of the paper is organized as follows: Sect. 2 describes the specification of the intersection traffic; Sect. 3 presents the simulation and analytical models; finally, Sect. 4 shows the experimental protocol and results.

II. ROAD INTERSECTION TRAFFIC SPECIFICATION

Let us consider a road intersection. Two roads, having two lanes each, intersect. In each road, lanes have opposing traffic...
flow (the roads are two-way traffic). The intersection traffic is managed by a set of traffic lights (See Fig. 1).

![Fig. 1 Traffic intersection with its arrival streams and crossing trajectories](image)

An intersection includes two axes. Each axis contains bidirectional traffic. At the intersection, motor vehicles are allowed to go straight ahead, to make a right turn, or to make a left turn. The traffic lights in a given axis have the same color: they are green simultaneously and switch to the red at the same time. The vehicles are stopped near the traffic light at the stop line when the traffic light is red. For the sake of simplicity, we assume there is no yellow light. Normally, left turning vehicles have a lower priority than that of the vehicles coming in opposite direction. To simplify the model, we assume no priorities for left turning vehicles. However, as the traffic is not chaotic—since vehicles are not allowed to make U turns, we distinguish several traffic classes: each class is determined by the traffic flow origin. Our model is, finally, a priority less multiclass queueing network.

### A. Description of the Intersection

The main parameters of a road intersection are the number of crossing roads, the number of lanes in each road per traffic direction, the arrival rate—the number of vehicles arriving at the intersection per time unit, the service rate. Ideally, for each parameter, we should have a set of values, each value corresponding to a period of time of the day, for instance, early morning, lunch time, late evening, and periods of commutation from and to work. This description defines the system its structural and the behavioral views.

#### 1. Structural View

It identifies the static elements of an intersection, in other words the input and output sectors corresponding to input and output lanes and the internal sectors, used by vehicles to cross the intersection. In our case, an intersection sector is an intersection part that can be occupied by only one vehicle at a time. One may consider larger intersections, where more than one vehicle may be in internal sectors. Sectors of the intersection illustrated in Fig. 1 are: input sectors (I1, I2, I3, and I4); internal sectors (S1, S2, S3, and S4); and output sectors (O1, O2, O3, and O4).

A vehicle entering an intersection first goes through an input sector, then through one or more internal sectors if the traffic light is green, and finally through an output sector.

![Fig. 2 Simplified structural view of the intersection shown in Fig. 1.](image)

To implement the fact that U turns are not allowed and to enforce trajectory routing, we can set a crossing trajectories matrix that tells the potential trajectory for each class of traffic. Table I shows the potential connections between input and output sectors. Each value in the matrix indicates if there is a route between input and output sectors. Each value in the matrix indicates if there is a route between input and output sectors. Each value in the matrix indicates if there is a route between input and output sectors. Each value in the matrix indicates if there is a route between input and output sectors. Each value in the matrix indicates if there is a route between input and output sectors.

<table>
<thead>
<tr>
<th></th>
<th>O1</th>
<th>O2</th>
<th>O3</th>
<th>O4</th>
</tr>
</thead>
<tbody>
<tr>
<td>I1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>I2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>I3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>I4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

### B. Vehicle Classes

Traffic classes or vehicle classes can be organized in different ways. We can consider a class of traffic as vehicles having the same origin—i.e. coming from the same input section, or as vehicles having the same destination—i.e. heading to the same output sector. We choose to define a traffic class as the flow of vehicles coming from the same input sector (with the same routing table, same routing probabilities, and same service rates). Consequently, four flow classes are identified:
Class 1: flow entering intersection through input I1 and leaving it on one of the following output sectors: O2, O3 or O4. The internal sectors of this class are S1, S3 and S4. 
Class 2: flow entering intersection through input I2 and leaving it on one of the following output sectors: O1, O3 or O4. The internal sectors of this class are S1, S2 and S4. 
Class 3: flow entering intersection through input I3 and leaving it on one of the following output sectors: O1, O2 or O4. The internal sectors of this class are S2, S4 and S3. 
Class 4: flow entering intersection through input I4 and leaving it on one of the following output sectors: O1, O2 or O3. The internal sectors of this class are S2, S4 and S3. 

Regarding potential routes (between sectors), each class of traffic has a routing probabilities table that determines the ratio of vehicles going to an output sector or another transitional internal sector. Routing probabilities are estimated and are mainly used in the simulation model.

III. SIMULATION AND ANALYTICAL MODELS

A. Simulation Model

Our simulation model consists of a set of connected service stations. A service station represents a resource in a real system (e.g., an intersection sector) and is composed of a queue where jobs wait for a server and a service. Service stations are cascaded in a way that the output of a server enters one or more servers’ queue as explained in the previous section. The links between stations reproduce the connections between resources of the real system (e.g., connections between input sector and internal sectors of an intersection). Vehicles are transferred between any two stations according to given routing probabilities. For example (see Table II), a class 1 traffic vehicle entering from input I1 goes to sector S1 with probability R_{11S1}=1 then it either leaves the intersection through output sector O2 with probability R_{S1O2} or it transitions to sector S4 with probability R_{S1S4}. From sector S4, it either leaves the intersection through output sector O3 with probability R_{S4O3}, or it transitions again to sector S3 with a probability of R_{S4S3}. Finally, from sector S3 it leaves the system by through output sector O2 with probability R_{S3O2}.

The simulation model takes as input the formal specification of the intersection—the set of input sector with their arrival and service rates, the set of internal sector with their service rates, viable routes between internal sectors (crossing trajectories matrix), routing probabilities to determine the ratio of vehicles turning right/left or going straight ahead. The simulation produces performance measures. Meaningful performance measures in the context of road traffic are the usage rate, the average queue length of input sectors, the average response time, the average waiting time, the average dwelling time, and the visit rate of a sector by a particular traffic class.

In the simulation model, the input buffers have virtually an infinite capacity. Conversely, internal sectors buffers have a capacity of 0 (only one vehicle is in the queue and it uses the station). During simulation, vehicle arrivals are generated at entry sectors. Each vehicle belongs to a traffic class. We assign to each generated event a route as prescribed by the routing probabilities table. When a route choice is possible, a random number is generated to decide which transition to choose (according to the routing probability). Vehicles are not allowed to change class within the system. Once the traffic light is green in the input sector, we check that the station of the next internal sector is free. If it is the case; a service time is generated according to the exponential distribution of parameter μ to serve the vehicle at the entry of the intersection. Once a vehicle is in an internal sector, we generate a routing probability. According to this probability, if the next sector is an output sector; the vehicle leaves the intersection by this sector. If not, it passes to the next internal sector and so on.

A variable representing the state of the traffic light, at the entry section, controls the propagation of arriving vehicles. When the traffic light is red, the arriving vehicles are put in the appropriate queue and their dwelling and waiting times are increased by the number of red light duration it witnesses.

B. Analytical Model

As queueing networks are used to model and analyze physical systems, they, in turn, make it possible to evaluate performance measures and to better understand the behavior of these systems. A queueing network may be open (Jackson’s), closed (Gordon/Newell’s), or mixed. The results of Jackson and Gordon/Newell were extended by Baskett, Chandy, Muntz, and Palacios [5,9]. This type of queueing network seems well adapted to model road traffic.

1. Open BCMP Network

We consider an open BCMP queueing network model that consists of M service stations, which are in our case sectors of an intersection. Each service station contains a single-server with first-come-first-served (FCFS) policy. Each queue in our model has only one server, arrival rates following Poisson distribution with parameter λ, and the service rates following exponential distribution with parameter μ. The queues having these characteristics are of type M/M/1 FCFS which implies

<table>
<thead>
<tr>
<th>Class</th>
<th>I1</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>R_{11S1}</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

As queueing networks are used to model and analyze physical systems, they, in turn, make it possible to evaluate performance measures and to better understand the behavior of these systems. A queueing network may be open (Jackson’s), closed (Gordon/Newell’s), or mixed. The results of Jackson and Gordon/Newell were extended by Baskett, Chandy, Muntz, and Palacios [5,9]. This type of queueing network seems well adapted to model road traffic.

1. Open BCMP Network

We consider an open BCMP queueing network model that consists of M service stations, which are in our case sectors of an intersection. Each service station contains a single-server with first-come-first-served (FCFS) policy. Each queue in our model has only one server, arrival rates following Poisson distribution with parameter λ, and the service rates following exponential distribution with parameter μ. The queues having these characteristics are of type M/M/1 FCFS which implies
that sectors of our queueing network are of type 1 (service policy is FCFS).

The service times of an FCFS node must be exponentially distributed and class-independent $\mu_{i1} = \mu_{i2} = \cdots = \mu_{iR} = \mu_i$ [6]. We assume that the number of vehicles in each class at each sector is always non negative. The unit entity (e.g., a car, bus, etc.) that is routed through the network is called a vehicle. Vehicles are grouped into $R$ different classes. For the sake of simplicity, and to be consistent with the simulation model, we assume that vehicles do not change classes while transitioning through the intersection.

In the BCMP model, there are $M$ stations and $R$ traffic classes. For each class, we must specify routing probabilities through the network. A class can either be open (vehicles enter from outside and eventually leave) or closed (vehicles loop and never leave the system) [7]. The traffic classes can differ in their service times and in their routing probabilities.

We will use the following notation to describe the open BCMP network applied to our road intersection model:

- $R$: the number of traffic classes in the network,
- $K_r$: the number of vehicles of the $r$th class at the $i$th sector.
- $\mu_{ir}$: the service rate of the $i$th sector of the $r$th class.
- $R_{ir,js}$: probability that a vehicle of the $r$th class at the $i$th sector is transferred to the $s$th class and the $j$th sector (routing probability)—these should be zero for $r \neq s$, since vehicles cannot change classes.
- $R_{0,js}$: the probability in an open network that a vehicle from outside the network enters to the $j$th sector of the $s$th class—or arrival rate at input sector $j$.
- $R_{ir,0}$: the probability in an open network that a vehicle of the $r$th class leaves the network after having been served at the $i$th station.
- $\lambda_{0,ir}$: the arrival rate from outside to the $i$th sector of the $r$th class.
- $\lambda_{ir}$: the arrival rate of vehicle of the $r$th class at the $i$th sector.
- $C_q$: Chain of classes.

The service rate of each single server is state-independent and the sectors are of type $M/M/1$. The vehicles of the intersection can enter and leave the system through input and output sectors respectively. Moreover, as provided by the theory, each queue of the network has an infinite buffer.

![Fig. 3 Model of an open BCMP queueing network for the intersection of two axes and two ways per axis](image)

2. Performance Analysis

In this section, we establish performance measures that we compute from the analytical model. Let $\rho_{ir}$ denote the utilization of the $i$th sector by $r$th class vehicles, $K_{ir}$ is the average number of vehicles of the $r$th class at the $i$th sector, $\overline{Q}_{ir}$ is the average queue length of class $r$ vehicles at the $i$th sector, and $\overline{W}_{ir}$ is the average waiting time of class $r$ vehicles at the $i$th sector. $\overline{T}_{ir}$ is the average response time of class $r$ vehicles at the $i$th sector.

After deriving the product-form equation, the determination of the performance measures described above is straightforward. These measures are given by equations (1) to (7), $i$ and $j$ indicating sector numbers, while $r$ indicates traffic class:

1. $e_{ir} = R_{0,ir} + \sum_{j=1}^{N} e_{jr} R_{jr,ir}$
2. $\rho_{ir} = \frac{\lambda_{ir}}{\mu_{ir}}$ [2]
3. $K_{ir} = \frac{\rho_{ir}}{1 - \rho_{ir}}$ [3]
4. $\overline{Q}_{ir} = \lambda_{ir} \overline{W}_{ir}$
5. $\overline{T}_{ir} = \frac{\overline{K}_{ir}}{\lambda_{ir}}$ [5]
\[ \bar{W}_{ir} = \bar{T}_{ir} - \frac{1}{\mu_{ir}} \]  
\[ \rho_i = \sum_{r=1}^{R} \rho_{ir} \]  

IV. EXPERIMENTAL RESULTS

Prior to show experimental results on the whole system, we propose to compare BCMP and simulations performance measures of a single M/M/1 queue, and a two cascaded M/M/1 queues BCMP network. This will help explain discrepancies between the analytical and the simulation models.

A. Single Service Station Queueing System

We consider an M/M/1 queue with \( \rho = \frac{3}{4} \) and \( \rho = \frac{3}{5} \). As the stability condition \( \rho < 1 \) is fulfilled, the system is then stable. Table III summarizes performance measures of the analytical and simulation model along with the absolute and relative errors. The relative error is a good indication of the accuracy of our simulator. Relative errors are around 6% which is a very acceptable error.

\[ \bar{Q} = \frac{\rho^2}{1-\rho} \]  
\[ \bar{T} = \frac{1}{\mu} \]  
\[ \bar{K} = \frac{\rho}{1-\rho} \]  
\[ \bar{W} = \frac{\rho}{\mu} \]  

B. Two Cascaded BCMP Service Stations Queueing Network

In this experiment, we show the performance a two cascaded BCMP service stations network. Similarly, we conducted a comparison of the analytical and simulations measures. We consider the one traffic class (\( r=1 \)) and two queues (\( i=1,2 \)). For \( \lambda = 1, \mu_1 = 2, \mu_2 = 3 \), arrival and service rates, Table IV summarizes the results. The relative error remains very low for the average waiting time \( \bar{W} \), still acceptable for the average response time \( \bar{T} \), and average for the average queue length \( \bar{Q} \). There are discrepancies between performance of queue 1 and queue 2. These disparities are mainly due to an insufficient number of simulation events.

<table>
<thead>
<tr>
<th>Performance measure</th>
<th>( \bar{Q} )</th>
<th>( \bar{T} )</th>
<th>( \bar{W} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical results queue 1</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Simulation results queue 1</td>
<td>0.6</td>
<td>0.88</td>
<td>0.53</td>
</tr>
<tr>
<td>Relative error</td>
<td>20%</td>
<td>12%</td>
<td>6%</td>
</tr>
<tr>
<td>Theoretical results queue 2</td>
<td>0.17</td>
<td>0.5</td>
<td>0.17</td>
</tr>
<tr>
<td>Simulation results queue 2</td>
<td>0.12</td>
<td>0.41</td>
<td>0.12</td>
</tr>
<tr>
<td>Relative error</td>
<td>29%</td>
<td>18%</td>
<td>29%</td>
</tr>
</tbody>
</table>

C. Full BCMP Queueing Network

In this section, we consider the performance comparison of the multiclass open BCMP network-based analytical model with the simulator—coded in C language. Before comparing the set of performance measures, we explain the following parameters and how they were adapted:

1. Arrival Rate

This rate defines the number of vehicles that arrive at a sector per time unit. This parameter is still under investigation to fit with real cases. We took \( \lambda = 0.33 \) vehicle per second as arrival rate in internal sectors. This shall be adapted as internal sectors should not have buffers.

2. Service Rate

Similarly, this rate defines the number of vehicles crossing a sector per time unit. We took \( \mu = 0.66 \) vehicle per second for input sectors and \( \mu = 1.0 \) vehicle per second for internal sectors.

3. Traffic Signal Duration

Traffic lights are characterized by their duration. Usually, there are three periods: green, yellow, red. For the sake of simplicity, we only consider green and red. We considered both color with the same duration of 30 seconds. However, as traffic lights are not modeled in the analytical model, we adapted the arrival rates at input sectors to account for the period of time where vehicles arrive but do not cross the stop line because of the red light.

The most important result in an intersection is to determine the number of vehicles waiting in input sectors since this measures conditions the occurrence of congestion or at least higher values are precursor of congestions. Fig. 4 shows the comparison the average queue length in input sectors from both analytical and simulation results. Already, with the parameters that we have just introduced, the size of the queues at input sectors obtained by simulation correspond to those obtained by the analytical model, which fairly indicates that the BCMP model might be appropriate for an intersection. The use of the conditional here is due to the fact that the other performance measures obtained from both models are significantly different from each other. We are still
investigating these disparities in order to either find an explanation or correct the analytical model.

![Graph showing comparison between simulation and analytical model](image)

**Fig. 4 Number of vehicles at the intersection input sectors.**

### V. CONCLUSION

The queueing theory is a technique to model systems with waiting phenomena. The theory allows to calculate systems’ performance measures and to determine their features.

In this paper, we proposed a model to represent road intersection traffic. The purpose is to analyze urban road networks in order to evaluate the network and assert that it is congestion-free, or to detect congestions along with their “when and whereabouts”. Unfortunately, urban road networks are large and it is out of question to try and write performance measures equations for the whole network: it is just overwhelming. Instead, one can proceed by simulating the network. Each intersection is represented by a node with known performance measures. Ultimately, it is interesting to know the stationary state performance measures as this minimizes discrepancies due to simulation—insufficient simulation events for example. For this purpose, the first step is to find a queuing model to represent the road intersection, which is sought here in this contribution. We proposed to consider BCMP queues type to build the BCMP queueing network of the intersection as we deal with different traffic classes. Each queue represents a sector of the intersection. In parallel, we programmed a simulator to emulate the vehicle behavior at the intersection and we computed average performance measures to use for comparison with the analytical model.

Possible crossing trajectories were randomly chosen: for each entering vehicle, all possibilities are kept open. For each traffic class, routing probabilities were also randomly defined so that the simulation is not biased by a particular choice. The experimental results showed some disparities between the analytical and the simulation models. These discrepancies are due to the fact that traffic lights are not modeled in the analytical model, although we managed to compensate for this by considering half the arrival rate at input sectors. Another difference is the zero capacity buffers for internal sectors used in simulation, and the infinite capacity buffers used in the analytical model. This is a significant difference that needs to be addressed.

The direct perspectives of this work are the consideration of the zero capacity buffers for internal sectors as this will bring the analytical and simulation model close to each other for thorough comparison and will enforce the well-fondness of the drawn conclusions. Another less important perspective is the modeling of the traffic lights in the analytical model to avoid tweaking input sectors arrival rates for compensation. Finally, the model should take care of traffic priorities to take into account left-turning vehicles. The introduction of priorities will allow evaluating the performance gain with the introduction of a separate traffic light and an extra lane for turning vehicles.

### ACKNOWLEDGMENT

J. Dad would like to thank Dr. Kaddour for helpful comments as well as for providing useful documentation pointers.

### REFERENCES


