Dynamic Traffic Modeling: Approximating the equilibrium for peak periods in urban areas

Raphaël Lamotte

Nikolas Geroliminis

LUTS, EPFL

April 2015

STRC | **15th Swiss Transport Research Conference**

Monte Verità / Ascona, April 15 – 17, 2015

LUTS, EPFL

Dynamic Traffic Modeling: Approximating the equilibrium for peak periods in urban areas

Raphaël Lamotte, Nikolas Geroliminis
Urban Transport Systems Laboratory
École Polytechnique Fédérale de Lausanne
GC C2 389, Station 18, 1015 Lausanne, Switzerland
phone: +41-21-693 24 81
fax: +41-21-693 53 97
{raphael.lamotte,nikolas.geroliminis}@epfl.ch

April 2015

Abstract

This article addresses the problem of the choice of departure time in cities subject to heavy congestion. The model originally developed by Vickrey and extended by many authors was developed for bottlenecks with constant capacity. While this assumption might be reasonable for physical bottlenecks submitted to light congestion, empirical evidence shows that in urban networks, the trip completion rate strongly depends on the density. Several analytical approaches have been proposed to address this shortcoming, but they all rely on some simplifications. In this work, the impact of some of these simplifications on the traffic flow dynamics is first evaluated using simulation. Then, a heuristic is proposed to approximate the equilibrium with more realistic assumptions and the results are compared with those that can be obtained with the widely used Method of Successive Averages (MSA). While analytical solutions assuming a constant capacity remain useful to design system optimum solutions, the heuristic proposed is complementary and allows addressing the user equilibrium problem. This talk concludes by presenting how this tool can be used to create strategies to alleviate peak-hour congestion in urban areas.

Keywords

Vickrey's bottleneck, Macroscopic Fundamental Diagram, Equilibrium, Heuristic

1 Introduction

Congestion during the peak hour on road networks represents a major part of the costs associated to transportation and it is at the same time one of the costs that can be most easily reduced with appropriate measures. However, designing such measures requires the combination of two fields of transportation engineering that have historically been considered separately: temporal demand modeling and congestion modeling.

On one hand, temporal demand modeling has mostly built on Vickey's bottleneck model. This model considers as given the knowledge of users' desired arrival time, value of earliness, lateness and travel time and aims at predicting their departure time, taking into account the resulting congestion. There have been many extensions but they almost always assume a simple bottleneck model with a capacity that is either a constant or a random variable, but does not depend on the demand. However, it is widely accepted in transportation engineering that the capacity of most facilities does depend on the demand, as evidenced by gridlock situations.

On the other hand, congestion modeling usually considers the inflow as given. At the very local level, the fundamental diagram relating the speed to the density of vehicles is probably the most widely used. However, the congestion observed during the peak hour is rarely a local phenomenon: it propagates along the highway and over urban networks. Thus, modeling the peak hour requires a network-wide approach. The most famous work on this issue is probably METROPOLIS, the software developed by De Palma and his colleagues de Palma and Marchal (2002). This software permits considering many different phenomena simultaneously and has been used in many cities across the world. However, since it models every link of a network individually, this approach is not only time-consuming but also requires a significant amount of data. These reasons motivated the use of an alternative approach, based on the Macroscopic Fundamental Diagram (MFD), an extension of the Fundamental diagram that accurately describes the overall performances of a network. By dramatically reducing the complexity of the network, the equilibrium can be approximated much more rapidly, which potentially allows for in-depth sensitivity-analyses or for the design of congestion-reducing measures.

The first part of this article introduces different models for both temporal demand and congestion while the second part introduces a heuristic to approximate the equilibrium that takes advantage of some ordering property of the equilibrium.

2 Modeling the dynamics of the peak hour

Modeling the dynamics of the peak hour requires modeling both the demand and the supply, two topics that have traditionally been addressed separately. In agreement with the related literature, this section introduces separately some models for both applications, highlighting their advantages and limitations, before rapidly reviewing the literature combining both supply and demands.

2.1 Vickrey's bottleneck model

Vickrey (1969) introduced a simple framework to model the choice of departure time when commuters need to pass a single bottleneck that has a constant capacity. If during some time interval there are more commuters that want to pass the bottleneck than allowed by the capacity, users start queuing and some of them do not arrive on time. This would then lead users to change their departure time to reduce their travel time (caused by queuing) or their schedule delay penalty (caused by an early/late arrival). The sum of the cost associated to travel time and of the schedule delay penalty represents a personal objective function, that every user seeks to minimize. Finding the equilibrium means finding a distribution of departure times such that, given the decisions of all the other users, one cannot improve one's cost by a change of departure time.

With the appropriate assumptions, the evolution of the demand over time can be very simply derived. Using the conventional assumptions and notations, let us assume that all users value travel time with the same linear function (coefficient α) and also have the same piece-wise linear schedule penalty function (with coefficient β for earliness and γ for lateness). The objective function to be minimized for each user is:

$$C(t) = \alpha tt(t) + \beta max(t^\star - t - tt(t), 0) + \gamma max(t + tt(t) - t^\star, 0), \quad (1)$$

where $tt(t)$ represents the travel time if departing at time t, and $t^\star$ represents the user's preferred arrival time. At equilibrium, $\frac{dC}{dt} = 0$. Thus, one can easily show that chosen departure times leading to early arrivals are characterized by $\frac{dtt}{dt} = \frac{\beta}{\alpha}$ while those leading to late arrivals impose $\frac{dtt}{dt} = -\frac{\gamma}{\alpha}$. Thus, by assuming only linear cost functions, Vickrey (1969) showed that the travel time should follow a triangular function during the peak hour. This result was later extended to allow for relaxed hypotheses (e.g. heterogeneous users (Newell, 1987), existence of unobserved variables (de Palma *et al.,* 1983)), although these often prevent obtaining analytical expressions.

Then, predicting the rate of departures and arrivals requires modeling congestion. For this purpose, Vickrey (1969) assumed a bottleneck with a constant capacity. While this model allowed him to obtain analytical results very simply, there are very few instances of isolated bottlenecks with constant capacity in the real world.

2.2 Congestion modeling: the MFD

The idea of an MFD relating the total accumulation in a network to the average speed is quite old (Godfrey, 1969). It has been reintroduced by Daganzo and Geroliminis (2008), and has attracted a growing interest since Geroliminis and Daganzo (2008) presented the first empirical results supporting this theory. In fact, real world measurements in the city of Yokohama showed two important results. First, Geroliminis and Daganzo (2008) found that the speed v, or equivalently, the production ($P = nv$), followed a well defined function of the accumulation n (the number of vehicles currently driving in the network). This first result will be referred to thereafter as the "production-MFD" or as the "speed-MFD". Second, the ratio of the production divided by the outflow was found to remain approximately constant throughout the measurement period. Thus, there is also a well-defined relation between the outflow and the accumulation. This second result will be referred to as the "outflow-MFD" and it allows us to formulate the following differential equation:

$$\frac{dn}{dt}(t) = \frac{dI}{dt}(t) - O(n(t)) \tag{2}$$

where $O(n)$ is the function described by the outflow-MFD and $I(t)$ is the integral of the inflow (also called the cumulative inflow). These findings motivated additional investigations and although similar results were obtained (Buisson and Ladier, 2009), it appeared that the conditions that led to the observation of a so well-defined MFD in the city of Yokohama in Geroliminis and Daganzo (2008) were very specific and that similar results should not be expected everywhere. In particular, a well-defined outflow-MFD should only be observed with slowly-varying inflows.

2.3 Applicability of the MFD for rapidly-changing conditions

2.3.1 Limitations of the outflow-MFD

As mentioned in the previous paragraph, the outflow-MFD is intrinsically limited to scenarii with slowly-varying inflow. To illustrate this limitation, let us consider the following situation. Assume that first, the inflow has a low value, much lower that the outflow at capacity. After

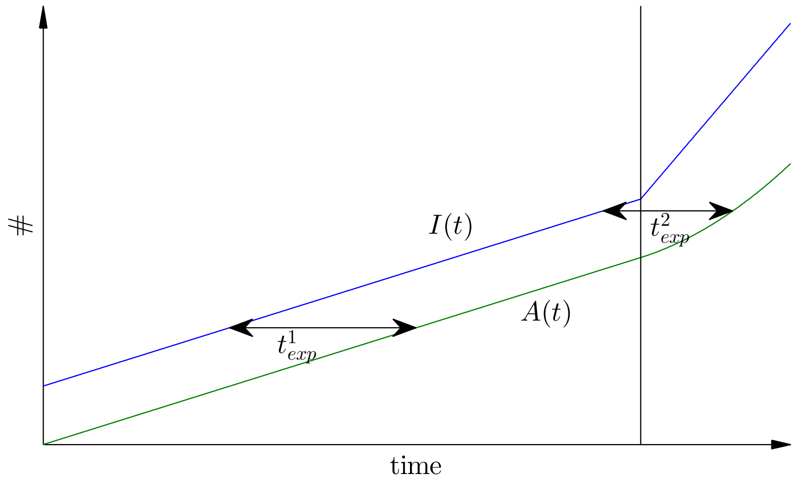

Figure 1: Illustration of the behavior of the outflow-MFD under rapidly-changing conditions

some time, the system will reach a steady-state, i.e. the outflow will be equal to the inflow, so that accumulation and travel time are constant. If then the inflow starts increasing, the accumulation will increase, and if the steady-state was not in the congested domain of the MFD, the outflow will increase as well, instantaneously. Thus, the experienced travel time will temporarily decrease, before it starts increasing to account for more congestion. This phenomenon is illustrated in Fig 1.

This un-natural behaviour results from the fact that the outflow-MFD is memory-less and does not model the time spent in the network, or the distance traveled. Intuitively, a peak in the inflow should increase the accumulation of vehicles instantaneously but the ouflow should start increasing only later, when the new users will start finishing their trips.

2.3.2 Exact formulation

Based on the observation above, it is possible to derive an exact formulation of the outflow for the dynamic case, assuming that the production-MFD (or, equivalently, the speed-MFD) remains valid.

$$\frac{dn}{dt}(t) = \frac{dI}{dt}(t) - \int_0^t \frac{dI}{dt}(\tau) f_l\left(\int_\tau^t v(n(u))du\right) d\tau \tag{3}$$

where n is the accumulation, $I(t)$ is the cumulative inflow at time t, f_l is the probability density function (pdf) of the trip length, $v(n)$ is the function defined by the speed-MFD and there are no users in the network for $t < 0$. Note that a similar formulation was already mentioned as an ideal model in a footnote by Arnott (2013) and is used in a working paper by Daganzo and Lehe (2014).

2.3.3 Impact on the dynamics

The exact formulation defined above and the outflow-MFD define two different dynamical systems. The response of these two systems to a peak in the demand was evaluated and compared by discretizing the differential equation 2 and 3 and the results are presented in Fig. 2. As shown in Fig. 2(a), the inflow was defined such that it first increased steadily to reach a steady state, and then a sinusoidal peak in the demand was applied, before returning to the initial low inflow. The 2^{nd} order polynomial model that was used for the speed-MFD is displayed in Fig. 2(b), while the outflow-MFD was simply obtained by multiplying the speed by the accumulation (we then have a 3^{rd} order model) and by dividing by the trip length. The analysis of Fig. 2(c) highlights how, for the same demand, the two dynamical systems described above can react differently. Note that the exact formulation leads to a much higher maximum accumulation and that therefore, the dynamics imposed by the outflow-MFD underestimate the risk of heavy congestion. In terms of outflow, the exact formulation leads to the anti-clockwise hysteresis observed in Fig. 2(d). Note that a clock-wise hysteresis has been reported by multiple authors on real measurements of the production-MFD so the hysteresis reported here would most likely be at least partly canceled out by other phenomena, such as a spatial heterogeneity in the density in the offset of congestion.

2.3.4 Combining Vickrey's demand model with congestion modeling

The most well-known work combining Vickrey's demand model with a more realistic demand model is most likely METROPOLIS, the software developed by De Palma and his team (de Palma and Marchal, 2002). METROPOLIS is an event-based simulator that was designed for large networks and has already been applied to different cities over the world. The simulator models both the choice of departure time and the route choice. Before each iteration (each day), agents choose a departure time based on the travel time experienced on the previous days by all users (perfect information is assumed). During their trip, agents choose the next link at each intersection based on the current conditions. Then, the travel time on each link is a function of the current conditions on this link only. Note that since the flow is directed on each link, it makes perfect sense that later arrivals should not impact the travel time of earlier agents (unlike in the MFD, in which users travel in all directions). This numerical approach is quite general since it allows considering non-linear cost functions and different types of users and still remains quite efficient, thanks to its event-base simulator. However, it inherits from the networks their

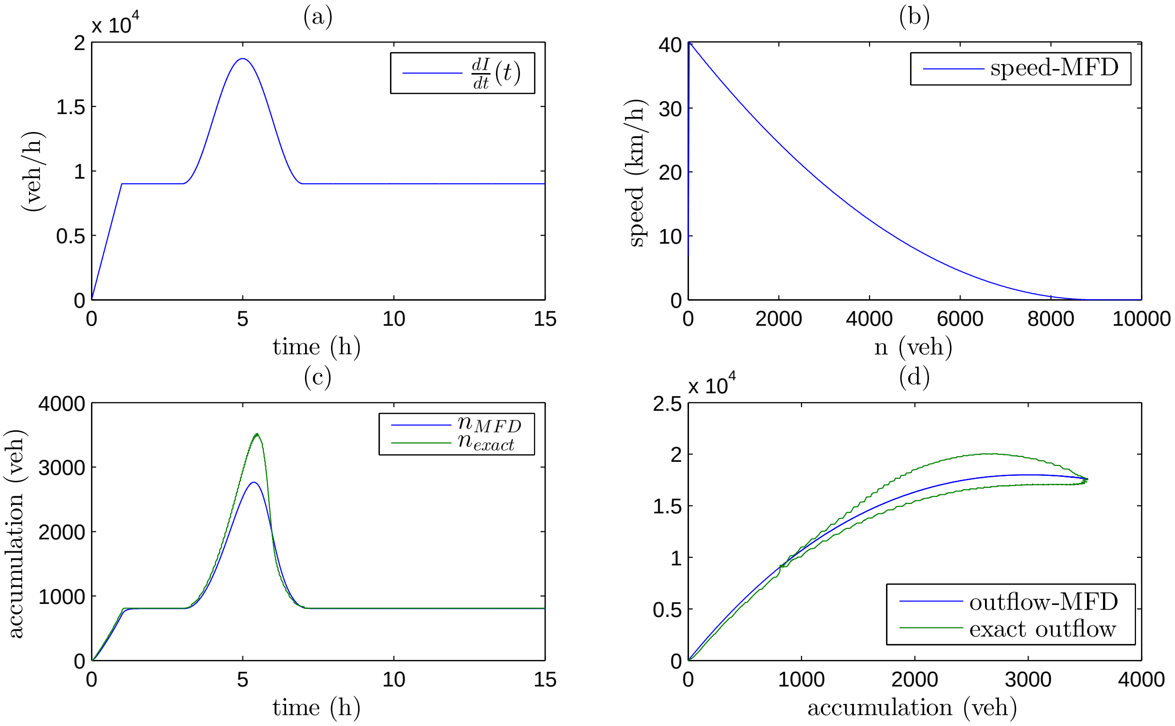

Figure 2: Comparison of the dynamics obtained with the outflow-MFD and with the exact formulation given by Eq. 3. (a): inflow over time ; (b): speed-MFD used ; (c): comparison of the resulting accumulations over time ; (d): comparison of the outflow obtained with the exact formulation and of the outflow predicted by the outflow-MFD, for the accumulation obtained with the exact dynamics.

complexity. Consequently, finding an equilibrium is time consuming, which may be a barrier for an in-depth analysis or for the design of congestion-reducing measures. Finally, such an approach requires modeling the entire network and having a precise picture of the origins and destinations, which makes its implementation heavy.

To better understand the impact of congestion on the choice of departure times, it was chosen in this work to use the MFD to model congestion, thus greatly reducing the complexity of the task. Geroliminis and Levinson (2009) and Arnott (2013) already suggested such an approach and even proposed a constructive and an analytical solution respectively. Nevertheless, this required again some strong hypotheses. Indeed, in order to derive an explicit expression of the accumulation from the travel time function obtained by Vickrey (1969), Geroliminis and Levinson (2009) used the instantaneous travel time at the arrival at work (when the user exits the network), defined by $tt(t) = \frac{n(t)}{O(n(t))}$. Arnott (2013) took a different approach, modeling the arrival of a user during an interval dt as a Poisson process, with probability $\frac{v(n)dt}{L}$. While both these expression are valid in steady state, they still have no "memory" and cannot model rapidly-evolving situations.

3 Approximating the equilibrium numerically

3.1 Problem description

Finding a distribution of departure times that leads to the equilibrium is a difficult problem. Before designing actual algorithms, one might first study analytically the existence and/or uniqueness of such an equilibrium. To give this problem a form that is more tractable, it is common to assume that the number of users is so big that the inflow, outflow and accumulation can be considered as continuous functions on the space of real numbers (this is known as the fluid approximation). With such an assumption and with a bottleneck of constant capacity, Smith (1984) and Daganzo (1985) showed respectively the existence and the uniqueness of an equilibrium for strictly convex and continuously differentiable schedule delay penalty functions. Thus, under reasonable conditions, there is a unique equilibrium in the simple case of a bottleneck with constant capacity but this result has not been extended yet to more general congestion models and intuition tells us that uniqueness would be difficult to obtain with an MFD.

In order to address this problem numerically, one should define an objective function that characterizes the equilibrium. One could for instance consider as an objective function the total number of users that are not happy with their departure time (i.e. that could reduce their cost by unilaterally changing their departure time). Alternatively, one could calculate how different the preferred and current departure times are for each individual and use as an objective function the sum of these differences over all individuals. This second approach was chosen in this work because it was found to be much more stable numerically (i.e. small changes in the departures have small impacts on the objective function).

Finally, the optimization method should somehow mimic the day-to-day adaptation of real drivers to changing conditions in order to identify a feasible equilibrium. Since there is no closed-form expression of the objective function, the derivatives of the objective function are not accessible. However, one can evaluate the travel times with the current departure rate function and identify how all users are likely to modify their departure time. In order to ensure some stability, most authors have adopted the method of successive averages (MSA), or some variation of this method - see e.g. Peeta and Mahmassani (1995) or de Palma and Marchal (2002). Applied on the departure time decisions, the MSA consists in updating the decisions of only a fraction of the population (for instance 5% or $\frac{100}{n}$%, where n is the iteration number). When the MSA is applied on the travel times, the decisions of all the agents are updated but the travel times that are used for their decision are obtained by averaging the last observation (with a weight equal for instance to 5% or $\frac{100}{n}$%) and the previous average. Note that it is also possible to apply MSA both on the travel times and on the departure times.

3.2 Evaluation of travel times through discretization

Before introducing the First-Wished-First-Passed, we rapidly explain how the travel times can be evaluated for a congestible facility that satisfies the First-In-First-Out (FIFO) property by discretizing Eq. 3.

Let us consider some discretized time $t = 0, 1, 2, ..., t_{ld}, ...t_{end}$, where t_{ld} is the last time that can be chosen for a departure and t_{end} is such that all users have finished their trips. Since users differ by their desired arrival time, they can be denoted by $i = 1, 2, ...n_i$ such that if $i < j$ then user i wishes to arrive earlier than user j. In the following, we will often use the same index i to refer to the individual and to her departure time. We will denote a choice of departure times for all individuals by the vector $\overrightarrow{d}$ of size $1 \times n_i$ such that d_i is the departure time of individual i.

Given the departure times at iteration n denoted by $\overrightarrow{d}(n)$, Eq. 3 can be discretized and the accumulation and outflow can be computed for each discretized time of a single day. Then, using the FIFO assumption, if we denote by $I(t)$ the cumulative inflow and by $A(t)$ the number of users that have been served at time t (i.e. the cumulative outflow), the travel time of the user entering the network at time t is simply given by $A^{-1}(I(t)) - t$ (cf. Fig. 2). Note that A is assumed to be invertible here, which is always true as long as the network is not empty or in a gridlock situation (these two extreme situations are not considered here). With this method, one can only measure the travel time for departure times that were actually used. In order to ensure that all travel times are known to update the decisions from one day to the next, a constant but negligible inflow was added to each departure time that can be chosen (from $t = 0$ to $t = t_{ld}$). These fictional users are then simply ignored when choosing departure times. Non-fictional users will be referred to as "active" users. Thus, the vector $\overrightarrow{tt}(n)$ having for components the travel times at iteration n and for the times $t = 0, 1, 2, ..., t_{ld}$ can be calculated. With these notations, updates methods such as MSA or the heuristic introduced in the next section aim at properly defining $\overrightarrow{d}(n + 1)$ based on $\overrightarrow{d}(n)$ and $\overrightarrow{tt}(n)$ and potentially also $\overrightarrow{tt}(m)$, $m < n$.

3.3 First-Wished-First-Passed property

This subsection builds on the work of Daganzo (1985), who showed that for a bottleneck of constant capacity, there cannot be more than one equilibrium and that in this equilibrium, drivers leave in the wished departure order. We will denote this property of the equilibrium by FWFP

(First-Wished, First Passed), after Daganzo (2013).

First, although Daganzo (1985) was specifically concerned with the bottleneck model, it should be noted that in the original paper, the proof of the following Departure Time Lemma only requires the congestion model to be FIFO and the schedule delay penalty functions to be identical for all users and strictly convex.

(Daganzo, 1985) Departure Time Lemma. *If there is equilibrium:*

$w' > w'' \Rightarrow l' \geq l''$,

where (l',w') are the actual and desired departure times for an individual, and (l'',w'') are the same variables for another one.

Second, in the very common case in which the schedule delay penalty functions are still identical for all users but only convex (piece-wise linear for instance), it can be shown that if there exists at least one equilibrium, then there exists also at least one equilibrium that verifies the FWFP condition:

Proposition. *Assume users differ only by their desired arrival time and that the schedule delay penalty function common to all users is convex. Assume $\overrightarrow{d}$ is a choice of departure times for all users such that there is an equilibrium.*

If $d_i > d_j$ and $i < j$, the departure times of users i and j can be exchanged and it is still an equilibrium.

The proof of this result is proposed in Appendix. It is directly inspired from a proof of Daganzo (2013) in a work on the System Optimum, that shows that swapping such users does not increase the social cost.

As a consequence of this property, one can restrict the search for an equilibrium to the sets of departure times for which the FWFP property is verified. The following iterative heuristic was developed to start from a solution that verifies the FWFP property (but is far from the equilibrium) and iteratively modify it while keeping conformity to the FWFP property.

In simple terms, this heuristic seeks to improve the global satisfaction by changing the departure times of those that block the most unsatisfied users since these unsatisfied users cannot be moved directly to their desired departure time without breaking the FWFP property. Note that other choices could have been made concerning the users for which the decisions should be updated and concerning the step size. Other similar strategies were tested and this one turned out to be both robust and quite efficient.

Algorithm 1 FWFP heuristic

initialize $\overrightarrow{d}(0)$ s.t. it satisfies the FWFP property
iteration = 0
while *iteration* < *MaxIteration* **do**
  iteration = *iteration* + 1
  Evaluate $\overrightarrow{tt_n}$ with the method described in Section 3.2
  Find for each active user the new preferred departure time and the associated gain
  Find the indexes $k_1, k_2, ...k_K$ of the K active users with the biggest potential gain and their preferred arrival times $w_1, w_2, ...w_K$
  $\overrightarrow{d}(n+1) = \overrightarrow{d}(n)$
  for i=1...K **do**
    if $w_i < d_{k_i}(n)$ **then**
      Find the smallest index j such that $w_i < d_j(n)$
      $d_j(n+1) = d_j(n) - 1$
    else
      Find the biggest index j such that $w_i > d_j(n)$
      $d_j(n+1) = d_j(n) + 1$
    end if
  end for
end while

Note also that if users differ not only by their desired arrival time but also by their values of α, β, γ, then the property does not stand. Thus, it is assumed in this work that users differ only by their desired arrival time. In practice, the previous heuristic could be applied on different subgroups of users such that the values of α, β, γ are the same for all users of a group.

In addition to the general heuristic described above, some simple techniques were adopted to improve the convergence. First, to avoid cycling phenomena, randomness was included in the choice of the K active users with the biggest potential gain for both MSA and the FWFP heuristic. In practice, the probability of being chosen was defined to be proportional to the potential gain raised to some high power (8 for instance). Second, when two departure times have equal or very similar costs, there is no reason to choose one rather than the other. Therefore the choice was made with a logit model in the MSA approach, as suggested in de Palma *et al.* (1983), except that in our application the scale parameter was chosen very small in order to obtain results that are consistent with Vickrey's theory (Vickrey, 1969) and to avoid an excessively flat peak hour. Such a technique was not compatible with the FWFP heuristic but instead, we allowed moving only a fraction of the users with the same desired arrival time. Finally, the maximum number of changes K can be adapted over iterations. To obtain the following results, K was progressively decreased as the number of iterations increased and as the objective function decreased.

3.4 Convergence results

In this subsection, MSA and the FWFP heuristic are first tested on a simple bottleneck model with constant capacity for which the analytical solution is known. Then, the two heuristics are tested on a MFD with light and heavy congestion. In all cases, the iterative search was started with free-flow travel times (0 for the bottleneck model) for all departure times.

3.4.1 On a bottleneck with constant capacity

The two heuristics were run 10 times for 600 iterations each on a bottleneck model with a symmetric trapezoidal distribution of the desired arrival time. The evolution of the objective function over the iterations is represented in Fig. 3 (a). Overall, the FWFP heuristic greatly outperformed MSA even though it is somehow slower at the beginning. MSA was found to be extremely unstable, regardless of the different parameters involved (the degree of randomness in the choice of the changes and the number of changes per iteration).

3.4.2 On a MFD with light congestion

The two heuristics were then run 5 times for 1000 iterations each on a MFD with light congestion (the maximum inflow was 5% bigger than the maximum possible outflow during a short time), with the same distribution of desired arrival times as for the bottleneck. The evolution of the objective is represented in Fig. 3 (b). Surprisingly, MSA was found to be much more stable with an MFD and it converged in average faster than the FWFP heuristic. One can explain this improvement by a “stabilizing” effect on the MFD. Indeed, while with the bottleneck model users are sensitive only to the number of users that arrived before them, in the MFD users are sensitive to the number of users throughout all the time they travel. If, with a bottleneck, user i “jumps” over user j (i.e. i switches his departure time from before j to after j , or vice versa), the cost of j is greatly impacted, which might trigger a chain reaction. With the MFD however, the cost of j would be just slightly impacted since both users will still be traveling simultaneously in the network most of the time.

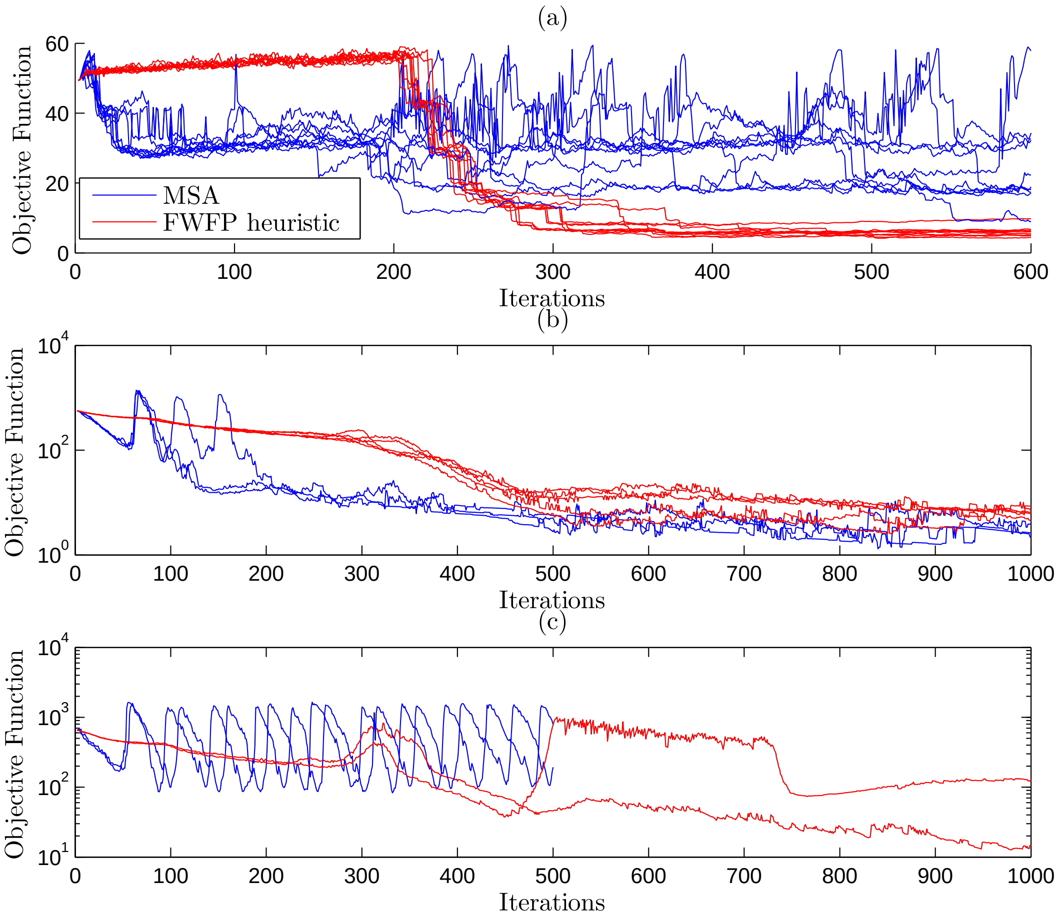

Figure 3: Convergence results of MSA and of the FWFP heuristic on: (a) a bottleneck for 10 approximation processes, (b) an MFD submitted to light congestion for 5 approximation processes, (c) an MFD submitted to heavy congestion for 2 approximation processes.

3.4.3 On a MFD with heavy congestion

Finally the two heuristics were run only twice with heavier congestion level (the maximum inflow was temporarily 10% bigger than the maximum possible outflow). The MSA was only run for 500 iterations since it entered a cycle and could not exit it. As the FWFP heuristic was still progressing after 500 iterations, it was run for 1000 iterations. The results are displayed in Fig. 3 (c). Note that in these conditions, MSA is unable to find the equilibrium while the FWFP is slightly unstable but can still approximate the equilibrium much more accurately.

4 Conclusion and future work

The article investigated the possibility to approximate the dynamic equilibrium during peak periods in urban areas with an MFD and a numerical approach. It was shown that with traditional assumptions, the decision space can be greatly reduced and a heuristic taking advantage of this property was proposed. This heuristic obtained better results on unstable networks such as a single bottleneck model or an heavily congested MFD, where the traditional MSA did not find any equilibrium. With an MFD submitted to lighter congestion, both algorithms converged to the equilibrium and MSA was generally faster. Thus, MSA remains useful for stable conditions but the heuristic proposed seems more promising for unstable networks. This heuristic could potentially be used to design congestion reducing measures or analyze the impact of different factors.

Future research should be carried out on the topic to test the heuristic with several classes of users that differ also by their values of α, β, γ and to compare the results with those that can be obtained analytically with less realistic hypotheses (Arnott, 2013, Geroliminis and Levinson, 2009).

5 References

Arnott, R. (2013) A bathtub model of downtown traffic congestion, *Journal of Urban Economics*, **76** (C) 110–121.

Buisson, C. and C. Ladier (2009) Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams, *Transportation Research Record: Journal of the Transportation Research Board*, **2124** (1) 127 – 136.

Daganzo, C. F. (1985) The uniqueness of a time-dependent equilibrium distribution of arrivals at a single bottleneck, *Transportation science : the publication of the Transportation Science Section, Operation Research Society of America*, **19** (1) 29–37, ISSN 0041-1655.

Daganzo, C. F. (2013) System optimum and pricing for the day-long commute with distributed demand, autos and transit, *Transportation Research Part B: Methodological*, **55** (0) 98 – 117, ISSN 0191-2615.

Daganzo, C. F. and N. Geroliminis (2008) An analytical approximation for the macroscopic fundamental diagram of urban traffic, *Transportation Research Part B: Methodological*, **42** (9) 771 – 781, ISSN 0191-2615.

Daganzo, C. F. and L. J. Lehe (2014) Distance-dependent congestion pricing for downtown zones, *Institute of Transportation Studies, Research Reports, Working Papers, Proceedings*, Institute of Transportation Studies, UC Berkeley.

de Palma, A., M. Ben-Akiva, C. Lefevre and N. Litinas (1983) Stochastic equilibrium model of peak period traffic congestion, *Transportation Science*, **17** (4) 430–453.

de Palma, A. and F. Marchal (2002) Real cases applications of the fully dynamic metropolis tool-box: An advocacy for large-scale mesoscopic transportation systems, *Networks and Spatial Economics*, **2** (4) 347–369, ISSN 1566-113X.

Geroliminis, N. and C. F. Daganzo (2008) Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings, *Transportation Research Part B: Methodological*, **42** (9) 759 – 770, ISSN 0191-2615.

Geroliminis, N. and D. Levinson (2009) Cordon pricing consistent with the physics of overcrowding, in W. H. K. Lam, S. C. Wong and H. K. Lo (eds.) *Transportation and Traffic Theory 2009: Golden Jubilee*, 219–240, Springer US, ISBN 978-1-4419-0819-3.

Godfrey, J. W. (1969) The mechanism of a road network, *Traffic Engineering and Control*, **11** (7) 323–327.

Newell, G. F. (1987) The morning commute for nonidentical travelers, *Transportation Science*, **21** (2) 74–88.

Peeta, S. and H. Mahmassani (1995) System optimal and user equilibrium time-dependent traffic assignment in congested networks, *Annals of Operations Research*, **60** (1) 81–113, ISSN 0254-5330.

Smith, M. J. (1984) The existence of a time-dependent equilibrium distribution of arrivals at a single bottleneck, *Transportation Science*, **18** (4) 385–394.

Vickrey, W. S. (1969) Congestion Theory and Transport Investment, *American Economic Review*, **59** (2) 251–60.

A Proof of the proposition

Proof. Let us denote by $C(d,k)$ the cost for user k when her departure time is d. $C(d,k) = \alpha tt(d) + p(d + tt(d) - k)$ where $tt(d)$ is the travel time for a departure at time da and $p(x)$ is the schedule delay penalty when arriving x time units after the preferred arrival time (if $x < 0$, $|x|$ is the advance). The function p is convex and should reach its minimum for $x = 0$. Let $\overrightarrow{d}$ be a choice of departure times for all users such that there is an equilibrium, and assume that there are users i and j such as $d_i > d_j$ and $i < j$ (i.e. j has a preferred arrival time that is later that i's preferred arrival time).

The equilibrium assumption imposes that

$$\begin{aligned}&\alpha tt(d_i) + p(d_i + tt(d_i) - i) \leq \alpha tt(d_j) + p(d_j + tt(d_j) - i)\\ &\quad\Leftrightarrow \alpha(tt(d_i) - tt(d_j))) \leq p(d_j + tt(d_j) - i) - p(d_i + tt(d_i) - i)\end{aligned} \tag{4}$$

$$\begin{aligned}&\alpha tt(d_j) + p(d_j + tt(d_j) - j) \leq \alpha tt(d_i) + p(d_i + tt(d_i) - j)\\ &\quad\Leftrightarrow \alpha(tt(d_j) - tt(d_i))) \leq p(d_i + tt(d_i) - j)) - p(d_j + tt(d_j) - j)\end{aligned} \tag{5}$$

By combining Eqs. 4 and 5

$$p(d_j + tt(d_j) - j) - p(d_i + tt(d_i) - j)) \leq p(d_i + tt(i) - j) - p(d_j + tt(j) - j) \tag{6}$$

Let $h(x) = p(x-i) - p(x-j)$. Since p is convex, p'' is positive and since $x-i > x-j$, $p'(x-i) > p'(x-j)$ so h is increasing. Besides, the queuing system is FIFO so $d_i + tt(d_i) > d_j + tt(d_j)$. Hence:

$$\begin{aligned}&h(d_i + tt(d_i)) > h(d_j + tt(d_j))\\ &\quad\Leftrightarrow p(d_j + tt(d_j) - j) - p(d_i + tt(d_i) - j)) \geq p(d_i + tt(i) - j) - p(d_j + tt(j) - j)\end{aligned} \tag{7}$$

Eqs. 6 and 7 impose that

$$p(d_j + tt(d_j) - j) - p(d_i + tt(d_i) - j)) = p(d_i + tt(i) - j) - p(d_j + tt(j) - j)$$

which imposes that $C(i,d_i) = C(i,d_j)$ and $C(j,d_i) = C(j,d_j)$ so users i and j can be exchanged and they are still at equilibrium. □