TRAFFIC OPERATION ON A ROAD NETWORK WITH RECURRENT CONGESTION

ANDRÁS SZELE1 & LAJOS KISGYÖRGY2
1Research Centre for Transport Development, Institute for Transport Sciences Non-profit, Budapest, Hungary
2Department of Highway and Railway Engineering, Budapest University of Technology and Economics, Hungary

ABSTRACT

Congestion is the natural consequence of intensive use of a road network in developed urban areas and is among the most critical negative factors affecting urban quality of life. There are several types of congestion, but the real damage is done by recurrent congestion during rush hours. Road networks with recurrent congestion are found in urban-suburban areas where the traffic volume is high enough and the road network is sufficiently dense to speak of network congestion. This congestion is caused by urban growth and long-term handling of it is more a matter of traffic management, rather than road design. According to our research, traffic operations of road networks with recurrent congestion remarkably differ from the traffic of both the non-congested road networks and the supposed mode of operation of the network, during the planning process. Our results are based on wide-range traffic data collection (traffic counts and trip time measurement) and their analysis. The formation of congestion on a road network with recurrent congestion plays out a typical scenario. In this scenario, some special traffic phenomena play important roles, for example, the wide-range use of alternative routes. The importance of the problem can be seen, as in some cases, within the most important cross-section, as the flow from local roads can reach 70–80% and drivers using alternative routes can reach 30–40% of the total flow. This research outlines a concept which integrates the interesting and important traffic phenomena of recurrent congestion into one theoretical system. The concept provides practical applicability in traffic management and offers the possibility of involving other sciences, such as psychology, sociology and game theory into congestion management. This new concept could fundamentally change the way we think about traffic congestion.

Keywords: congested roads, rush hour, traffic congestion, traffic management, traffic volume.

1 INTRODUCTION

One of the greatest challenges of transportation planning and traffic management in urban areas is the handling of recurrent congestion. The typical rush hour traffic landscape in urban areas is defined by the congestion that is simply generated by too much traffic. A congestion charge or fee is the only method that has proven successful against congestion, but it has also been proven that a congestion charge sometimes does not reduce the traffic, just that drivers avoid the protected area. For example: in Milan, due to the congestion fee, traffic in the protected area decreased by about 30%. Detailed measurements have shown that behind a decline is primarily the driver’s intention to avoid the city centre and a payment [1]. The congestion charge has a relatively small impact on traffic outside the protected area, and its introduction in a city is a very complex and politically sensitive process. The most important motivation behind its introduction is a rapidly deteriorating traffic situation experienced by urban dwellers [2].

In alleviating congestion, there is a lot expected from the spread of vehicle-to-vehicle (V2V) communication and autonomous vehicles, but this probably will not solve the problem of recurrent congestion. The spread of autonomous vehicles, because of the change in the car-use behaviour, will be likely to increase the number of vehicles on the roads, which will surely increase the presence and extent of congestion. Regarding V2V communication, it is regularly raised that more widely available information will help a driver find the fastest route in the congestion; however, this information should now be available through navigation...
applications. Moreover, **use of the alternative routes by the drivers is not a solution, but a part of the problem.** Based on our research, the use of alternative routes and underlying human behaviour play an important role in evolving and operation of recurrent congestion, which creates traffic dynamics not yet treated in traffic planning and management.

Almost every road network with recurrent congestion has reserve capacity, traffic management is improvable and could achieve a significant modal shift to public transport. But the overall growth in passenger car traffic all around the world is simply much faster than the development of possibilities on the road networks with recurrent congestion. In addition, the congestion causing cross-sections in urban-suburban areas can be found often in densely populated and built-up areas where the capacity for expansion is not just very expensive, but can hardly be justified. In the long term, perhaps with the spread of autonomous vehicles and V2V communication, the emergence of strict corridor and network management solutions is expected [3], we already have the precursors [4]. These solutions can actually alleviate congestion, but at the same time the drivers will lose their ordinary freedom of driving. The long-term solution will require the learning of traffic operation to manage recurrent congestion.

2 LITERATURE

The most important information on traffic congestion is summarized in a 2007 OECD publication [3]. In addition to the small number of comprehensive publications, congestion can be encountered in two very different ways in transport research literature. One of the ways is the theoretically, especially through traffic modelling work. The other field of research is traffic management, which typically involves examining a particular road segment. A significant problem is that the results of theoretical work are rarely usable in practice (one good example is the congestion charge/fee) and that the traffic management tools used in practice vary a small number of long-known solutions, with no great success.

In theoretical and modelling research, congested traffic flow has long been at the forefront of interest [5], [6]. The most important theories describing the flow of traffic include hydrodynamic and kinematic models, car-following models and queueing models [7], [8]. The kinematic model is easy to use in traffic planning [9]. One of the relatively new research domains (related to Daganzo and Geroliminis [10]), models the uniformly congesting traffic in urban areas at the aggregate level, using the Macroscopic Fundamental Diagram, where the accumulation and speed of vehicles play a leading role [10]. The theory was widely discussed and further developed for special circumstances [11]. Later on, the urban network gridlock issue was also studied and modelled on this base [12].

There is a long tradition of describing congestion as a 2-phase system [13]–[15], where the traffic on the main network elements is a mixture of standing and moving vehicles [16]. Critics of the models used and the benefits of the 3-phase system are emphasized by Kerner [17]. Other authors also criticize the theoretical models [7], [17]; for example, because of the weak correlation between the models and observed traffic, orienting designers more towards empirical models.

Traffic management tools are the other great part of congestion research. In addition to the congestion fee, these traffic management measures are those actually implemented in some road section. These traffic management programs aim to remove bottlenecks, maintain the operation of the network, make better use of existing capacities and manage demand.

The measures used are the following [18]–[23]:

- Maintain Network Operation: Dynamic speed control, traffic control, optimize signal timing plans, enforce road traffic laws, speed regulation.
Better use of existing capacities: lane management (HOV/HOT/Tidal lanes), parking management, providing traveller information, ramp metering, variable speed limit, rerouting of traffic, arterial/corridor management.

Demand management: improving the pedestrian, cycling, and public transport infrastructure and services; refining park-and-ride; road pricing and variable tolling; rationalise distribution; manage driver behaviour (making a less congested route or travel in a less congested time); telecommuting, home office; and inbound flow control [24].

The successful operation of traffic management solutions has been trying to prove many times. According to the coordinated adaptive traffic system modelled in Sydney, the system reduces travel time by some 28% and traffic stops by 25% [25]. A similar system on the M25 expressway around London delayed the formation of congestion by 10-20 minutes and improved the distribution of traffic between lanes [19]. In evaluating the results, it is worth taking into account that behind the results in congestion alleviation almost always the better utilization of the existing capacities, which is ultimately capacity increasing [26]. So far, long time series about the operation of traffic management systems have not been investigated. It is believed that growing/rerouting of traffic eventually found and used these new capacities.

Our research can be found between theoretical and management research. We introduce unfamiliar traffic phenomena that offer new opportunities and tools for both researchers and designers, in both theory and practice.

3 METHODOLOGY

Our research aimed to familiarize the traffic operation strategies of road networks with the characteristics of recurrent congestion. Deeper knowledge was provided by wide-range traffic measurements on several sections of the southwestern road network of Budapest, mainly on roads that are leading into the city. We conducted traffic counts and travel time measurements. All traffic in all intersections on the selected network was counted from the beginning to the end of the congestion period. During the same period, the travel times of recurrently congested arterials were also measured. To examine some traffic phenomena, we used the results of traffic data collection for other purposes, and the “typical traffic” data of Google Maps, which are available to everyone.

According to our original intention based on these data, a traffic simulation would have been built to explore and demonstrate the effects of various congestion management solutions; however, during data analysis we found interesting results, which, among other things, suggest surprising alternative route usage, and the complete rearrangement of the traffic hierarchy of road networks with recurrent congestion. After reviewing some parts of the measurements for audit purposes, we re- evaluated old traffic data on other road networks. These tests confirmed that we had found new, unrecorded traffic phenomena. Examining these phenomena requires the definition of new concepts, such as the vitality of traffic directions [27], and an estimate of the proportion and the number of users of alternative routes.

Though the extensive use of alternative routes has not been previously unknown (Fig. 1), the magnitude of the phenomenon or the rearrangement of the network hierarchy have not been investigated so far.

To work out a new concept about the traffic phenomena, we first introduce our terminology (Table 1), which is illustrated in Fig. 2.

The arterial and local road concepts provide information about the location of the traffic, the main flow and local flow concepts clarify the origin of the traffic.
Behind the traffic phenomena experienced on the road network with recurrent congestion, the next traffic operation was assumed as a framework: The congested traffic flow on the arterial, towards the key cross-section, passes through intersections. In these intersections, significant flow joins the congested main flow from local roads.

In the road network with recurrent congestion, the flow from the local roads increases significantly during congestion. A significant part of this flow at this time is the traffic from the alternative routes. Because the congested arterial does not have a reserve capacity, each joining flow from local roads causes stops, delays and traffic disturbances to the flow on the arterial. These disturbances form shock waves that go upstream in the queue [9]. The travel speed of the arterial further from the key cross-section is getting lower. As the drivers on the congested arterial are getting closer to the key cross-section and passing through the traffic disturbances due to joining flows, and the queue formation on the arterial is accelerating.

Seen from another point of view, the joining flows from the local roads take capacity away from the congested arterial. Some drivers regularly use the road network with recurrent congestion, knowing the network and its operation, and plan their route to avoid the primary congestion; and using the alternative routes, join into the queue near the key cross-section, thus contributing to the formation of congestion. The lack of capacity of the key cross-section causes significant damming on the arterial. This damming effect done by the traffic of the arterial is much smaller towards the local roads.

4 TRAFFIC OPERATION ON THE ROAD NETWORK WITH RECURRENT CONGESTION: THE SCENARIO

According to the results of our research, the formation of congestion follows a typical scenario on a road network with recurrent congestion. First, this scenario will be introduced; then, the most important steps will be supported in detail. The scenario consists of six consecutive steps:

Figure 1: Spatial spread of rush-hour traffic. (Source: ECMT, 2007.)

Figure 2: Elements of the concept.
Table 1: Terminology of the traffic research.

<table>
<thead>
<tr>
<th><strong>Recurrent congestion</strong></th>
<th>Congestion could be categorized by the root cause. Recurrent congestion is generated by factors that affect the road network on a regular basis, such as commuting [3].</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Road network with recurrent congestion</strong></td>
<td>An urban or suburban road network with re-occurring vehicle congestion, typically where the dominant part of the analysed traffic has a particular destination, such as the city centre.</td>
</tr>
<tr>
<td><strong>Arterial network</strong></td>
<td>Capacitive and interconnected part of the urban or suburban road network dedicated to carrying the highest volumes of traffic. Its elements are called the arterials.</td>
</tr>
<tr>
<td><strong>Local road</strong></td>
<td>Road joining the arterials.</td>
</tr>
<tr>
<td><strong>Main flow</strong></td>
<td>The massive traffic flow during the peak period, moving towards the most desired destinations (such as the city centre). This flow creates the recurrent congestion. The dominant part of the main flow comes from outside of the road network with recurrent congestion.</td>
</tr>
<tr>
<td><strong>Local flow</strong></td>
<td>Traffic flow generated by local demands.</td>
</tr>
<tr>
<td><strong>Key cross-section</strong></td>
<td>Recurrent congestion is typically caused by a lack of capacity in a cross-section of the network. This capacity-causing cross-section is referred to as the key cross-section. The key cross-section cannot be avoided on the road network with recurrent congestion. In the key cross-section, the available capacity is smaller than the flow that wants to pass, so queues are created. After the key cross-section, there is no more congestion until the impact area of the next key cross-section. The volume of flow passing through the key cross-section is also the capacity of the key cross-section, which changes within tight boundaries during congestion. The key cross-section of a road network with recurrent congestion is typically an intersection with a traffic signal.</td>
</tr>
<tr>
<td><strong>Primary congestion</strong></td>
<td>Congestion resulting from the lack of capacity of the key cross-section of the arterials. The generated queues can be found on the arterials.</td>
</tr>
<tr>
<td><strong>Alternative route</strong></td>
<td>A route to avoiding primary congestion on road networks with that have recurrent congestion. Two types of alternative routes are distinguished: the first one comes out of the primary congestion and returns before the key cross-section, thus skipping some part of the queue (alternative route type A in Fig. 2). This operation is clearly visible in the congestions. The other alternative route type is less spectacular, but more significant in its impact. In this case, the drivers from a busy local road decide to avoid the primary congestion (alternative route type “B” in Fig. 2). Traffic that uses alternative routes wants to skip the queue for as long as possible queue, so drivers looks for the closest joining option to the key cross-section.</td>
</tr>
<tr>
<td><strong>Secondary congestion</strong></td>
<td>Congestion on an alternative route. This congestion is partially generated by the drivers whom are avoiding the primary congestion.</td>
</tr>
<tr>
<td><strong>Widely congested state</strong></td>
<td>State of the road network with recurrent congestion, where the secondary congestion occurs along with the primary congestion, and the traffic operation of the network is determined by the drivers who want to avoid the most the congestion areas.</td>
</tr>
</tbody>
</table>
1. In urban areas, a significant part of drivers is trying to get to their destination before the congestion occurs. This is a significant traffic volume on the road network, actually higher than during the congestion (Fig. 3); however, key cross-sections, which are the least capacitive elements of a road network, can still pass surprising traffic volume at that time. This period is the most effective time for a road network with recurrent congestion. One of the reasons for this is that these flows are on the capacitive arterials, and there are no signs of alternative route use. At this time the flow from local roads carries significantly less traffic than during the congestion; however, this state of grace lasts only for a short time. The huge traffic volumes quickly expire the capacity of the key cross-sections, creating the first queues.

2. In the key cross-section, the queue transforms into congestion in a very short time. This is the primary congestion. It is important that in this case, during the formation of the congestion, that road construction work or accidents play no role, even though there is a greater likelihood of accidents at these high traffic volumes. The key cross-sections are on the arterials, so the primary congestion is also located on the arterials. In essence, these primary congestions are the roots of the widely congested state that will be forming in the next steps.

3. A significant portion of drivers don’t want to drive in the primary congestion. Most of them, in knowing the network and its operation, plan their trip on alternative routes. Other drivers are looking for an alternative route when realising the congested situation, and many drivers are just following the instructions of a navigation application (Table 2 and Fig. 4).

4. A significant part of the main flow using alternative routes avoids the primary congestion and joins the queue near the key cross-section; thus, completely transforming the traffic operation (Fig. 4). The ratio of the flow from the arterial into the key cross-section decreases considerably, and its place is taken up by the flow from the alternative routes. It is typical that alternative routes affect residential streets and intersections that were designed for much lower traffic.

5. Secondary congestion on alternate routes is occurring, and the traffic on the whole network is transforming into a widely congested state. An important feature of the road network with recurrent congestion is that in the peak period every arterial is congested due to the lack of capacity of the key cross-sections. This network traffic operation is unfavourable; and based on the steps outlined so far in the scenario, the congested network extends, and significant part of the main flow from the capacitive arterials transfers to local roads. Controls in the intersections typically bring the arterials to an unfavourable situation. Compared to this situation, the traffic operation before the congestion is clearly visible, as is the powerlessness of the flow on the arterials; and consequently, the degradation of the traffic operation.

6. The incoming traffic flow slowly begins to decrease (Fig. 3), queues shorten, and the congestion ceases and starts its normal, non-congested traffic operation.

To support the first step of the scenario, the traffic volumes of three busy roads in the south-southwestern (S-SW) agglomeration at the border of Budapest are presented (Fig. 3). At this step, there was no congestion on the arterial sections presented, and there were no alternative routes as well, so traffic volumes show essentially the number of vehicles moving towards the city, in other words, the traffic demand. This traffic demand dominates the traffic operation of the road network with recurrent congestion, they are typically much greater than the local flows. Some drivers from the congested arterial area later used alternative routes, but in this step, these routes were not used. The chart clearly shows that the number of arrivals
in the city is significantly higher before the congestion, than during the congestion, which indicates the extreme volume of traffic occurring before congestion has an important role in forming that congestion.

In the third step of the scenario, we could not quantify the three types of drivers using alternative routes, but we could estimate the number of all drivers using the alternative routes in order to determine the magnitude of the problem (Table 2 and Fig. 4).

The increase of traffic flow on the alternative routes during congestion is significant: in the three cases, this is 21%, 50% and 73%, respectively.

5 THE EFFECTS OF TRAFFIC ON A ROAD NETWORK WITH RECURRENT CONGESTION, ON THE TRAFFIC PLANNING PROCESS

The research and analyses of congestion have so far been focused on describing the formation of queues; congestion has been treated mostly as a linear 1-dimensional phenomenon; however, based on the results of this research, it is more than queues becoming longer and occasionally interfering with traffic at other intersections. Within the network operation presented, the congestion analysis could change from the 1-dimensional approach of the queues to the examination of the entire road network covered by alternative routes.

Conducting systematic research on the presented base opens the way for exploring the relationship between network operation and psychological and sociological factors, using competition, decision-making and game theory. For example, driver’s psychological functioning could be found behind the decisions between taking the arterials and the alternative route; and in addition to the psychological factors, the sociological aspects were found to play a role in the act of joining the queue. Since operation of the road networks with recurrent congestion has not been sufficiently investigated, the elements of contemporary

Figure 3: Traffic volumes of three agglomeration arterials at the border of Budapest, on a working day morning in the fall of 2017. (Source: Institute of Transport Sciences.)
Table 2: The estimated traffic volume of the alternative routes.

<table>
<thead>
<tr>
<th></th>
<th>Northeastern part of the Grand Boulevard, Budapest</th>
<th>Egér Road, Budapest</th>
<th>Balatoni Road, Diósd</th>
</tr>
</thead>
<tbody>
<tr>
<td>The increase in the traffic volume of the flow from the local roads at the joining point to the arterial, including local traffic and the traffic on the alternative routes (vehicle/hour)</td>
<td>361</td>
<td>1320</td>
<td>870</td>
</tr>
<tr>
<td>The increase in traffic volume in the flow from local roads, relative to the traffic volume of the key cross-section</td>
<td>21%</td>
<td>50%</td>
<td>73%</td>
</tr>
<tr>
<td>The estimated traffic volume of the alternative routes (vehicle/hour)</td>
<td>180</td>
<td>790</td>
<td>520</td>
</tr>
<tr>
<td>The traffic volume of the key cross-section (vehicle/hour)</td>
<td>1720</td>
<td>2650</td>
<td>1190</td>
</tr>
</tbody>
</table>

Figure 4: Changes in traffic operation in three road networks with recurrent congestion, in and around Budapest, Hungary.
traffic planning methods from research to road design assume significantly different network operation than the actual one. As a result, the planned and created network operation is barely comparable to the actual operation. After all, it has been proposed to develop new traffic planning procedures, to forecast the traffic on the road network with recurrent congestion.

New design procedures should be developed along the following principles:

- The operation of a road network with recurrent congestion is radically different from the operation of a non-congested network. Not only are the speeds lower, but completely different the routes are used and there is a radical change in the hierarchy of the road network. These factors cause well-functioning traffic management tools from the non-congested period, including signalisation, road markings and signposts, to be inadequate during traffic congestion.

- Traffic on a road network with recurrent congestion uses masses of road sections and intersections that are not designed for such a traffic volume. Protecting residential streets or the handle of the intersections from becoming dangerous because of increasing traffic is a very important and unavoidable task. These questions involve some philosophical problems in network design, such as the right of the limitation of free routing on the roads. The basic principle of today's network planning is that (almost) all traffic flow is legitimate. Looking ahead a bit farther into the era of autonomous vehicles, the right of the limitation of free routing may become an essential problem. Until then, network control, which is not yet elaborate in its tools, can be the solution for controlling and managing congested network traffic. The first step is the inbound control [24]. It will also be important to handle the policy implications of limitation, as some drivers will consider it a significant disadvantage to limit alternative routes.

- The operation of the road network with recurrent congestion is not fair. Users of some alternative routes gain time, compared to the users of the arterials. In other words, travel times do not equalize on a network level. The question is whether this process has an available social optimum or is worth achieving.

- Alternative routes are powered by the easy joining into the queue. Limiting local roads with excessive vitality, and thus reducing the benefit of the use of alternative routes, would in itself be a major step toward improving network traffic quality, as it would provide a way to accelerate and reclaim the flow on arterials, thereby mitigating the impact of adverse network processes. There is a need to make the decision, in principle, of whether we want to strengthen the flow on arterials or not. The designation of the arterials has an important purpose in transportation planning. The arterials are capacitive elements: usually there is public transport on them and these are the main channels for freight transport. If a road is designated an arterial, then the consequences of this act must be considered and the arterial has to be able to fulfil this role, even in the times of congestion.

6 CONCLUSION

The most important result of the presented research is the six-steps-scenario for how traffic operates on road networks with recurrent congestion. In the first step, before the congestion, there are huge traffic volumes on the arterials, without any sign of congestion or the use of alternative routes. During the congested period, an intensive use of the alternative routes can be experienced. These traffic flows join the arterials near the congestion-causing cross-section, and these joining flows from local roads take significant capacity away from the congested arterial. From the point of view of road network planning, use of the alternative
routes by drivers is not a solution, but a part of the problem. The special traffic phenomena discovered by research are not side effects, but the essential properties of road networks with recurrent congestion.

The traffic flow in the non-congested and the congested networks is fundamentally different. The outlined traffic concept points to the unfavourable processes taking place. It is very important that human activity is the driving force behind the presented traffic activity. The unfairness of this traffic operation is based on a contradiction between individual and social interests. It is worthwhile to clarify the potential role of human sciences in network planning. In the future, with the emergence of autonomous cars, the operation of such networks could be expected to deteriorate even further, if the control algorithm of the system will seek the shortest way past all the vehicles. It is an important matter of principle that the navigation applications are now the inevitable actors in the use of alternative routes. Their role, their potential regulation needs to be clarified and their limitations need to be taken into account.

As long as our cities operate the way they do today, there will always be cross-sections without enough capacity and road networks with recurrent congestion that cannot be fixed, so further research is necessary to find better solutions. We hope our research will encourage further experiments.

REFERENCES


