The MULTITUDE Traffic Simulation Primer

How to use traffic simulation – the most urgent issues.

MULTITUDE: develop, implement and promote the use of methods and procedures for supporting the use of traffic simulation models, especially regarding model calibration and validation

Date 5. November 2013
Version 1.0

Authors Peter Vortisch,
Karlsruhe Institute of Technology, Germany
Peter Wagner,
German Aerospace Centre, Germany

The COST action MULTITUDE has dealt with the current state of the praxis when it comes to the application of traffic simulation tools, among other topics. As has been proved in the past, sometimes such studies have given false or misleading results which led to huge costs in finally correcting those short-comings. One of the aims of MULTITUDE is to raise awareness that there are well-known procedures and methods that help in avoiding such pitfalls.

This traffic simulation primer approaches a slightly different, but related issue. At the heart of any microscopic model and simulation tool are a couple of important concepts that are sometimes made not clear enough in the guidelines written so far. Therefore, this text does not try to reproduce all those work but acts as a guide through some of the most important and pressing issues when it comes to the best demonstrated practice of simulation studies, and especially of microscopic simulation studies.

When do we need simulation?

Simulation is used whenever there is something that is too complicated for humans to deal with analytically. Sometimes, the examples when this happens are surprisingly simple. Consider e.g. a T-intersection as shown in figure 1, located between two signalized intersections. How long does the red vehicle have to wait before it can turn left?

Of course this depends on the traffic volume on the main road. If there were no signals near the T-junction, the average waiting time could be calculated using analytical formulae typically given in road design manuals. But a close-by signal will change the arrival pattern of the vehicles strongly. During red-times, no vehicle arrives, while during green, they came “bunched”, thereby forming platoons of vehicles on the main road. Obviously, this will influence the situation at our T-junction; typically it becomes easier for vehicles to turn between the platoons.

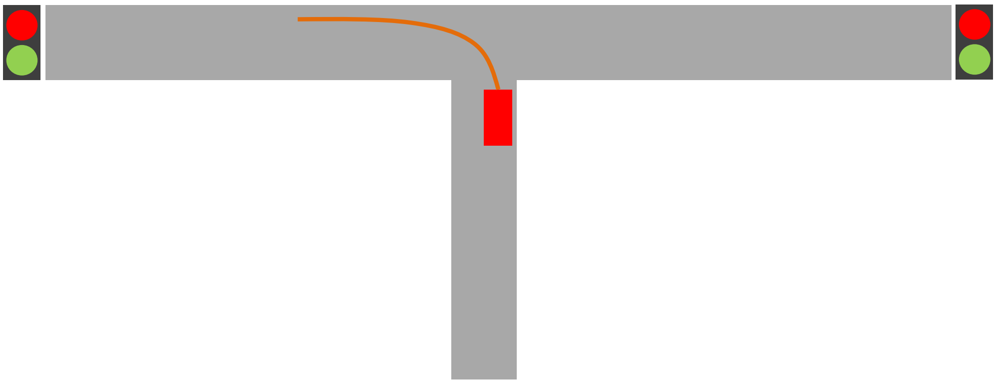

Figure 1: Vehicle at T-intersection.

But what if we have nearby signals on both sides? Depending on the signal timings we could have platoons from left and right at the same time at the T-junction, what is good because it allows the left turners to enter the main road after the platoons, or we can have the platoons alternating, so that the left-turners cannot enter at all since there is always either the platoon from the left neighbour junction or from the right. If we have different cycle times control at the signals, the situation at the T-junction will continuously change over time. And in case of vehicle actuated signals or public transit priority, the situation cannot longer be handled analytically.

This is the point when we need simulation. To assess capacity and level of service, design guidelines typically contain calculation procedures making simplifying assumptions.

Whenever the interaction between transport network elements or their temporal dynamic is significant, microscopic simulation is the appropriate tool.

Obviously, this is a very simple example. More complex real-world examples are complicated weaving sections, more than one round-about in succession and in general any case where spill-back becomes important, to name but a few.

…and when not

On the other side, simulation is not needed in every case. Still, there will be a lot of cases, where it is straightforward to use the knowledge that is laid down in works like the Highway Capacity Manual or the corresponding frameworks of other countries. However, we are quite sure that due to the increasing complexity of traffic and the ever better simulation software, the number of those cases will decrease in the future.

Stochasticity: It's random, stupid!

Have you ever tried to find out the occupation number of vehicles? The occupation number is this magic number that tells us, that on average 1.3 persons travel in each vehicle. To do so, one positions oneself at a place of interest and compiles a simple list: number of vehicles past, and, let's simplify matters even further, the number of vehicles with more than one passenger. Obviously, it is not possible to infer this number from just one vehicle observed. So let us say we observe ten vehicles. There is still a chance that we observe only vehicles with one person in it, and an even smaller chance that we observe only vehicles with more than one person. But "normally" we will have a mix of both and the average number of the persons in the cars will be a better estimate than just looking at one car. If we observe even more vehicles, e.g. 100 or 1,000, our estimate will become more and more reliable. See Figure 2 for six example experiments, which however have been done on a computer.

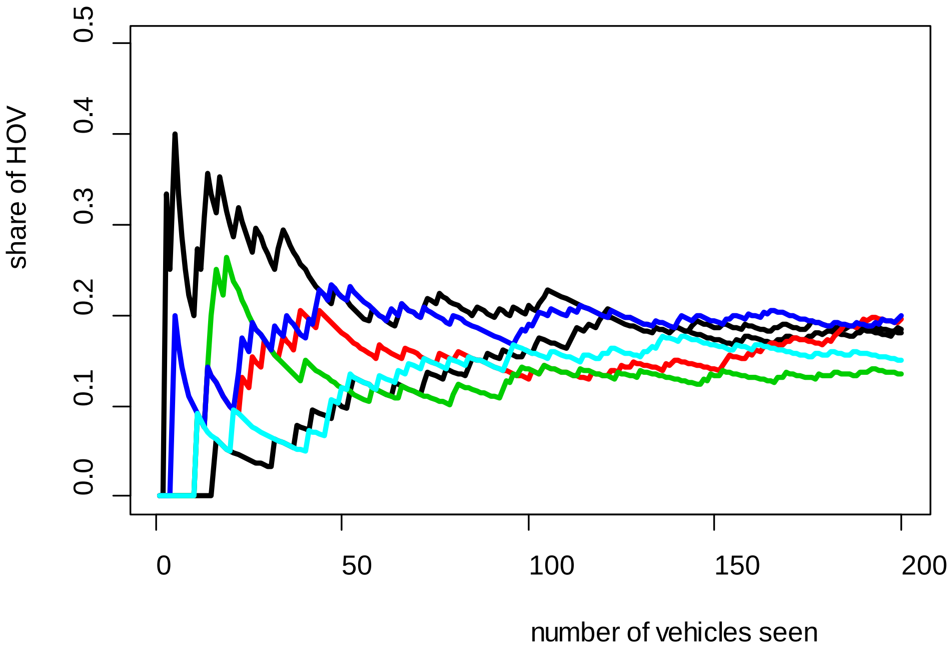

Figure 2: Share of vehicles with more than one occupant as function of the number of vehicles seen so far. The different curves are for different measurement campaigns (different days) and show the variation stemming from the randomness of the underlying process.

Let us look at another example: How long does it take you to go by car from home to work? If you commute by car, you have conducted that "experiment" very often. You might have a pretty exact idea what travel time you expect, but would you bet on the exact time for the trip tomorrow morning?

The point here is that traffic flow is random in several aspects. Although the traffic patterns are quite repetitive every day, you cannot know the exact sequence of vehicles, their exact individual speeds etc. for the next day. This randomness is replicated in microsimulation tools, but in a controllable manner. Computers can produce sequences of random numbers, and each sequence need a starting number, a so called "random seed" which starts off such a sequence. Another seed yields another sequence of random numbers, the same seed leads to the same sequence of random numbers. Most tools let you specify the random seed number, which determines which random sequence of vehicles and their attributes will be generated and how the vehicles will drive through the simulated road network. Essentially, running simulations with different seeds mimic different days in reality, with the import addition that in a simulation one can repeat even the same day, different from reality.

But which random seed number will then generate the correct result? Obviously, there is not one "correct" random seed, as there is neither a "correct" day in real traffic. To make a reliable statement about the observed traffic, average values from several days are needed, the more the better. For simulation this means that you need several simulation runs with different random seed numbers and then average the result.

So it is wrong to run the simulation only once. But how many runs do we need? That depends on how certain you want to be about the result and how much the output value under observation varies between the runs. A qualitative approach can be to run the simulation several times (five to ten) to get a feeling and a first approximation of the variation of the results, and then decide if more runs are needed. For instance, if the travel time in the simulation you are interested in fluctuates only by 1% of the mean value, this mean value is already known with a large degree of certainty. If the fluctuations are big (nothing forbids them to become larger as the mean value itself), then it is statistically needed to perform much more runs in order to trust the mean value computed from these simulations. Furthermore, in such cases, it is a good advice to check what is going on in your simulated system – large fluctuations should always send the user on alert, since it may stem from a system with multiple regimes: on one day, the intersection gets completely stuck, while on another day a "good" combination of vehicles keeps the intersection non-jammed. There is precise statistical formula that catches this dependence between fluctuations, the quality with which the result is needed, and the number of necessary runs, but there is no need to display this here.

But running many simulations is costly

Don't believe this, in general this is not true and even if it is now, it is subject to change. What does compute time costs? Right now, it is fairly cheap, and this price is expected to go further down in the future. So, obviously, there is more work needed to handle the data from multiple simulation runs. Yep, that is in fact true, but it is true only if it is done for the first time. Then, a good consultant has written all the necessary scripts to do this analysis, and is absolutely the same amount of work needed that has to be done to analyse the output from a single simulation run. There is only one exception from this rule, and this happens for simulation programs that do not have the means to control them from extern. Such programs exist, in this case all the runs have to be done manually, which is obviously a night-mare if you do that more than ten times or so. However, we are quite sure that such software will rapidly go extinct, it does not fit any longer in a world brimming with multi-core, multi-cpu machines.

Calibration

Consider again the issue of travel times. Now, we want to now not only your personal travel time from home to office, but we want to know the travel time of all the travellers in the city, together with an estimate of the emission caused by them. Furthermore, one would like to compare the two different approaches to minimize congestion and emission production in the city, since the better of the two will be going to be implemented in the future. In this case, we have jumped into the most complex task for the calibration of the simulation needed to compare the two different scenarios. Of course, the result one will finally obtain depends on a huge number of different parameter, data, and so on, but we will only focus here on something that at a first glance it not very important: the reaction-time of the drivers. However, own experience tells you that if there are a lot of drivers who are reacting slowly to a traffic light that just switched to green, the number of vehicles that can pass a traffic signal during green are seriously impaired. So, it is clear that this number matters, and it is: not known. To deal with this is now the task of calibration. In the best of all worlds, you would go to the psychological department of your university and ask the psychologists whether they can give you this number. However, they often do not know it, and even if they do, they have measured it in the laboratory, so that there is not a single chance that this number is also valid in front of the traffic lights of this city. To make matters even worse, this number is not a single number for all of the population, but different drivers have different reaction times.

Here, calibration comes to the rescue and saves the day. The only thing which finally matters is how many vehicles do pass on average a given traffic signal during green time. And most of the simulation models have a single parameter called reaction time that can be used to adopt the simulation results to the results observed in reality. Small reaction time gives a large number of vehicles, large reactions times gives a small number of vehicles. At some magic number, which is to be found by the methods to actually do the calibration, the simulated number of vehicles during green exactly fit to the measured number of vehicles. Then, this model is said to be calibrated with the data at hand. And this is a bit magic, and it turns out to be a truly powerful approach: for instance, it is not really clear, that the reaction time in the model is in fact similar to the reaction-time in reality. It just act as a handle to actually do the calibration, and it wonderfully smoothens away small or even not too large uncertainties and modelling errors in the used simulation approach.

So, if anything turned out as planned, one finally ends up with a calibrated simulation of the city under consideration which gives the correct travel times and their fluctuations around the mean value.

Unfortunately and not surprisingly, there is a price to pay, or actually two prices. The first is that data are needed, and the second is that it must be made sure that the adaptation of the simulation to the real-world data is not due to some particularly, very rare statistical fluctuation, so to speak, the one and only number in the lottery, which lets you crack the jackpot. This process that is closely related to calibration is named…

Validation

To make sure, that the result obtained by calibration is general enough to allow to be used in slightly or even strongly different situations, it is needed to run the simulation with another, independent set of data and to determine, how well it describes this changed reality. In the example above, the data for calibration may come from a measurement campaign (which may even simply consist of reading data from a traffic management centre) on one day, while the data for the validation relates to another day. In case of doing the validation, it is not allowed to change the reaction time obtained in the calibration step. It is taken as fixed and used as it is in the new simulation study. It should come as no surprise, and in almost all of the cases we have seen so far this was true either, that the validation does not work as well as does the calibration. This is to be expected, since calibrating the parameters is usually a really bold endeavour expected to lead to a really good reproduction of reality by the model at hand.

No data

It has been seen so far that calibration and validation need data. One point that has been learned during the course of MULTITUDE is that a large number of all studies are done without ever using any data. Although there might be rare occasions where to do a study without any data, we strongly advise against this. If there are no data available, get them. If there are still no data available, increase your efforts to get them. Even if you learn from these data that the parameters in your model (as put into it by the manufacturer of the model) fit the data you have obtained quite well without any calibration, your study is now much stronger and more trustful than it was without data.

However, if, after all these warnings, you are still want to go ahead without data, then here are the few exceptions from “use data!” claim.

For instance, a lot of research papers go without looking into data. They simply compare a new model and its performance against another model already described to progress the state of the art. This can be done by using generic situations and try to compare the models’ performances for as many different situations as possible. An example might be a new traffic-actuated traffic light algorithm, which is to be tested at an idealized intersection for a large number of different demands for all the traffic streams. In this case, completeness is exchanged for having real data, and in this case this is definitely a good idea as long as it can be made sure that the input data are within certain regimes.

Many other things

Apart from randomness and calibration, there is a large bunch of more detailed and more technical stuff that turns out to be important for the success and correctness of a traffic simulation study. This consists of making sure to have the right network and infrastructure data, to have the right definition and data regarding the demand for travel, to know the important parameters of route choice and modal choice, to ensure that the simulation covers the appropriate area and time. If only a part of the day is simulated, it must be ensured that the simulation is started early enough before the actual study time interval under consideration: this avoids the simulation of the situation with a simulation network filled only partially. And be aware: all, with the exception of the simplest, network and infrastructure data do have errors.

Of course, all these and much more details are needed to do a valid and sensible simulation study that truly helps and that avoids costly mis-planings. These are described in the much more detailed simulation guide-lines and manuals, and they are described in another, much longer document produced by this COST action, that is MULTITUDE’s “Case for Guidelines” document.