International Journal of Modern Physics C, Vol. 9, No. 1 (1998) 113–132

RELAXATION CRITERIA FOR ITERATED TRAFFIC SIMULATIONS

TERENCE KELLY*
Los Alamos National Laboratory and University of Michigan
E-mail: tpkelly@eecs.umich.edu

KAI NAGEL†
Los Alamos National Laboratory and Santa Fe Institute
E-mail: kai@lanl.gov

Received 24 October 1997
Revised 12 December 1997

Iterative transportation microsimulations adjust traveler route plans by iterating between a microsimulation and a route planner. At each iteration, the route planner adjusts individuals' route choices based on the preceding microsimulations. Empirically, this process yields good results, but it is usually unclear when to stop the iterative process when modeling real-world traffic. This paper investigates several criteria to judge relaxation of the iterative process, emphasizing criteria related to traveler decision-making.

Keywords: Traffic; Transportation; Large Scale Simulations; Shortest Paths; Relaxation.

1. Introduction

An individual who wants to go from location A to location B is faced with the questions of whether to go at all (maybe it is too far), when to go, and which mode and route to select. The problem is complicated by other travelers: Congestion will make some routes slow and thus influence the decision.

From the point of view of transportation planning, the question is how to assign transportation demand to the transportation infrastructure. Since answering all the above questions simultaneously is a hard problem, it often gets reduced to the route choice part alone. Current transportation simulation projects approach this problem by day-to-day iterations: Each individual selects a route, the microsimulation is run and congestion is recorded, some individuals select a new route which would have been better, the microsimulation is run again, etc.[1–6]

It is empirically clear that this has a beneficial effect on the structure of the traffic patterns.[4,6] It is less clear how to describe exactly where such iterations

*Advanced Technologies Laboratory, University of Michigan, 1101 Beal Avenue, Ann Arbor, MI 48109-2110 USA.
†Los Alamos National Laboratory, TSA-DO/SA, MS M997, Los Alamos, NM 87545 USA.

lead: Do they converge? If so, do they converge towards a fixed point, towards a periodic attractor, or towards something else? Are there possibly several basins of attraction? Most importantly: Does this have anything to do with what the real system does?

This paper will start with a short review of traditional assignment (Sec. 2), leading to the issue of stopping criteria in iterative assignment, both for traditional methods based on link performance functions and for simulation-based methods (Sec. 3). Sections 4–7 describe the data that we used: the geographical context, the microsimulation that was used, the replanner that was used, and a description of which particular simulation output was used. In general, all investigations compare three situations: (i) uncongested; (ii) congested "unrelaxed," i.e., at the beginning of the iteration process; (iii) congested "relaxed", i.e., at a later stage during the iteration process. Section 8 describes velocity distributions for cars and on links for these three situations, and Secs. 9–12 describe various results concerning the structure of fastest paths and their alternatives. This is followed by a discussion of the results (Sec. 13), a short section on computational considerations (Sec. 14), and a summary.

2. Traffic Assignment

In static traffic assignment, we assume time-independent demand and allocate traffic flows in some "optimal" way. Optimal here means that some cost, say time, is minimized, either for each individual user (User Equilibrium, Nash Equilibrium, Wardropian Equilibrium), or for the system as a whole (System Optimum). To include the effects of congestion, one needs a mechanism which reflects the higher cost of congested links. Traditionally, this is done using "link performance functions", i.e., functions which return the speed on a link as a function of the number of vehicles that use it, $V(\rho_i)$, where i is the number of the link and ρ is the density of vehicles on it. Given these ingredients, there are many ways to solve this problem.[7,8]

One of the problems with this approach is that its handling of strong congestion is dynamically inconsistent. If in a real-world network more vehicles attempt to use a link (say a bridge) than the bridge can handle (i.e., demand is higher than capacity/supply), the system reacts with queue build-up at the entry to the bridge. And now, the travel-time one needs to get across the bridge depends on when one gets there: If one gets there just when the queue build-up starts, the waiting time is still short; if one gets there an hour later, one may have to wait a long time.

Of course, the problem with traditional assignment is that the cost function is not history dependent. Making it history dependent would be possible, but it would make the optimization problem mathematically much harder. And then, the deeper problem is that the approach in general makes it complicated to deal with added complexity, such as queuing in turn pockets, interactions between slower and faster vehicles, bus schedules, individually different preferences, non-rational behavior, etc.

This explains why using micro simulations, which only recently became possible on a large enough scale due to advances in hardware and algorithms, are an attractive alternative: A microsimulation can, at least conceptually, add all the complicated elements of reality and generate a dynamically consistent behavior. However, since none of the mathematical methods works any more, one resorts to iteration as described in the introduction, i.e., use the information generated by the microsimulation to adjust some of the route plans, run the simulation again, adjust some more of the route plans, etc. Note that one of the earliest and simplest steady-state assignment algorithms, the Frank–Wolfe algorithm, is very similar in spirit, although geared towards steady-state situations.

3. Stopping Criteria for Iterative Assignment

When doing iterations, one needs a stopping criterion. In deterministic steady-state assignment, this is easy: Just monitor the system, and stop when changes (according to a pre-defined measure) are smaller than some pre-defined ϵ. That requires a convergence proof, which is available for deterministic steady-state assignment.[7,8]

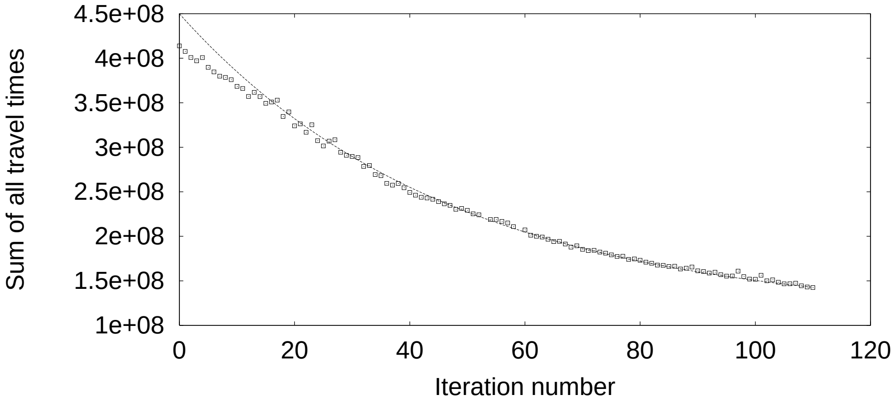
The microsimulations that we are using are neither steady-state nor deterministic, and it is unclear if any of the above holds. To get some intuition, one can for example plot the sum of all individual travel times versus the iteration number (Fig. 1). One sees that there is some roughly exponential convergence in that quantity; yet, the plot says nothing about if the underlying traffic patterns remain the same from day to day or if there are strong fluctuations which just average out on the macroscopic level.

For comparison, Fig. 2 shows an often-used stopping criterion for traditional assignment applied to our data series (described later). The function on the y-axis is defined as

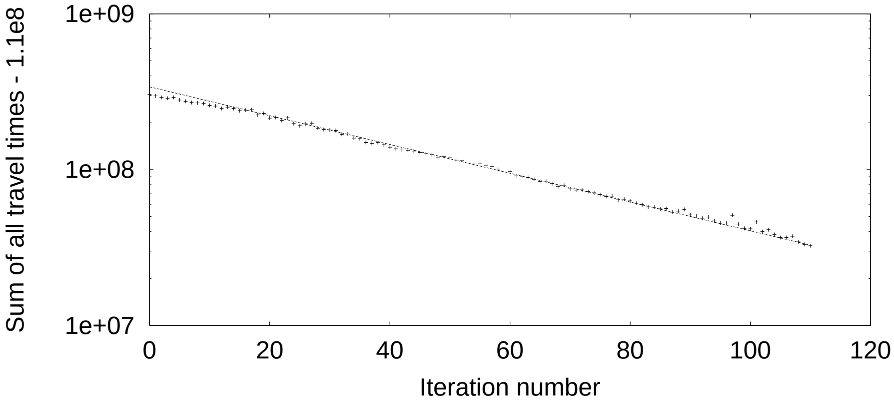
$$F = \frac{\sqrt{\sum_a (x_a^{n+1} - x_a^n)^2}}{\sum_a x_a^n}, \tag{1}$$

where x_a^n denotes the flow on link a at iteration n; link flows here are always positive; this measures changes from one iteration to the next on an individual link basis. [Ref. 7, p. 119] The iterations are supposed to stop when this quantity decreases below a certain level. Yet, note that, in Fig. 2, its value during the first iterations is not much higher than what is reached near the end; and the plot indicates that it will probably never converge to zero (indicating that there are indeed considerable variations between iterations, which just cancel out in the aggregate variable used above). Equation 1 is not appropriate for our purposes.

Thus, at best one could demand that some function converges *in the average.* By the strictest definition this would mean to run, say, the plan set of iteration $n-1$ for a couple of times with different random seeds, then the plan set of iteration n for a couple of times, then average the link flows for both iterations, and then compare them according to the above criterion.



(a)



(b)

Fig. 1. (a) Sum of all individual trip times, S, as a function of the iteration number, and possible exponential fit to the tail of the data, $S = S_\infty + (S_0 - S_\infty) \cdot \exp(-n/\nu)$, with parameters $S_\infty = 1.1 \cdot 10^8$, $S_0 = 4.5 \cdot 10^8$, $\nu = 47$. n is the iteration number. The particular value of S_∞ is also justified by results from other iteration series which indeed seemed to "converge" at that value. S_0 and ν were fitted afterwards, giving more emphasis to the tail of the data. Clearly, at the beginning of the iteration process, heavy congestion makes the sum of all trip times large. Better trip distribution across links relieves congestion. Note that the curve is not yet flat after 110 iterations, indicating that the system is not yet completely "relaxed." (b) Same data, but y-values reduced by $1.1 \cdot 10^8$ and y-axis logarithmic. Data from M. Rickert.

Yet, this seems like a waste of computational resources. This is especially true since, as pointed out in the introduction, there is no reason to believe that real world transportation systems go to this state. For the remainder of this paper, we want to look for more "structural" quantities, i.e., quantities that:

- reflect more directly how people behave (travel times are a very indirect result of their behavior)
- are robust with respect to calibration (travel times depend on, e.g., speed limits)
- could, at least in principle, be obtained in the real world.

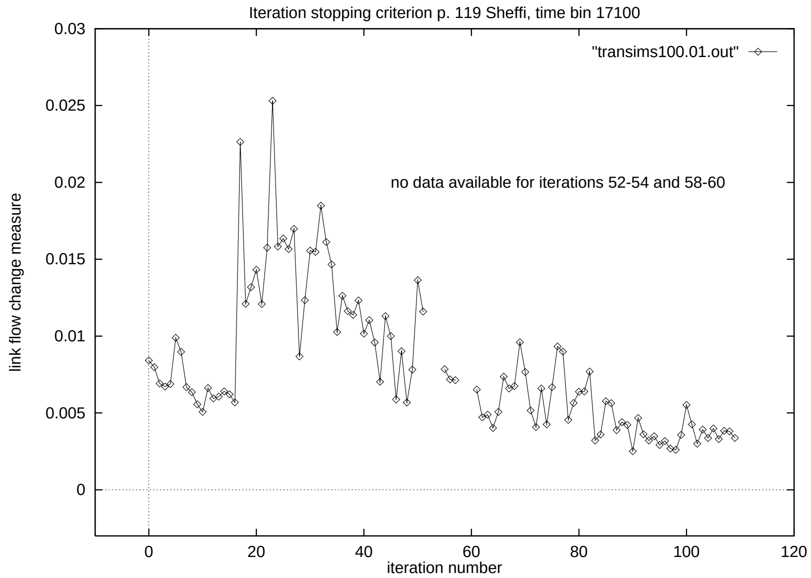


Fig. 2. A conventional stopping criterion (Eq. 1) as a function of the iteration number. Data points are unavailable for iterations 52–54 and 58–60. Data from M. Rickert.

As a guiding theme, we will use the following idea: When traffic fills up the system as the rush hour approaches, vehicles will first use the fast links such as freeways, in consequence slowing them down. Eventually, freeways will not be any faster than major arterials, and people start using major arterials. Once major arterials are sufficiently slowed down, minor arterials will start filling up, etc. Thus, a “relaxed” system should somehow reflect a different balance between fast and slow than an unrelaxed or an empty system.

4. Context

The context of this study is the TRANSIMS Dallas/Fort Worth case study, as described in Ref. 9. Most of the information relevant for the present paper can also be found in Ref. 10. We focus on a “study area” which is an approximately five miles by five miles area around the busy intersection of the LBJ freeway and the Dallas North Tollway. This is the only area for which all streets, including local streets, are in our data base. With increasing distance from the study area, less and less information is provided; for example, Fort Worth is only represented by its freeways. The whole network consists of 14751 (mostly bi-directional) links and 9864 nodes.

The initial travel demand input to the case study consists of production-attraction (PA) matrices provided by the Dallas/Fort Worth regional planning authority. From these matrices, individual trips (i.e., lists of: starting time, starting location, and ending location) between 5 am and 10 am are generated.[9,11] A trip distribution as function of time-of-day is used here. For the so-called *initial plan set*, these trips are routed using the fastest path based on free speeds. All routes which do *not* go through the study area are discarded. Subsequent iterations are run on the remaining set of routes; i.e., the re-planner can route trips around the study area, but these trips are not discarded.

5. The Microsimulation

The microsimulation module itself has been described in Ref. 12. It is based on the cellular automata (CA) technique, that is, the road is divided into cells each 7.5 meters long, each cell is either empty or occupied by exactly one vehicle, and as a result vehicles move by hopping between cells.[13,14] The simulation includes the correct number of lanes plus lane changing because of slower vehicles ahead, speed limits, and plan following (i.e., each vehicle follows a precise, link by link, route plan, pre-computed by the route planner). Intersections are fairly abstracted; for example, vehicles can move from any incoming lane into any outgoing lane. Since that means the simulation does not differentiate between directional links, "average" light cycles across all lanes are used. As implied above, the microsimulation used for this study runs on "pre-computed" route plans, i.e., at the beginning of each microsimulation run, for each vehicle the starting time, starting location, and the precise, link-by-link route plan are already known and cannot be changed during the simulation. For further information, see Refs. 6 and 12.

6. The Replanner

As explained above, we use the information from the previous microsimulation run to adjust the plan set. This is done by selecting a certain fraction, X, of route plans randomly and replacing them with (time-dependent) fastest paths based the link travel times computed by the previous microsimulation. For the results in this paper, the re-planning fraction was one percent, i.e., one percent of the travelers per iteration had the option to change their route plans. More details will be published elsewhere[15,6]; the general approach is similar to the one described in Ref. 10.

Note that this approach is neither "rational" nor behaviorally well-justified.

- Rational behavior would imply that agents optimize their behavior so the the "expected" (equivalent, in our case, to "average") travel time would be minimized. Our algorithm, when selected for re-routing, only minimizes with respect to the last iteration; for a stochastic microsimulation, this most probably will not lead to an optimized "average" behavior for the individual.

  Note that when using a deterministic microsimulation, some of these problems vanish and mathematical statements concerning convergence are possible.[16–18] Also note that forcing the algorithm to generate averaged behavior does not necessarily lead to good results; e.g., in the context of departure time choice see Ref. 19.[a] This fact has at least intuitively been known for a long time.[20–22]

- From a behavioral perspective, people in reality do not have access to as much information (travel time on all links) as they have in the simulations. It is actually possible to make the iterations more realistic with respect to the second aspect[22]; also, the Intelligent Transportation Systems (ITS) initiative may provide more information in the future.

7. Methodology

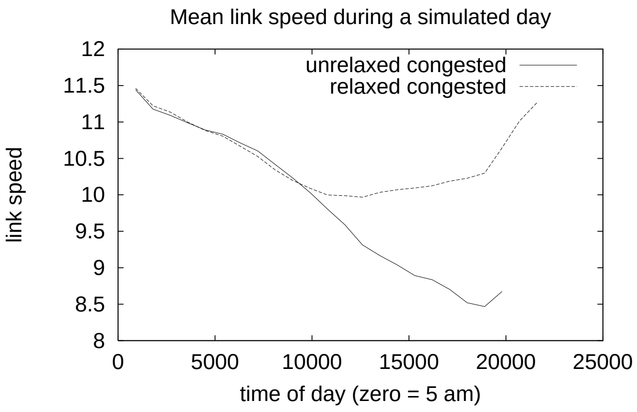
We begin with the assumption that our simulated network is far from relaxation as the planner–microsimulation iterative process begins. At the initial run ("0th iteration") of the microsimulation, when each motorist has chosen routes under the false assumption that she has the road network to herself, the freeways quickly grid-lock due to over-use, and in many respects the traffic patterns do not resemble real traffic. After many iterations of the planner and microsimulation (100 iterations at 1% re-planning in the examples discussed here, "run5" in Rickert's terminology,[15,6]) the network has relaxed to the point where link flows and turn counts resemble those of the real-world network. We therefore assume that the zero iteration network is far from relaxation, and the 100th iteration is more relaxed (although not completely). We seek an equilibration measure which distinguishes clearly among three network regimes: an uncongested network, an "unrelaxed" congested network, and a "relaxed" congested network. Unless otherwise noted, we use (i) "free" speeds on the links for the uncongested case, (ii) speeds averaged between 8 am and 8:15 am from the initial run ("0th iteration") for the "unrelaxed congested" case, and (iii) the mean of 100th through 110th iteration link speeds between 8 am and 8:15 am to represent "relaxed congested" network link speeds.

8. Velocity Distributions

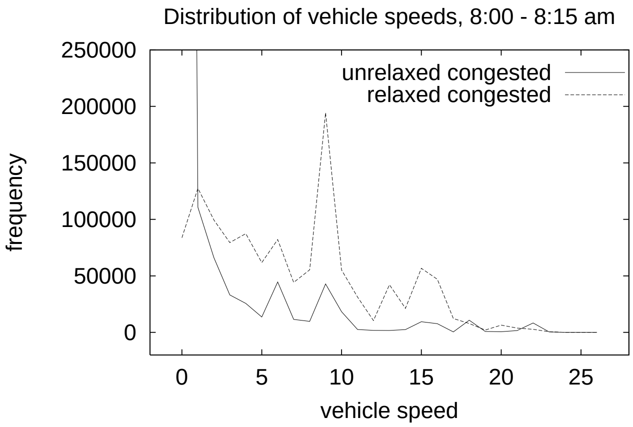
First, we look at speed statistics. Figure 3 first shows how the average link speed, i.e., $V = \dfrac{1}{M}\sum_{i=1}^{M}\langle v\rangle_i$ with $\langle v\rangle$ the average speed on a link and M the number of links, develops as a function of day. The curve for the relaxed situation (110th iteration) shows the expected form: First decrease of average speed due to the rush period; later increase of average speed when the rush period is over. In contrast, for the non-relaxed situation, the average link speed just keeps decreasing during most of the simulation period.

To test the hypothesis that congestion slows down the fast streets, we also look at speed distributions, both by vehicle and by link (Figs. 3(b) and (c)). Clearly, in the relaxed situation, speeds are much more concentrated around an average

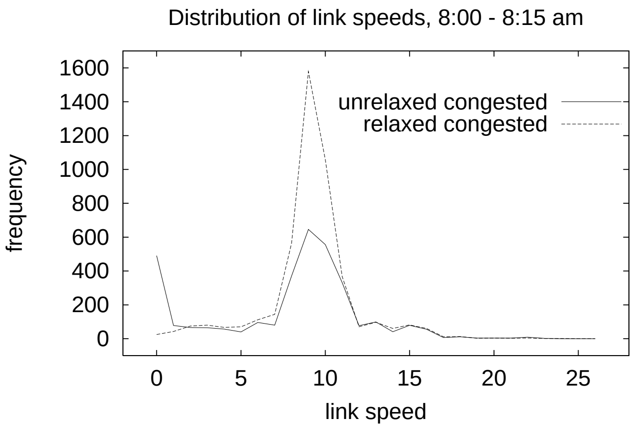
[a] The reason for this instability is, loosely speaking, the following: Too many people choose more or less the same solution; the system reacts by generating a very broad distribution of performances for each individual for this solution, i.e., the performance becomes highly unpredictable; but the expectation value over this distribution still favors this solution so everybody sticks with it.



(a)



(b)



(c)

Fig. 3. (a) Mean of link speeds as function of time-of-day for the 0th ("unrelaxed") and the 110th ("relaxed") iteration. (b) Distribution of vehicle speeds during a congested time-of-day. (c) Distribution of link speeds during a congested time-of-day.

value (around 9 mph), whereas in the unrelaxed situation, the distribution for both statistics is more flat and spread out, and has strong contributions at zero speed, indicating gridlock (see the discussion under Sec. 13).

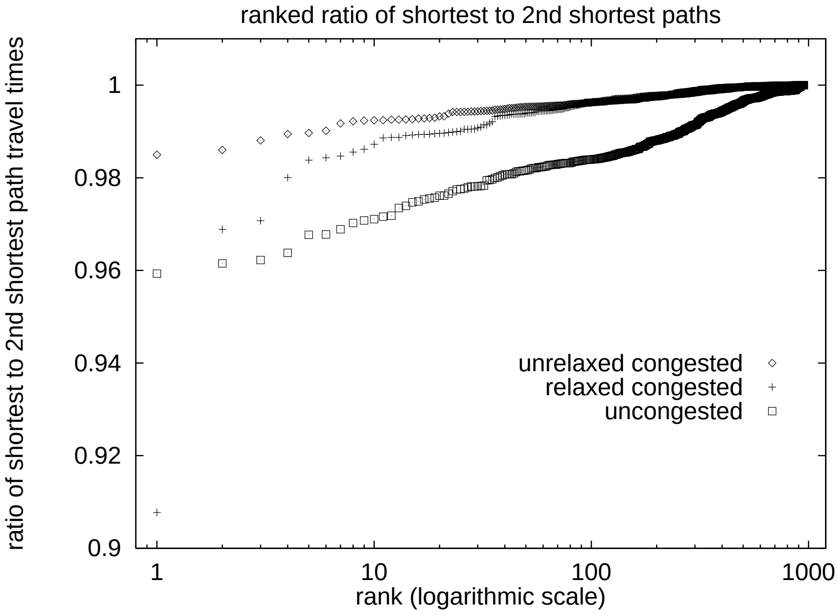
9. Second Fastest versus Fastest Path

Next, we want to look at measures which reflect the structure of the decision-making of the individual user. The above argument implies that, in the relaxed situation, the decision landscape should be "flat" near the minimum: Many different route choices should result in similar travel times. This leads us naturally to consider K-fastest paths between randomly-chosen origin–destination pairs. First, we consider only the fastest ($K = 1$) and second-fastest ($K = 2$) paths between a large set of randomly-chosen O-D pairs. Technically, we use 955 O-D (origin-destination) pairs in the study area which are between 7 km and 7.01 km apart. We then calculate the fastest and the second-fastest path based on three link costs: (i) free flow,[b] (ii) link speeds provided by the initial ("0th iteration") microsimulation, and (iii) link speeds given by the mean of the 100th to 110th iteration simulations.

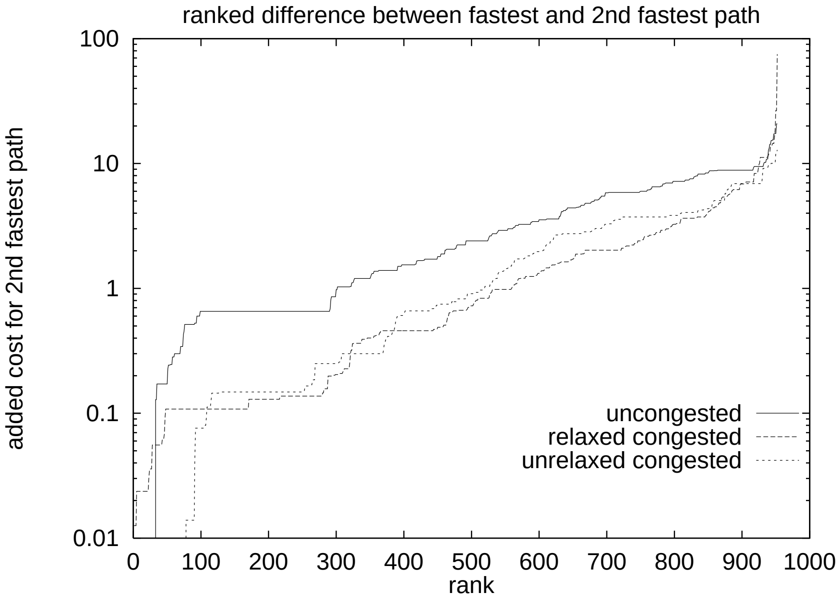
In cases (ii) and (iii), paths start at 8:01 am, and as link costs we use link travel times averaged from 8:00 am to 8:15 am. In general, we compute K-fastest simple (i.e., loopless) paths using an algorithm described in Ref. 23. Note that this algorithm does *not* compute fastest paths in time-dependent networks, i.e., networks with varying link travel times. In all of the fastest paths considered here we use link travel times taken from a (15-minute long) "snapshot" of our network. We were unable to locate an efficient (i.e., polynomial-time) algorithm in the literature which computes K-fastest simple paths in graphs with time-varying link travel times. Yet, most of our paths reach their destination before 15 minutes are over so that this is only a minor concern for the present study. For more computational details, see Sec. 14.

Our intuition is, again, that in a congested network in equilibrium, routes will be distributed more evenly among paths than in a free-flow network, and therefore the ratio of fastest to 2nd-fastest path travel times will be nearer to unity than in the free-flow network. To put it differently, the disincentive to deviate from a fastest path will be weaker in a relaxed congested network than in an uncongested network. As we see in Fig. 4, this is indeed the case. In other words: Small deviations from the fastest path are most expensive in the uncongested network. We also notice that the zero-iteration line lies *above* the curve representing the relaxed network, indicating that under unrelaxed congested conditions small deviations are even less expensive than under relaxed congested conditions. Thus, this measure seems to be capable of differentiating between uncongested and congested conditions, but not between relaxed and unrelaxed congested conditions. It thus does not appear to be a promising means of measuring a network's proximity to equilibrium.

[b] For the present study, we use the free speeds as provided by the Dallas regional planning authority. These are not necessarily the same as the free speeds generated by the microsimulation. The effect of this needs to be tested in future work.

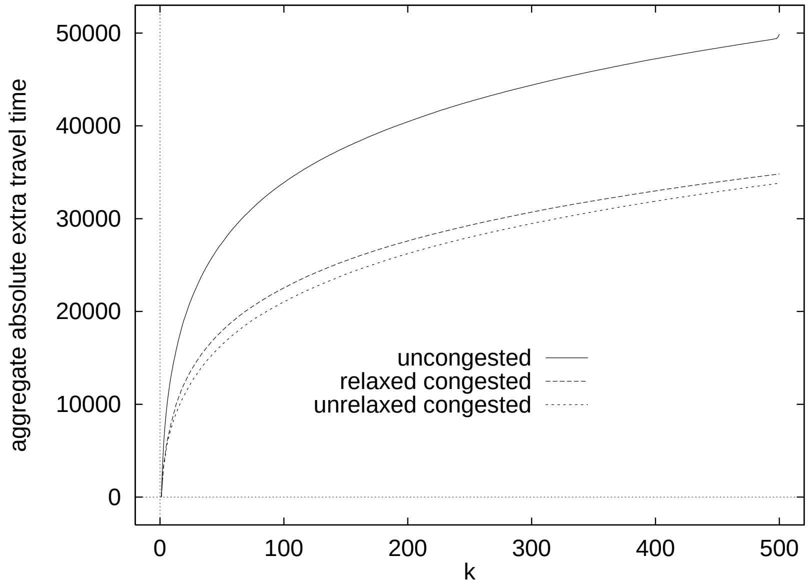


(a)

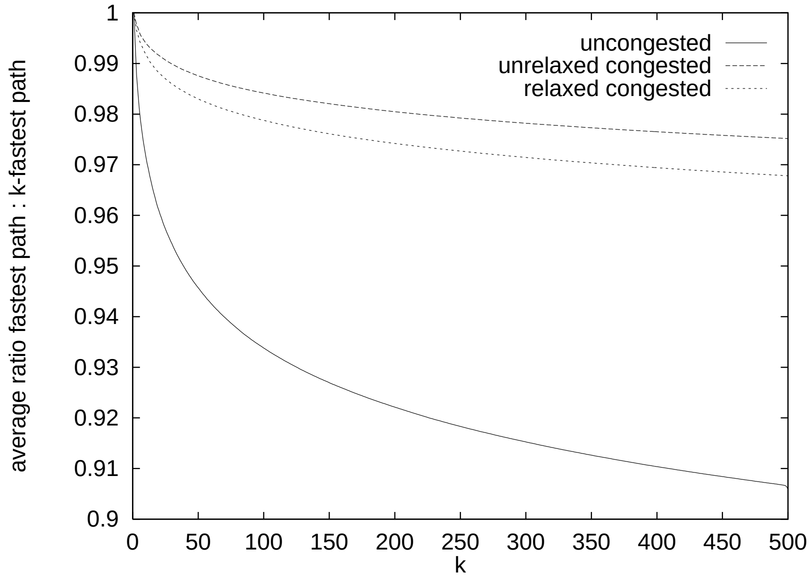


(b)

Fig. 4. For 955 O-D pairs at distance 7 km in Dallas, we compute the ratio (a) or the difference (b) of fastest path to second-fastest path trip time and sort these values in increasing order. The "relaxed congested" curve has one outlier at a ratio of about 0.91 and at an added cost of about 80.



(a)



(b)

Fig. 5. Aggregated rank-cost profiles of 955 Dallas O-D pairs at 7 km distance at 8:01 a.m. (a) Aggregate extra travel time for the 2nd, 3rd, ..., K-fastest path. The three sets of rank-cost profiles were aggregated by simple addition, i.e., for each rank we plot the sum of travel time minus fastest path travel time. The three lines are significantly different from exponentials. (b) Average ratio between the K-fastest and the fastest path. In both figures, we see again that the penalty for deviating from the fastest path is much higher on an uncongested network than on either the zero iteration network or the relaxed congested network.

10. *K*-Fastest versus Fastest Path

To explore further our intuition that disincentives to deviate from a fastest path should be weaker in relaxed congested networks, we consider the rank-cost profile of K-fastest paths, i.e., of excess travel time for fastest, 2nd-fastest, 3rd-fastest paths, etc. The results of this investigation are summarized in Fig. 5 which shows the aggregated rank-cost profiles of 500 fastest paths between each of a large random sample of O-D pairs computed using uncongested, unrelaxed congested, and relaxed congested link travel times. We see that these aggregated rank-cost profiles can again distinguish between uncongested and congested networks — note the large gap between the free-flow curve and the others — but they are less useful in distinguishing between equilibrated and unequilibrated networks.

11. Average Similarity to Fastest Path as a Function of Extra Travel Time

So far, our measures fail to reveal any systematic difference between unrelaxed congested and relaxed congested conditions. Yet this is not surprising. The above measures look at additional absolute or relative travel time only as a function of the number K (from the K-fastest path). This is a purely computational quantity and has little meaning in the real world. A second fastest path typically is only a tiny variation of the fastest path, such as leaving and re-entering a freeway at a ramp.[24] This is not what we mean by "real alternatives"; real alternatives should be significantly different from the first choice.

In order to access this problem quantitatively, we need a measure of path similarity. Many measures of path similarity have been proposed. Reasonable definitions

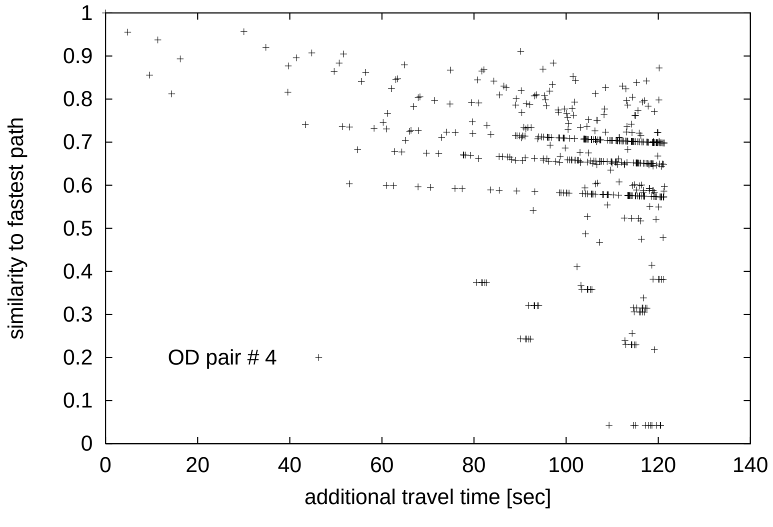
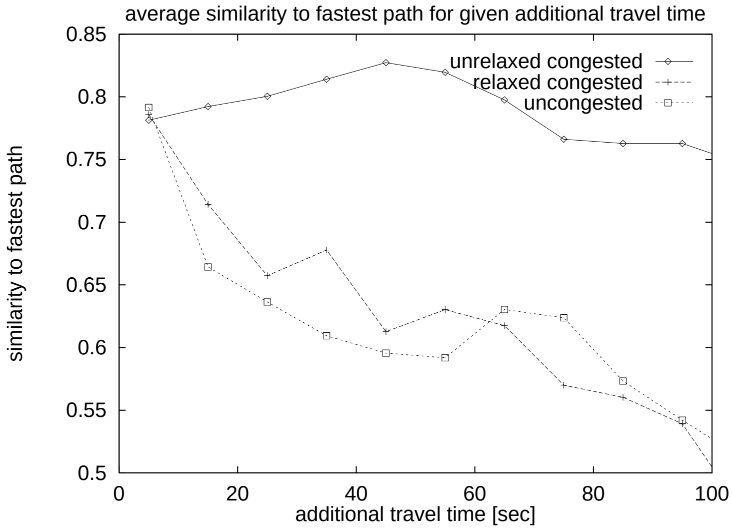
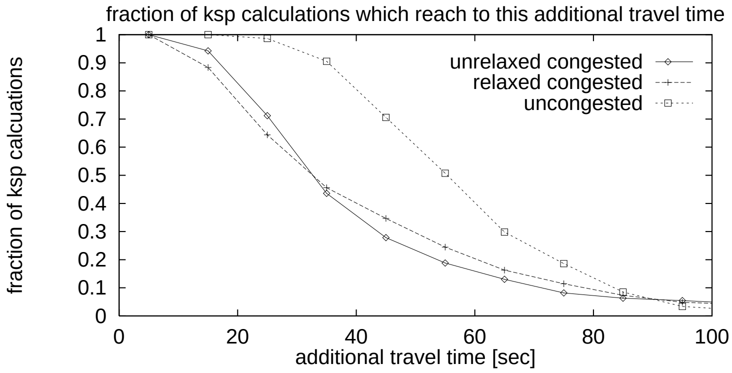


Fig. 6. Similarity to respective fastest path as function of added absolute travel time for a particular O-D pair. The nearly horizontal traces come from "robust" alternatives and small variations around those.



(a)



(b)

Fig. 7. (a) Average similarity to fastest path as function of added absolute travel time. Averaged over 955 O-D pairs. For any of the situations (uncongested, unrelaxed congested, congested), the similarity measures of all K-fastest paths that had additional travel times between 0 and 10 seconds, 10 and 20 seconds, etc., were averaged. (b) Fraction of calculations that reached out to that value of additional travel time.

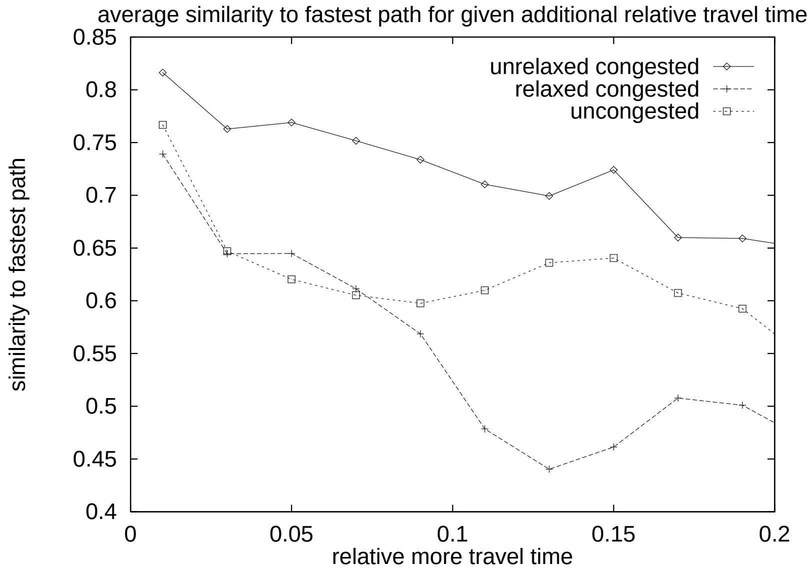
exist for "detour" paths[25] and paths with a bounded number of edges in common.[26] For the present purpose, we use the "common travel time ratio" Φ defined in Ref. 27:

$$\Phi_j(i) = \frac{\sum_{a=1}^{n} t_a d_{ij}^a}{\sum_{a=1}^{n} t_a d_{ii}^a}, \tag{2}$$

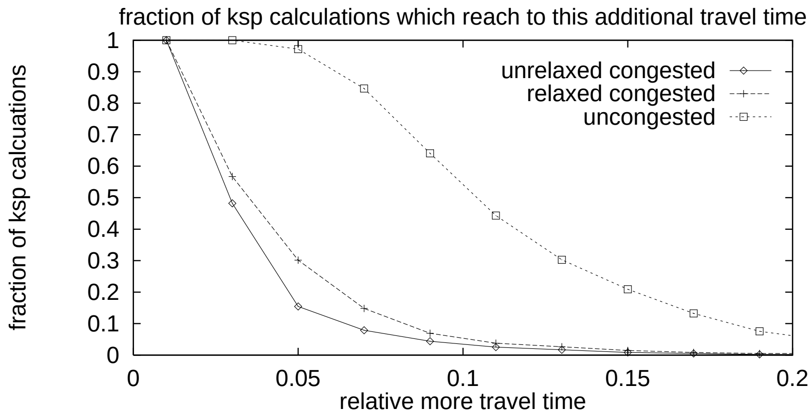
where $\Phi_j(i)$ is the common travel time ratio of path i with respect to path j, $d_{ij}^a = 1$ if link a is used by path i and path j, $d_{ij}^a = 0$ otherwise, t_a is travel time on link

a, and n is the number of links in the network. This is simply the fraction of path i's travel time spent on links shared with path j. Note that $\Phi_j(i)$ must lie in the range $[0, 1]$ and need not equal $\Phi_i(j)$.

Figure 6 shows a plot of this similarity measure as a function of additional travel time for a particular O-D pair and the first 500 shortest paths.[c] In general, one seems to find more paths that are very different from the fastest one with increasing



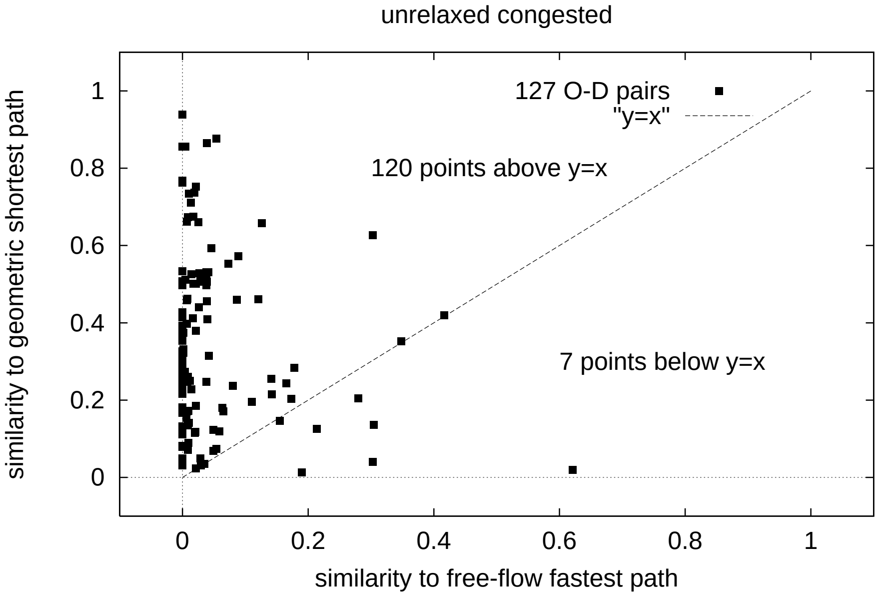
(a)



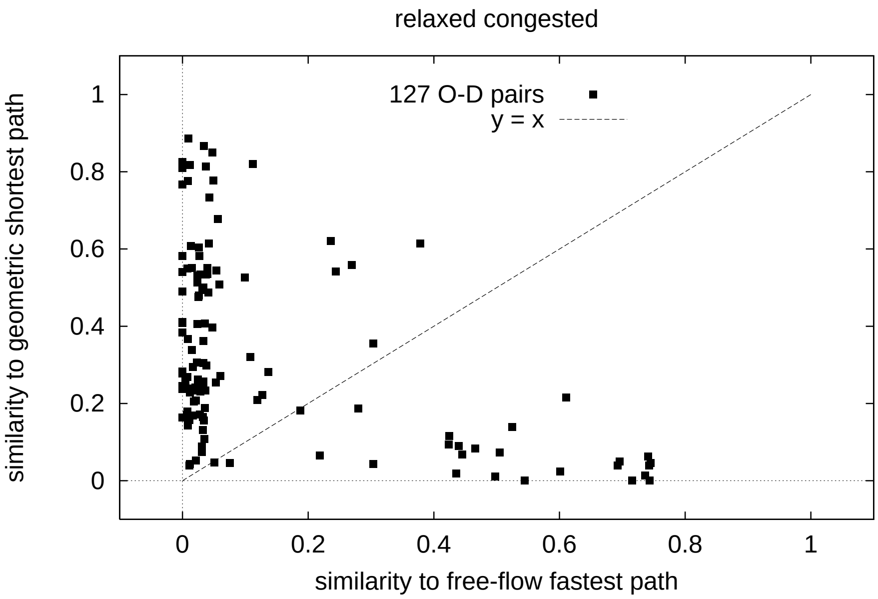
(b)

Fig. 8. (a) Average similarity to fastest path as function of added relative travel time. Averaged over 955 O-D pairs. (b) Fraction of calculations that reached out to that value of additional relative travel time.

[c]Similar plots have been made for proteins: Similarity to the ground state as a function of the additional energy.[28]



(a)



(b)

Fig. 9. We selected 127 O-D pairs at distance 7 km in Dallas such that the common travel time ratio measure $\Phi_f(g) < 0.05$ where f is the free-flow fastest path and g is the geometrically shortest path. For each of these pairs we compute a fastest path r using unrelaxed congested link travel times and relaxed congested link travel times. Above we plot $\Phi_f(r)$ and $\Phi_g(r)$. (a) shows the result for the unrelaxed congested fastest paths while (b) shows that for relaxed congested fastest paths.

additional travel time; yet, the values are strongly fluctuating. Nevertheless, averaging over all 955 O-D calculations shows an interesting result (Fig. 7): For a given additional travel time; yet, the values are strongly fluctuating. Nevertheless, averaging over all 955 O-D calculations shows an interesting result (Fig. 7): For a given additional travel time, both uncongested and relaxed congested K-fastest paths show a much larger difference from the fastest path than under *un*relaxed congested conditions. That is, if one is ready to accept a certain additional travel time, the alternatives one finds under *un*relaxed congested conditions are less *different* from the best alternative. Plotting in the same way the average similarity as a function of the additional *relative* travel time (Fig. 8) shows that there may even be a difference between all three regimes, uncongested being in the middle between unrelaxed congested and relaxed congested.

Underneath the similarity plots are plots which show the fraction of K-fastest path calculations that have reached particular values of additional absolute and relative travel times. It would have been better to use the absolute/relative additional travel time as stopping criterion for the computations; with the current data material, we cannot be sure that the effect is a result of the fact that some calculations have high additional travel times within the first 500 fastest paths and others do not. This, plus the fact that the plots show highly fluctuating results, indicates that more elaborate computations should be made.

Yet, the same information also explains why K-fastest paths alone reveal a different information: Given a certain amount of additional (absolute or relative) travel time, there are many more options under congested conditions than under uncongested conditions. It is that effect that dominates many of the measures and clouds up the differences between the two congested regimes.

12. Similarity to Geometrically Shortest Paths

If link speed variance is low in relaxed congested networks, one implication is that during the peak period freeway links are not inherently more attractive to motorists than arterials or local streets. Congestion reduces the distinction between different functional class link categories. One consequence is that a fastest path computed on a relaxed congested network will resemble a straight-line geometrically shortest path far more closely than it will resemble a fastest path computed using uncongested link travel times. This effect was evident to us when we visually inspected fastest paths in a congested network. We again used the similarity measure $\Phi_j(i)$ of Eq. (2) to quantify the similarity of fastest paths in a congested network with fastest paths in the uncongested network and geometrically *shortest* paths computed using link lengths rather than link travel times.

We selected a large sample of O-D pairs such that for each O-D pair the uncongested fastest path was very different from the geometric shortest path. We then computed the Φ similarity measure between the relaxed congested network fastest path and the uncongested and geometric paths. The results are shown in Fig. 9. We see clearly that fastest paths in the relaxed congested network resemble geometric shortest paths far more closely than they resemble uncongested fastest paths.

13. Discussion

With respect to the distinction between regimes, our results can be summarized as follows:

- All methods except "similarity to fastest path as a function of extra travel time" (Sec. 11) differentiate between congested and uncongested regimes.
- Only the speed statistics (Sec. 8) and "similarity to fastest path as a function of extra travel time" (Sec. 11) differentiate between non-relaxed and relaxed congestion.

Remember that we started out in the search of a measure of "equilibration" of a traffic system, and we assumed that, in the simulations, the system should be more "equilibrated" after many iterations. For that reason, the second item is more interesting. Yet, it only means that the methods enable an observer to distinguish relaxation in the same system; without looking at different systems it is not possible to decide if any of the methods is capable to generate an empirically "universal" number, i.e., a number which only depends on the actual "relaxedness" and not on network characteristics.

For that reason and in general it is probably clear from our results that more extensive calculations will be necessary to settle the problem. Our analysis shows that there is structure in the data; yet, to which extent the results are robust is an open question. For example, it is unclear how much of the "relaxed versus unrelaxed" results depend on the fact that the unrelaxed situation included grid-lock. Grid-lock in itself is not detrimental to our interpretation since it is simply a sign of a strong "imbalance" (i.e., non-relaxedness); yet, practical use of the results would be limited if most of the quantitative signal would be generated by it.

Our findings also have some implications on the value of real-time congestion information. Under conditions of very light congestion, a motorist's optimal policy is to compute a fastest path using the posted speed limits of network links. Real-time congestion information is of no value when there is very little congestion; this is hardly surprising. But our results indicate that the same is true under conditions of extremely heavy congestion: when a network is *fully* congested (i.e., when all links are essentially slow-moving parking lots), the optimal policy is to compute a *shortest* path simply using the *lengths* of network links. This will also be a *fastest* path, because under extreme congestion the speeds on links will be roughly equal regardless of the functional class of the links. As in the case of very light congestion, a motorist's optimal strategy may be able to safely ignore real-time congestion information. The value of real-time information seems to be highest when the system operates near capacity; the observation that the system is probably least predictable in that regime has been made before.[29]

Another implication may be that, regarding the initial plan set of the iteration process, it may make sense to have a certain amount of drivers choose geometrically shortest paths during peak periods in order to speed up the relaxation process.

14. Computational Considerations

For computing the K-fastest paths for one O-D pair in a time-*in*dependent network, we used a two-phase process as described in Ref. 30 and analyzed by Ref. 23. In the first phase, the complete Dijkstra tree is computed. For a sparse graph (as we have) and using a heap implementation of the Dijkstra algorithm, this has time-complexity $O(E \log_2 N)$ where E is the number of edges and N the number of nodes in the graph. This returns the fastest path. In the second phase, the Dijkstra-tree is used to systematically generate the second-shortest, third-shortest, etc., path. The computational complexity of this second phase seems to be difficult to find.[23,31]

In practice, we needed, for each given traffic situation, approximately 8 h on a 250 MHz UltraSparc CPU in order to calculate the 500 fastest paths between 955 O-D pairs in a network of 6124 links and 2292 nodes.

It would have been desirable to do these computations in a time-dependent network. However, as stated earlier, we were unable to locate a polynomial algorithm that calculates the K-fastest paths in a time-dependent network. We tested a brute-force algorithm of exponential complexity, which first computed the time-dependent fastest path, then started a series of time-dependent fastest path calculations with one link of the original fastest path removed, etc. For our problem, we were able to compute up to $K = 5$ in 2.5 h per O-D pair; the 6th fastest path would, in the average, have needed several days to compute.

15. Summary

In iterated transportation simulations, the result of one iteration is fed into the route planner which adapts some portion of the routes to the congestion encountered in the last iteration. This paper analyzes a series of such iterations which started from a set of route plans where every driver assumed that the network was empty. This starting configuration clearly leads to too much traffic and thus heavy congestion on the freeways. We call this situation "unrelaxed congested." After the iterative process, traffic has moved from the freeways to the arterials, and traffic overall is much faster. We call this situation "relaxed congested." This paper analyzes several measures to distinguish the two different regimes and also an "uncongested" regime for comparison. Link speed variances are much higher in the "unrelaxed congested situation." The relation between fastest and K-fastest paths are similar in both congested regimes. When looking at the similarity to the fastest path as a function of additional travel time, unrelaxed congested alternative paths are much more similar to the fastest path than relaxed congested alternative paths. Or in other words: Under relaxed congested conditions, solutions with roughly equivalent performance are more diverse than *un*relaxed congested conditions. Lastly, we found that, under congestion, the fastest path resembles *geometrically* shortest paths.

In short: Between the criteria we tested, the most suitable to check for relaxation were: (i) velocity distributions, and (ii) similarity to fastest path.

Acknowledgments

We thank Chris Barrett, Steven Eubank, Madhav Marathe, Riko Jacob, and Stephane Lafortune for discussions and helpful suggestions. Special thanks go to Marcus Rickert who, besides giving valuable comments on a draft of this paper, provided the input data for this study and was very helpful at every step of the way. This work was performed at the Los Alamos National Laboratory, operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36.

References

1. D. E. Kaufman, K. E. Wunderlich, and R. L. Smith, "An iterative routing/assignment method for anticipatory real-time route guidance," IVHS Technical Report 91-02, University of Michigan Department of Industrial and Operations Engineering, Ann Arbor MI 48109, May 1991.
2. D. J. Reaume, "A modular implementation of the SAVANT anticipatory route guidance algorithm," Technical Report 95-27, University of Michigan Department of Industrial and Operations Engineering, Ann Arbor MI 48109, November 1995.
3. M. Van Aerde, B. Hellinga, M. Baker, and H. Rakha, "INTEGRATION: An overview of traffic simulation features." *Transportation Research Records*, in press.
4. K. Nagel and C. L. Barrett, *Int. J. Mod. Phys.* **C8**(3), 505 (1997).
5. G. L. Chang, H. S. Mahmassani, and R. Herman, "A macroparticle traffic simulation model to investigate peak-period commuter decision dynamics," *Transportation Research Record* **1005**, pages 107–120, 1985.
6. M. Rickert, *Traffic simulations on distributed memory computers.* PhD thesis, University of Cologne, Cologne, Germany, 1998.
7. Yosef Sheffi, *Urban Transportation Network* (Prentice Hall, 1985) ISBN:0-13-939729-9.
8. M. Patriksson, *The Traffic Assignment Problem: Models and Methods* (VSP, 1994) ISBN: 90-6764-181-2.
9. R. J. Beckman *et al.*, TRANSIMS Dallas/Fort Worth case study report. Los Alamos Unclassified Report (LA-UR), Los Alamos National Laboratory, TSA-Division, Los Alamos, NM 87545, 1997. To be released.
10. K. Nagel and C. L. Barrett, *Int. J. Mod. Phys.* **C8**(3), 505 (1997).
11. R. J. Beckman, Personal communication.
12. M. Rickert and K. Nagel, *Int. J. Mod. Phys.* **C8**(3), 483 (1997).
13. K. Nagel, "Freeway traffic, cellular automata, and some (self-organizing) criticality," eds. R. A. de Groot and J. Nadrchal *Phys. Comp. '92* (World Scientific, 1993), p. 419.
14. K. Nagel and M. Schreckenberg, *J. Phys. I France* **2**, 2221 (1992).
15. M. Rickert and K. Nagel, in preparation.
16. D. E. Kaufman, R. L. Smith, and K. E. Wunderlich, "Dynamic user-equilibrium properties of fixed points in iterative routing/assignment methods," IVHS Technical Report 92-12, University of Michigan IVHS program, Ann Arbor MI 48109, July 1992.
17. D. E. Kaufman, R. L. Smith, and K. E. Wunderlich, "User-equilibrium properties of fixed points in iterative dynamic routing/assignment methods," *Transportation Research C*, forthcoming.
18. A. Garcia and R. L. Smith, "On the convergence of iterative routing-assignment procedures in dynamic traffic networks," Technical Report 95-26, University of Michigan Department of Industrial and Operations Engineering, Ann Arbor MI 48109, November 1995.
19. T. Kelly, *Physica* **A235** (3–4), 407 (1997).

20. T. Hogg and B. A. Huberman, *IEEE Trans. Sys., Man. and Cyber.* **21**(6), 1 (1991).
21. B. Arthur, *American Economic Review* (*Papers and Proceedings*) **84**, 1994.
22. K. Nagel, *Int. J. Mod. Phys.* **C7**(6), 883 (1996).
23. A. Perko, *Networks* **16**, 149 (1986).
24. A. Leliveld, Personal communication.
25. T. Shibuya, T. Ikeda, and H. Imai, *Second Intelligent Tranportation Systems* **4**, 2037 (1995).
26. K. Scott, G. Pabon-Jimenez, and D. Bernstein, "Finding alternatives to the best path," Preprint 970682, The Transportation Research Board, January 1997. From the 76th annual meeting of the TRB.
27. D. Park and L. R. Rilett, "Identifying multiple and reasonable paths in transportation networks: A heuristic approach," Draft dated April 1997. Forthcoming in *Transportation Research Record*.
28. U. Bastolla, H. Frauenkron, E. Gerstner, P. Grassberger, and W. Nadler, "Testing a new Monte Carlo algorithm for protein folding," Preprint cond-mat/9710030 1997.
29. K. Nagel and S. Rasmussen, "Traffic at the edge of chaos," In R. A. Brooks and P. Maes, eds. *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, p. 222–235 (MIT Press, Cambridge, MA, 1994).
30. S. Clarke, A. Krikorian, and J. Rausen, *J. Soc. Indust. Appl. Math.* **11**(4), 1096 (1963).
31. Y. Y. Yen, *Mgt. Sci.* **17**(11), 712 (1971).