Systems Modelling & Simulation: The Sharing of Traffic Knowledge

André Cruz, Edgar Carneiro, and Xavier Fontes

Faculty of Engineering of the University of Porto
Rua Dr. Roberto Frias s/n 4200-465 Porto, Portugal
{andre.ferreira.cruz, edgar.carneiro, xavier.fontes}@fe.up.pt

Abstract. Advanced traveller information systems (ATIS) have seen a recent surge in popularity among urban users. These systems have the ability to considerably increase traffic flow, but are limited by their penetration ratio among the city's population. In this work we simulate a road network where different levels of information percolation are tested as well as their usage policy. The analogy with real life is to assess the extent to which an ATIS can improve total travel time and road utilization, and quantify this usefulness by means of improvement of traffic flow on several metrics.

Keywords: Traffic Modelling · Simulation · Advanced Traveler Information Systems

1 Part I - Project Introduction

1.1 Context of the application domain

In large urban cities the problems generated by traffic can not be disregarded. The stress it causes to the people that get stuck in it and the inefficiencies it brings in time wasted are undesired. In a similar fashion, over-utilization of infrastructures can bring about the need for more monetary investments in maintaining the safety of the roads. In this context, some applications surged as a way to tackle the stress of road users, like *WAZE* [Mobile, 2019]: a GPS application that enables sharing of information among users of the app as a way to better suggest alternatives to desired target locations.

The success of applications like *WAZE* raises curiosity into how strong this information percolation must be in order to actively improve traffic network usage as well as travel time for its users, along with other defined metrics. In this paper we explore how different percolation rates and operating policies of an Advanced Traveler Information Systems (ATIS) can impact the aforementioned metrics, while abstracting the problem as an event-based simulation on top of a graph structure.

1.2 Problem Statement

In this paper, we make an analysis of the possible impacts of having an urban community that makes use of an ATIS and how the influencing variables affect its performance. Several influencing variables can be named, such as the traffic density through a certain route, the percolation rate of an ATIS and the dissemination policies of information.

However, gathering traffic data from a large urban center network or even from several distinct urban centers can prove to be a very exhausting and resource-consuming task. Therefore, this problem will be approached using a modelling and simulation technique using a macro environment approach, where variables that model the road network as a whole are used in lieu of individual characterizing variables.

Even though the main focus of this paper is to study how different percolation rates of an ATIS affect its prediction success rate, there is also the intention to study how the aforementioned remaining factors can also affect the network performance, specially dissemination policies, from peer-to-peer to centralized distribution and proximity based.

1.3 Motivation to tackle the problem

Currently, we still see traffic inefficiencies in road networks. Greedy approaches whose one goal is to optimize the travel time of a given user have no regards for cooperation between users, or the notion of sacrificing the shorter distance to improve the travel time of everyone else involved indirectly. In the end, this leads to users occupying the some traffic links in an unsustainable manner, creating long lines, increased travel times and if that wasn't enough, more pollution from combustion of fossil fuels while standing still.

To the authors, this is more than enough reason to investigate the bare minimum of percolation one would need to have to improve on the current status quo. We will tackle the problem by abstracting some of its details while making sure no insight-discovery opportunities are lost.

1.4 Research/Simulation questions & hypothesis

In this paper we test the hypothesis that increasing levels of information percolation have a positive impact on a traffic grid's flow. Additionally, we propose by intuition that this relation is non-linear, following a curve similar to a logistic function. We'll also assess whether the discovered relation holds for different network topology instances.

In order to test our hypotheses, we use different simulation scenarios and performance metrics. For measuring performance, we use several known traffic performance metrics [Kaparias et al., 2011]: average travel time to points of interest, average origin-destination travel time, average origin-destination travel distance, average travel speed, average fuel consumption, time spent in dense traffic, cost per vehicle miles of travel (related to fuel consumption).

1.5 Expected contributions

From a realistic approach we can see our work supporting our hypothesis and at the same time provide an easy way to simulate this kind of problems. Depending on what kind of insights we derive and how strongly we prove (or for this matter, disprove) our hypothesis it would make sense to present this work to authorities responsible for infrastructure management of our region's traffic networks.

1.6 Aim and goals of the project

Our goal is to find an abstraction as a way to model traffic networks where we can simulate information percolation with respect to different policies of sharing such knowledge. The primary goal will be to test our hypothesis and verify what kind of insights this kind of work can provide, specially from a context of simulation. Our secondary goal will be to develop an infrastructure to simulate these kind of network problems. This last goal might be unfeasible within our time constraints but we believe some groundwork can be laid out, nevertheless our main focus will be at the primary goal.

2 Part II - Problem Formalisation

We will formulate our problem using graph and tree diagrams, as well as diagrams that show the logic behind our simulation processes. These formulations are still subject to change but as of this writing represent the author's desired direction with this work.

2.1 Variables and respective domains

In table 1 we present most of the variables in our problem formulation.

2.2 Assumptions and premises

We will assume the decision of vehicle users can be abstracted to a decision tree similar to the one in figure 1. At the same time we assume we can abstract a traffic network using graphs, similar to the one in figure 2. We start from the premise that some vehicles are inclined to take longer routes if it means a better travel time or less time spent in dense traffic. We also start from the principle that when a vehicle has determined a route to take, that route is not abandoned half-way.

2.3 Constraints and limits

We defined some constraints and limits in tables 2 and 3, respectively.

Table 1. Variables and their domains.

Variable	Description	Domain
U_i	Vehicle i	Obeys a probabilistic decision tree.
TD_i	Departure time of U_i	Non-negative Float
TT_ij	Travel time of U_i on L_j	Non-negative Float
R_i	Route taken by U_i	List of ID's that represent the links taken by U_i from its origin to its destiny.
G	Graph representing the traffic network	Input from the simulation user
L_i	Link i, an edge of G	Non-negative Integer ID
C_i	Capacity of L_i per unit of time	Non-negative Integer
FF_i	Free flow travel time on L_i per unit of time	Non-negative Integer
V_i	Volume of traffic on L_i per unit of time	Non-negative Integer

Table 2. Some constraints for the problem.

Description	Constraint
We can always make it to the destiny from the origin	Network graph $G = (V, E)$ with origin V_Ori and destiny V_Dest, we can follow some edges of E to go from V_Ori to V_Dest

2.4 Cost/utility functions

Our TT_i will be calculated using the following congestion function:

$$S_a(v_a) = t_a\left(1 + 0.15\left(\frac{v_a}{c_a}\right)^4\right)$$

Where $S_a(v_a)$ is the average travel time for a vehicle on a given link a; t_a is the free flow travel time on link a per unit of time; v_a is the volume of traffic on link a per unit of time and c_a is the capacity of link a per unit of time.

Table 3. Limits for the problem.

Description	Limit
Time limit of the simulation	5000
Number of vehicles to generate	300

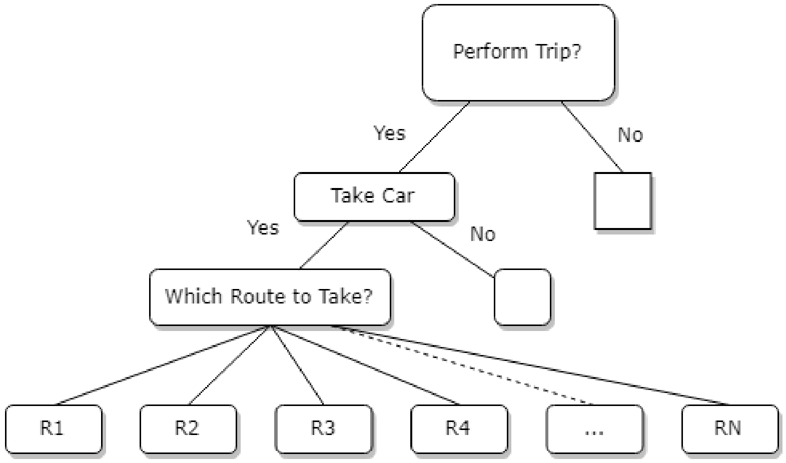

Fig. 1. An example decision tree for the behaviours of users.

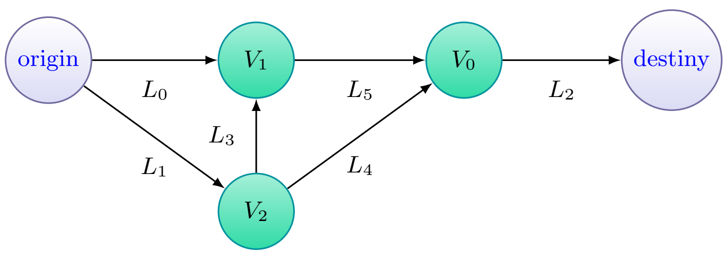

Fig. 2. An example traffic network abstracted as a directed graph.

Our V_i will be updated every time a new vehicle U_i enters L_i and every time a vehicle leaves L_i, in an event-based fashion.

2.5 Optimisation functions

Our goal is to check which rate of information percolation and which policy for percolation works best when improving the following metrics:

- Vehicle travel time – the sum of all the travel times from the links in its route.
- Link congestion – the difference between the maximum capacity of a given link and the volume undergoing the same link in a given time.
- Vehicle travel distance – sum of distances the vehicle traveled when following its desired route.
- Time spent in dense traffic – using a threshold to define what should be considered as dense traffic, calculate the time a vehicle was subject to such a situation.

2.6 Thresholds and intervals

To simplify our first approach at the problem, we defined the following thresholds and intervals in tables 4 and 5, respectively.

Table 4. Thresholds for the problem.

Description	Threshold
Congested link	$\frac{V_i}{C_i} >= 0.9$

Table 5. Intervals for the problem.

Description	Interval
Capacities	$C_i \sim \mathcal{N}(200,\, 20^2)$
Free flow travel times	$FF_i \sim \mathcal{N}(20,\, 5^2)$
Average Speeds	$FF_i \sim \mathcal{N}(90,\, 15^2)$

3 Part III - Methodological Approach

At a high abstraction level, our simulation work will follow the diagram in figure 3. We will use the decision tree to simulate the demand of the network, whose supply will be given by the graph that abstracts the traffic roads.

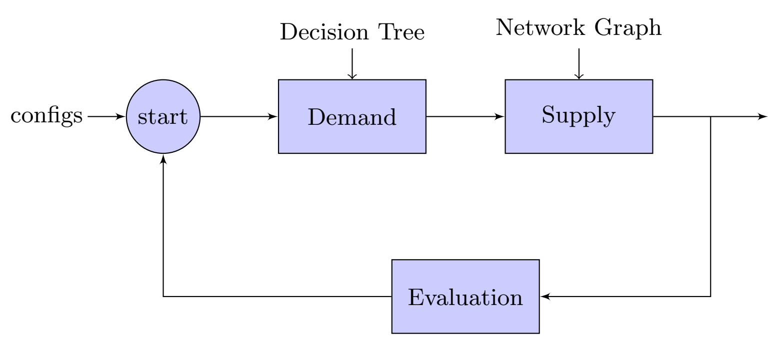

Fig. 3. Our simulation process, at a high abstraction level.

3.1 System Model

To understand how our system will work we can take a look at figure 3 where a high level take is present. The demand comes from a decision tree, of which a simplified example is present in figure 1, while the supply, this is, what the vehicles will consume in a broader sense of the word, is given by a network graph like the one in figure 2.

Modelling Metaphor As one can notice from reading sections above, we are using a Discrete-Event Simulation (DES) strategy. A vehicle is generated at a particular instant in time t and added to the system while other vehicles might leave the system after their routes are completed, all this taking into account that in any particular time delta t_Δ the system is well defined in the sense that we are aware of how many vehicles are in every link of the network.

Input Variables Besides the decision variables mentioned in Section 3.1, we employ the following variables in the simulation: distribution/frequency of accidents, distribution/frequency of weather phenomena, and distribution of traffic intensity.

Output Variables For each agent, we have the following output variables: total trip time, average velocity, and its chosen path, whether it followed the ATIS's suggestion or not.

Metrics Several metrics are calculated for the entire population: average trip time, average velocity, percentage of occupation of links, peak occupation of links.

Performance Measures All aforementioned metrics (Section 3.1) can be interpreted as performance measures.

Benchmarks As benchmark data we will focus on mimicking loads that arise from the inquiry answers as well as search simulation models in the literature as a way to test similar types of loads in our scenarios.

Decision Variables As decision variables go, we can take into account that the percolation rate $Perc$ can be assumed as a decision variable, as well as the policy Pol of information dissemination.

Operating Policies As mentioned in several parts of this document, we are looking to evaluate interaction of different percolation rates of information with a traffic network system. Not only that, this percolation will be made using three distinct policies. First policy is peer-to-peer where each vehicle that adheres to the protocol of information percolation shares information with every other vehicle. Second, we have a policy where every adherent vehicle communicates with a central authority that then informs the vehicles, this policy can be subdivided into different sub-policies, as many as the different strategies we have to disseminate the information (both what information and to which vehicles). Lastly we have a third policy where the vehicles exchange information strictly on a proximity basis, similar to the dissemination pattern of an infectious disease.

Entities of the System Entities participating in the system: multiple travelling agents, and the ATIS. Regarding relationships: depending on the policy, each agent may exchange information with the ATIS or with other agents, and may also choose to obey or ignore incoming advice.

3.2 Conceptual & Logical models

For the development of the models we need to later define how we can *verify*, *validate* and *calibrate* them as best as possible.

We will verify that our models are sound by experimenting with small enough values such that we are able see possible flaws in their design. For validation we will use publicly available traffic data from certain regions and lastly we will calibrate the network graph using OpenStreeMaps (OSM) data and the decision tree using answers from people to an inquiry.

3.3 Coding

From an implementation perspective we will use Python [Rossum, 1995] for its ease of use and amount of available tools. We will also use *matplotlib* [Hunter, 2007] for information plotting and *networkx* [Hagberg et al., 2008] for the design and management of graph structures. We'll also use *SciPy* [Jones et al., 01] due its plentiful utilities for scientific computing. Python is a general-purpose language and we are not foreseeing the use of simulation-specific tools at the moment.

3.4 Data requirements (input)

Our whole work could be made to use synthetic data, like decision trees made by us and artificial graph networks. However, trying to balance direct applicability of our results we are collecting data through inquiries as a way to create a decision tree (or several) that encompasses the preferences of a sample of the population. At the same time our traffic network will be made taking into account real origins and destinies from Portugal, while abstracting as nodes important sections of different sub-trajectories as a way to model different routes that connect the same origin and destiny.

What we are not collecting from these two fronts we will synthesize based on other approaches from the literature and probability distributions. Using synthetic data also allows us to have more control about the variables of the system.

3.5 Data requirements (output)

Due to our choice for implementation we will be able to display several metrics while the simulation is running and even a visualization of the network in question.

One idea we are still pondering is gather as much data from each simulation configuration as possible and explore the data from a knowledge discovery point-of-view. It would be interesting to see if we can design a model that *learns* how different parameters in the configuration relate to each other.

3.6 Simulation scenarios

Our goal is to test our hypothesis using different percolation rates as well as percolation strategies. Despite this, we would also like to experiment different scenarios from the point of view of road infrastructure changes and demand characteristics (both of these can be directly or indirectly influenced by governmental policies). To do this, we will use the inquiry data and publicly available data as our reference scenario baseline and create different scenarios based on the reference one, with incremental changes.

Our simulation plan will be simple and akin to other simulation works. Every scenario will be run a minimum if 100 times and the results will be averaged with its standard deviation calculated to see just how different the runs are from each other. We are considering warm-up strategies at the beginning of the runs but we believe these will not be necessary, because we are evaluating the system along the time instead of first waiting for the system to reach some sort of equilibrium (even though this equilibrium is something we will also measure). The termination criteria for the runs falls together with the duration of the run, as specified in table 3. We are thinking of repeating the scenario until the distribution of the average of the metrics can be approximated with a normal distribution with small enough standard deviation.

4 Planning

We present our expected planning for this work in figure 4.

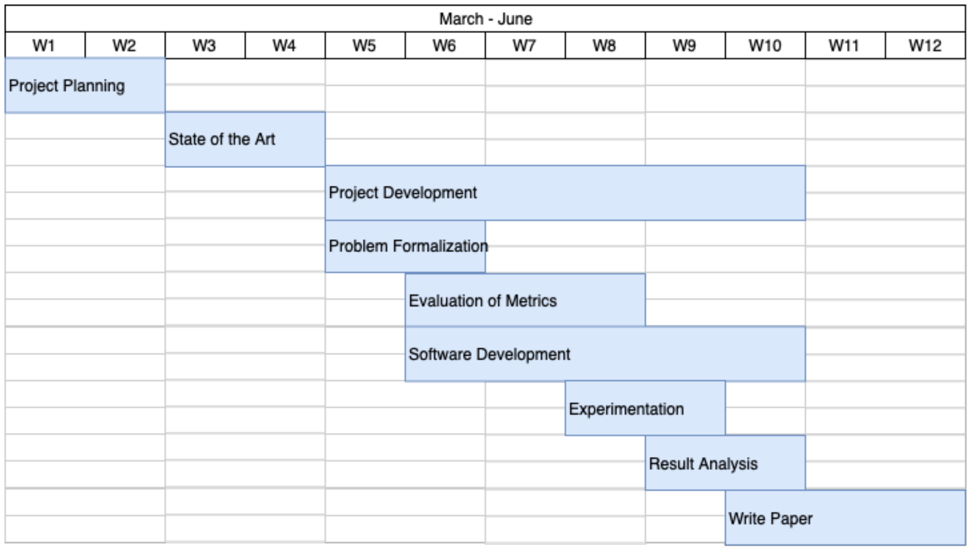

Fig. 4. A Gantt Chart of our expected work development.

References

[Hagberg et al., 2008] Hagberg, A. A., Schult, D. A., and Swart, P. J. (2008). Exploring network structure, dynamics, and function using networkx. In Varoquaux, G., Vaught, T., and Millman, J., editors, *Proceedings of the 7th Python in Science Conference*, pages 11 – 15, Pasadena, CA USA.

[Hunter, 2007] Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing In Science & Engineering*, 9(3):90–95.

[Jones et al., 01] Jones, E., Oliphant, T., Peterson, P., et al. (2001–). SciPy: Open source scientific tools for Python. [Online; accessed ¡today¿].

[Kaparias et al., 2011] Kaparias, I., Bell, M., and Tomassini, M. (2011). Key performance indicators for traffic management and intelligent transport systems. *ISIS*, 14(06).

[Mobile, 2019] Mobile, W. (2009–2019). Waze. http://www.waze.com/.

[Rossum, 1995] Rossum, G. (1995). Python reference manual. Technical report, Amsterdam, The Netherlands, The Netherlands.