Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

MODELING TRAFFIC FLOW USING SIMULATION AND BIG DATA ANALYTICS

Casey N. Bowman

Department of Mathematics
University of North Georgia
1201 Bishop Farms Pkwy
Watkinsville, GA 30677, USA

John A. Miller

Department of Computer Science
University of Georgia
415 Boyd Graduate Studies Research Center
Athens, GA 30605, USA

ABSTRACT

Improving the efficiency, safety, and cost of road systems is an essential social problem that must be solved as the number of drivers, and the size of mass transit systems increase. Methodologies used for the construction of traffic simulations need to be examined in the context of real world big traffic data. This data can be used to create models for vehicle arrivals, turning behavior, and traffic flow. Our work focuses mainly on generating models for these concepts and using them to drive microscopic traffic simulations built upon real world data. Strengths and weaknesses of various simulation optimization techniques are also considered as a methodology issue, since the nature of traffic systems weakens the effectiveness of some optimization techniques.

1 INTRODUCTION

All over the world, populations are rising, and the need to create safe and efficient road systems becomes more and more prescient. Developing nations have long lagged behind industrialized nations in the amount of vehicles in use on a day-to-day basis, but this is changing, as many of these nations begin to truly enter the ranks of the first world (Cervero 2013). In the United States, where large tech corporations are moving toward self-driving vehicles (Levin 2015, Fisher 2013), the need for coordinated and adaptive road systems, i.e. Intelligent Transportation Systems (ITS) is essential (Sussman 2005, Ran, Jin, Boyce, Qiu, and Cheng 2012). However, the sheer volume of vehicles on our roads forces the need for Big Data techniques in acquiring, processing, and utilizing traffic data for the purpose of traffic system optimization (Vlahogianni 2015).

Much work is needed to model the behavior of vehicles in such a system, and we discuss our approach in detail. We have created a method for modeling vehicle arrivals based on real data using time-series/regression techniques (Lippi, Bertini, and Frasconi 2013). We also introduce a new model for controlling the behavior of vehicles within the system, so that it is determined by the data, and is accomplished at simulation time, and not by a pre-simulation route-planner. The behavior of vehicles relative to acceleration and velocity also needs to be modeled, and the flow of vehicles within the traffic system needs to be realistic. A major contribution of our work has been to expand the capabilities of our simulation engine to handle these modeling challenges. The engine itself is a part of ScalaTion, a system for simulation, analytics, and optimization. (Miller, Han, Hybinette, et al. 2010) We expanded the capabilities of our software to handle large-scale simulations, which is needed for simulating and optimizing traffic systems. And we have incorporated mathematical models for car-following behavior, as well as free driving behavior, so as to create a realistic flow of vehicles once they are in the system and traveling along roads. In the near future, with the potential mass proliferation of autonomous vehicles, there will be a tremendous need to use real world traffic data to create high fidelity traffic simulation models. Our work is a first step in an attempt to create such models.

978-1-5090-4486-3/16/$31.00 ©2016 IEEE

The structure of the paper is as follows: Section 2 presents a summary of previous work in the field of traffic simulation and modeling. In section 3, we present the structure of our traffic system, including the models for vehicle arrivals, vehicle behavior and traffic flow, as well as some detail about the structure of our simulation system. In section 4 we briefly investigate optimization techniques which could be used to optimize characteristics of a traffic system, and give some thoughts on the appropriateness of each technique in the context of traffic simulation. Section 5 gives an accounting of our results, and section 6 presents our conclusions, and some of the ideas and avenues of research we plan on pursuing in the future.

2 RELATED WORK

There have been many efforts made at creating models and simulations of traffic systems within the microscopic and macroscopic simulation paradigms. Macroscopic models are generally constrained network flow models that assume continuous streams of traffic flow through nodes. Individual vehicles and their behavior are not considered. This approach requires less computational cost than other approaches, but concedes a lesser amount of detail in the results of the simulation. Tampère, Corthout, Cattrysse, and Immers (2011) applied dynamic network loading to create models of traffic system using simple merge and diverge models to represent the different types of connections roads can make inside a traffic system. These models consider the traffic system as a network of nodes and edges, with vehicle volumes equating to network flow, and the overall goal being to optimize this network flow over the entire system. They focus mainly on deriving generic requirements and constraints that such models must fulfill. Flötteröd and Rohde (2011) add to this work by building a more robust model for representing traffic flow.

Car-following models have been used to model traffic flow and the behavior of vehicles for a long time. One of the earliest models was proposed in Gipps (1981) and computes accelerations based on the differences between successive vehicles' velocities and locations. Gipps updated his models for his work on the MULTSIM traffic simulation system (Gipps 1986). Other car-following models include the Optimal Velocity Model (Bando, Hasebe, Nakayama, Shibata, and Sugiyama 1995), the Generalized Force Model (Helbing and Tilch 1998), the Full Velocity Difference Model (Jiang, Wu, and Zhu 2001), and the Intelligent Driver Model (Treiber, Hennecke, and Helbing 2000). Somewhat recently (Li, Sun, Liu, Zhang, Zhao, Liao, and Tang 2011) formulated a car-following model based on the headways, velocities, and accelerations of multiple preceding vehicles.

The majority of car following models are time-step driven, and update mathematical formulas each time-step, however, there are some models that use event-based methods. In Wiedemann (1991) Wiedemann devised a psychophysical car-following model which combines mathematical concepts with observed psychological phenomena in drivers' reactions to events while driving. Another psychophysical model was presented in Schulze and Fliess (1997), where accelerations are only updated when certain thresholds in distances and speeds with leading vehicles are are crossed.

Microscopic simulation models provide a much greater amount of detail than macroscopic models, since individual vehicles and their behavior are represented with much more complex algorithms to control their movement and decisions. The obvious trade-off is that this requires a much greater computational cost, as simulations will usually contain hundreds or even thousands of vehicles in the system at the same time. The open-source traffic simulation platform SUMO-Simulation of Urban MObility was introduced in 2001 (Behrisch, Bieker, Erdmann, and Krajzewicz 2011) which has been used by many researchers to validate their own models, and to optimize characteristics of traffic systems. Another microscopic simulation platform is VISSIM, which is time step based and was used by researchers at the Georgia Institute of Technology (Hunter, Fujimoto, Suh, and Kim 2006) to create traffic simulations based on real-world data. They generate vehicles using a Poisson counting process to produce random interarrival times.

Many simulation systems, including SUMO and VISSIM, are capable of using Open Street Maps to generate road networks. This feature makes modeling real world traffic systems much easier, and it lends more credibility to the traffic simulations themselves.

Modeling traffic flows and vehicle arrivals is essential to simulating traffic systems in all three paradigms, and there has been much effort put into researching methods to create these models. Lippi, Bertini, and Frasconi (2013) used time series analysis and support vector regression to forecast traffic flows for short-term time periods.

Many different approaches have been used to optimize traffic light timings. Spall and Chin (1997) applied neural networks to the problem using the simultaneous perturbation stochastic approximation (SPSA) algorithm in the context of macroscopic simulation. Ezzat, Farouk, El-Kilany, and Moneim (2014) used the third party simulation software ExtendSim to create, execute, and optimize their traffic models. The software uses an evolutionary approach to optimization. They based their system performance on both queue lengths and vehicle waiting times. Osorio and Chong (2012) used metamodels to optimize simulations of transportation systems. Their metamodel is based on a system of linear and nonlinear equations, which they test for suitability in reducing traffic congestion in a large-scale traffic system.

3 SYSTEM STRUCTURE

In a real world road network, there are a few specific events that occur as vehicles move around the system:

Arrivals In a simulation, only a restricted area of the road system is included and there must be a model that controls the arrival of vehicles to the network.

Traffic Flow Once vehicles are in the network, their behavior as they move along roads should be as realistic as possible.

Turning When vehicles arrive at intersections, they must choose a direction to continue their travel.

In our system we have formulated several essential models that govern the movement, arrival, and decisions of vehicles in the network. A model for the flow of traffic and the behavior of vehicles as they either drive freely or follow other vehicles is given. Also, models for vehicle arrivals and turning behavior, both of which are based on real-world vehicle count data, are presented.

3.1 Simulation Structure in ScalaTion

Our simulation model was built upon the ScalaTion system (Miller, Han, Hybinette, et al. 2010). Figure 1 shows a sample traffic system built in ScalaTion. The system uses several different types of components from the ScalaTion architecture, as well as components that were created specifically for traffic simulation purposes. ScalaTion's simulation system is discrete-event, however, since most car-following models are time-driven, we created a component to regularly schedule the car-following formulas to be updated.

Sources generate vehicles at interarrival times using a predefined random variate. In our system, this random variate is powered by real world data so that vehicle arrivals are realistic. The construction of these random variates is described in a later section.

The concept of a transport is that of a component which moves actors from place to place in a simulation. We advanced our available methods of movement by creating a Road component. A Road is used to move vehicles from one intersection to another with the motion controlled with the formulas outlined above. Roads are lightweight components that function as a guide so that vehicles have an easier time knowing their locations within the system. Functionally, the motion of the vehicles is controlled from the predefined formulas.

Traffic signals are simulated using a Gate component, which can be used to control the flow of traffic by cycling between red and green phases. When a Gate is shut, this means the traffic light is in a red phase, and so the motion formulas produce a deceleration until the velocity of the lead vehicle is reduced to zero. There is no need for following cars to care about the state of the traffic signal, as they are merely reacting to the behavior of the car in front of them. When the Gate opens, signifying a green light, the lead vehicle begins to accelerate and move across the intersection. At this point the vehicle will either turn or go straight, which depends on a turn choice algorithm described later in the paper.

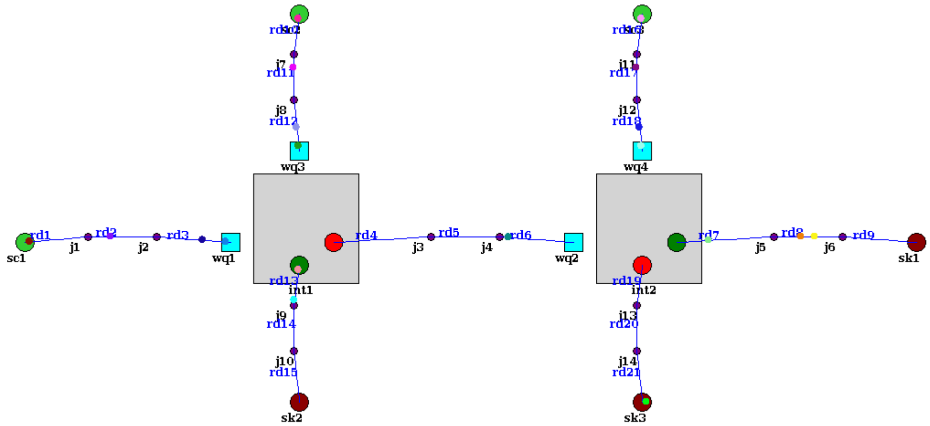


Figure 1: Sample Traffic System.

The cars themselves are modeled using a Vehicle component, which is an specific example of a SimActor component. A Vehicle records its acceleration, velocity, and location along a Road at each time increment, and also holds a reference to the vehicle immediately preceding it. If there is no preceding vehicle, then it is the lead vehicle, and can drive freely, which means it will asymptotically approach the maximum speed. This behavior will only change if a traffic signal turns from green to red in front of it, which will necessitate a deceleration.

Sinks receive vehicles that are exiting the system. They also record the amount of time a vehicle spent in the system, as well as how many vehicles exited through it, which are both important metrics for analyzing, and eventually optimizing, characteristics of a traffic system.

Each of the components can be located in a realistic fashion using GPS coordinates so that distances between landmarks is realistic. Currently our model uses lines or simple curves for Roads, but we plan to implement more complex curvature of roads in the future. The data necessary for such models is harder to come by, but mapping services such as Google Maps, and OpenStreetMaps can often be used to achieve such constructions.

3.2 Traffic Flow and Car-Following Models

A car which is either in the lead, or is far enough behind the immediately preceding car can be thought of as a free-driving vehicle. These vehicles will only be affected by the distance to, and state of, a traffic signal. If the traffic signal is red and the vehicle is close enough, then it must begin to brake. Our current braking model for freely driving vehicles is given by using the basic physics formulas

$$s = s_0 + vt + \frac{1}{2}at^2 \tag{1}$$

$$u = v + at \tag{2}$$

Assuming $s_0 = 0$, and substituting $t = (u - v)/a$ from (2) into (1), rearranging the result yields the formula

$$a_{new} = -v^2/2s \tag{3}$$

where v is the vehicle's current velocity and s is the distance between the vehicle's current location and the traffic signal. If the lead vehicle does not have to brake for a red traffic signal, then we use the following free-driving acceleration model

$$a_{new} = \min\{|\omega a_f + (1-\omega)a|, |\delta(v_f - v)|\} \tag{4}$$

where a_f is the maximum free acceleration, ω is a weight parameter, a is the current acceleration, v_f is the maximum free velocity, v is the current velocity, and δ is a scaling parameter. This formula takes a weighted average of the maximum free acceleration and the current acceleration, which has the effect of gradually increasing the velocity toward the maximum. However, if the current velocity is very close to the maximum velocity, then the new acceleration should be based on this, which will keep the lead car from going faster than the speed limit. The second formula is based on Gipps' basic model (Gipps 1981), however there is no component for the distance between cars since our formula is for freely driving vehicles.

Vehicles which are following another vehicle, and are close enough that they cannot drive freely use the Intelligent Driver Model (Treiber, Hennecke, and Helbing 2000) to update the acceleration, velocity, and location of the vehicles.

3.3 Modeling Arrivals

Traffic systems are extremely complex, and creating a realistic model of vehicle arrivals is essential for simulations of traffic. In order to create such a model, real-world data collection is vital. In the U.S.A. there are a growing number of municipalities and states which are setting up data collection stations on roads and highways. The data being collected is largely in the form of vehicle counts, that is, the data collection station counts all vehicles which pass by it in a set time interval. Quite often, the actual times of the vehicle pass-bys is not kept. Therefore the granularity of the time intervals is very important.

3.3.1 Data Collection and Processing

In order to create valid and realistic simulations of traffic systems real world data must be collected and analyzed. A large number of municipalities around the globe have started collecting data on traffic, which provides researchers a great opportunity to create accurate models of traffic systems. The data used in our project is in the form of vehicle counts collected at multiple sensors along a suburban road in Kenmore, Washington, U.S.A. The sensors provide vehicle counts for every 5-minute interval of the day during a 17-week period starting in September of 2013 and ending in January of 2014. The data set contains information from both directions of the main roadway, broken down by lane, as well as information about most of the side streets.

We believe that the choice of road and intersection in our data reflects a typical suburban traffic system, and that our work can be applied to most any similar system. This also highlights the need for Big Data techniques for modeling more complex traffic systems. A high ratio of drivers will move through progressively busier traffic systems as they go to work each day, which means different traffic models will be needed to represent their entire commute. This will require tremendous amounts of data, and processing of that data, to create simulations with which to work.

3.3.2 Vehicle Generation

There is a very large amount of vehicle count data available from around the world, and this data can be used to create mathematical models of vehicle arrival rates. It seems the majority of previous work in the area of traffic simulation uses vehicle arrival models which are based on simple Poisson counting processes, which are not necessarily even appropriate for modeling the arrivals of vehicles into a traffic system. We chose to model arrivals using a modeling approach which creates vehicles in times that are close to real world data.

There are two approaches which are typically used. The first is to model the interarrival times of vehicles by modeling the time headway of vehicles on a road, which is the defined as the distribution of times between vehicles passing the same geographical location on the road. The second method of vehicle generation is to model the number of cars that should enter the traffic system in any given interval of time.

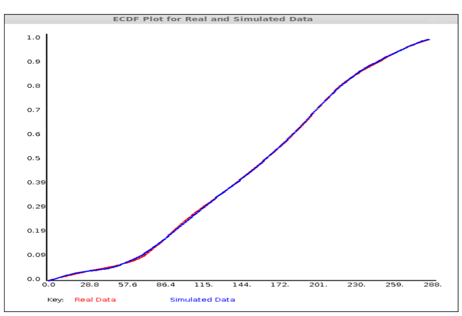
Poisson counting processes, which are based on exponential distributions, are quite often used to model the time headway for vehicles entering a traffic system. Meng and Khoo (2009) suggest that the Poisson process is actually not appropriate for use in modeling vehicle arrivals, and that a better approach is to use self-similar processes which are used to model network traffic that exhibits fractal characteristics. (Leland, Taqqu, Willinger, and Wilson 1994) Other distributions that have been suggested include the Log-Normal distribution (Li, Fa, Rui, Jian-Ming, and Yan 2010), and the Gamma distribution. (Dey and Chandra 2009)

We must account for the ebb and flow of vehicle volume, as most busy intersections display a bimodal distribution of vehicle counts over the course of a normal business day. The two peaks correspond to the usual busy periods of morning and afternoon rush hours when the majority of workers are traveling to and from work, respectively. Leemis (2003) suggests that Non-Homogeneous Poisson Processes (NHPP) can be used to model arrivals in systems that exhibit multiple busy periods.

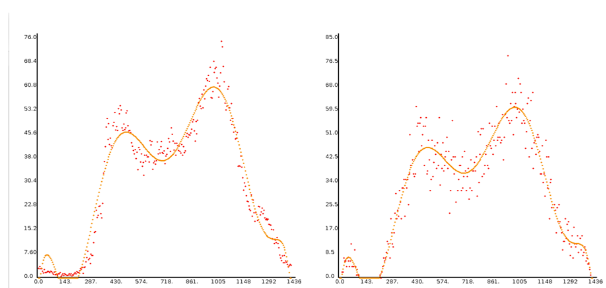
Our approach is to fit each day's vehicle count distribution with a polynomial, which is then used to drive an NHPP that generates interarrival times of vehicles to the system. Another option which we will explore in the future is using Poisson regression to create a fit to the data, and use this instead of a polynomial. It is likely that we would need to use a Non-Homogeneous version of Poisson regression in our work.

Standard Poisson processes use a constant rate parameter λ, which is not appropriate for use in a model which will contain fluctuating arrival rates throughout the life of that model. The NHPP solves this problem by allowing for either a rate function $\lambda(t)$, or a vector of rates. Since we are estimating the counts for a whole hour, we use the latter approach, where we generate a discrete approximation to our polynomial curve and take these function values for our rate vector. This rate vector is used to build a cumulative rate vector, which represents how the arrivals build up over time. This cumulative rate vector is a piecewise-constant approximation to the cumulative intensity function. Using an exponential random variate, arrival times can be generated using linear interpolation. Since we have data from each direction heading into an intersection, it is possible to generate realistic interarrival times that display the varying traffic densities between main roads and side streets.

We validated the NHPP using a Kolmogorov-Smirnov (KS) test, which is a widely used method to compare samples. Based on the results of the test we believe that the vehicle counts are distributed as an NHPP. Figure 2(a) shows that our vehicle arrivals, generated by an NHPP, are very close to the vehicle counts from our real data set. The two samples have a KS-statistic of 0.0113026952 which passes a 95% confidence test. Figure 2(b) shows a comparison of simulated vehicle arrivals with actual vehicle arrivals.



(a) KS Test of simulated and real vehicle counts.



(b) Visual comparison of simulated and real vehicle counts.

Figure 2: Validation of Vehicle Arrivals.

3.4 Turning Behavior of Vehicles

When it comes to route choice, there are several different approaches that have been used. Esser and Schreckenberg (1997) chose to give their vehicles predefined routes that detail the roads they will travel along within the system, usually based on origin-destination tables. Another common technique is to use simple randomizers that choose the next road based on a discrete random variable. A simple discrete random variable, which only uses constant probabilities to generate values, is not really appropriate since the probabilities of turn choices is certainly affected by the time of day, and even which day of the week you are representing.

We feel these methodologies can be improved upon by using real lane vehicle count data. Our approach is to use this lane data to create a probabilistic choice model that vehicles use to decide their route when they reach an intersection. This is fairly straightforward to do when you have access to the individual lane vehicle counts, including turn lanes.

First, we use the lane data to decide the vehicle counts for each of the three choices of turn left, go straight, and turn right. Second, we convert these vehicle counts into a percentage by dividing by the total vehicle counts for all lanes. This gives us an estimate of the percentage of cars turning left, going straight, or turning right during each time interval throughout the day. We then generate polynomials $p_{left}(t)$ and $p_{straight}(t)$ using the percentage values for turning left and going straight. Finally, we create a discrete random variable based on these polynomials and a $U(0,1)$ uniform random variable:

Algorithm 1 Turn Choice Algorithm.

procedure GENERATE TURN CHOICE(t)
    $u \leftarrow U(0,1)$ ▷ Uniform random variable
    if $u < p_{left}(t)$ **then**
        output 0 ▷ Turn left
    else if $u < p_{left}(t) + p_{straight}(t)$ **then**
        output 1 ▷ Go straight
    else
        output 2 ▷ Turn right
    end if
end procedure

This method of creating turn choices is, we believe, a novel approach to the problem. It allows for much more flexibility to represent complex intersections, where the relative percentages of turn decisions can change drastically throughout the day. Not taking this issue into account leads to less accurate simulations which weakens their effectiveness as tools to analyze the real world.

4 SIMULATION OPTIMIZATION

There are many options to choose from when attempting to optimize characteristics of traffic systems. These are stochastic simulations which makes it possible for the same input vector to yield different results from two independent simulation runs. Also, very small changes in the characteristics of traffic systems will likely have no real effect on the outcomes of simulations. So only somewhat large changes in these characteristics are particularly useful. However, it can be difficult to overcome the issue of noise generated by using real world data, which can lead to massive differences in results when common sense would imply there should not be. All of these issues make the selection of optimization technique a difficult one. Below, we briefly discuss strengths and weaknesses of a few optimization techniques.

4.1 Gradient Techniques

Very small changes to the values of traffic system characteristics, such as traffic signal timings and speed limits, will not realistically result in significant changes in the flow of traffic. This makes gradient-based techniques difficult to use since gradients are computed using very small changes in each variable. Finding a proper scale for these perturbations is an optimization problem in its own right, which adds an additional layer of complexity to the optimization.

BFGS The BFGS (Broyden 1970, Fletcher 1970, Goldfarb 1970, Shanno 1970) quasi-Newton method makes use of the gradient and an approximation to the Hessian (matrix of second-order partial derivatives) of a function to iteratively move toward a solution. Line search algorithms are usually employed to decide how large of a step to take in each iteration. One problem with quasi-Newton methods is that the computation of the gradient requires many simulation runs. In fact, if the gradient is being computed using a symmetric difference quotient, there will be two simulation runs for each variable in the input vector. Some of this inefficiency can be removed through parallelization, but not all of it.

SPSA The Simultaneous Perturbation Stochastic Approximation algorithm (Spall 1998) is an effort to remove much of the inefficiency of quasi-Newton methods by simultaneously perturbing all variables at once, resulting in only two simulation runs regardless of the number of variables in the input vector.

4.2 Gradient-Free Optimization

Nelder-Mead Simplex This method (Nelder and Mead 1965) is an unconstrained, derivative-free, direct-search optimization algorithm based on evaluating the objective function at vertices of a simplex. Each iteration typically requires only a few objective function evaluations, and so can be computationally less expensive than many other methods. The goal is to gradually decrease the function values at the vertices of the simplex as it is transformed. Barton and Ivey Jr (1996) show that there are some potential problems with applying the Nelder-Mead algorithm to stochastic objective functions in simulations, mainly due to the fact that changes to the simplex can take place erroneously, based on the stochastic nature of the responses.

Tabu Search Another technique is to apply the Tabu search algorithm (Glover 1989) to an integer domain of input values, as this search guarantees that you will not revisit input vectors already deemed to be sub-optimal. Dengiz and Alabas (2000) used the Tabu search algorithm for simulation optimization and found that it clearly outperformed a random search, giving credibility to its use. According to Fu, Glover, and April (2005), with the cost of running each simulation being so high, the tabulation of the search space greatly improves the time-efficiency of the optimization. With a stochastic objective function, as in simulation, there is always the chance that points previously considered will give different results if you revisit them.

Genetic Algorithms If we restrict the timings of traffic signals to be integers, then the set of all feasible timing combinations is a discrete set. However, for most traffic systems this set is still too large to do an exhaustive search. Genetic algorithms (Holland 1975) have been shown to have good optimization abilities in many applications. Wang (2005) introduced a hybrid approach to simulation optimization using GAs and Neural Networks (Werbos 1974) when there is an unknown form to the objective function.

Genetic Algorithms are well suited for optimizing many characteristics of traffic systems because many of the values involved are discrete. It may not be helpful to think of traffic light timings as a continuous space of possibilities since small changes in timings are unlikely to cause noticeable differences in real world vehicle behavior. Other characteristics such as the number of lanes on a road, and the speed limit of the road are also going to be discrete sets, which are well suited for use with genetic algorithms.

4.3 Response Surface Methodology

Response surface methodology (Box and Wilson 1951) holds some promise for traffic simulation optimization because all response values are computed before the optimization algorithm is applied, so no simulation runs are required during the optimization process itself. The process starts with a predefined lattice of input values and a simulation run is computed with each lattice point. This creates an implied surface of function values which can then fit by more well-defined surface using regression techniques. This surface can be any type of surface, though typically quadratic surfaces are chosen for their ease of use. This more well-defined surface can then be optimized using standard techniques. The process can also be repeated to identify better areas of the surface on which to focus.

This method is attractive because the stochastic nature of the simulations is removed during the optimization phase. Also, the choice of optimization technique is only dependent on the type of surface that has been created. Gradient and non-gradient techniques alike are all possibilities for this optimization. There are some potential drawbacks to RSM though. Response surfaces usually need to be fairly large in scope, which might require a large lattice of points on which to conduct simulation runs. If a multi-stage RSM process is being used, then the number of simulation runs needed to build the response surfaces might be too large.

5 RESULTS

We believe our method of using an NHPP model for generating vehicles closely matches with observed arrival times from real world data. Changes in vehicle congestion as time passes, are handled naturally using our technique. This process can be done ahead of time so that the simulations rely only on the random variate constructed from the data. Our model for deciding turning behavior, which was based on real data, seems to closely align with what can be observed at normal intersections. And our traffic flow models show realistic movement of vehicles as lead cars drive freely, and following vehicles decide new acceleration and velocity changes based on the vehicles in front of them.

The implementation of these three models, namely, a vehicle arrival model, a turning model, and a traffic flow model, have greatly improved the capabilities of our traffic simulation system. While there is still work to do, we believe we have created a system that can accurately simulate many real-world traffic scenarios. The validation plan is to compare traffic counts over an entire day produced by the simulation model with those recorded by the sensors for the road system under study. Counts are affected by the sources, vehicle speeds, the leading car model and following car model, the car turning model, the duration of traffic lights, and the synchronization of multiple lights.

We have also researched the possibility of optimizing the various characteristics of traffic systems through simulation optimization techniques. Several common techniques are examined and critiqued in the context of traffic simulation, and specifically the optimization of traffic light timings. It appears to be a challenge to use gradient-based techniques in this context, as the issues with scaling and noise make it difficult to find good search directions and step sizes.

6 CONCLUSIONS AND FUTURE WORK

In this paper we presented a model for simulating traffic systems in a microscopic simulation paradigm. Realistic models for vehicle arrivals and turning behavior were created which closely match real world data. We also presented formulas for the motion of vehicles within the system. Specifically, we implemented car-following behavior as well as free-driving behavior, which improves the quality of the overall model.

There is still much work to be done to improve our system. Currently there is little facility for including multiple lane roads, and no ability for vehicles to change lanes. There is also no current way in our system to model different modes of traffic as all vehicles are treated as exactly the same. Our future work will include differentiating between different types of vehicles such as large trucks, buses, and possibly even bicycles and pedestrians. There is also a need to improve the mathematical formulas governing the flow of vehicles.

In the future we plan to implement the use of open source mapping information to create traffic simulations of real road systems. Open Street Maps is already being used by many traffic simulation systems, and its inclusion in our system is a definite goal.

We also plan to refine our model so that we can apply it to the many ”what-if” scenarios in the domain of traffic research. For example, special events near a traffic system can greatly increase vehicle counts in that area. Another idea is to study the response of a traffic model to traffic accidents, and perhaps to formulate automated response plans for signal timing that can handle such unforeseen occurrences. We also plan on exploring other types of traffic system structures such as roundabouts, amongst many others.

Lastly, we would like to investigate a growing area of interest in traffic modeling, which is autonomous vehicles. The future of driving is likely to include such vehicles, and traffic simulations which include them will be needed. Traffic systems with such vehicles will require the use of Big Data Analytics techniques in order to deal with the extremely large amounts of real-time data, and to react to changing conditions in an effective and timely manner. Autonomous vehicles provide the opportunity to create truly cooperative traffic systems, where the vehicles actually work together to improve efficiency and provide safe travel to human occupants.

Comprehensive real world traffic planning will require sophisticated software systems in order to maximize the benefits of traffic simulation. The ability to integrate our models with other traffic simulation software, in an effort to broaden the capabilities of the entire system, could greatly aid in the use of traffic simulation for city planning, which is an increasingly important endeavor.

ACKNOWLEDGMENTS

We would like to thank Mustafa Nural for his help in finding and securing vehicle count data. We would also like to thank Michael Cotterell for his help in developing coroutine software within ScalaTion.

REFERENCES

Bando, M., K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama. 1995. "Dynamical Model of Traffic Congestion and Numerical Simulation". *Physical Review E* 51 (2): 1035.

Barton, R. R., and J. S. Ivey Jr. 1996. "Nelder-Mead Simplex Modifications for Simulation Optimization". *Management Science* 42 (7): 954–973.

Behrisch, M., L. Bieker, J. Erdmann, and D. Krajzewicz. 2011. "SUMO–Simulation of Urban MObility". In *The Third International Conference on Advances in System Simulation (SIMUL 2011), Barcelona, Spain*.

Box, G. E., and K. Wilson. 1951. "On the Experimental Attainment of Optimum Conditions". *Journal of the Royal Statistical Society. Series B (Methodological)* 13 (1): 1–45.

Broyden, C. G. 1970. "The Convergence of a Class of Double-Rank Minimization Algorithms 1. General Considerations". *IMA Journal of Applied Mathematics* 6 (1): 76–90.

Cervero, R. B. 2013. "Linking Urban Transport and Land Use in Developing Countries". *Journal of Transport and Land Use* 6 (1): 7–24.

Dengiz, B., and C. Alabas. 2000. "Simulation Optimization Using Tabu Search". In *Proceedings of the 2000 Winter Simulation Conference*, edited by J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, 805–810. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Dey, P. P., and S. Chandra. 2009. "Desired Time Gap and Time Headway in Steady-State Car-Following on Two-Lane Roads". *Journal of Transportation Engineering*.

Esser, J., and M. Schreckenberg. 1997. "Microscopic Simulation of Urban Traffic Based on Cellular Automata". *International Journal of Modern Physics C* 8 (05): 1025–1036.

Ezzat, A. A., H. A. Farouk, K. S. El-Kilany, and A. F. A. Moneim. 2014. "Optimization Using Simulation of Traffic Light Signal Timings". In *Proceedings of the 2014 International Conference on Industrial Engineering and Operations Management Bali, Indonesia*.

Fisher, Adam 2013. "Google's Road Map to Global Domination". http://www.nytimes.com/2013/12/15/magazine/googles-plan-for-global-domination-dont-ask-why-ask-where.html?_r=0.

Fletcher, R. 1970. "A New Approach to Variable Metric Algorithms". *The Computer Journal* 13 (3): 317–322.

Flötteröd, G., and J. Rohde. 2011. "Operational Macroscopic Modeling of Complex Urban Road Intersections". *Transportation Research Part B: Methodological* 45 (6): 903–922.

Fu, M. C., F. W. Glover, and J. April. 2005. "Simulation Optimization: A Review, New Developments, and Applications". In *Proceedings of the 2005 Winter Simulation Conference*, edited by M. E. Kuhl, N. M.

Steiger, F. B. Armstrong, and J. A. Joines, 83–95. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Gipps, P. 1986. "Multsim: a Model for Simulating Vehicular Traffic on Multi-Lane Arterial Roads". *Mathematics and Computers in Simulation* 28 (4): 291–295.

Gipps, P. G. 1981. "A Behavioural Car-Following Model for Computer Simulation". *Transportation Research Part B: Methodological* 15 (2): 105–111.

Glover, F. 1989. "Tabu Search-Part I". *ORSA Journal on Computing* 1 (3): 190–206.

Goldfarb, D. 1970. "A Family of Variable-Metric Methods Derived by Variational Means". *Mathematics of Computation* 24 (109): 23–26.

Helbing, D., and B. Tilch. 1998. "Generalized Force Model of Traffic Dynamics". *Physical Review E* 58 (1): 133.

Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems: an Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.* U Michigan Press.

Hunter, M. P., R. M. Fujimoto, W. Suh, and H. K. Kim. 2006. "An Investigation of Real-Time Dynamic Data Driven Transportation Simulation". In *Proceedings of the 2006 Winter Simulation Conference*, edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 1414–1421. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Jiang, R., Q. Wu, and Z. Zhu. 2001. "Full Velocity Difference Model for a Car-Following Theory". *Physical Review E* 64 (1): 017101.

Leemis, L. 2003. "Estimating and Simulating Nonhomogeneous Poisson Processes". In *International Conference in Reliability and Survival Analysis (ICRSA), The University of South Carolina*.

Leland, W. E., M. S. Taqqu, W. Willinger, and D. V. Wilson. 1994. "On the Self-Similar Nature of Ethernet Traffic (Extended Version)". *IEEE/ACM Transactions on Networking* 2 (1): 1–15.

Levin, Doron 2015. "Googles Self-Driving Cars are Coming Sooner Than You Think.". http://fortune.com/2015/10/02/google-self-driving-cars-soon/.

Li, L., W. Fa, J. Rui, H. Jian-Ming, and J. Yan. 2010. "A New Car-Following Model Yielding Log-Normal Type Headways Distributions". *Chinese Physics B* 19 (2): 020513.

Li, Y., D. Sun, W. Liu, M. Zhang, M. Zhao, X. Liao, and L. Tang. 2011. "Modeling and Simulation for Microscopic Traffic Flow Based on Multiple Headway, Velocity and Acceleration Difference". *Nonlinear Dynamics* 66 (1-2): 15–28.

Lippi, M., M. Bertini, and P. Frasconi. 2013. "Short-Term Traffic Flow Forecasting: An Experimental Comparison of Time-Series Analysis and Supervised Learning". *IEEE Transactions on Intelligent Transportation Systems* 14 (2): 871–882.

Meng, Q., and H. L. Khoo. 2009. "Self-Similar Characteristics of Vehicle Arrival Pattern on Highways". *Journal of Transportation Engineering* 135 (11): 864–872.

Miller, J., J. Han, M. Hybinette et al. 2010. "Using Domain Specific language for Modeling and Simulation: Scalation as a Case Study". In *Proceedings of the 2010 Winter Simulation Conference*, edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yucesan, 741–752. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Nelder, J. A., and R. Mead. 1965. "A Simplex Method for Function Minimization". *The Computer Journal* 7 (4): 308–313.

Osorio, C., and L. Chong. 2012. "An Efficient Simulation-Based Optimization Algorithm for Large-Scale Transportation Problems". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, 423. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ran, B., P. J. Jin, D. Boyce, T. Z. Qiu, and Y. Cheng. 2012. "Perspectives on Future Transportation Research: Impact of Intelligent Transportation System Technologies on Next-Generation Transportation Modeling". *Journal of Intelligent Transportation Systems* 16 (4): 226–242.

Schulze, T., and T. Fliess. 1997. "Urban Traffic Simulation with Psycho-Physical Vehicle-Following Models". In *Proceedings of the 1997 Winter Simulation Conference*, edited by S. Andradottir, K. J. Healy, D. H. Withers, and B. L. Nelson, 1222–1229. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Shanno, D. F. 1970. "Conditioning of Quasi-Newton Methods for Function Minimization". *Mathematics of Computation* 24 (111): 647–656.

Spall, J. C. 1998. "An Overview of the Simultaneous Perturbation Method for Efficient Optimization". *Johns Hopkins APL Technical Digest* 19 (4): 482–492.

Spall, J. C., and D. C. Chin. 1997. "Traffic-Responsive Signal Timing for System-Wide Traffic Control". *Transportation Research Part C: Emerging Technologies* 5 (3): 153–163.

Sussman, J. M. 2005. *Perspectives on Intelligent Transportation Systems (ITS)*. New York : Springer Science+Business Media, 2005.

Tampère, C. M., R. Corthout, D. Cattrysse, and L. H. Immers. 2011. "A Generic Class of First Order Node Models for Dynamic Macroscopic Simulation of Traffic Flows". *Transportation Research Part B: Methodological* 45 (1): 289–309.

Treiber, M., A. Hennecke, and D. Helbing. 2000. "Congested Traffic States in Empirical Observations and Microscopic Simulations". *Physical Review E* 62 (2): 1805.

Vlahogianni, E. I. 2015. "Computational Intelligence and Optimization for Transportation Big Data: Challenges and Opportunities". In *Engineering and Applied Sciences Optimization*, 107–128. Springer.

Wang, L. 2005. "A Hybrid Genetic Algorithm–Neural Network Strategy for Simulation Optimization". *Applied Mathematics and Computation* 170 (2): 1329–1343.

Werbos, P. J. 1974. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Ph. D. thesis, Harvard Univ., Cambridge, MA. Ph.D. dissertation.

Wiedemann, R. 1991. "Modelling of RTI-Elements on Multi-Lane Roads". In *Drive Conference (1991: Brussels, Belgium)*, Volume 2.

AUTHOR BIOGRAPHIES

CASEY N. BOWMAN is an Assistant Professor of Mathematics and Computer Science at the University of North Georgia, and a Ph.D. student in the Department of Computer Science at the University of Georgia. His research interests include Modeling and Simulation, Simulation Optimization, Big Data Analytics, and Traffic Modeling. Mr. Bowman received a B.S. in Mathematics in 2003, and an M.A. in Mathematics in 2005, both from the University of Georgia. His email address is casey.bowman@ung.edu.

JOHN A. MILLER is a Professor of Computer Science at the University of Georgia. He has been the Graduate Coordinator for the department for 9 years and is currently the Associate Head. His research interests include Modeling & Simulation, Web Services/Workflow, Database Systems, Big Data Analytics and Bioinformatics. Dr. Miller received a B.S. in Applied Mathematics from Northwestern University in 1980 and an M.S. and Ph.D. in Information and Computer Science from the Georgia Institute of Technology in 1982 and 1986, respectively. As part of his co-operative undergraduate education program, he worked as a Software Developer at the Princeton Plasma Physics Laboratory. In his areas of interest, Dr. Miller has authored of over 180 research papers. He is an Associate Editor for the ACM Transactions on Modeling and Computer Simulation, SIMULATION: Transactions of the Society for Modeling and Simulation International, and previously IEEE Transactions on Systems, Man and Cybernetics as well as an Editorial Board Member for the Journal of Simulation and International Journal of Simulation and Process Modelling. His email address is jam@cs.uga.edu.