*Proceedings of the 2015 Winter Simulation Conference*
*L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds.*

# OPTIMIZING AN AGENT-BASED TRAFFIC EVACUATION MODEL USING GENETIC ALGORITHMS

Matthew Durak
Nicholas Durak
Erik D. Goodman

BEACON Center for the Study of Evolution in Action
Michigan State University
East Lansing, MI 48824, USA

Robert Till

Department of Security, Fire and Emergency Management
John Jay College, City University of New York
New York, NY, USA

## ABSTRACT

Computer simulations are commonly used to model emergencies and discover useful evacuation strategies. The top-down conceptual models typically used for such simulations do not account for differences in individual behavior and how they affect other individuals. To create a more realistic model, this study uses Agent-Based Modeling (ABM) to simulate the evacuation of an urban population in case of a chlorine spill. Since the agents (each a car and driver) in this model do not behave uniformly, and the initial traffic and spill locations are randomized, optimizing traffic lights is challenging. A commercial evolutionary optimizer controls execution of the simulator, seeking to optimize the control of traffic lights in order to minimize deaths and injuries. ABM for a traffic evacuation could prove useful in the real world, when the threat is at a known location such as a power plant or a specific railway segment.

## 1 INTRODUCTION

Emergencies requiring local evacuation can sometimes be mitigated if a rapid, pre-planned response is available. Computer simulation is an important tool for preparing for such emergencies. Existing methods have typically used top-down conceptual models. For example, airplane evacuation is modeled as fluid flow from a container, using equations borrowed from physics (Thompson and Marchant 1995). This ignores individual behaviors that affect such an evacuation. A useful model should consider the non-random interactions between passengers and that not all behaviors are consistent across individuals.

Agent-Based Modeling (ABM) has been developed to deal with such problems (North and Macal 2007), and takes a bottom-up approach, with agents interacting with each other and their environment. They typically behave according to set rules with randomly varied parameters to represent the diversity of behaviors. The overall behavior of the system becomes an emergent property of the individual behaviors.

ABM has found frequent application in such situations as subway station or stadium evacuation in a fire (Batty 2005). One issue it raises is how emergency responders can discover in advance the optimum approach to handling the emergency, and design, plan and train accordingly. Solving for the best control strategy is the task presented to the designer of the system. However, these parameters are difficult to optimize because they may all be interdependent. The large number of parameters creates a search space that cannot possibly be exhaustively tested, as complex ABM simulations may require long run times. Systematic design of experiments (DOE) can reduce the number of experiments needed to explore the interactions among parameters, but model nonlinearities make it important to do the DOE near an optimal operating point to avoid misleading results. ABM is a good tool to use with a stochastic search algorithm

978-1-4673-9743-8/15/$31.00 ©2015 IEEE 

that begins by randomly looking at the search space, then progressively investigating the "better" areas more closely. The solution need not be globally optimal—a "very good" solution is fine in the real world.

## 2 PROBLEM

This study looks at a chlorine spill in an urban area. It models the drivers in that area and aims to optimize the timing of traffic lights in order to evacuate those drivers with minimum casualties. The intention is not to provide an online algorithm that optimizes the evacuation in real-time, but to simulate various emergency scenarios ahead of time, so that corresponding traffic light timings can be planned. The measure of a good design is minimizing the number of casualties due to drivers remaining in the hazardous area for too long. The goal is to demonstrate the practical application of evolutionary computation to optimize an ABM of a critical incident as an illustration of such an approach.

This type of evacuation scenario has been studied in the past, but not by systematic optimization of traffic signal controls using an ABM. Chen et al. (2007) summarize early work. These studies did not include the effects of the hazard on drivers and compared several pre-selected strategies rather than optimizing the strategy. Chen (2011) provides a NetLogo-based model with spatially-aware and time-aware agents in an evacuation scenario, but does not optimize the evacuation strategy. Chen does consider factors that are not considered in the current study, including the influx of emergency vehicles, roadside parking, and public transportation. Kimm and Maassen (2011) optimize cell size and evacuation strategy using a cell-level simulation that differs in not modeling the evacuation hazard (e.g., chlorine).

## 3 METHODS

### 3.1 Model

The model is created in NetLogo (Wilensky 1999). It includes a "world" modeling a portion of a city represented as a grid of roads meeting at traffic-light-controlled intersections, the timing of which is to be optimized. The model simulates the emergency of a chlorine spill and the resulting gas plume, as described initially in (Till, Durak, and Goodman 2012). To increase realism, parking garages have been included. A second component is the system for notifying drivers about the emergency and instructing them to evacuate the city. Finally, the agents themselves—in this case, the people driving cars—are modeled in some depth. Code and documentation of the model are posted at (https://github.com/njd911/Agent-Based-Traffic-Evacuation-Model).

Unfortunately, there are no concrete data on which to base much of this modeling, so many parameters have been chosen based on conceptual models and data from similar situations or on grounds of reasonability. This model, then, should not be expected to produce precise predictions, but rather to represent scenarios that may be more realistic than those produced using vastly oversimplified models. Fortunately, these parameters are configurable and can be adjusted using data that may become available. Emergency responders using an ABM system can tune the parameters as they see fit.

#### 3.1.1 The World (An Urban Area)

The urban area modeled has a grid of eight east-west "streets" by eight north-south "avenues" and covers a 1.333 by 0.833 mile area. The roads are unevenly spaced but perpendicular. They consist of a variable number of lanes with both one-way and two-way roads. In NetLogo, they are represented as "patches," which are square units spanning all roadways. The patches themselves are squares measuring 22 by 22 feet (but lanes are also significant, and may be less than 22 feet wide in actuality). There are two main "evacuation roads", which are the six-lane roads running east-west. Additionally, there are two four-


Figure 1: Simulation at 1800 ticks (15 minutes). The arrows denote the number of lanes in each direction. Black numbers show the contents of parking garages. Tiny colored numbers atop cars represent their AEGL states. A blue cone illustrates the visible region of a sample sign.

lane roads running north-south. The remaining roads are two-lane side streets (Figure 1).

Simulation begins with 5000 cars. Before the emergency, the world is a toroid—cars drive off the map at one side and re-enter at the opposite side, simulating steady-state traffic. Once the evacuation starts, that stops. Drivers learning about the evacuation are told to evacuate east or west because the plume is moving north. Those leaving to the east or west are counted as having evacuated. It is also assumed that roads are blocked so cars no longer enter from the east or west. However, cars that leave via the north or south turn around after a time and reenter the map to continue their evacuation.

### 3.1.2 Intersections

There are 64 intersections on the map (Figure 1). In the simulation, these intersections act as controllers of the traffic lights. During non-emergency operation, a uniform light timing is used to get a realistic traffic distribution from a random initial distribution. Then during the evacuation phase, the lights change to use the new evacuation timing. The cycle transitions from green in the east-west direction, to amber, to red in all directions, to green in the north-south direction, to amber, to red in all directions again. As discussed below, drivers may not follow traffic laws in a time of emergency. The times for the lights to be all red are fixed at 1 second (or 2 simulation steps). The amber phase is fixed at 5 seconds in this experiment. The green phase in either direction is configurable by the optimizer. The traffic cycle timer at each intersection has a "base" temporal offset from time 0. This allows the lights to be timed such that a car can drive east without hitting a red light. Offsets are proportional to the distances between lights.

### 3.1.3 The Plume

Implicit in this study is the need to anticipate accidents/events at a number of point locations or small bounded regions and have pre-designed evacuation strategies to deal with such events. In order to make the system as manageable as possible, each event should require only a rough localization (for example, for a rail line, a spill within a region about 200m long). The authors determined experimentally that increasing the size of the span to three times that size had an adverse effect on the resulting strategy, as moving the precise location of the event around at random within such an expanded span required quite different scenarios for optimizing the evacuation strategies within the simulated area.

The emergency situation is a chlorine gas plume caused by a tanker spill. The spill occurs two blocks south of the visible map, anywhere (uniform random) within a 27-patch span (about 200m) and each "run" of the simulator uses a different randomly chosen location for the spill within that span. Figure 2 shows the plume progressing on the map. Chlorine concentrations come from the ALOHA software, version 5.4.1.2 (Office of Emergency Management 2009), using the "threat at a point" feature, with a 3mph wind blowing north. Chlorine concentrations are estimated at approximately every intersection. ALOHA provides 36 points at 5-minute intervals for 1 hour, interpolated here to 30-second intervals.

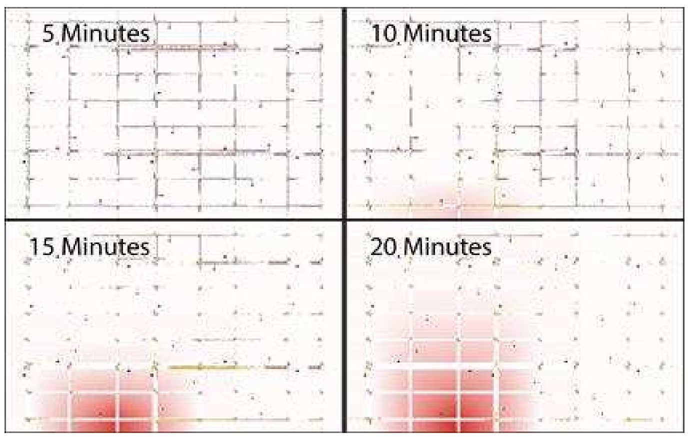


Figure 2: Plume progression in five-minute intervals. The red intensity represents chlorine concentration.

The simulation stores the chlorine concentration at each patch, using an air exchange model to generate driver exposures in vehicles. The model uses the Acute Exposure Guideline Levels (AEGL), where AEGL-1 is the level at which 50% of people can detect the gas, AEGL-2 is the level at which 50% of people are affected by the gas and in some way disabled, and AEGL-3 is the point at which the gas is fatal to 50% of people (National Research Council 2004). Driver exposure is modeled assuming 7 air exchanges per hour inside the vehicle (Ott, Kleptis and Switzer 2007). This results in an internal concentration $C(t+1) = C(t) + R * \left(C_{external}(t) - C(t)\right)$, where $R = \frac{7}{7200}\frac{air\ exchanges}{tick}$ . This concentration is used to measure the human exposure to the chemical in the AEGL model: $E(t+1) = E(t) + [0.5 * C^{1.5}(t)]$ (National Research Council 2004). Each driver is randomly assigned a "susceptibility" in [0.1, 0.9] in a truncated normal distribution with mean 0.5 and standard deviation 0.2, determining how prone the driver is to reaching each AEGL level. This distribution of AEGL values among drivers makes the AEGL the point at which agents with susceptibility > 0.5 are affected. For example, a driver is in AEGL-2 when $E > (AEGL2) * 2s$, where s is susceptibility.

The exposure model is used to affect driver behavior, with different consequences for AEGL-1, AEGL-2, and AEGL-3. At AEGL-1—detection—drivers become aware of the gas. They are more likely to turn on their radios and to learn more quickly about the instructions to evacuate. Pedestrian exposure level is calculated using the plume concentration as the concentration in the exposure formula above.

Pedestrians will thus reach AEGL-2 earlier and show outward symptoms that can affect driver behavior. Pedestrian behavior and mortality are not tracked, except for this single purpose. Drivers who see people in the street exhibiting symptoms of chlorine gas poisoning will become more aware of the emergency. Drivers who reach AEGL-2 will stop their vehicles. These drivers could potentially survive if their exposures never rise to AEGL-3, but they will continue to be exposed to gas in their stopped positions and their vehicles can block traffic. The final level of AEGL-3 only differs in the simulation in that the person is considered a fatality and thus the impact of that particular design is higher due to the loss of life.

### 3.1.4 Emergency Notification

It is assumed that the city has some way of sending a signal to people nearby via sirens and that it will take emergency responders some time to find the chlorine leak and notify people to evacuate. This evacuation phase in the model begins 75 seconds after the chlorine starts leaking, at which point the sirens begin to alert drivers, likely implying some level of automation in the warning system. The alert as well as the 12 emergency signs shown in Figure 1 should make people more likely to turn on their car radios in order to receive further instructions. These instructions could be broadcast over all stations, so the model assumes that anyone with a radio on is receiving the evacuation message. Initially, 40% of drivers have the radios turned on. At each 0.5-second step in the emergency phase, we assume a 0.1% chance that drivers will turn their radios on or off. However, drivers who have read an electronic sign telling them to tune in, or drivers with an awareness level beyond 50% (see Section 3.1.7) will not turn off their radios.

### 3.1.5 Car-Following and Intersection Behavior

The most basic agent behavior is a simple car-following model. This includes all of the rules that agents use for how fast to go, when to accelerate or brake in response to other cars or intersections ahead to avoid accidents, and the physics that affect car velocity based on acceleration. The basic concept is that the roads have a speed limit of 30 miles per hour. However, some people drive faster or slower than the speed limit so each car has its own maximum speed normally distributed around 33mph (although a high sense of urgency may increase a driver's maximum speed). At each time step, a driver's personal speed limit varies by up to +/- 1.5 mph from his nominal limit. This is to model the fact that without cruise control, there will be some variation in driving speed even when a person intends to go a specific speed. Each car analyzes the car in front of it at every time step in order to meet its goal of driving at its speed limit so long as it has a safe stopping distance to other cars in its path.

The other consideration for the speed of the car is the traffic light ahead. Assuming the driver is not already braking due to other vehicles ahead, he or she will stop before a red or amber light provided it is within braking distance. Left-turning drivers wait for a break in oncoming traffic (if any) before turning. Drivers also follow the "don't block the box" rule. This means that cars will not enter the intersection when they cannot safely exit the intersection without blocking traffic. This prevents gridlocks and is a law in some areas, such as New York City (*N.Y. VAT. LAW § 1175*). In certain situations where the driver has a high sense of urgency, he or she will break these rules as described in the psychology section.

### 3.1.6 Parking Garages

The parking garages simulated (Figure 1) add a source of traffic as people begin leaving these garages. There are 15 garages in the study area, with 50-200 cars each for a total of 1500 cars. It is assumed that not everyone with a car in a garage will (or should) choose to evacuate. The concentration of and exposure to chlorine in the garage is calculated as it is for vehicles on the map.

### 3.1.7 Goal Seeking

Each agent uses a goal-seeking model in order to navigate the world. Agents can have different types of goals depending on their states and the state of the model. During the pre-evacuation phase, all drivers are pursuing normal goals, which are either to drive to an intersection (local) or drive north, south, east, or west off the visible map (remote). Cars with remote goals are assigned the direction (east, west, north, or south) with equal probabilities. To keep things simple, cars do not drive into parking garages, which are kept static until the emergency phase. Each car has a 50% chance of being assigned a local goal and 50% chance of being assigned a remote goal. Upon reaching a destination, cars are given a new random goal.

During the evacuation phase, not all drivers seek to evacuate immediately, even if they know the situation. Sufficiently aware drivers (at least 50% awareness) have a chance of choosing to evacuate at each step, which is highest the moment they reach their goal. The probability of choosing to evacuate when selecting a new goal goes up as the driver becomes more aware. By default all drivers evacuate to the east, because the plume is west of the center of the map. However, drivers knowledgeable about the city and aware of the plume's location may know that they are heading into the plume and may instead drive to evacuate to the west. Evacuating cars will favor multilane roads when possible.

### 3.1.8 Other Factors – Agent Perception and Interpretation of Emergency

The final part of the model involves other motivating factors for the agents, and is, by its nature, the most speculative and least validated component. However, the benefit of capturing diverse behaviors of drivers predicated on different psychological profiles and different levels of experience and situational information seems clear, even if the models contain inaccuracies. A conceptual model for human behavior in fires by Kuligowski (2009) identifies many influential factors in the behavioral process in emergencies. Three variables were distilled from her extensive list to describe agent states during the emergency, each in the range [0, 1]. As described in Canter (Canter 1980) "panic" is not used to describe this behavior.

The first variable is *knowledge*, representing driver knowledge of the city. The second aspect is *awareness*, measuring driver awareness of the emergency situation. In general, awareness increases over time due to the radio messages and instructional signs the driver hears/sees. The final variable is *sense of urgency*, capturing the amount of stress felt due to the emergency. People with a higher sense of urgency are modeled as driving more aggressively, which can assist in their evacuation. However, as urgency increases, the "tunnel vision" effect can cause drivers to make risky decisions (Hamdar 2007, 13-19).

Knowledge is initially distributed uniform randomly and changes relatively slowly from the initial distribution. It increases by 0.01 at every intersection, so that a car that is driven around the city for a long time will be more knowledgeable about the geography of the area. In this model, knowledge is used only in some decisions governing which direction to drive when evacuating.

Awareness is an important factor in causing a driver to decide to evacuate. It is initially 0, and increases due to several factors. A driver in AEGL-1 has a 5% chance of becoming more aware at every time step, and a 10% chance upon seeing a pedestrian being affected by AEGL-2 at a given step. However, a driver can only become up to 50% aware without receiving more data over the radio. Drivers listening to the radio while in AEGL-1 or witnessing the effects of chlorine on pedestrians will become aware of the situation twice as fast as those with just the radio on. Notification of emergencies via the cellular telephone network have not been modeled here, although such systems are proliferating rapidly.

Sense of urgency captures how stressed an individual is in response to a threat. This value is initially set to 0 and increases over time due to awareness, AEGL effects, and getting stuck in traffic (Hamdar 2007). A driver's sense of urgency decays without any of these stressors. Each agent is affected differently by sense of urgency. Each has a threshold of urgency at which the driver will experience "tunnel vision" and not pay attention to certain things (Silverman et al. 2005). Tunnel vision may lead to an accident (a casualty that blocks traffic) in certain situations. For example, a driver with a sense of urgency above his/her "law compliance threshold," who has been waiting at a red light for more than 5 seconds, will run the red light if traffic is clear. However, if also above the tunnel vision threshold, the driver may, probabilistically, make an error in checking whether traffic is clear and thus may cause an accident by running a red light. An agent with a sense of urgency above its law compliance threshold may also ignore the "don't block the box" rules and enter an intersection even without room to exit properly.

A driver who is stuck because drivers ahead are dead or incapacitated may decide to pass in an opposing lane if traffic is clear—the higher the sense of urgency, the smaller the wait time before deciding to do so. If a car is stopped in traffic for more than 1 minute and the driver's sense of urgency is high enough, then the driver may try to make a U-turn. An effect of this nature seems important to capture, however crudely, because it appears unlikely that drivers would remain stopped in a dangerous situation if they have a way to get out. Drivers may exceed the speed limit by increasing their personal speed limit by a multiplier of their sense of urgency. The multiplier has a mean of 0.11 and standard deviation of 0.05 bounded to [0, 0.3], which is used to augment personal speed limits. Thus, drivers with a very high sense of urgency may speed by up to 30% faster than their normal personal speed limits.

Drivers also use the model of knowledge, awareness, and sense of urgency when choosing evacuation routes during an evacuation, or to drive in the opposite direction if it means avoiding the gas plume. If a car is evacuating but driving away from its goal, it may continue to its goal if it means it will avoid the plume. This happens if the driver has awareness of at least 0.5 and combined awareness and knowledge greater than 1. The driver will not do this after reaching the tunnel vision threshold. If the driver's knowledge is greater than 0.25 or their knowledge combined with awareness is greater than 1, and they don't have tunnel vision, they will attempt to drive to an evacuation route (see (Silverman et al. 2005)).

### 3.2 Evolutionary Computation

Since this agent-based traffic model has a search space that cannot be exhaustively tested, a stochastic search algorithm is used to seek high-performance designs. This stochastic optimization is implemented using the MO-SHERPA Multi-Objective Optimizer in the HEEDS Multidisciplinary Design Optimization Software (Red Cedar Technology Inc. 2013). MO-SHERPA was chosen because it was written by one of the authors of this paper and is a high-performance algorithm (Chase et al. 2009). The goal is to minimize both deaths and impaired drivers. A weighted sum of the two casualty terms yielded a single objective function called "impact." The impact (objective 1) definition is: $Impact = (1 * AEGL2) + (10 * AEGL3)$, reflecting the more serious nature of fatalities. However, because the model is stochastic, the same model may produce a low-impact (high-performance) design in some runs, but a high impact in other runs. To take advantage of the superior diversity-maintenance properties of 2-objective search, and because low-variance solutions are clearly superior to high-variance solutions, the standard deviation in impact of each design (from run to repeated run) was introduced as a second objective. Such "multi-objectivization" (Saxena and Deb 2007) has been recognized as a useful tool in many situations, even if the second objective is not actually required. In this case, the two objectives (mean impact and variance of impact) are not necessarily in conflict, but they often conflicted. The study is thus a two-objective Pareto optimization. MO-SHERPA was used to run the simulation repeatedly for each candidate design, seeking to minimize the two objective functions by altering the traffic light timings used in the evacuation phase. It begins by randomly generating parameter values to form several different "chromosomes" in a "population" of individuals. The NetLogo input files defining each traffic control design (chromosome) are written to a named directory by HEEDS, and NetLogo is invoked in each such directory, generating output files that are "mined" by HEEDS for the required outputs. An evaluation runs the simulation ten times with a given set of assigned parameters (in that design's directory) and determines which individuals performed better by calculating Objective 1 and Objective 2. At this point, the algorithm chooses the best individuals according to non-dominated sorting (Deb et al. 2002) and forms a new generation based on these "parents." This new generation is typically formed by two operations, mutation and recombination. Mutations are generally done by varying a single or multiple parameters by some random amount. Recombination generally involves taking some parameters from one chromosome and some from another to form a new individual. MO-SHERPA does this autonomously.

During this MO-SHERPA optimization, each generation produces 140 individuals. Figure 3 shows what the chromosome that MO-SHERPA is modifying looks like. As seen in this figure, each of the 64 intersections that are altered during the optimization has 3 parameters that are directly changed: offset, green time in the east-west direction, and green time in the north-south direction. Each of these can be set to whole-second values. To make the optimization simpler (with fewer parameters to change), the amount of amber light time in each direction is fixed at 5 seconds per cycle. There are also other ways the formulation reduces the number of possible parameter combinations it must test. The offset value for each adjacent intersection along a given street differs by the same amount for that street. This offset difference is limited to being between 0 and 360 seconds. The optimization also works with a single north-south offset anywhere from -360 seconds to 360 seconds. This value is simply the difference between offset values at adjacent intersections along the east-most avenue. Additionally, the amount of green light time in any direction must be between 10 and 180 seconds. Each street has the same green light durations for every intersection it runs through. Thus, the total number of light-timing (or design) variables that MO-SHERPA must manipulate is 25. All of these limitations placed on the parameters of the traffic evacuation file allow the HEEDS optimization to perform faster and avoid traffic light timings unlikely to perform well.

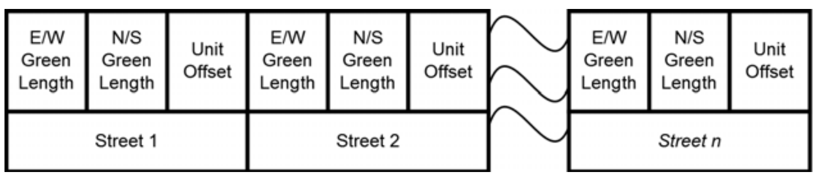


Figure 3: Chromosome being modified during the MO-SHERPA optimization.

### 3.3 Standard Traffic Light Design

For comparison with optimized solutions, a “standard” traffic light control for evacuation was needed. This standard design was found by trial-and-error exploration of designs from (Chen, Chen and Miller-Hooks 2007), augmented by simple designs conceived by the authors. It was important that the “standard” traffic light designs be given repetitive testing as were the designs generated by MO-SHERPA (as described later). A design with a 150-second cycle and green light times the same in each direction for all intersections (design 150_50) was finally selected as the control. After 200 runs, the impact measure for this design had a mean of 628.3 and a standard deviation of 852.9.

## 4 RESULTS

### 4.1 Optimized Design

The optimization of the simulation was run for 69 generations, including evaluation of 10,076 individual designs (occasionally a local failure caused a generation to evaluate 280, or in one case, 276 individuals). This took roughly 15 days in 140 nodes of an Intel10 cluster in Michigan State University’s High Performance Computing Cluster. A scatter plot of Objective 1 vs. Objective 2 for the entire population of individual designs is shown in Figure 4. The majority of points with low Objective 1 also had a low Objective 2 and vice versa, indicating that there was often no conflict (trade-off) when optimizing for the mean and standard deviation of the impact measure in this model. These data are based on the ten simulation runs HEEDS performed for each design (each run had a different random spill location among the 27 patches, as well as different initial traffic).

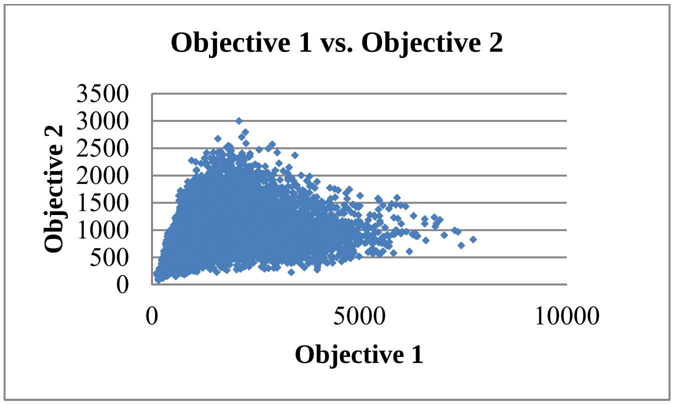


Figure 4: Scatter plot of each design's Objective 1 and Objective 2 values from the HEEDS optimization.

The difficulty is that since the problem is stochastic (with random placement of plume origin and randomly initialized traffic) the ten simulation runs per individual design during the evolutionary process are not enough to guarantee that a given best performer would remain the best if many more runs were done. In an effort to see which design truly was best, the six HEEDS designs with the lowest mean impact were tested with 50 new random seeds. The mean impact from these 50 runs can be seen in Table 1. Design 4143, with a mean impact of 408.2, was found to be the most robust design of the top six individuals (Table 1). It is also worth noting, but is not surprising, that the mean impact of the 10-run results for a design was significantly better than the mean impact of that same design after 50 additional runs. For example, Design 9089 had a mean Impact of 112.4 based on the initial 10 HEEDS runs, yet had a mean Impact of 877.7 across an additional 50 runs (Table 1). It was simply "lucky" in the first 10 runs.

Table 1: Top six individuals based on the mean impact of the ten runs performed by HEEDS.

| **Generation / Design** | | 62 / 9089 | 56 / 8155 | 28 / 4143 | 64 / 9369 | 51 / 7422 | 18 / 2646 |
|---|---|---|---|---|---|---|---|
| ***Mean Impact*** | **HEEDS** | 112.4 | 126 | 131 | 133.1 | 148.9 | 157.7 |
| | **50 additional runs** | 877.7 | 607.4 | 408.2 | 718.8 | 717.5 | 646.5 |

Stochasticity makes it possible for a design that is actually Pareto non-dominated based on many trials to be lost before the final generation. To see if this was the case in our runs, the top six individuals from generations 69, 68, and 67 were tested with 50 random seeds. The best of these additionally tested individuals was design 9849 from generation 68, with an average Impact of 488.9. However design 4143 remained better, with an Impact (being minimized) of 408.2. The testing of additional individuals from late generations was not done further back than generation 67, since the best design from generation 67 had an Impact of 622.8, making it already worse than the best design from generation 68.

Figure 5 shows a scatter plot of Objective 1 vs. Objective 2 for these designs run 50 times. The two best designs based on this plot are design 4143 and design 9849. Design 4143 had a smaller Objective 1 value than design 9849, but design 9849 had a smaller Objective 2 value. However, after 200 runs, design 4143 dominated design 9849 by having both lower Objective 1 and Objective 2 values.

### 4.2 Comparing Optimized and Control Designs

After selecting design 4143 as the best optimized design and design 150_50 as the best standard traffic light design according to their mean impacts (Objective 1), each was run a total of 200 times. Table 2 shows that after these 200 runs, the average AEGL-3 and impact of the optimized design remained better than those of the control design. The optimized design averaged 12 fewer deaths than the control, while

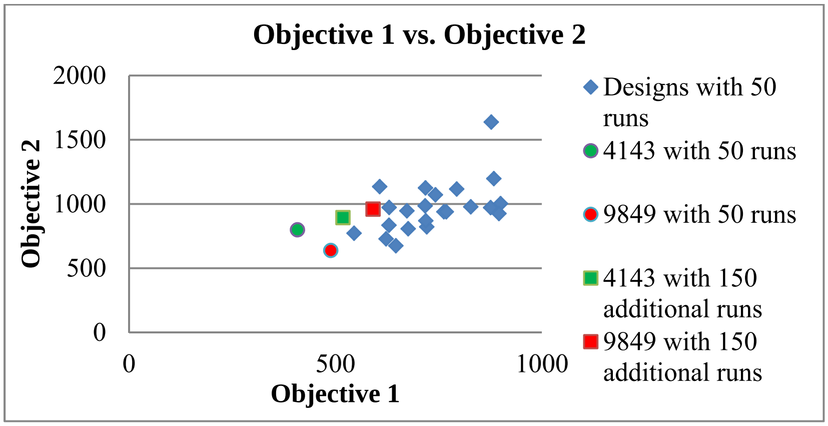


Figure 5: Scatter plot showing the Objective 1 and Objective 2 values of selected "best" designs.

the control averaged 14 fewer drivers impaired to any extent by the chlorine. While the optimized design had a lower mean impact measure, the control design, having lower standard deviation for all values (Table 2), was not strictly dominated by the optimized design. However, since the standard deviation was added mainly to promote search diversity, design 4143 is judged superior to the best "standard" design.

Table 2: Results of the optimized and control designs after 200 additional runs.

| Design | Aegl-2 (impaired) | | Aegl-3 (deaths) | | Impact | |
|---|---|---|---|---|---|---|
| | ***Mean*** | ***SD*** | ***Mean*** | ***SD*** | ***Mean*** | ***SD*** |
| Optimized 4143 | 129.1 | 165.6 | 38.9 | 81.7 | 518.4 | 893.2 |
| Control 150_50 | 114.8 | 155.3 | 51.3 | 77.6 | 628.3 | 852.9 |

Both the "best" optimized design and the "best" standard design had standard deviations for AEGL-2, AEGL-3, and impact higher than their respective averages (Table 2). This implies that statistical testing is needed to determine whether or not the difference in the two designs' results was significant. The Wilcoxon Rank Sum Test was performed on these three dependent variables. The test is non-parametric as it only deals with the rank of each value rather than the actual value.

When comparing the AEGL-2 between the two designs, the null hypothesis was that the two designs have the same AEGL-2. The alternative hypothesis was that AEGL-2 of the control was lower than AEGL-2 of the optimized design. This resulted in a left-handed Wilcoxon Rank Sum Test with a p-value of 0.6623, failing to reject the null hypothesis. For comparing AEGL-3 (deaths) of the two designs, the alternative hypothesis was that the optimized design had a lower AEGL-3 than the control. The p-value in this Wilcoxon Rank Sum test was 0.000005, supporting that the fewer mean deaths with the optimized design was not due to chance. Testing significance of the overall impacts of the two designs gave a p-value of 0.00189 supporting the hypothesis that the optimized design's outperforming of the control design was not due to chance.

## 5 CONCLUSION

The mean impacts of optimized and control timings indicate that the optimized design is better than the control design (Table 2). Although it would have been ideal for the optimized design to have fewer drivers impaired than the control, having fewer deaths is more important, as indicated by a lower *impact*, which weights deaths 10 times more than impairments. The difference in AEGL-2 (impairment) rates was not statistically significant. Since the MO-SHERPA algorithm used by HEEDS was told to minimize the impact measure and its standard deviation, impact is the most important parameter when comparing the optimized and control solutions. The slightly lower standard deviation of the control implies that the optimized design is more dependent on the actual traffic pattern than is the control, even though it does better on average.

Practical implementation of ABM-simulation-based traffic light timings to speed evacuation in the case of particular events depends on the capability to localize the source of the threat and its proliferation. In this case, the source of the spill (threat) was randomly distributed along the length of one city block, as might be done with an industrial rail line passing through an urban area. Additional simulations have shown that the performance of the evolved traffic light timings depends on how precisely the threat can be localized when the timing is put into place. It would be quite feasible to run optimizations for evacuations necessitated by accidents at chemical plants, on segments of rail lines, and other high-risk scenarios. It would be much more difficult to run such optimizations for hundreds of potential accident sites. However, it is also possible that further studies could allow inference of more general control patterns, less effective than individual optimizations, but superior to the timings in the "control" cases tested here. The gains, in terms of reduced fatalities, were modest, but that may reflect the fact that the "control" evacuation timing was already optimized to a fairly high degree. Gains against less capable "controls" might be much greater.

Agent-based models can clearly be used for optimizing evacuations in many other scenarios (subway stations, buildings, etc.). However, stochasticity and run times require that strategies based on these models be optimized in advance of an actual evacuation event. If a particular location is to be evacuated, the modeling of the physical infrastructure may be much more straightforward than the modeling of the awareness and psychological state and the resulting behavior of individual drivers.

## ACKNOWLEDGMENTS

This material is based in part on work supported by the National Science Foundation under Cooperative Agreement No. DBI-0939454. Any opinions, findings, and conclusions or recommendations expressed here are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

Batty, M. 2005. "Agent-Based Technologies and GIS: Simulating Crowding, Panic, and Disaster Management." In *Frontiers of Geographic Information Technology*, edited by S. Rana and J. Sharma, 81-101. Berlin: Springer.

Canter, David. 1980. *Fires and Human Behaviour*. New York, NY: John Wiley & Sons.

Chase, N., M. Rademacher, E. Goodman, R. Averill, R. Sidhu. 2009. "A Benchmark Study of Multi-Objective Optimization Methods." BMK-3021, Rev. 06.09. Accessed June 28, 2015 http://www.redcedartech.com/pdfs/MO-SHERPA_paper.pdf

Chen, M., L. Chen, and E. Miller-Hooks. 2007. "Traffic Signal Timing for Urban Evacuation." *Journal of Urban Planning and Development* 133, no. 1: 30-43.

Chen, X. 2011. "Agent-Based Modeling of Emergency Evacuation in a Spatially-Aware and Time-Aware Environment." In *Proceedings of the 25th International Cartographic Conference*. Paris, France. July 3-8, 2011. ISBN: 978-1-907075-05-6.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002. "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II," *IEEE Trans. on Evol. Computation*, 6(2), 182-197.

Hamdar, S. H.. 2004. "Towards Modeling Driver Behavior Under Extreme Conditions." Ph.D. thesis, Department of Civil Engineering, University of Maryland, College Park, Maryland. http://drum.lib.umd.edu/handle/1903/2141 [Accessed July 8, 2015].

Kimm, A. K.-C. Maassen. 2011. "Extended Cell-Transmission-Based Evacuation Planning in Urban Areas." *Pesquisa Operacional* 31(3): 405-441.

Kuligowski, E. 2009. "The Process of Human Behavior in Fires." NIST Technical Note 1631, National Institute for Standards and Technology, Gaithersburg, MD.

National Research Council. 2004. *Acute Exposure Guideline Levels for Selected Airborne Chemicals.* Vol. 4. Washingtion: National Academies Press.

North, M. J., and C. M. Macal. 2007. *Managing Business Complexity: Discovering Strategic Solutions with Agent-Based Modeling and Simulation*. New York: Oxford University Press.

N.Y. VAT. LAW § 1175 : NY Code – Section 1175: *Obstructing traffic at intersection.*

Office of Emergency Management, EPA; Emergency Response Division, NOAA; National Oceanic and Atmospheric Administration. 2009. "ALOHA 5.4.1.2." EPA; NOAA.

Ott, W. R., N. E. Klepeis, and P. Switzer. 2007. "Air change rates of motor vehicles and in-vehicle pollutant concentrations from secondhand smoke." *Journal of Exposure Science and Environmental Epidemiology*, 1-14.

Red Cedar Technology Inc. 2013. *HEEDS' Exclusive Search Technologies.* Accessed August 12, 2013. http://www.redcedartech.com/products/heeds_mdo/heeds_exclusive_search_technologies.

Saxena, D. and K. Deb. 2007. "Trading on Infeasibility by Exploiting Constraints' Criticality through Multi-Objectivization: A System Design Perspective." In *Proceedings of the 2007 IEEE Congress on Evolutionary Computation*, Singapore, 919-926.

Silverman, B. G., N. I. Badler, N. Pelechano, and K. O'Brien. 2005. "Crowd Simulation Incorporating Agent Psychological Models, Roles and Communication." Technical Report, Computer and Information Science, Center for Human Modeling and Simulation, University of Pennsylvania, Philadelphia, Pennsylvania.

Thompson, P. A., and E. W. Marchant. 1995. "Computer and fluid modeling of evacuation." *Safety Science* 18(4): 277-289.

Till, R. C., M. Durak, and E. D. Goodman. 2012. "Modeling Critical Incidents to Improve Response." Chap. 6 in *Handbook of Critical Incident Analysis*, edited by R. W. Schwester, 99-111. Armonk, New York: M.E. Sharpe, Inc.

Wilensky, U. 2003. "NetLogo Traffic Grid model." Northwestern University, Evanston, IL: Center for Connected Learning and Computer-Based Modeling.

Wilensky, U. 1999. "NetLogo." Northwestern University, Evanston, IL: Center for Connected Learning and Computer-Based Modeling.

## AUTHOR BIOGRAPHIES

**MATT DURAK** (mattdurak@gmail.com) is a Software Development Engineer at Microsoft with a B.S. in Computer Engineering from Michigan State University. He began this research during his undergraduate studies. He currently develops wireless networking features for Windows.

**NICK DURAK** (nickdurak@gmail.com) is pursuing a B.S. in Computer Science at Michigan State University.

**ERIK D. GOODMAN** (goodman@msu.edu) is a professor of electrical and computer engineering at Michigan State University He directs the BEACON Center for the Study of Evolution in Action, an NSF Science and Technology Center. He received the Ph.D. in Computer & Communication Sciences from the University of Michigan, and was the founding chair of ACM's SIGEVO.

**ROBERT TILL,** Ph.D. (rtill@jjay.cuny.edu), is Associate Professor and Deputy Chair for the undergraduate Fire Science program and the graduate program in the Department of Protection Management, John Jay College of Criminal Justice, City University of New York. He studies fire protection engineering, emphasizing firefighting and protection of transport infrastructure projects. He received the Ph.D. from Worcester Polytechnic Institute.