The Importance of Modelling Realistic Human Behaviour When Planning Evacuation Schedules

Chaminda Bulumulla
RMIT University
124 La Trobe St
Melbourne, Australia
chaminda.bulumulla@rmit.edu.au

Lin Padgham
RMIT University
124 La Trobe St
Melbourne, Australia
lin.padgham@rmit.edu.au

Dhirendra Singh
RMIT University
124 La Trobe St
Melbourne, Australia
dhirendra.singh@rmit.edu.au

Jeffrey Chan
RMIT University
124 La Trobe St
Melbourne, Australia
jeffrey.chan@rmit.edu.au

ABSTRACT

In planning evacuation schedules various projects work on devising optimal schedules for evacuation of the population. These assume that people do as they are asked in terms of when and by what route they leave the area. However, we know from numerous case studies, that this is not so. People have their own priorities and concerns and will address these before doing as they are asked. We know from many emergency situation examples that a key behaviour that people engage in is to assemble the family group, and to check on relatives or close friends. These behaviours are highly likely to affect how an optimal schedule that is provided plays out in practice. In this work we obtain an actual optimised evacuation schedule for a large area of over 35,000 households, which is prone to wildfires. We then compare an agent based simulation that adheres to this schedule, with ones that use it as a base, but where some of the agents first engage in the behaviours mentioned. We analyse the differences on 17 different configurations, exploring both the statistical significance of differences, as well as the importance of the differences for ensuring successful evacuation.

1. INTRODUCTION

Planning for potential evacuation of large numbers of people from an area due to an emergency or a threat, is an application where agent based modelling has been used in simulations to aid in understanding potential issues (e.g. [6,13,28]). Some excellent work has been done on exploring optimal management of evacuation scenarios, in some cases directed towards potential use in real-time during an emergency incident [8–10,21]. However, this optimisation approach assumes that people do as they are told, whereas in fact we know that they may not do this. In this work, we explore how adding some typical individual behaviours to an optimised plan may compromise that plan and we assess the importance of explicitly modelling such behaviours.

It is well documented that in emergencies people will first gather or ensure the safety of their family members [2,11]. A small number of studies consider picking up behaviours when modelling evacuations [14,18,19], and several argue that exclusion of such behaviours may result in predicting excessively optimistic evacuation travel durations while failing to capture complex traffic flow patterns. In this work, we investigate how picking up behaviours affect an independently developed globally optimised evacuation simulation. In contrast to other simulation work exploring pickup behaviours, we use an agent based approach which is useful exactly because it is able to individually represent heterogeneous agents, simulating the many interactions and observing the system level effects.

Our baseline is an optimised timing and routing schedule, developed for a relatively large geographical area, where wildfires are a potential threat. We then introduce a behaviour where some agents pick up a child or relative before doing anything else. In some cases this delays the time at which they start evacuation, or causes them to connect to the planned evacuation route from the point at which they pick up a family member, rather than from the home. We explore 17 different configurations, differing in the percentage of the population that exhibits these extra behaviours, the maximum distance they can drive to pick up a family member, and the amount of time the pickup takes. We then assess the effects on timing as compared to the optimised baseline without such behaviours. We find that there are statistically significant differences between almost all of the test simulations and the baseline version. However, although they are statistically significant, depending on the domain details, they may not be practically important. The biggest effect is in delays to agents that do adhere to the optimised schedule.

In the next section we provide some background regarding the approach and systems we used. We then describe the experimental setup, followed by a description of the specific experiments and the key questions we wished to explore such as to what extent clearance time or evacuation rate are impacted. We then analyse the results, compare our work with related work and conclude that certainly there are

Appears in: *Proc. of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

statistically significant effects once likely human behaviour is modelled. However, the differences on key dimensions may be small enough to justify the use of optimised scheduling without concern for these effects.

2. BACKGROUND

Our comparison simulation uses the BDI-MATSim system to represent and explore key behaviours and their effects on evacuation dynamics. This system has been successfully used in other applications [25], and allows us to naturally combine individual agent behaviours with an efficient agent based traffic simulator, capable of managing large numbers of agents.

2.1 The BDI Framework

Belief-Desire-Intention (BDI) is a philosophically grounded cognitive framework which balances an agent's commitment to pursue specific goals (long-term and short-term), with the ability to react to the changes in the environment [3]. This framework has been extensively used for developing intelligent agent systems and lies at the core of well-known agent implementation platforms like JACK [5], Jadex [4] and JASON [1].

The goal-directed approach in BDI systems is suitable for modelling complex human behaviours in dynamic environments (e.g. [24], [23], [27]). Essentially, a BDI representation is of the form $G : \psi \leftarrow P$, which means that the plan P is suitable to achieve goal G when its context condition ψ is believed true. In other words, it is an AND-OR tree of goals (AND) and their potential plans (OR). This representation is quite similar to human reasoning and also provides an intuitive understanding of the behavioural model, even for non-programmers [25].

2.2 MATSim

MATSim is a mature agent based traffic simulator, which is capable of simulating many hundreds of thousands of agents [22]. MATSim was originally developed to understand how traffic patterns would evolve as individual agents independently attempted to achieve their goals on a typical weekday. Agents have predefined travel plans for a typical day and these plans are optimised over multiple iterations using a co-evolutionary algorithm towards a user equilibrium [12]. Evacuation is not something which is repeated many times with individuals iteratively modifying their schedule based on previous experience. Consequently we use only the module which simulates a single day, and not the co-evolutionary algorithm. In [12] we have described how we modified the single day module to allow reactive changes to plans, necesary for this domain.

A MATSim travel plan is a list of activities and legs. An activity represents a task that an agent is pursuing at a specific location until a defined end-time. A leg describes the plan to travel from one location to another by defining a route, which is a list of links (i.e., road-segments) that should be traversed in the given order. The baseline system consists of a MATSim travel plan for each agent that first has them engage in a wait activity until the defined departure time, followed by a route, finishing with a wait at the destination.

2.3 The BDI-MATSim System

The BDI-MATSim system is an integration of a BDI framework with MATSim, which facilitates "within-day replanning" in MATSim [20]. This integration allows agents to proactively make decisions to change their original plan, depending on both environmental situations and agent goals. Conceptually, the "brain" of a MATSim agent is modelled in the BDI system (as a BDI agent) while the "body" remains inside MATSim. The communication between these agent counterparts is defined based on standard agent concepts, *percepts* and *actions*. A MATSim agent sends percepts to the BDI counterpart, which conducts high-level reasoning and issues a (BDI) *action* for the MATSim agent to execute [20]. *Percepts* from the MATSim counterpart can be either information about it's own state (e.g. location), or an observation from the MATSim environment. Basically, a BDI action modifies the travel plan of a MATSim agent using low-level MATSim functions. For example, the BDI action *driveTo(x,y)* will use MATSim functions to find the closest road-segment near coordinates(x,y), creating a leg with a route and then inserting this leg as the next step in the travel plan.

The integrated simulation is a synchronisation of time-based MATSim simulation and an event-based BDI system. This synchronisation is achieved by passing the control to the BDI system at the end of each MATSim time step. Control is returned when reasoning about actions for the next step concludes.

3. EXPERIMENTAL SETUP

The experimental setup compares the baseline optimised simulation scenario with various configurations of the comparison simulation scenario, incorporating pickup behaviours into the optimised schedule.

The optimised schedule for this experiment was obtained from NICTA[1] (now Data61), which was generated for the Hawkesbury region in North-West of Sydney, Australia. This is a MATSim configuration consisting of 38,434 agents covering 80 evacuation zones and 5 safe destinations. Agents in each evacuation zone start their evacuation from a "Central Evacuation Point" (CEP), where each agent has it's own departure time and evacuation route, produced to obtain globally optimal results. As it is unrealistic to expect people/agents to all start from a central evacuation point, we make a number of changes to enhance the realism and avoid artificial congestion around the central points. We first describe Scenario 1, the optimised schedule baseline scenario, then Scenario 2 with the addition of our behavioural model.

3.1 Scenario 1 (S1) Baseline Model

To increase realism and to ensure comparability to our test case, we have made two modifications to the original NICTA configuration:

1. Instead of every agent starting off at a CEP, we dispersed each agent to a random home location within the radius of 4.049km from the corresponding CEP. This radius is determined from the circular area of an average size SA1 [2] region in Hawkesbury.

2. The NICTA optimised configuration contains a simplified road network as it only includes the specific nodes and links required for the optimised evacuation. This hinders the incorporation of realistic traffic behaviours

[1]https://www.nicta.com.au/about-nicta/

[2]An SA1 is the smallest area for which census data is provided in Australia.

into the simulation as these behaviours require a more complete network of locations and roads in the Hawkesbury area. However, the optimised evacuation schedule is defined on the simplified network which necessitates preserving the network. Therefore, we expanded the road network by extracting a comprehensive road network of the region from OpenStreetMap[3], converting it to a MATSim road network using MATSim utilities and then merging the two networks by linking each node of the simplified network with the closest node of the comprehensive network.

This scenario is considered as the baseline for our experiments.

3.2 Scenario 2 (S2) Added Behaviours

To explore the effects of pickup behaviours on an optimised schedule, we consider a wildfire evacuation during school hours. The behavioural model is developed using known behaviours during evacuations: residents detouring to collect children (from schools) and checking on relatives, before driving to the safe destination [24]. Geographical locations of 17 schools in Hawkesbury were mapped into the model. Households were then assigned one of the school locations or a location of relatives, randomly assigned within a certain distance range.

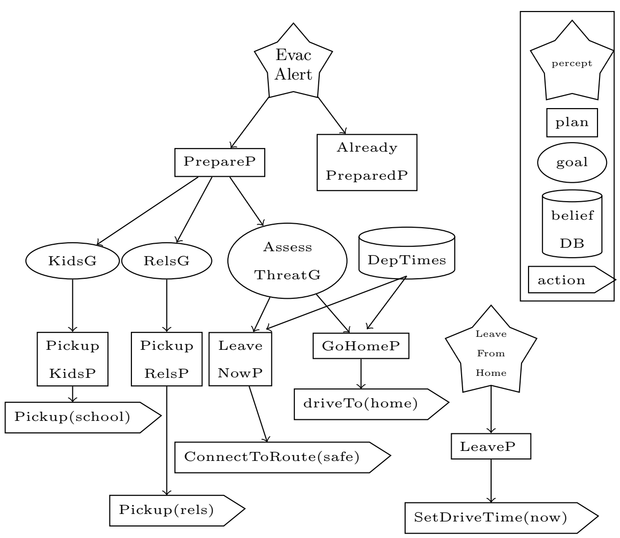


Figure 1: BDI agent design for Scenario 2

Figure 1 depicts the design of the BDI agent model. The percept *EvacAlert* is notified to all BDI agents by the simulation environment. This percept triggers two plans: *Prepare* plan is chosen if the context condition pickupkids or pickupRelatives is believed true, *AlreadyPrepared* plan is selected otherwise. If the latter is selected, their MATSim counterparts will execute the predefined MATSim plan adhering to the optimised schedule. The Body of *Prepare* plan consists of three subgoals, *KidsG*, *RelsG* and *AssessThreatG*, all of which must be processed successfully for the plan to succeed. Primarily, the *AssessThreatG* goal assesses the current situation by deciding whether an agent should return back to home and wait for the scheduled departure time, or, in the case that time is too short for this, start evacuating immediately from the current location. The plans *LeaveNowP* and *GoHomeP* are associated with this goal. As contextual information for these plans, a reasoning agent uses acquired travel time for pickup/s from its MATSim counterpart (T_{travel}), and the time left for the optimised departure (T_{left}) which is measured using it's belief *depTime*. If the context condition $T_{travel} > T_{left}$ is believed true, the agent selects *LeaveNowP* plan, else selects *GoHomeP* plan. If an agent is waiting at home, it perceives *LeaveFromHome*, which is an alert notified from the BDI system when the scheduled departure time arrives. In addition to the existing *driveTo* BDI action, three new BDI actions are introduced to the model with functionalities described as follows:

- *Pickup(loc,pickup-time)*: plans a route from the current location to the destination (*loc*), and spends a defined time duration (*pickup-time*) at the destination.
- *ConnectToRoute(route)*: finds the shortest distance "entry point" to the scheduled *route* from the current location. When this action is initiated from the *LeaveNowP* plan, it finds the closest "entry point" to the optimised route defined from CEP to the safe destination in the MATSim plan.
- *SetDriveTime(endTime)*: sets the end time for the "wait" MATSim activity to allow the agent to start evacuation from home.

Integrated Simulation

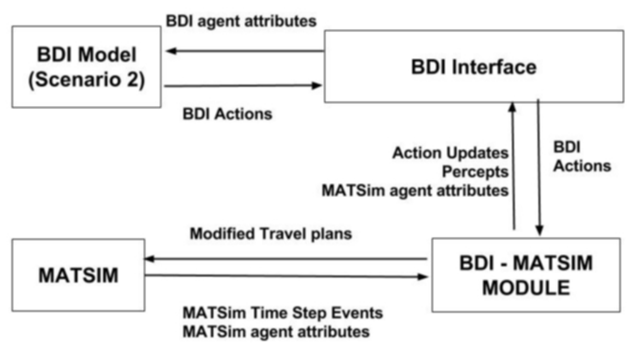


Figure 2: Information flow of the integration architecture

Figure 2 shows the information flow of the BDI-MATSim system. The modular architecture of this system allows the integration of different BDI agent models via the BDI interface. Initially, a BDI counterpart is instantiated for each MATSim agent, linked with the same agent ID. As part of this process, essential information (e.g. home location coordinates, departure time, safe destination coordinates) is extracted from the travel plan and stored as beliefs and attributes of the BDI agent. The BDI agents determine high level actions, which the MATSim agents (by using MATSim utilities) then convert into activities and legs for the MATSim activity plan. For example, for *Pickup(loc,pickup-time)* BDI action, the travel plan is modified by creating a leg with a route from current location to *loc* and an activity with endtime of *pickup-time*, and inserting them as the next steps of the MATsim plan. When a BDI high level action

[3]https://www.openstreetmap.org

is successfully completed, the BDI agent is informed by MATSim. The current location of the MATSim agent is then updated by the BDI agent and is later used for reasoning.

4. EXPERIMENTS

Experiments are conducted by comparing Scenario 2 (added behaviours) to Scenario 1 (baseline). In this comparison, we focus on 3 optimised evacuation benchmarks in the baseline: departure time (T_{opt_dep}), time from departure to arrival at the safe destination (T_{opt_evac}) and the arrival time (T_{opt_arr}). An agent's departure time (T_{dep}), evacuation time (T_{evac}) and arrival time (T_{arr}) in S2 is compared with the corresponding benchmark in the baseline. As differences at the level of seconds are clearly irrelevant we defined a variance limit δ and consider the interval [T_{opt} - δ, T_{opt} + δ] as equivalent to the optimised benchmark. We used a variation of 10 mins (δ=10 mins). While the value of δ is somewhat arbitrary, it is clearly necessary to have some equivalence range rather than a single point at 1 second granularity.

measurement	definition
ΔearlyDep	$(T_{dep} - T_{opt\_dep}) : T_{opt\_dep} - \delta > T_{dep}$
ΔlateDep	$(T_{dep} - T_{opt\_dep}) : T_{opt\_dep} + \delta < T_{dep}$
ΔshortEvac	$(T_{evac} - T_{opt\_evac}) : T_{opt\_evac} - \delta > T_{evac}$
ΔlongEvac	$(T_{evac} - T_{opt\_evac}) : T_{opt\_evac} + \delta < T_{evac}$
ΔearlyArr	$(T_{arr} - T_{opt\_arr}) : T_{opt\_arr} - \delta > T_{arr}$
ΔlateArr	$(T_{arr} - T_{opt\_arr}) : T_{opt\_arr} + \delta < T_{arr}$

Table 1: Definitions of the output measurements

Our objective here is to understand the impact of realistic behaviours on the optimised schedule. The impact on arrival and departure times can either be negative (early) or positive (late), while the impact on the travel time duration can be shorter (negative) or longer (positive)[4]. In order to measure these impacts, we derive 6 time difference measurements from the optimised benchmarks, which are defined in table 1. Each agent can have a maximum of three of these measurements, as they can be either early or late (and shorter or longer) compared to the benchmark, but not both. Figure 3 exemplifies how the time difference measurements (ΔearlyDep, ΔlongEvac and ΔlateArr) can be calculated for an agent by comparing the timelines of the two scenarios. The agent starts the evacuation early and arrives at the destination later than the optimised arrival time. T_{evac} - T_{opt_evac} gives the amount by which travel time is extended, namely ΔlongEvac.

The input configuration is based on three input parameters: fraction of the population with a child or relativec to pickup (%kidsRels); duration of the pickup activity (pickupTime); and maximum distance to the pickup location (distance). These parameters are input into a Latin Hypercube Sampling

[4]Negative/positive refers only to the direction from the optimum, not to whether it is desirable or undesirable.

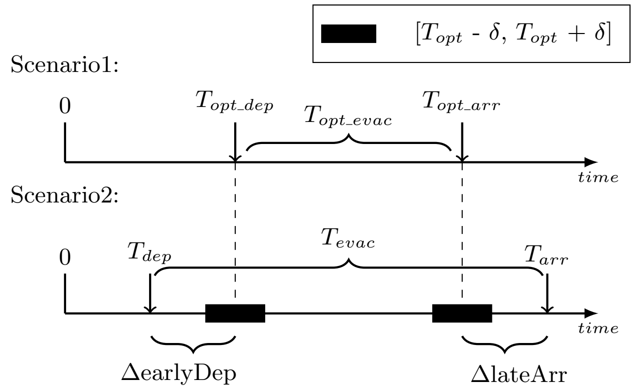


Figure 3: A sample timeline of an agent in S1 vs S2.

(LHS) tool [15] to generate input configurations for S2. LHS is a widely used technique for ensuring even sampling across a multi-dimensional parameter space. The LHS tool generated 17 input configurations for the three-dimensional parameter space with value ranges: %kidsRels(20-40%), pickupTime (15-60 mins) and distance (10-30 km).

We ran the baseline case once (as it is fully deterministic), and ran 40 replicates of each of the 17 S2 configurations. These were run on a supercomputing facility, and used MATSim 0.7.0 and JACK 5.6. The output data is collected for each simulation run. The measurements are averaged across the 40 replicates for each of the 17 input configurations. We have analysed the data to explore 6 main questions, as follows:

Which inputs influence the outcomes of interest?
To understand the behaviour of the simulation model and the relationship between input parameters and simulation outputs, we performed a sensitivity analysis. Specifically, we conducted the Spearman's rank correlation test for the inputs and outputs of the S2 simulation. As inputs, we considered the 17 configurations generated using the LHS approach. For the outputs we calculated mean values of the early and late departure and arrival times (ΔearlyDep, ΔlateDep, ΔearlyArr and ΔlateArr) as well as the mean value of the shorter and longer travel times (ΔshortEvac and ΔlongEvac), as defined in table 1.

What is the effect on departure and arrival times of agents?
To assess this we classify ΔearlyDep, ΔlateDep, ΔearlyArr and ΔlateArr into 30 min intervals and count the number of agents in each of these categories, as well as the number departing/arriving within 10 mins of the time given by the benchmark scenario for each of the agents [5]

What is the effect on evacuation travel time of agents
The analysis of evacuation travel time differences provides further understanding of the deviation from the baseline from a traffic-related perspective. We categorised agents according to their departure: *optimised* (starts departure by the T_{opt_dep}), *early* (departs earlier than the T_{opt_dep}) and *late* (departs later than the T_{opt_dep}). In this way, we are able

[5]For the 30 mins interval on either side of T_{opt_dep}/T_{opt_arr}, only 20mins will be counted as early/late due to 10 mins either side of optimal being regarded as equivalant to optimal.

to analyse the impact of agents who adhere to the optimised schedule as well as the ones that did not. ΔshortEvac or ΔlongEvac difference is determined for each agent and is grouped into 30 mins intervals similar to the analysis of arrival and departure times.

How is the rate of evacuation affected?
We examine the cumulative fraction of the agent population that has successfully evacuated in 30 min intervals. We did not distinguish the fraction based on differing safe destinations as we wanted to compare the overall evacuation rate between the scenarios.

What is the delay in Clearance Time?
Arrival of the last agent to a safe point is referred to as the Clearance Time (CT) of that safe point. We consider each of the 5 safe points in the baseline. Here we are interested in what the effect is for those agents most adversely affected by the change - i.e. those who arrive last at the safe point(s). For each safe point, we look at the difference in time between the last arrival in the baseline and the average last arrival time in each of the 17 configurations. We also look at the last arrival across all 5 destinations, i.e. the Evacuation Clearance Time (the time required to evacuate the whole population).

How statistically significant are the differences from the baseline?
We have tested the statistical significance of differences from the baseline scenario using two different methods. Firstly, we take the 40 iterations for each of the 17 S2 configurations, and establish how far the optimised baseline is from the mean (of the 40 iterations) for that configuration (in terms of standard deviations from the mean). If it is more than two standard deviations from the mean this is significantly different at the level of $p<0.05$.

Secondly, we consider the distributions across the agent population for average departure, travel time and arrival (for the 40 iterations). As these S2 distributions are non normal, we use the Wilcoxon Signed Rank Test to compare each configuration with the baseline. We note that statistically significant difference is necessary but not sufficient for the differences to be considered meaningful. An additional analysis of whether a statistically significant difference is meaningful or important is dependent on the particular domain.

In the next section, we discuss in turn the results with regard to each of the questions of interest we have identified here.

5. RESULTS

Configuration	%kidsRels	PickupTime	distance
min	0.23	35	11
median	0.3	38	20
max	0.35	57	29

Table 2: Parameter values of selected input configurations

The results we present are based on 681 simulation runs (40 iterations of each of the 17 configurations, plus the baseline run). When discussing a particular input configuration, we use the averages from the 40 iterations. Consequently output values for a particular agent, such as its arrival time at the destination, are a result of the average of that agent's arrival time across 40 iterations for the particular configuration. In order to summarise the results from the 17 input configurations explored, we choose what we call *min, med* and *max* input configurations. The min configuration is the one which exhibits the smallest effects on each of the output variables, the max exhibits the largest effects and the med configuration has effects approximately midway between these two. The actual inputs for these three configurations are shown in Table 2.

5.1 Influence of inputs on outputs

In order to clearly see which of our three inputs most influence our six identified outputs of interest we did a sensitivity analysis using Spearman's rank correlation test, which measures the monotonic relationship between inputs and outputs of the simulation model. The results of this are shown in Figure 4 as a correlation plot. Each cell contains the correlation coefficient between the input parameter and the output parameter. Correlations that are not statistically significant (at p=0.05) are marked with a cross. An interesting observation is that all the input variables have a positive correlation (though not always significant) with all outputs. That is in all cases, as the inputs increase in value, so do each of the outputs.

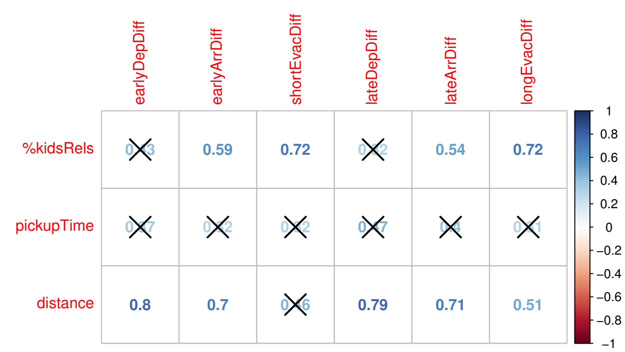
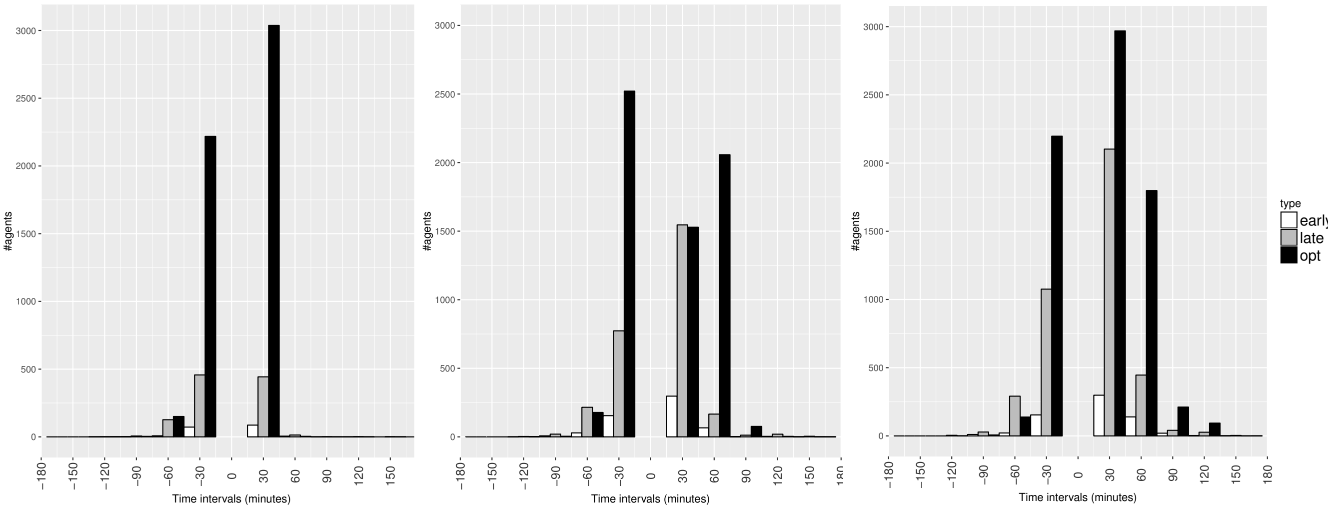


Figure 4: Spearman's Correlation Plot

The greatest influence is that of distance on departure time, with the effect on early and late departure time being almost the same. This influence of distance travelled to pick up children or relatives affecting early departure times was an unexpected effect. On analysis this happens because the further the pick-up point is from the home, the greater the chance that after pickup there is insufficient time to return home before the scheduled evacuation time, even though that scheduled time is in the future. This results in immediate start of (early) evacuation.

As would be expected, the effect on early and late start to evacuation flows onto effects on early and late arrivals. Interestingly there is also a positive correlation with length of travel time (EvacDiff), both shorter and longer. We were initially surprised to see shorter travel times, but it appears that this is due to decrease in congestion for some agents. That is, a fraction of the population may have a shorter T_{evac} due to less traffic congestion while others have longer T_{evac}



(a) min (b) median (c) max

Figure 5: Evacuation travel time differences in terms of $\Delta shortEvac$ (-) and $\Delta longEvac$ (+). Colours refer to the categorisation of the agents based on departure time: those who left earlier than scheduled (early), those who left on-time (opt), and those who left later than scheduled (late). The differences are classified into 30 mins intervals, for instance, 60 marks each category of agents having $\Delta longEvac$ within 30-60 time interval. Noteworthy, 30 represents the 10-30 mins time interval as we consider an equivalence (δ) of 10 mins.

due to higher traffic congestion than in the baseline scenario. The percentage of agents picking up children or relatives is also positively correlated with both early and late arrivals as well as shorter and longer travel times. Presumably this is because as more people do these extra pickup tasks, the greater the chance for changes in traffic congestion compared to the baseline, thus affecting travel time and arrival time, in both directions.

5.2 Departure and Arrival times

Config	$\Delta lateDep$	$\Delta lateArr$	mean($\Delta lateArr$)
min	2726	5427	9mins
median	5723	9051	27.2mins
max	7682	12482	36mins

Table 3: Summary of departure and arrival time differences

As can be expected based on the above sensitivity analysis, we see increasing numbers of both early and late arrivals and departures, as the scenarios become more challenging. In all cases the effects on arrival times are substantially greater than the effects on departure times. Table 3 summarizes these results. In the min configuration the number of late arrivals is 99% more than the number of late departures (2726 to 5427), with 58% increase in the median configuration (5723 to 9051) and 62% increase in the max configuration (7682 to 12482). The last column of table 3 shows the mean of $\Delta lateArr$ differences for each configuration. While the average delay in arrival time ranges from 9 mins to 36 mins, there is a substantial spread, the worst case is actually a 4 hour delay. For instance, an agent who evacuates in 32.5 mins (i.e., travel time from home to the safe destination) in baseline, encounters a delay of 3 hours and 45 mins in S2 by driving to the pickup location after travelling for around 2 hours (due to traffic congestion), engaging in pickup for 57 mins and driving to the safe destination for 1 hour 15 mins.

The most interesting effect on arrival times is actually the effect of delayed arrivals on those agents who obeyed the optimised schedule and left on time. For those agents who left on time (i.e., category opt), increase in travel time implies also late arrival. Figure 5 shows, for the three configurations, min, median and max, the number of agents (y-axis) with shorter or longer drive times to the evacuation point (x-axis). Those with negative x values have shorter travel times than in the baseline, whereas those with positive have longer. We see that substantial numbers of agents leaving on time (opt) had increased travel times - and therefore will have arrived late. In the worst case there are 94 agents that leave on time, and eventually arrive between 90 and 120 mins later than the baseline, due to the non-optimal behaviour of other agents.

5.3 Evacuation travel times

Config	$\Delta shortEvac$	$\Delta longEvac$
min	early=80 opt=2372 late=580	early=93 opt=3041 late=458
median	early=191 opt=2706 late=1011	early=360 opt=3667 late=1753
max	early=186 opt=2342 late=1300	early=461 opt=5073 late=2634

Table 4: Summary of evacuation travel time differences

Table 4 summarises the effect on travel times for agents in three different categories: those who left early (early), those who left on time (opt), and those who left late (late). For each configuration $\Delta shortEvac$ shows the numbers in each category with a shorter travel duration while the column $\Delta longEvac$ shows the numbers with longer travel times. So the first entry indicates that in the min configuration 80 agents who left early had a shorter travel time than in the baseline.

Looking at both Table 4 and Figure 5 it is apparent that the pickup behaviours have mostly affected the evacuation travel times of agents with optimal departures, resulting in higher numbers in $\Delta longEvac$ than $\Delta shortEvac$. A difference in travel time for an agent in the optimal time category, either short or long, is caused primarily by the traffic level in the road network. Some of the agents in this category (2371 in min to 2342 in max) experience less traffic congestion than in the baseline, gaining shorter travel times. However, a larger number of the opt category (3041 in min to 5073 in max) face increasing levels of traffic congestion, thereby taking longer time than expected to reach the safe destination. The mean evacuation travel time delay ($\Delta longEvac$) rises from 16 mins in the min configuration to 37 mins in the max configuration. As shown in figure 5*c*, the dispersion of $\Delta longEvac$ differences in the max configuration is such that there are 1798 agents in 30-60 mins delay interval, 211 agents in 60-90 mins interval and 94 in 90-120 mins interval. Overall, in the max configuration, 32% of the population have a difference in evacuation travel time while the rest travel to the safe destinations with travel times similar to the baseline scenario.

5.4 Evacuation rate

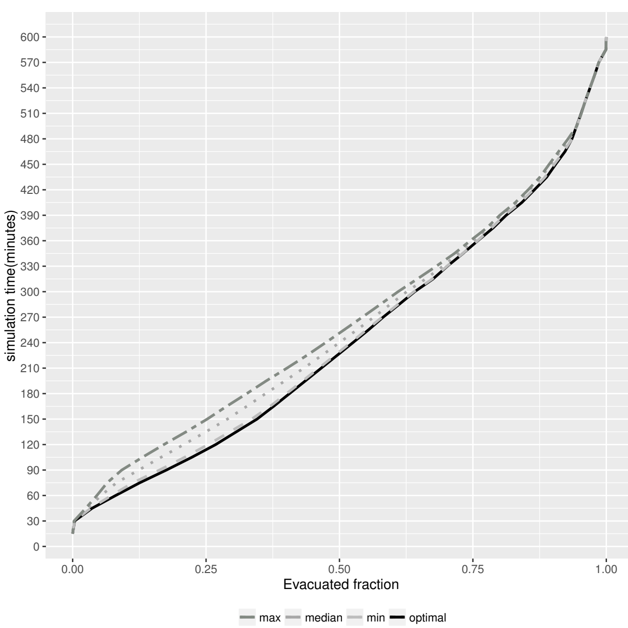


Figure 6: Evacuation rate: optimised vs min vs median vs max

Figure 6 shows the evacuation rate of the min, med and max configurations in comparison to the baseline. We can see that 25% of the population have been evacuated by about 105 mins in the baseline, whereas it takes 150 mins to evacuate the same % in the worst case.

The effect of the introduced behaviours is greatest between 30 mins and 150 mins. That is, in each 30 min interval from 30 to 150 mins, the number of agents reaching the safe destination is 400 less in the maximum configuration than in the baseline evacuation schedule.

After the 150 mins, the baseline evacuation rate decreases. Evacuation rate of the max configuration remains approximately the same. As a result, the rates converge after about 6 hours. This also shows that the introduced behaviours can impact the evacuation rate for up to 6 hours, at which time 75% of the agents have been evacuated.

5.5 Clearance Times

Safe	min	median	max	CT-baseline
Safe0	12mins	57mins	1h 15mins	6h 7mins
Safe1	8mins	48mins	1h	6h
Safe2	0	0	0	9h 8mins
Safe3	2mins	22mins	27mins	7h
Safe4	5mins	16mins	21mins	7h 8mins

Table 5: Time differences of the last arrivals at the safe points

We have compared the differences in Clearance Time (CT) of each safe point with respect to the baseline CTs. The results of this analysis are listed in table 5. The column CT-baseline contains the CTs recorded for all safe points in the baseline scenario, and the other columns show the respective delays in time for each configuration. Safe2 has the highest CT in the baseline, and apparently there is no impact on its CT from the behavioural modifications. That is, Evacuation Clearance Times of all configurations are equal to the baseline (9h 8mins). We also measured the number of agents reaching safe points later than the baseline CT. For Safe0 and Safe1, the delays are higher, but fewer agents reach the safe points after the CT of the baseline (in max configuration, 28 agents in Safe0 and 78 agents in Safe1). In contrast, Safe4 has comparatively low delays, but higher numbers of agents are delayed (in max configuration, 463 agents reach the safe point during the 21 mins delay). The delay in Safe3 involves only a single agent.

5.6 Statistical Significance

Comparing the output values T_{opt_dep}, T_{opt_evac} and T_{opt_arr} in the baseline to the distribution of the 40 runs within each of the 17 configurations, we found that for all but three of the configurations the optimised values were more than 2 standard deviations from the mean: i.e. significantly different ($p<0.05$). The 3 configurations with the lowest %KidsRels were between one and two standard deviations from the mean.

Using the Wilcoxon Signed Rank Test comparing the agent departure times, arrival times and travel duration in the baseline to the average of each configuration, we found the difference to be significant in all cases ($p<0.01$).

Clearly the differences are statistically significant. However, this does not necessarily imply that they are of a magnitude to be considered of importance, given the domain. The magnitude of many of the differences are such that it seems likely they would not be important in practice. A few are larger and would require further analysis.

6. RELATED WORK

In this section, we provide a comparison in terms of pickup behaviour modelling approaches and behavioural effects on evacuation measures.

Pickup behaviours are modelled by generating trip chains for households initially, which are then fed into traffic simulators. Optimisation techniques are often used for this, particularly linear integer programs (e.g. [14, 18, 19]). These programs are solved determining meeting locations, pickup sequences and routes for each household, but not departure times. Another approach is the use of an activity-based model to identify temporal and spatial locations of individuals for pickup trip generation [16]. For trip chains, the minimum cost route is selected for each household, meaning that there is no sense of optimisation at a population level. One limitation in all these techniques is that the generated travel plans are static throughout the simulation. In comparison, our agent-based approach focuses on decision making at an individual level, which results in dynamic change of (optimised) travel plans (i.e., change of routes and departure times) based on the actual situation as it evolves and is experienced in the environment. This is a more realistic way of modelling human decision making than having static travel plans upfront.

Several studies conclude significant deviations in evacuation measures as a result of pickup behaviours. Murray-Tuite et. al. claim that in their case, it is necessary to have about 150% of the travel demand to give the increased Clearance Times predicted by the model including pickup behaviours [19]. However, they use a simplified road network similar to a grid-based structure. They also consider 51% of the population engage in picking up 1–3 children while we considered a range of 20%-40% of the population who commit to a single pickup. Liu et. al. [17] argue that evacuation rate (i.e. the number of evacuees reaching safe points) is 50% less at certain time thresholds when pickup behaviour is introduced to a no-notice evacuation. These results are based on a non-optimised evacuation (as it is a no-notice evacuation) while our results are based on an optimised evacuation schedule.

There are also other aspects which may account for the differences between our results and theirs. The above simulations model the whole population beginning their pickup or evacuation activity within a shorter time duration (within 30-60 mins in [19] and within 2 hours in [17]), whereas our baseline departure times are distributed over 8 hours and 42 mins. In the other models, those agents that engage in pickup activity immediately start evacuation after completing the pickup. In our model, agents wait at home after the pickup activity, until their scheduled departure, if time allows.

Some simulation models take into account various transport modes when modelling pick up behaviours, for instance, evacuating a car-less population using public transport [26] and considering multiple transport modes such as taxi, bicycle and carpool [17]. We assume that residents use private vehicles for wildfire evacuations. Moreover, results of one social survey reveals that 84% of the participants prefer private vehicles for evacuation [7].

7. DISCUSSION & CONCLUSION

In this work we have taken as a baseline a large simulation of an optimised evacuation schedule developed for an actual geographical region, using the correct road network and accurate numbers of households within the region as determined by census data (with each household represented as a single car). We have then compared this with a simulation which retains the optimised scheduling to the extent possible, after prioritising behaviour to pick up a child/relative: a well known human behaviour that is top priority in emergency situations. The introduction of this behaviour can result in agents being both delayed or early in their departure. If pick-ups are not accomplished before the scheduled evacuation, agents will be delayed in their departure. If after picking up there is insufficient time to return home before scheduled departure, they will immediately connect to their intended departure route. This can result in early departure as well as some initial route modification. Disruptions in the departure times have flow-on effects to the travel time and the arrival time. Effects on the road network of greater or lesser congestion also affect agents leaving on time, and indeed these are the largest number of agents affected, with a greater number of delayed arrivals than early arrivals.

In almost all the configurations tested the results were statistically significant with 95% certainty. However the magnitude of the delays were far less than were suggested by previous work, and in fact may be short enough to be considered unimportant. There are a number of possible reasons for these differences. Firstly, agent based modelling provides a finer granularity and more accurate results than mathematical modelling. Secondly in our work we started with an optimised global schedule (as our aim was to measure how much outcomes of this were affected by inclusion of behaviour modelling). Thirdly even in our maximum configuration only 40% of agents engaged in one pick-up, whereas in other work higher proportions of the population engaged in possibly multiple pick-ups.

A conclusion we can draw from this is firstly that, contrary to expectations, it may be justifiable to use optimised scheduling as an adequate approximation, without specific consideration of likely deviations. This is important as one aims for the simplest model suitable for the task. This also demonstrates the importance of careful analysis and testing before adding more realistic complexity to a model. Incorporating behaviours into the optimisation strategy would possibly add unnecessary detail. However we also note that further exploration is needed with potentially larger numbers of agents engaging in possibly multiple pick-ups. Also in this work all agents were assumed to attempt to adhere to the optimised schedule although some had one higher priority task. In fact human decision making and behaviour is more complex than this and further investigation, modelling and analysis is needed to understand the appropriate use of optimised schedules in real world evacuation management.

Acknowledgement

This research was supported by a NICTA/Data61 scholarship and Australian National Computational Infrastructure (NCI) computing resources.

REFERENCES

[1] R. H. Bordini, J. F. Hübner, and M. Wooldridge. *Programming multi-agent systems in AgentSpeak using Jason*, volume 8. John Wiley & Sons, 2007.

[2] J. Boylan, C. Cheek, and T. Skinner. Preliminary report on the january 2013 fires in the south-eastern tasmania research project. *University of Tasmania Rural Clinical School, Burnie*, 2013.

[3] M. Bratman. *Intention, plans, and practical reason.* Harvard University Press, Cambridge, MA, 1987.

[4] L. Braubach, A. Pokahr, and W. Lamersdorf. Jadex: A bdi-agent system combining middleware and reasoning. In *Software agent-based applications, platforms and development kits*, pages 143–168. Springer, 2005.

[5] P. Busetta, R. Rönnquist, A. Hodgson, and A. Lucas. Jack intelligent agents-components for intelligent agents in java. *AgentLink News Letter*, 2(1):2–5, 1999.

[6] X. Chen, J. W. Meaker, and F. B. Zhan. Agent-based modeling and analysis of hurricane evacuation procedures for the florida keys. *Natural Hazards*, 38(3):321–338, 2006.

[7] D. Deka and J. Carnegie. Analyzing evacuation behavior of transportation-disadvantaged populations in northern new jersey. In *Transportation Research Board 89th Annual Meeting*, volume 10-1584, 2010.

[8] C. Even, V. Pillac, and P. Van Hentenryck. Convergent plan for large-scale evacuations. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI-15)*, 2015.

[9] C. Even, V. Pillac, P. Van Hentenryck, et al. Nicta evacuation planner: Actionable evacuation plans with contraflows. In *ECAI*, pages 1143–1148, 2014.

[10] C. Even, A. Schutt, and P. Van Hentenryck. A constraint programming approach for non-preemptive evacuation scheduling. In *International Conference on Principles and Practice of Constraint Programming*, pages 574–591. Springer, 2015.

[11] J. Heath, C. Nulsen, P. Dunlop, P. Clarke, P. Burgelt, and D. Morrison. The february 2011 fires in roleystone, kelmscott and red hill. *Bushfire CRC, Melbourne*, 2011.

[12] A. Horni, K. Nagel, and K. W. Axhausen. *The multi-agent transport simulation MATSim.* Ubiquity Press, 2016.

[13] G. Lämmel, M. Rieser, and K. Nagel. Large scale microscopic evacuation simulation. In *Pedestrian and evacuation dynamics 2008*, pages 547–553. Springer, 2010.

[14] S. Liu, P. Murray-Tuite, and L. Schweitzer. Relocating children in daytime no-notice evacuations: methodology and applications for transport systems of personal vehicles and buses. *Transportation Research Record: Journal of the Transportation Research Board*, 2234:79–88, 2011.

[15] S. Sanchez. NOLHdesigns spreadsheet, 2011. 30.10.2016.

[16] D.-Y. Lin, N. Eluru, S. Waller, and C. Bhat. Evacuation planning using the integrated system of activity-based modeling and dynamic traffic assignment. *Transportation Research Record: Journal of the Transportation Research Board*, 2132:69–77, 2009.

[17] S. Liu, P. M. Murray-Tuite, and L. Schweitzer. Incorporating household gathering and mode decisions in large-scale no-notice evacuation modeling. *Comp.-Aided Civil and Infrastruct. Engineering*, 29(2):107–122, 2014.

[18] P. Murray-Tuite and H. Mahmassani. Model of household trip-chain sequencing in emergency evacuation. *Transportation Research Record: Journal of the Transportation Research Board*, 1831:21–29, 2003.

[19] P. Murray-Tuite and H. Mahmassani. Transportation network evacuation planning with household activity interactions. *Transportation Research Record: Journal of the Transportation Research Board*, 1894:150–159, 2004.

[20] L. Padgham, K. Nagel, D. Singh, and Q. Chen. Integrating bdi agents into a matsim simulation. In *ECAI*, pages 681–686, 2014.

[21] V. Pillac, P. Van Hentenryck, and C. Even. A conflict-based path-generation heuristic for evacuation planning. *Transportation Research Part B: Methodological*, 83:136–150, 2016.

[22] M. Rieser, C. Dobler, T. Dubernet, D. Grether, A. Horni, G. Lämmel, R. Waraich, M. Zilske, K. W. Axhausen, and K. Nagel. Matsim user guide, 2013.

[23] A. Shendarkar, K. Vasudevan, S. Lee, and Y.-J. Son. Crowd simulation for emergency response using bdi agent based on virtual reality. In *Proceedings of the 38th conference on Winter simulation*, pages 545–553. Winter Simulation Conference, 2006.

[24] D. Singh and L. Padgham. Community evacuation planning for bushfires using agent-based simulation: Demonstration. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 1903–1904. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[25] D. Singh, L. Padgham, and B. Logan. Integrating BDI agents with agent-based simulation platforms. *Autonomous Agents and Multi-Agent Systems*, pages 1–22, 2016.

[26] R. Swamy, J. E. Kang, R. Batta, and Y. Chung. Hurricane evacuation planning using public transportation. *Socio-Economic Planning Sciences*, 2016.

[27] H. Van Truong, E. Beck, J. Dugdale, and C. Adam. Developing a model of evacuation after an earthquake in lebanon. In *Proceedings of the ISCRAM Vietnam 2013 Conference*, 2013.

[28] M. J. Widener, M. W. Horner, and S. S. Metcalf. Simulating the effects of social networks on a populations hurricane evacuation participation. *Journal of geographical systems*, 15(2):193–209, 2013.