DSCC2019-9222

AN AGENT-BASED MODEL OF LOST PERSON DYNAMICS FOR ENABLING WILDERNESS SEARCH AND RESCUE

Amanda Hashimoto, Nicole Abaid*

Engineering Mechanics Program
Virginia Tech
Blacksburg, Virginia, 24061
Email: ahashimo, nabaid@vt.edu

ABSTRACT

In this paper, we introduce an agent-based model of lost person behavior that may be used to improve current methods for wilderness search and rescue (SAR). The model defines agents moving on a landscape with behavior considered as a random variable. The behavior uses a distribution of four known lost person behavior strategies in order to simulate possible trajectories for the agent. We simulate all possible distributions of behaviors in the model and compute distributions of horizontal distances traveled in a fixed time. By comparing these results to analogous data from a database of lost person cases, we explore the model's validity with respect to real-world data.

INTRODUCTION

Every year, thousands of people go missing in the United States. According to the National Crime Information Center, nearly 610,000 people were reported as missing in the United States in 2018 [1]. Coordinated search and rescue (SAR) operations provide the best chance to locate a missing individual alive, and the key to a successful search is locating the lost person (LP) as rapidly as possible. As time elapses, the size of the search area grows geometrically and searchers are tasked with an increasingly daunting task of finding a person in a large area in a short amount of time. Because these searches can involve high risks to humans, the lost person SAR can be redefined by enabling teams of human searchers and autonomous unmanned

aerial vehicles (UAVs) to collaborate towards improving search outcomes. In order to optimize the search process, mathematical models of lost person behavior with respect to geophysical and transient landscape features could be used in conjunction with observable data from the UAVs. The goal of this paper is to introduce a model of lost person behavior that could be used to make UAV SAR searches more efficient.

Existing models of human behavior have been used extensively to study pedestrian dynamics, including force-based models showing collision avoidance, vision-based guidance, or goaloriented behaviors [2-6] and agent-based models based on behavioral heuristics [7], as well as the motion of animal groups [8]. Many of these models are used to study social interactions between individuals, but the behavior of a walking lost person can also be modeled analogously to a pedestrian, that is, as a selfpropelled particle. Among the limited academic literature for modeling the behavior of a LP, there are two models that implement a human movement algorithm, including a Bayesian approach to estimate the location of the LP [9] and an agent-based approach [10]. The Bayesian method discretizes the search area and computes probabilities of transitioning between cells based on previous behavior and landscape features [9]. The agentbased model in [10] simulates different behaviors for a LP on a landscape. However, both of these models fail to include specific types of lost person behavior (e.g., people with dementia or autism), which are known to significantly change behavior [11]. Thus, creating a spatially-resolved, agent-based model will allow for predicting specific paths taken by the lost person that can

^{*}Address all correspondence to this author.

be quickly explored by UAVs. In this paper, we will also take into account the different types of lost person behavioral profiles that have been collected and analyzed in the International Search & Rescue Incident database (ISRID), which amalgamates data from over 50,000 lost person searches [12].

When using an agent-based model, the rules and interaction behaviors are defined with respect to the individual. In this work, we model the behavior of a lost individual in the wilderness using a variety of different behavioral lost person types (LPT) as defined by [11]. By knowing the type of individual that is missing, the strategies used by that person may be predicted, and in turn optimize the SAR process. Specifically, once we have information on the type of person who is lost, we can predict the trajectories of the LP for SAR teams to utilize. Our model will allow us to generate a family of trajectories that define a spatial distribution where a lost person can be expected as a function of time. Conversely, by simulating possible distributions of behaviors, we can also compare the LPTs taken from real-world data to the behaviors used in our model to inform the profiles with information from the dataset.

MODELING Agent dynamics

In this model, an agent is considered to be a self-propelled particle moving in discrete time on a 2D grid of $L \times W$ square cells of fixed side length L_c . Cells of the grid are considered to be on a so-called linear feature or not, which is discussed in detail in the Map data processing subsection below.

The agent moves from its position to one of the eight adjacent cells by executing an algorithm which defines the strategies that it may use. We define four possible strategies based on [11]:

- 1. Random Walking (RW): An individual moves randomly. One may often have no apparent purpose other than finding something that looks familiar.
- 2. *Route Traveling* (RT): An individual travels on some unknown trail, path, drainage, or other aid. These specific paths are defined by linear features and are calculated based on elevation at a certain point in the model.
- 3. *Direction Traveling* (DT): An individual moves crosscountry, often ignoring trails and paths. They will sometimes cross railroad tracks, power lines, highways, and backyards
- 4. *Staying Put* (SP): An individual makes an active decision to stay put, often attempting to signal or build shelter.

We model the LPT as a random variable whose probability mass function (PMF) is a 4D vector that captures the probability of an agent using each strategy at a given time. For example, an agent with LPT defined by a probability (RW, RT, DT, SP) = (0.5,0,0.4,0.1) has a 50% chance of random walking, a 40% chance of direction traveling, and a 10% chance of staying put

at a given time step. To update the agent's position in time, independent realizations of this distribution are generated and the agent's position is updated according to the model of the randomly selected strategy.

We select an initial position $x(1) \in \mathbb{N}^2$ for the agent and randomly generate a second position $x(2) \in \mathbb{N}^2$ from the eight cells surrounding x(1). We compute v(1) = x(2) - x(1) as an initial velocity. To generate motion from the four strategies, we first consider the agent placed in the center of a 3×3 grid, which is a subset of the larger discrete map in body coordinates with respect to the agent. At time step t, the 3×3 grid with x(t) at its center is oriented so that x(t) + v(t) is the cell above x(t), which is accomplished by appropriately rounding coordinate values when this orientation is not orthogonal to the map axes.

We consider how the four strategies above define updates in this 3×3 grid and x(t+1) in global coordinates, using Figure 1 as a reference and the PMFs in Table 1. When the individual walks randomly (RW, strategy 1), the chance of moving into any of the other adjacent cells is the same, including staying in place. When route traveling (RT, strategy 2), the agent will check each of the surrounding cells for a linear feature and, if found, the updated position will be randomly selected from positions 1, 2, or 3 in body coordinates if they are on the linear path, see Figure 1. If a feature is not found, the agent will walk randomly following strategy 1. When direction traveling (DT, strategy 3), the agent will only move forward, that is, to position 2 in body coordinates. Lastly, when staying put (SP, strategy 4), the only possible update is the previous position, that is position 5 in body coordinates. At time step t, the agent generates a realization of

TABLE 1. PMF FOR EACH BEHAVIOR STRATEGY, GIVEN IN THE ORDER OF POSITIONS (1,2,3,4,5,6,7,8,9) FOR THE 3×3 GRID COORDINATES RELATIVE TO THE AGENT.

Strategy	PMF for updated position
RW	$(\frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9})$
RT	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0, 0, 0, 0, 0)$ if on a linear feature
DT	(0,1,0,0,0,0,0,0,0)
SP	(0,0,0,0,1,0,0,0,0)

the LPT PMF, which defines a strategy to use. Then the updated position x(t+1) is computed using this strategy, as well as the resulting velocity v(t+1). The model is iterated to generate the agent's trajectory for T time steps.

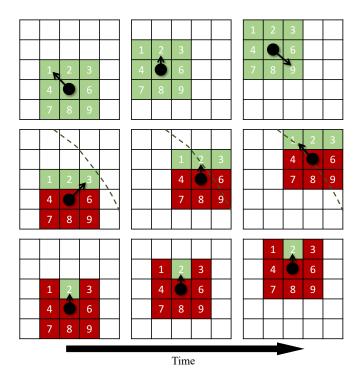


FIGURE 1. SCHEMATIC OF AN AGENT'S TRAJECTORIES FOR THREE BEHAVIOR STRATEGIES EVOLVING OVER TIME FROM LEFT TO RIGHT. THE AGENT'S INITIAL POSITION IS THE CIRCLE, EACH POSSIBLE NEXT STEP IS GREEN, AND LINEAR FEATURES ARE REPRESENTED AS THE DARK GREEN DASHED LINE. RANDOM WALKING (TOP) WILL HAVE EQUAL PROBABILITIES FOR THE NEXT POSITION IN THE ADJACENT CELLS. ROUTE TRAVELING (MIDDLE) WILL MOVE FORWARD ONTO A LINEAR FEATURE. DIRECTION TRAVELING (BOTTOM) WILL ALWAYS MOVE FORWARD INTO POSITION 2.

Map data processing

The possible trajectories for a lost individual are simulated on a map generated from USGS geographic information system (GIS) data. We collect this data using the MATLAB command webmap, which accesses the ESRI database [13]. The map gives the elevation of the landscape with respect to sea level at a given resolution. Linear features, such as drainages and roads, are generally visible as critical points in the magnitude of the gradient of the elevation map. As such, we calculate the gradient field using the derivative of a Gaussian filter. We then compute the gradient magnitude and find linear features using the Canny Edge Detection method [14]. This function is able to find edges by looking for local maxima and minima of a scalar field.

When searching for a lost individual in real applications, searchers need to have certain information about the LP to begin the investigation. An important piece of data is initial location of the LP, which could be either the point last seen, given by an



FIGURE 2. SATELLITE PHOTO OF A REGIONAL MAP OF BLACKSBURG, VIRGINIA USED AS A DOMAIN FOR SIMULATED LP.

eyewitness, or the last known point where there is often substantial evidence to place the subject. Since both of these points can move as the search progresses, the initial planning point (IPP) is the recommended choice to use to measure the progress of the search [11]. In the model, the initial positions are assumed to be the IPPs and the horizontal distance from initial to final point

$$D_h = ||x(T) - x(3)|| \tag{1}$$

is the metric used to compare the different behaviors. Notice that we use x(3) as the initial position to enable D_h to go to zero when only SP is used, since agents are necessarily forced to move in the first time step to compute initial velocity.

SIMULATION RESULTS

The model was simulated on a $L \times W = 2000 \times 4000$ gridded map of Blacksburg and the surrounding area with center coordinates $37'16''12^{\circ}N$ $80'21''36^{\circ}W$. A satellite image of the map is shown in Figure 2, which covers an area of $13.3 \text{ km} \times 26.7 \text{ km}$. The resolution of each square block is approximately $L_c = 6.7 \text{ meters}$. Simulations were run for the time it would take for a person to walk an hour with an average pace of 4 kilometers an hour and with a time step of 6 seconds, that is T = 600.

The elevation of the map was procured from the GIS data and is seen in Figure 3 with the locations of IC1 = [1406, 906] and IC2 = [1967, 918] marked. We used MATLAB version 2018a to analyze the map and simulate the model. Using the Canny Edge Detection method in MATLAB, the linear features were determined based on the intensity of the gradient of elevation. The black points in Figure 4 represent the resulting linear features used in the simulations.

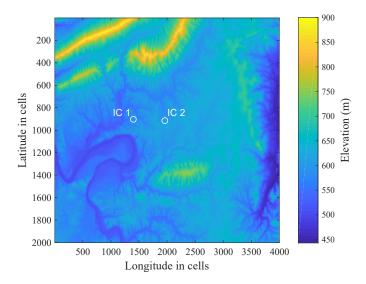


FIGURE 3. ELEVATION MAP IN METERS WITH TWO INITIAL CONDITIONS MARKED, IC1 AND IC2, AS WHITE CIRCLES. THE FIRST INITIAL CONDITION IS LOCATED NEAR A PROMINENT LINEAR FEATURE AND THE SECOND INITIAL CONDITION IS NOT.

To be able to realistically model the behavior of a lost person, all possible LPT PMFs for the four behaviors need to be simulated, since a single person might employ a variety of different strategies while lost. We increment the probability of each strategy between 0 and 1 by steps of 0.1, and retain valid PMFs (that is, which sum to 1). The result is 266 unique LPTs with varying proportions of random walking, route traveling, distance traveling, and staying put.

We run the simulation over 100 replicates, and we compute D_h for all 266 distributions and two initial conditions IC1 and IC2. In Figure 5, the different distributions for the four strategies are shown for each of the 266 different types in the top plot. The second and third plots depict the distances from the initial condition averaged over 100 replicates with corresponding error bars for the standard deviation. By comparing the distances to their distributions, we can look for patterns in the behavioral types.

In Figure 5, the horizontal distances for each type of LP are shown below the first plot. The patterns in this figure show the correlation between the different distributions and how far the LP will travel from the IPP. As a sanity check, we observe that, when the only strategy used is staying put (all yellow in the top plot), the distance from the initial point is zero. In contrast, when the distribution is only direction traveling (all green in the top plot), the distance jumps to over 4000 meters, which is much higher than any of the other D_h values. Another pattern shows that the distance decreases as route traveling increases with respect to distance traveling and staying put decreasing (increasing pattern of light blue in the top plot). However, when only route traveling

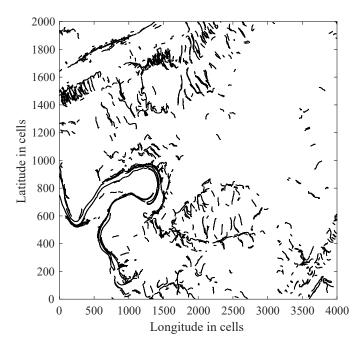


FIGURE 4. RESULTS OF THE CANNY EDGE DETECTION USED TO FIND THE LINEAR FEATURES OF THE DOMAIN.

is used, we see that the distance is increased in IC1 in comparison to IC2 due to the fact that the first initial condition is placed near a linear feature and the second is not. In this case, the agent near IC2 actually performs random walking since route traveling is not possible. In general, as more random walking occurs (more dark blue in the top plot), the distances tend to decrease to a more constant distribution in contrast to the saw-tooth patterns seen in the earlier distributions. These saw-tooth configurations occur when staying put decreases while direction traveling increases, causing the distance to increase and abruptly drop off once staying put becomes dominant.

Validation

In order to evaluate the validity of the model, a metric is needed to compare different distributions of behaviors. In his book, Koester has categorized types of lost person subjects from ISRID along with statistics from incident reports [11]. The tables list distances between where the LP is found and the IPP for 25%, 50%, 75%, and 95% quartiles of cases, with 50% representing the median. These distributions denote a cumulative distribution (CDF) of distance for each type of lost person and can be directly compared to the PMFs of D_h generated by repeatedly simulating our model. We transform the CDFs of the person types into PMFs on the same support as our model PMFs for comparison.

To define a metric of distance between the two distributions, the Kullback-Leibler (KL) divergence metric is used [15]. For

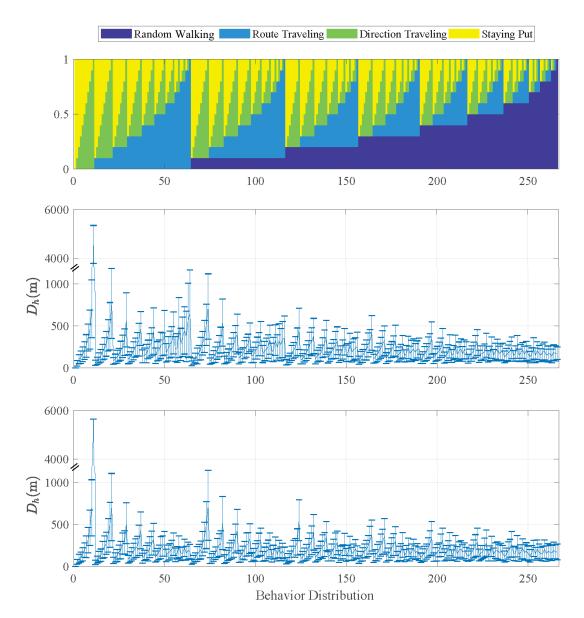


FIGURE 5. (TOP) PROBABILITY DISTRIBUTIONS FOR ALL 266 LPTS WITH CORRESPONDING D_h IN METERS AVERAGED OVER 100 REPLICATES BELOW. THE DISTANCES ARE SHOWN FOR INITIAL CONDITION 1 (MIDDLE) AND INITIAL CONDITION 2 (BOTTOM). ERROR BARS SHOW \pm ONE STANDARD DEVIATION OVER REPLICATES.

two discrete distributions P and Q, the KL divergence is defined as:

$$D_{\mathrm{KL}}(P \parallel Q) = \sum_{x \in \mathcal{X}} P(x) \log \left(\frac{P(x)}{Q(x)} \right) \tag{2}$$

where $x \in \mathcal{X}$ are realizations of the distributions. Because the KL divergence is applied to PMFs, it is necessary to compute the probability mass functions for the CDFs from [11]. First, we

normalize the model distributions by making the set of values of D_h standard normal and compute a normalized PMF for each of the 266 types. Then, we define the database LPT as a CDF of D_h by interpolating the CDF over a linear space with 100 steps spanning the largest possible distance reported in [11]. We differentiate this interpolated CDF to find the unnormalized database PMF. After making this PMF standard normal by subtracting the mean and dividing by the standard deviation, we interpolate both the model and database distributions over their overlapping bins.

This is to ensure that neither distribution has any zero entries, which may cause the KL divergence to go to infinity. Finally, we compute the metric

$$D_{\mathrm{KL}}(P \parallel Q) + D_{\mathrm{KL}}(Q \parallel P) \tag{3}$$

where P is each database PMF and Q is each of the LPT distributions from the model. Notice that this quantity is always positive, unlike the KL divergence in (2). We use this measure of difference to show which model LPT fits the best with the distributions given in [11].

Comparison with real LP data

The five different behavior categories selected from [11] are Autism, All Terrain Vehicle (ATV), Child, Dementia, and Vehicle. A LP in the autistic LP category will often show impairment in social interaction and difficulty in communication. They will often travel randomly to follow the path of least resistance, stay put by hiding in structures for extended periods of time, or travel in a direction due to a fascination with water, lights, reflections, and an intense interest in transportation [11]. The ATV LP category includes subjects who were driving an ATV for recreation or some activity and became lost. Some of the possible strategies used are route traveling, direction traveling, and many times, staying put. Most cases reported are due to trauma which can disable the ATV. The Child LP category describes children from the ages 7-9 years old. At this age, directional skills are much more developed than children under six, and they are usually able to give directions and draw maps. Children in this age range become lost because they will often try to take shortcuts, and will frequently use the route traveling strategy. The Dementia LP category includes LPs with a propensity to keep traveling straight ahead until they get stuck. They appear to lack the ability to turn around and have also been seen to just stay put when they are stuck. In more severe cases, a LP with dementia will be seen to random walk. The Vehicle LP category refers specifically to LPs originating at an abandoned vehicle. This category represents individuals who leave a vehicle to proceed on foot, which may be the result of an accident, the vehicle becoming stranded, or running out of fuel. Many of the LPs will follow roads or linear features when traveling on foot [11].

In Figure 6, the five different database PMFs for the chosen behavior categories are shown plotted against the best fitting model PMFs for the two initial conditions. Using the KL divergence in (3), we found the model LPTs with the smallest difference from each of the database, or experimental, PMFs. The corresponding distributions for the best fits in Figure 6 are listed in Table 2 for IC1 and IC2. In the figure, we see that the model PMFs are consistent between the ICs and some of them match the experimental data better than others. The Autism, Child, and Dementia categories all tend to have similar shapes, including

light tails. When we look at Table 2, we see that even though the D_h plots from Figure 5 seemed to match well between the different ICs, different strategies are used for the experimental LPTs. For example, in the Autism category for IC1, 90% of the behavior comes from route traveling and 10% from staying put. In contrast, for IC2, while 50% is still route traveling, we now see some random walking and direction traveling as well. In the ATV category, the distributions are more similar with 50% direction traveling in both IC1 and IC2 with the other behaviors also alike in percentages. For the Child category, we see the same distributions between the ICs as the Autism LPT. There is a large difference in the ICs for the Dementia category with 90% of the behavior coming from route traveling in IC1 and 0% of route traveling in IC2. Lastly, for the Vehicle LPT, we also see a large difference in strategies used with 60% direction traveling in IC1 and 0% of the same strategy in IC2. Even though the D_h plots looked similar for these initial conditions, the results from these best fit computations tell us otherwise.

TABLE 2. BEST FIT BEHAVIOR PMFS FOR LOST PERSON BEHAVIOR TYPES (RW, RT, DT, SP). ALSO SHOWN IN FIGURE 6.

LPT	Init Cond 1	Init Cond 2
Autism	(0,0.9,0,0.1)	(0.1, 0.5, 0.2, 0.2)
ATV	(0.2, 0.2, 0.5, 0.1)	(0.1, 0.1, 0.5, 0.3)
Child	(0,0.9,0,0.1)	(0.1, 0.5, 0.2, 0.2)
Dementia	(0,0.9,0.1,0)	(0.2,0,0.3,0.5)
Vehicle	(0.3, 0.1, 0.6, 0)	(0.2, 0.6, 0, 0.2)

DISCUSSION

The agent-based model developed in this paper attempts to model the dynamics of a lost person in the wilderness by using different behavioral types and statistics from a database of lost person cases [11]. In search and rescue operations, knowing the type of lost individual can help predict the strategies that may be utilized to find one's way and gain insight into the goals of the LP. Using this model in conjunction with UAVs will hopefully make the SAR process more efficient and reliable.

By using the distance from the IPP as a metric, we can see patterns in the way a LP might behave and the total displacement they will travel. Generally, when traveling randomly, a LP would likely travel a shorter distance versus when they are traveling in one direction. Based on the description from [11], a LPT for Autism will frequently travel randomly or in one direction. However, from the comparison with the model PMFs, we

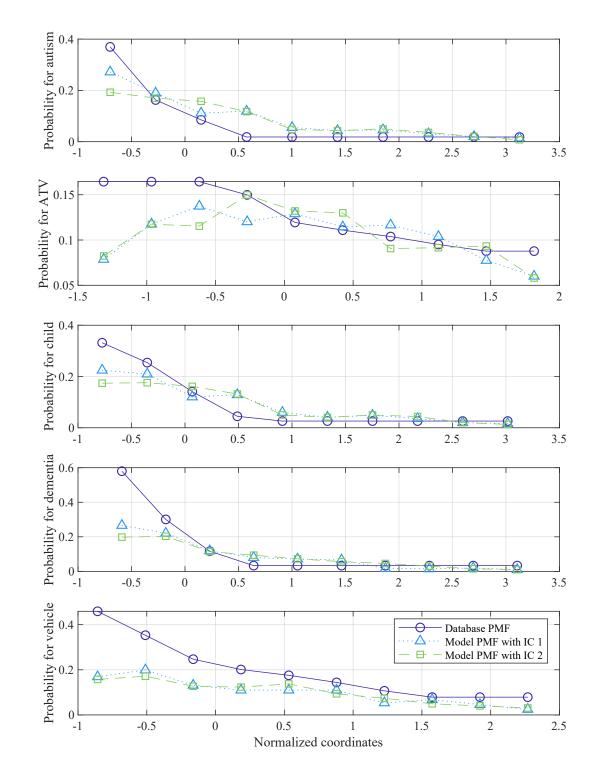


FIGURE 6. NORMALIZED PROBABILITY MASS FUNCTIONS FOR EACH OF THE FIVE CHOSEN LOST PERSON TYPES FROM THE DATABASE (AUTISM, ATV, CHILD, DEMENTIA, VEHICLE) PLOTTED WITH THE BEST FITTING MODEL PMFS FOR BOTH INITIAL CONDITIONS.

see that the best fit behaviors are route traveling and staying put. Similarly with a LPT for Dementia, the typical behavior expected by the SAR community will often be direction traveling, but in the comparison in Table 2, the best fit PMFs have little direction traveling.

These discrepancies between the model and the ISRID statistics make it difficult to draw hard conclusions from this study. This lack of correspondence may signify that this model is still in an early form and can be improved with more behaviors that LPs are known to use, like view enhancing where the LP moves uphill to visually orient herself or backtracking where the LP follows a previously traversed route. In addition, we must consider that the statistics from the database report the distributions where LPs were found, which may not be consistent with where they would be after an hour of travel, which is what we simulated. Incorporating the length of searches into this research in the future may enable a more fair comparison.

An interesting aspect of the model is that the plots of the distances from the IPP tend to be the same even though different behavioral strategies are used. Thus, the distributions of strategies are not unique since a LP can get to the same place by doing different things. To improve this model, the ability to change behaviors while traveling should incorporate more realistic behaviors, especially if considering elevation data and the tendency for people to move downhill.

CONCLUSIONS

In this paper, we introduce an agent-based model of lost person behavior that defines agents moving on a landscape with behavior considered as a random variable. By using a distribution of four known lost person behavior strategies to simulate possible trajectories for the agent, we compute distributions of horizontal distances traveled in a fixed time and compare these results to analogous data from a database of lost person cases in order to explore the model's validity with respect to real-world data.

The next steps for the model are clear. Including more lost person strategies which consider the landscape beyond only linear features, accounting for the effect of fatigue, and performing simulations over longer time intervals will paint a more realistic picture of the actual behaviors used by lost people. Making these improvements will give more accurate trajectories and hopefully decrease the amount of time a SAR operation takes.

ACKNOWLEDGMENT

This work is supported by the National Science Foundation under grant 1830414.

REFERENCES

- [1] 2018 NCIC Missing Person and Unidentified Person Statistics. https://www.fbi.gov/file-repository/2018-ncic-missing-person-and-unidentified-person-statistics.pdf/view. (Accessed on 04/10/2019).
- [2] Seyfried, A., Schadschneider, A., Kemloh, U., and Chraibi, M., 2011. "Force-based models of pedestrian dynamics". *Networks and Heterogeneous Media*, 6(3), pp. 425–442.
- [3] Moussaïd, M., Helbing, D., and Theraulaz, G., 2011. "How simple rules determine pedestrian behavior and crowd disasters". *Proceedings of the National Academy of Sciences*, 108(17), pp. 6884–6888.
- [4] Fuchs, A., and Jirsa, V. K., 2007. *Coordination: neural, behavioral and social dynamics*. Springer Science & Business Media.
- [5] Karamouzas, I., Skinner, B., and Guy, S. J., 2014. "Universal power law governing pedestrian interactions". *Physical Review Letters*, *113*(23), p. 238701.
- [6] Seyfried, A., Steffen, B., and Lippert, T., 2006. "Basics of modelling the pedestrian flow". *Physica A: Statistical Mechanics and its Applications*, **368**(1), pp. 232–238.
- [7] Bonabeau, E., 2002. "Agent-based modeling: Methods and techniques for simulating human systems". *Proceedings of the National Academy of Sciences*, *99*(3), pp. 7280–7287.
- [8] Couzin, I. D., Krause, J., James, R., Ruxton, G. D., and Franks, N. R., 2002. "Collective memory and spatial sorting in animal groups". *Journal of Theoretical Biology*, 218(1), pp. 1–11.
- [9] Lin, L., and Goodrich, M. A., 2010. "A Bayesian approach to modeling lost person behaviors based on terrain features in wilderness search and rescue". *Computational and Mathematical Organization Theory*, 16(3), pp. 300–323.
- [10] Mohibullah, W., 2017. "Agent-based lost person movement modelling, prediction and search in wilderness". PhD thesis, UCL (University College London).
- [11] Koester, R. J., 2008. Lost Person Behavior: A search and rescue guide on where to look for land, air and water. dbS Productions LLC.
- [12] International Search and Rescue Incident Database (IS-RID). http://www.dbs-sar.com/. (Accessed on 04/10/2019).
- [13] Environmental systems research institute (ESRI). https://www.esri.com/en-us/home. (Accessed on 04/10/2019).
- [14] Find edges in intensity image MATLAB edge. https://www.mathworks.com/help/images/ref/edge.html. (Accessed on 04/8/2019).
- [15] MacKay, D. J., and Mac Kay, D. J., 2003. *Information the-ory, inference and learning algorithms*. Cambridge University Press.