

Modeling the Dynamics of Hurricane Evacuation Decisions from Twitter Data: An Input Output Hidden Markov Modeling Approach

Kamol Chandra Roy

Ph.D. Student

Department of Civil, Environmental, and Construction Engineering
University of Central Florida
12800 Pegasus Drive, Orlando, FL 32816
Phone: 321-318-2067
Email: roy.kamol@knights.ucf.edu

Samiul Hasan

Assistant Professor

Department of Civil, Environmental, and Construction Engineering
University of Central Florida
12800 Pegasus Drive, Orlando, FL 32816
Phone: 407-823-2480
Email: samiul.hasan@ucf.edu
(Corresponding Author)

*[This is the pre-print version of the following article: Roy, K. C., & Hasan, S. (2021). Modeling the dynamics of hurricane evacuation decisions from twitter data: An input output hidden markov modeling approach. *Transportation Research Part C: Emerging Technologies*, 123, 102976, which has been published in final form at <https://doi.org/10.1016/j.trc.2021.102976>.]*

ABSTRACT

Evacuations play a critical role in saving human lives during hurricanes. But individual evacuation decision-making is a complex dynamic process, often studied using post-hurricane survey data. Alternatively, ubiquitous use of social media generates a massive amount of data that can be used to predict evacuation behavior in real time. In this paper, we present a method to infer individual evacuation behaviors (e.g., evacuation decision, timing, destination) from social media data. We develop an input output hidden Markov model (IO-HMM) to infer evacuation decisions from user tweets. To extract the underlying evacuation context from tweets, we first estimate a word2vec model from a corpus of more than 100 million tweets collected over four major hurricanes. Using input variables such as evacuation context, time to landfall, type of evacuation order, and the distance from home, the proposed model infers what activities are made by individuals, when they decide to evacuate, and where they evacuate to. To validate our results, we have created a labeled dataset from 38,256 tweets posted between September 2, 2017 and September 19, 2017 by 2,571 users from Florida during hurricane Irma. Our findings show that the proposed IO-HMM method can be useful for inferring evacuation behavior in real time from social media data. Since traditional surveys are infrequent, costly, and often performed at a post-hurricane period, the proposed approach can be very useful for predicting evacuation demand as a hurricane unfolds in real time.

Keywords: Hurricane evacuation, Hurricane Irma, Social media, Input output hidden Markov model, Twitter, Florida.

1. INTRODUCTION

Extreme weather events have become more common these days due to climate change and other related causes. These extreme events have caused significant physical and socio-economic losses (Guha-sapir et al., 2017; Hasan and Foliente, 2015; Re, 2013; Tousignant Lauren, 2017). A real-time demand-responsive evacuation system is essential to save human lives and minimize losses (Murray-Tuite and Wolshon, 2013). Traditionally, post-hurricane surveys are conducted to collect data to understand and predict the evacuation behavior of a population. Such surveys are costly, time consuming, and not effective to manage evacuation when a hurricane unfolds in real time (Chaniotakis et al., 2017). With the ubiquitous use of social media platforms (e.g. Twitter, Facebook etc.), a massive volume of real-time data is available. Such data can provide valuable insights on individual behavior during extreme events such as a hurricane (Sadri et al., 2017a; Wang and Taylor, 2014; Xiao et al., 2015).

Thus, large-scale social media data can be used for a better understanding of evacuation behaviors during hurricanes (Martín et al., 2017). However, one of the major challenges of using social media data is to reliably model evacuation decisions from such data. To date, the studies investigating social media data are limited to inferring evacuation choices. These studies (Chaniotakis et al., 2017; Martín et al., 2017) have mainly adopted clustering approaches that locate a user during pre-evacuation and evacuation periods. A recent case study (Kumar and Ukkusuri, 2018) on hurricane Sandy Twitter data shows the relationship between social connectivity and evacuation decision without specifically modeling the real-time dynamics of evacuation decision-making. Using geotagged Facebook data from hurricane Irma, Harvey, and Maria, another recent study (Metaxa-Kakavouli et al., 2018) has analyzed the influence of social ties on evacuation behavior. Although these studies have demonstrated the significant potential of using location-based social media data in an evacuation context, they have not developed any modeling framework that can answer *what*, *when*, and *where* users participate in different activities during a hurricane.

In this paper, we present a modeling approach for understanding the dynamics of hurricane evacuation from social media data. In particular, we have developed an input-output hidden Markov model (IO-HMM) to infer evacuation behavior from Twitter data. We have gathered large-scale Twitter data during hurricane Irma and used the spatio-temporal and contextual sequences from this data to run the proposed model. Hurricane Irma, the largest storm ever recorded in the Atlantic Ocean, made its landfall on the southern coastal areas of Florida. The storm generated a massive amount of social media posts nationwide, especially in Florida. This paper has the following contributions:

- We implement a process to gather hurricane evacuation information from geo-tagged Twitter data. We validate the results by manually checking locations and tweet texts of the users. As traditional survey data is costly and often confined with small geographic region, this type of data can be used for understanding evacuation behavior during hurricane to complement traditional approaches.
- We develop a Word2Vec model to extract contexts based on the tweets collected from multiple hurricanes (Sandy, Matthew, Harvey, and Irma). The model has been trained using

more than 100 million tweets having about 882.54 million words (after filtering out the stop words, punctuations, emoticon, URLs). This model can contribute in future research to determine disaster contexts from Twitter data.

- We develop an input output hidden Markov model from the sequences generated from user tweets. To the best of our knowledge, this is one of the first studies that use social media data for modeling the dynamics of hurricane evacuation decisions. The model can capture the dynamics of hurricane evacuation by answering what, when, and how users participate in different activities during a hurricane.

2. LITERATURE REVIEW

During a hurricane, timely evacuation is critical to reduce hazard risks and save human lives (Baker, 1979). Despite the importance of evacuation, some people choose not to evacuate (Whitehead et al., 2000). Therefore, a thorough understanding of the determinants of evacuation behavior is needed to protect the loss of lives, especially for the vulnerable communities (Hasan et al., 2011b). Many studies have investigated population response during hurricanes from different perspectives, particularly focusing on evacuation choices (Murray-Tuite and Wolshon, 2013). These topics include: evacuation decision making (Gladwin et al., 2001; Hasan et al., 2011b; Kang et al., 2007; Yang et al., 2019), evacuation time (Hasan et al., 2013; Lindell, 2008; Rambha et al., 2019), evacuation demand (Xu et al., 2016), destination choice (Mesa-arango et al., 2013), and mode and route selection (Sadri et al., 2015, 2014). However, most of these studies are based on post-disaster household surveys collecting information on population behavior instead of real-time dynamics. Studies (Lin et al., 2009; Ukkusuri et al., 2017; Yin et al., 2014) have developed high fidelity agent-based models to predict population responses in future hurricanes. One of the major shortcomings of these models is that factors influencing evacuation decisions do not change over time. Although a few models (Fu and Wilmot, 2004; Sarwar et al., 2018) considered the dynamics of evacuation decision-making process, these models depended on post-disaster surveys, mainly focusing on household characteristics with limited transferability (across regions, communities, and disaster contexts) (Hasan et al., 2011a; Martín et al., 2017). Survey data have limitations in capturing the dynamic nature of the evacuation decision-making process (Murray-Tuite et al., 2019).

However, hurricane response is a dynamic event with significant changes and uncertainties involving parameters beyond household characteristics. During a hurricane, emergency agencies and weather services issue frequent advisories providing information on the hurricane's projected trajectory and category, wind speed, rainfall, storm surge, evacuation warning etc. Local and national news channels disseminate information on the present condition of the hazard and traffic situation. Context awareness, considering all these dynamic factors, plays a critical role for a large number of populations to decide whether to leave or not. Lee et al. explored the dynamics of visiting patterns to the weather-related websites during Hurricane Katrina (Lee et al., 2009). Yabe et al. developed a web-search query-based evacuation prediction model (Yabe et al., 2019). These studies mainly focused on understanding risk perceptions without modeling the spatial-temporal dynamics of evacuation behavior. As an alternative to relying on static post-disaster surveys, dynamic predictive models can be built employing real-time information received from multiple sources including individuals, transportation facilities, and emergency services. For instance,

Meyer et al. studied the dynamics of risk perception by using survey data collected during an approaching hurricane (Meyer et al., 2014). Studies also developed a physics-based hazard modeling approach to simulate evacuation uncertainty considering the physical interaction among multiple hazard components (Blanton et al., 2020; Davidson et al., 2020). However, these studies were based on simulated environments and did not use real-time information available from different sources. Evacuation models can utilize the vast amount of streaming data available from social media, giving us real-time insights on individual actions during evacuations (Chaniotakis et al., 2017; Kryvasheyeu and Chen, 2015; Martín et al., 2017; Sadri et al., 2017b).

Recently, the role of social media in a disaster management context has gained a significant attention, mainly from the perspectives of crisis communication (Lachlan et al., 2016; Roy et al., 2020; Sadri et al., 2017b, 2017a), human mobility analysis (Beiró et al., 2016; Pan et al., 2013; Roy et al., 2019; Yanjie Duan et al., 2016), nowcasting damage assessment (Kryvasheyeu et al., 2016), and event detection (Dong et al., 2015; Kryvasheyeu and Chen, 2015). However, its potential in understanding evacuation behavior is still underexplored. Existing studies on inferring evacuation decisions from social media data found home locations and displacements to determine if a user has evacuated or not. Chaniotakis et al. (Chaniotakis et al., 2017) used a density based clustering approach to identify home and geotagged tweet counts during an evacuation order to identify evacuation decision. Using hurricane Matthew data, Martin et al. (Martín et al., 2017) showed that Twitter data can be used to understand evacuation compliance behavior. This study considered user median locations during a normal period as their homes and median locations during a hurricane as their evacuation destinations. Using similar approach on hurricane Sandy twitter data, Kumar and Ukkusuri (Kumar and Ukkusuri, 2018) studied the evacuation decision of New York City residents in relation to the social connection of the users, distance from coastline, and time to evacuation. They have found that higher number of social ties (number of friends, followers) decrease the likelihood to evacuate. A recent study using Facebook data of hurricane Irma, Harvey, and Matthew found a similar result that social ties decrease the likelihood to evacuate (Metaxa-Kakavouli et al., 2018). However, these studies did not capture the dynamics of individual evacuation decisions requiring a modeling framework that can infer evacuation choices from geo-location data.

In this paper, we present an input output hidden Markov model to infer evacuation behavior from Twitter data. Hidden Markov models (HMMs) relate a sequence of observations to a sequence of hidden states that explain the observations (Ghahramani and Jordan, 1996). HMMs have been widely used in speech recognition (Rabiner, 1989), protein topology (Krogh et al., 2001), social science (Eagle and Pentland, 2006), and activity modeling (Yin et al., 2017). HMMs have been used to classify activity categories considering spatiotemporal features (Ye et al., 2013) and to determine activity-location sequence from geo-location data (Hasan and Ukkusuri, 2017). Duong et al. (Duong et al., 2005) introduced a switching hidden semi-Markov model for online activity recognition and abnormality detection. Input-output hidden Markov model is an extension to the standard hidden Markov model for using the HMM in a supervised fashion (Bengio and Frasconi, 1995). IO-HMM has shown the added advantages over HMM to map the output sequences with the inputs in studies such as audio-visual mapping (Bengio and Frasconi, 1995),

price forecasting (González et al., 2005), hand-gesture (Marcel et al., 2000) etc. Yin et. al (Yin et al., 2017) proposed an IO-HMM based modeling framework to infer urban activity patterns.

3. DATA PREPROCESSING AND DESCRIPTION

In this study, for inferring evacuation choices from social media posts, we have used Twitter data from hurricane Irma. Using its streaming API, we collected around 1.81 million tweets made by 248,763 users between September 5, 2017 and September 14, 2017. We collected the data using a bounding box covering Florida, Georgia, and South Carolina. To obtain user activities during a pre-disaster period, we also collected user-specific historical data using Twitter's rest API which allows to collect the most recent 3,200 tweets for a given user. We collected user specific data for 19,000 users who were active for at least three days between the day the first evacuation order was issued and the landfall day, so that we have enough data for capturing the activity dynamics during the evacuation.

For our analysis, we have considered only the tweets with geo-location information. The geolocation information is provided either as a point (latitude, longitude) or a bounding box (area defined by two latitude and longitudes pairs). The point location is the exact location whereas the bounding box has different level of precision of where a tweet has been posted. We use the center point of a bounding box as the latitude and longitude of that place. To convert all the locations to a region under a geocoding system and to protect the privacy of the users, we have used geohash geocoding system with a precision of ~5 kilometers. Geohash converts a latitude, longitude pair into a short string of letters and digits depending on the precision (length of the strings) (Balkić et al., 2012). In our study, we have used a geohash of length 5, which is equivalent to a region surrounded by $\sim 5 \text{ km} \times 5 \text{ km}$ area and has a reasonable resolution to capture the spatial dynamics.

3.1 Preparing Evacuation Data

From the historical tweets of a user, we extracted the most visited place during office hours (9:00 AM to 6:00 PM) on weekdays and the most visited place during nighttime (10:00 PM to 7:00 AM). For each user, we assigned the most frequent office hour place and night hour place as office location and home location, respectively. For some users, the office and home location can be same because users may not be a worker or may have their offices within 5 km from home.

Every year Florida attracts millions of visitors from home and abroad. We adopt several steps to remove the users who came from outside of Florida (international visitors and domestic users coming from states other than Florida). Through the filtering steps, we consider only the users whose home and office locations are within Florida, whose evacuation distance is less than 2,400 km (chosen based on the literature (Cheng et al., 2008; Han et al., 2019)), and who have returned to their home after the landfall.

In this study, we have focused on capturing the evacuation demand that is most likely to affect traffic flows on highways. Short distance evacuations (e.g., going to a nearby shelter) are not likely to impact highway traffic. Also, previous studies found that short distance evacuations are only a small percentage of the total evacuation count. During hurricane Floyd, very few evacuations were found less than 50 miles ($\sim 80.5 \text{ km}$); about 3.5% of the respondents chose a shelter or a church as an evacuation destination (Cheng et al., 2008). Based on hurricane Matthew

Twitter data, a recent study (Han et al., 2019) has found that evacuees are likely to move more than 200 km for an evacuation. During hurricane Irma, only 4% of the respondents were found to evacuate to a shelter (Wong et al., 2018, 2020). Moreover, some of the geotagged tweets do not have the necessary granularity (tweets with locations as a bounding box) to detect short distance evacuation. Thus, we select a threshold of 200 km to identify evacuation. After returning home, a user may not have any tweets posted from her home but may have posted from nearby locations. Thus, we select a 20 km distance threshold from someone's home to identify the return of an evacuee.

Starting from the beginning (9 days prior to landfall) of the location sequence to the landfall day, if a user has not tweeted from home or office but tweeted from somewhere else with a displacement of at least 200 kilometers, we consider that the user evacuated and the corresponding time as evacuation time. After landfall, a return is considered as the time when an evacuated user is first seen within the 20 kilometers from her home or office. We collect the information on evacuation orders from the official Twitter account of each county. We have considered the timings of the evacuation orders issued by each county. So, if someone evacuates before the first official order, it is considered as an evacuation without an official order. We have found that 252 users have evacuated among 2,571 identified Florida users.

We have manually checked the results of the above approach of identifying a user's home location and evacuation (if any), it's destination, and timing. Please see the supporting information section for details of the manual checking process. We compare the results from the manual checking process with the results obtained from this approach. We find that both the results match with respect to evacuation time and displacement traveled during evacuation. We use this resulting data as labeled dataset for the purpose of model estimation and validation. After all the processing, our final dataset contains 38,256 geotagged tweets, posted by 2,571 users from Florida. For each user, we created a sequence of his/her tweets and the corresponding locations posted between September 2, 2017 and September 19, 2017.

3.2 Data Exploration

Figure 1 shows the origins and destinations of the evacuated users. Here the identified home location of an evacuee is considered as the origin and the evacuation destination place is considered as the destination. Figure 1 shows the result of 252 Florida-based users after filtering out the tourists/visitors. Residents of Florida evacuated to Georgia (Atlanta was one of the major destinations), Alabama, South Carolina, and North Carolina. Some users (at right bottom of Figure 1, near coast) moved to places that are closer to the coast than before. This is reasonable as the projected path of hurricane Irma changed overnight on September 8, 2017. Initially, Irma was expected to hit from the east coast of Florida, but later it changed its path and was predicted to hit from the west coast. These results seem plausible according to the news updates from different sources during hurricane Irma (Luz Lazo, 2017; Marshal, 2017). The majority of the evacuees were from Miami, Tampa, West Palm Beach etc. (see Figure 1), where mandatory evacuations were ordered.

Figure 2 shows the distribution plot of evacuation time and return time of the users who evacuated during Irma. Figure 2(a) shows the marginal and joint frequency distributions of

evacuation time and return time; the top histogram along x axis shows the distribution of evacuation count in 24-hour intervals; the right histogram along y axis shows the distribution of return time in 24-hour intervals; each cell in the heatmap shows both evacuation count and return count with respect to the corresponding 24-hour evacuation interval on x axis and return interval on y axis.

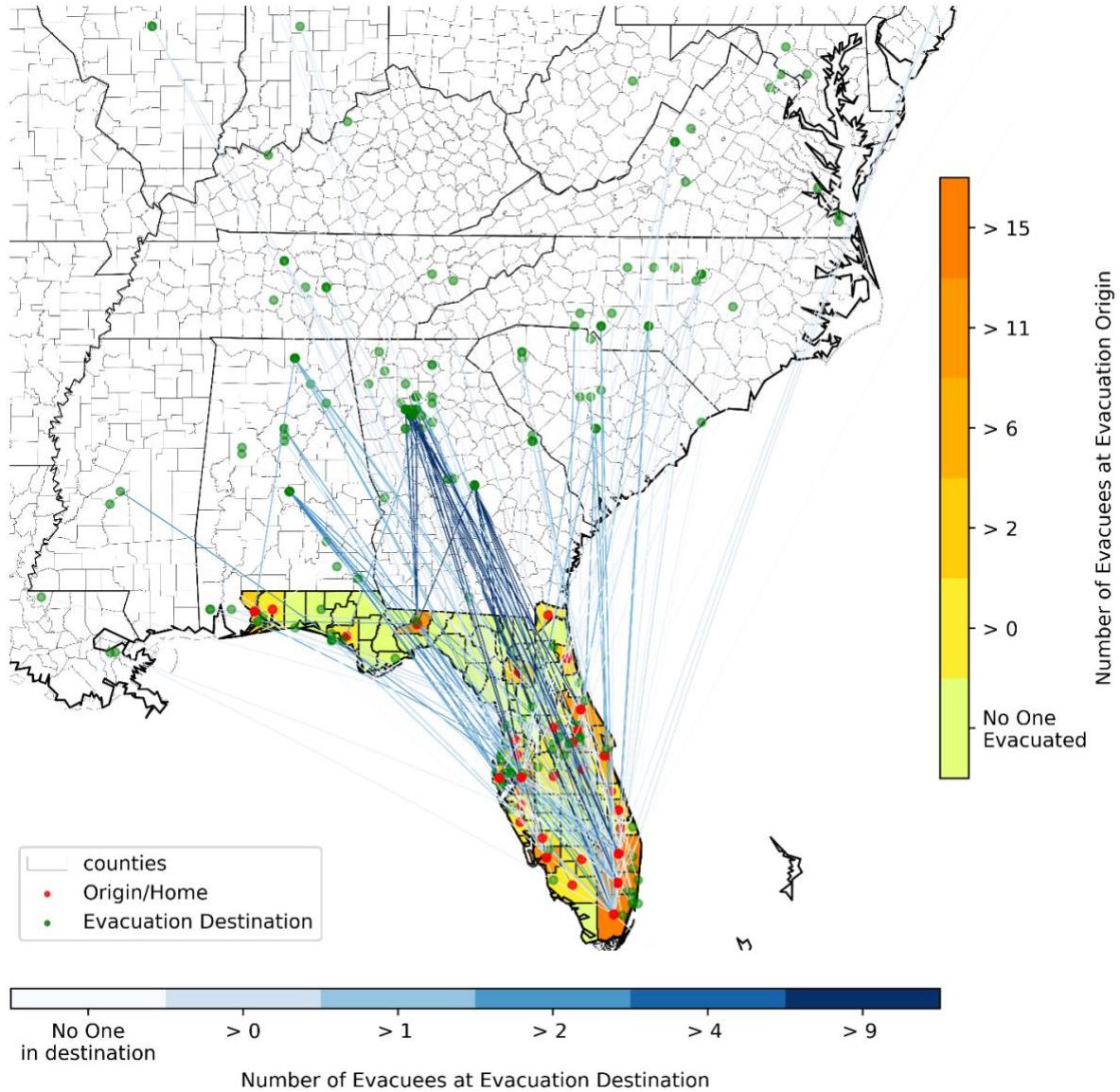
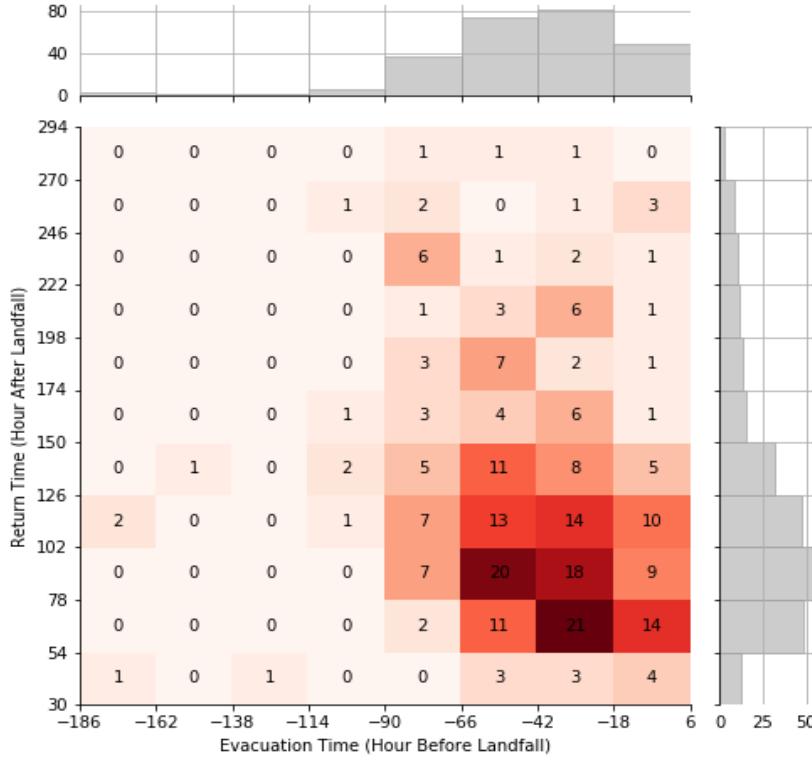


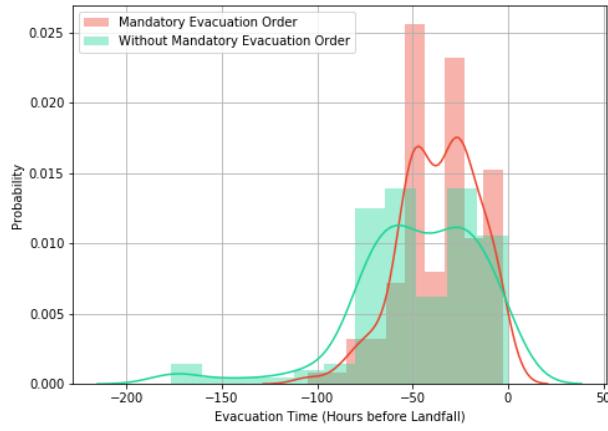
FIGURE 1 Evacuation Origin and Destination

Figure 2(b) and 2(c) show the probability distributions of evacuation and return time considering the type of evacuation order received. The evacuation time and return time are expressed as the time difference from landfall time (September 10, 2017), a negative value indicates a period before the landfall and a positive value indicates a period after the landfall. Most evacuees left within 100 hours before the landfall (September 10, 2017); 18 to 42 hours before landfall was the most frequently chosen evacuation time window. On the other hand, 78 to 102

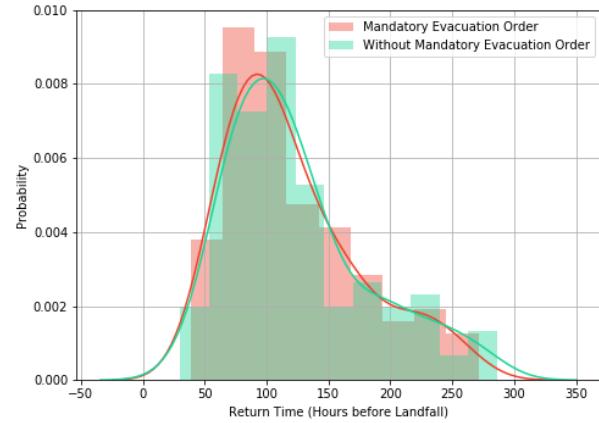
hours after the landfall was most frequently chosen return time window. People started evacuating before the official evacuation order (see Figure 2 (b)). Although the pattern of evacuation time is different for voluntary and mandatory orders, the patterns of return times are almost similar (see Figure 2 (c)). The resulting distributions are aligned to the actual evacuation time and return time according to the concurrent news reports during hurricane Irma (ABC News, 2017; FLKEYSNEWS, 2017).



(a)



(b)



(c)

FIGURE 2 Distributions of Evacuation Time and Return Time during Hurricane Irma,

(a) Joint distribution of evacuation time and return time (b) Probability distribution of evacuation time for mandatory and without mandatory evacuation order, and (c) Probability distribution of return time for mandatory and without mandatory evacuation order.

4. METHODOLOGY

We have used an Input Output Hidden Markov Model (IO-HMM) to identify activity sequence during a hurricane. We compare the results with a standard Hidden Markov Model (HMM). The model structures are shown in Figure 3. The IO-HMM is similar to HMM, but it maps the input sequence to output sequences and applies the expectation maximization algorithm (EM) in a supervised fashion.

In an HMM modeling framework, the system being modeled follows a Markov process with unobserved (i.e., hidden) states. Figure 3(a) shows a graphical representation of an HMM. The solid circles represent the observed information and the transparent circles represent the hidden state latent variables, in our case the activity types of a user. Here, the hidden states, $(H_1, H_2 \dots H_T)$ are assumed to follow a Markov process that means a hidden state, the probability distribution of H_t depends only on the previous state, H_{t-1} ; i.e., $H_t = f(H_{t-1})$. On the other hand, for the observations $(O_1, O_2, \dots O_T)$, an observation, the probability distribution of O_t depends only on its current hidden state, H_t ; i.e., $O_t = f(H_t)$.

Unlike the standard HMM, in IO-HMM, the probability distribution of hidden state H_t at time t , depends on the previous state H_{t-1} and the input I_t at time t ; i.e., $H_t = f(H_{t-1}, I_t)$. The probability distribution of observation O_t at time t depends on both the hidden state H_t and I_t at time t ; i.e., $O_t = f(H_t, I_t)$ (see Figure 3(b)).

Here, $I_t \in R^m$ is the input vector at time t . $O_t \in R^m$ is an output vector, and $H_t \in \{1, 2, \dots, T\}$ is a discrete state. Similar to HMM, IO-HMM has three set of parameters (θ): initial probability parameters (α), transition model parameters (β), and emission model parameters (γ).

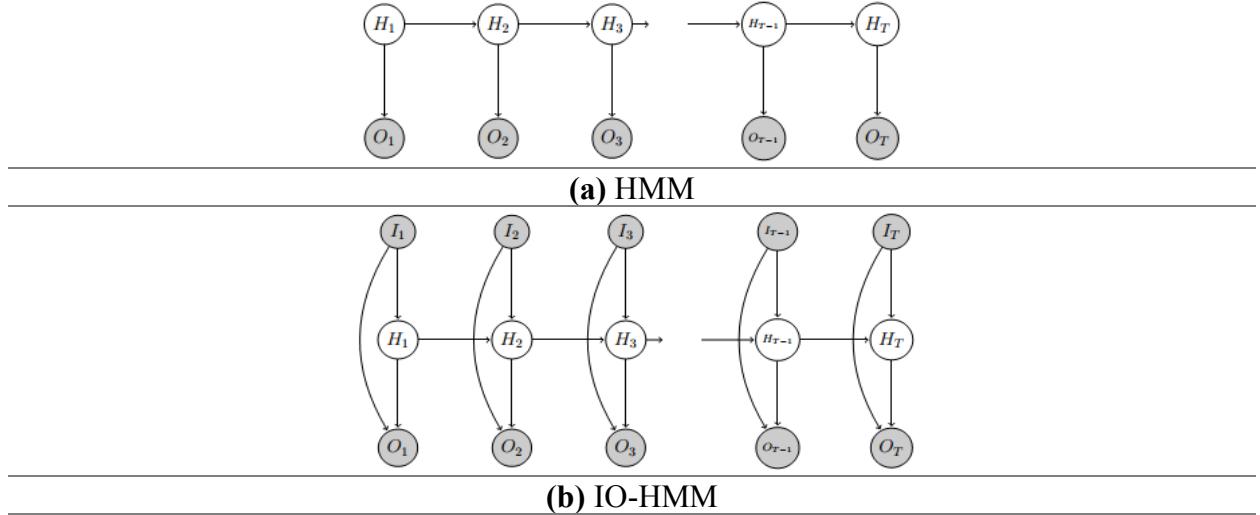


FIGURE 3 Graphical Model Specifying Conditional Independence Properties (a) For a Hidden Markov Model (b) For an Input Output Hidden Markov Model

The likelihood of a data sequence given the model parameters (θ) is given by:

$$L(\theta, O, I) = \sum_H \left(\Pr(H_1 | I_1; \alpha) \cdot \prod_{t=2}^T \Pr(H_t | H_{t-1}, I_t, \beta) \cdot \prod_{t=1}^T \Pr(O_t | H_t, I_t; \gamma) \right) \quad (1)$$

The model parameters are estimated by an expectation maximization algorithm (McLachlan and Krishnan, 2007). For initial and transition models, we have used a multinomial logistic regression model. If we assume that there are k hidden states, the equation of initial probability model becomes the following:

$$\Pr(H_1 = i | I_1; \alpha) = \frac{e^{\alpha^i I_1}}{\sum_k e^{\alpha^k I_1}} \quad (2)$$

where α is a coefficient matrix for initial probability model with α^i represents the coefficients for the initial state being at state i .

The transition from the state i to the state j can be modeled as:

$$\Pr(H_t = j | H_{t-1} = i, I_t; \beta) = \frac{e^{\beta_i^j I_t}}{\sum_k e^{\beta_i^k I_t}} \quad (3)$$

where β represents the transition probability matrices with the (β_i^j) being the coefficients for transitioning to next state j given the current state is i .

For the output model, we have used a linear model for a continuous outcome:

$$\Pr(O_t | H_t = i, I_t; \gamma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(O_t - \gamma_i I_t)^2}{2\sigma_i^2}} \quad (4)$$

where, γ_i represents the emission coefficient when the hidden state is i . For a hidden state i , γ_i and σ_i denote the arrays of model coefficient and standard deviation of the linear model.

And a logistic regression model is used for a categorical outcome:

$$\Pr(O_t | H_t = i, I_t; \gamma_i) = \frac{e^{\gamma_i I_t}}{\sum_k e^{\gamma_k I_t}} ; \text{if } O_t \text{ is categorical} \quad (5)$$

where, γ_i denotes the model coefficient when the hidden state is i .

Detailed descriptions of HMMs and the associated inference algorithms can be found in ref (Rabiner, 1989). The IO-HMM model architecture and its formulation can be found in this ref (Bengio and Frasconi, 1995).

5. MODEL DEVELOPMENT

An IO-HMM model considers data as sequences of inputs and outputs for each user. For that purpose, we need to process the data from raw tweets in that specific form. Figure 4 shows the sequence generation process. For a user, $\{t_1, t_2, \dots, t_T\}$ represent the times of the tweets posted at locations $\{l_1, l_2, l_3, \dots, l_T\}$, respectively. We have collected hurricane related information for each of the location (county level) such as whether the location had a mandatory evacuation order

or not, whether the location had a voluntary order or not, whether the time is before landfall or after landfall. This information is encoded as a binary variable in the sequence. Other information associated with each location is distance from home and time difference from landfall, similarity score of the posted tweet text with the evacuation context words. For simplicity, all information is not shown in Figure 4. We calculate the similarity score by training a word to vector model using tweets from 4 hurricanes (Irma, Matthew, Harvey, and Sandy).

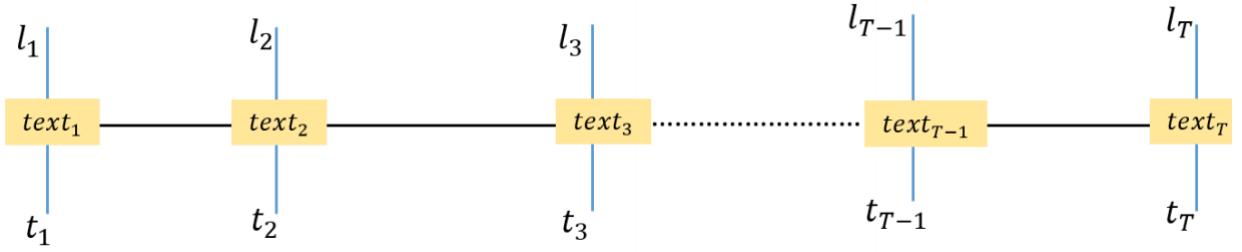


FIGURE 4 Schematic diagram of sequence generation. Here, l_i =location of the user when posting tweet i ; $text_i$ = texts of the tweet i ; t_i = time of the tweet i ; and T = total number of tweets posted.

5.1 Inferring Evacuation Context from a Tweet

In general, the text of a tweet may reflect the underlying context such as hurricane awareness, evacuation intent, information sharing/seeking, power outage etc. We use a similarity score to quantify how similar a tweet is to an evacuation context (e.g., words such as ‘evacuate’, ‘evacuating’, ‘sheltering’). We have used a vector space model called word2vec to learn the word vectors of an evacuation-related tweet.

Vector Space Model (D. E. Rumelhart, G. E. Hinton, 1986) is a natural language processing tool to represent texts as a continuous vector where words that appear in the same contexts share semantic meaning (Sahlgren, 2008),(Baroni and Dinu, 2014). A detailed description of how the model works is given in the supporting information. Once a model is trained, every word in the vocabulary will have a vector representation of a length equal to the vocabulary size (see supporting information). We train a word2vec model using CBOW architecture (please see the supporting information for details) on a corpus of 100 million hurricane-related tweets, collected during multiple hurricanes (Hurricanes Sandy, Harvey, Matthew, and Irma). We calculate the cosine similarity (Huang, 2008) between two word vectors $Word_i$ and $Word_j$ by the following equation:

$$Similarity = \cos(\theta) = \frac{Word_i \cdot Word_j}{||Word_i|| ||Word_j||} \quad (6)$$

Here, $ord_i, Word_j \in \{vocabulary\ in\ the\ model\}$, $k = dimension\ of\ the\ word\ vector$. In our study, to calculate the similarity of a word to an evacuation context, $word_i \in \{'evacuation', 'evacuating', 'sheltering'\}$ and $word_j \in \{words\ in\ a\ tweet\}$. If a window size of n is selected, then score of a tweet is calculated by summing up to n^{th} top score for the words

present in the tweet. We have used the top one score to represent the similarity of a tweet with respect to an evacuation context. Similarity scores of the tweets posted at locations $l_1, l_2, l_3 \dots \dots \dots l_T$ are denoted by $\{Score_{text_1}, Score_{text_2}, \dots \dots \dots, Score_{text_{T-1}}, Score_{text_T}\}$.

6. MODEL ESTIMATION

As we obtain the sequences of all the information needed, we need to specify the inputs and outputs. In IO-HMM, both inputs and outputs are available at the training stage; but after training, the model should infer the outputs given its inputs. In general, the inputs are known before the start of a transition to a new state/activity, but the outputs are not known. In our model, we choose input variables that are likely to influence the decision of a user's next activities. We select 5 input variables including: the time difference from landfall in hours represented as negative to positive where negative means a pre-landfall period (I^1), a binary variable representing a pre-landfall or a post landfall period (I^2), an interaction variable representing the time difference from landfall only for a pre-landfall period (I^3), a binary variable representing if the user's home location is under a mandatory evacuation order (I^4), and a binary variable representing if the user's home location is under a voluntary evacuation order (I^5). As outputs, we choose two variables such as: current location's distance from home (O^1) and evacuation similarity score (word2vec score) of the tweets posted in the location (O^2).

We make several assumptions for selecting the dependencies among the initial, transition, and output models of the IO-HMM structure. We assume that the transition (see Figure 3b) between the hidden states depend on the current state and the input variables I^1 (*time difference from landfall*), I^2 (*post landfall*), I^4 (*user's home is under mandatory evacuation*), I^5 (*user's home is under voluntary evacuation*). We did not explicitly study the research question of what factors impact an individual's evacuation behavior. Rather we model this behavior as part of an activity dynamics process, where evacuation is considered as an activity type. The coefficients of these five input variables corresponding to an evacuation activity transition indirectly capture the factors impacting evacuation behavior. Existing literatures have also found that variables like mandatory evacuation order, voluntary evacuation orders, time of landfall significantly affect people's evacuation decisions (Pham et al., 2020; Whitehead et al., 2000; Wong et al., 2018). We also assume that output variables include O^1 (*home distance*) and O^2 (*word2vec score*). These output variables depend on the current state and the input variables including I^1 (*time difference from landfall*), I^3 (*time difference from landfall during pre-landfall*), I^4 (*user's home is under mandatory evacuation*) and I^5 (*user's home is under voluntary evacuation*). Moreover, no input is chosen for the initial probability model and thus parameters will be learned by the EM algorithm only. Multinomial logistic regression is used as the transition and initial models. Since both the outputs distance from home (O^1) and word2vec evacuation similarity score (O^2) are continuous, linear regression models are used as output models.

In the given setting of IO-HMM, to unfold the dynamics of hurricane evacuation, we choose four types of activities as hidden states: home activity, office activity, evacuation, and other activity. Home activity and office activity represent the activities when the user stays at home and office, respectively; any activities participated at other locations are defined as "other" activities.

Considering these four states would allow us to better capture the dynamics in user activities as a hurricane approaches to make a landfall. For instance, it would enable us to capture the differences between activity transitions in a zone with a mandatory evacuation order and those transitions in other areas without a mandatory order (e.g., in a mandatory order zone, an individual is likely to end office activity earlier).

Starting from the first evacuation order to the landfall day, if a user has only other activity but no home/office activity and is found at a location of 200 kilometers or more away from home, we labeled it as an evacuation activity. We train the IO-HMM model using the labeled sequences of 80% ($n=202$) of the evacuated users and 80% ($n=1855$) of the non-evacuated users. To validate the model, we use the data from the 20% ($n=50$) of the evacuated users and 20% ($n=463$) of the non-evacuated users. We implement the models in Python programming language using the IO-HMM package developed by (Yin et al., 2017), available at <https://github.com/Mogeng/IO-HMM>.

7. RESULTS

In this section, we present the results of our evacuation dynamics model. First, we apply a standard HMM model to find the learned distributions of the selected outputs. Then we interpret the results of IO-HMM.

7.1 HMM Results

In the HMM structure, we have four latent states/activities considering the output variables as mixtures of gaussian distributions. Table 1 presents the posterior distributions of home distance and word2vec score for each latent activity. Mean distances from home are 0, 22.64, 129.18, 699.97 kilometers for home activity, office activity, other activity, and evacuation, respectively. The model has estimated higher average distance of 699.97 kilometer for evacuation activities. Average distance for office activity is 22.64 kilometers with dispersion of 32.64 kilometer. We choose other activities as a broad category for simplicity of the model; it may include grocery shopping before hurricane, eating at restaurants, short or long trips etc., thus 129.18 kilometer of average distance with the highest dispersion of 340.44 kilometer seems reasonable.

We are interested in learning how users respond to evacuation warning in their tweets. We find the word2Vec evacuation similarity scores as 0.54, 0.48, 0.48, and 0.50 for home activity, office activity, other activity, and evacuation, respectively. Although the difference is not that much, we see a higher score during home activity and evacuation activity. This means that users have tweeted about evacuation more during evacuation or home activity (0.54 and 0.50 word2vec similarity score). It is expected since evacuated users are more likely to share posts about evacuation. Also, during a hurricane, people are more likely to share evacuation related updates from their homes.

Table 1 Posterior Distributions of Output Variables. Here $\mathcal{N}(\mu, \sigma)$ represents a normal distribution with mean μ and standard deviation of σ .

Latent States (Activity Types)	Distribution of Output Variables	
	Distance from Home	Word2Vec Score
Home	$\mathcal{N}(0.00, 0.00)$	$\mathcal{N}(0.54, 0.2)$
Office	$\mathcal{N}(22.64, 32.64)$	$\mathcal{N}(0.48, 0.19)$
Other	$\mathcal{N}(129.18, 340.44)$	$\mathcal{N}(0.48, 0.2)$
Evacuation	$\mathcal{N}(699.97, 444.5)$	$\mathcal{N}(0.5, 0.22)$

7.2 IO-HMM Results

Although an HMM can learn the latent activities from the observed tweets, it does not allow to incorporate contextual input variables to infer the latent activities and their relationship with the outputs. Table 2 shows the coefficients of the output model when applied an IO-HMM structure. We have considered other variables such as friends count and follower count; but the estimated coefficients for these variables are not significant. We have excluded these variables from our final model. The output variable, *distance from home*, given the current state is a home activity has no coefficient. This is plausible since, for any user at home, distance from home is always zero. As expected, among all activities, the evacuation activity has the highest intercept for the *distance from home* output variable. The coefficients of I^1 and I^3 , for an evacuation activity, are found statistically insignificant, indicating a lack of evidence in the data that evacuation distance depends on the time difference from landfall. However, negative values of these coefficients indicate that an increase in the time difference from landfall (e.g., a time closer to the landfall in a pre-landfall period) would decrease evacuation distance. Positive coefficients for both mandatory (I^4) and voluntary order (I^5) indicate that evacuated users from mandatory and voluntary evacuation zones will travel longer than a user from a zone with no evacuation order. Furthermore, users from voluntary evacuation zones would travel longer than the users from mandatory evacuation zones.

TABLE 2 Coefficients of the Output Models for IO-HMM

Output Variables	Latent Variables	Intercept	Input Variables			
			Time difference from landfall (hour), I^1	Time difference from landfall*Pre-landfall period (hour), I^3	Home location under mandatory evacuation order, I^4	Home location under voluntary evacuation order, I^5
Distance from Home	Home Activity	0	0	0	0	0
	Office Activity	22.565***	0.015***	-0.021*	1.893*	-5.724***
	Other Activity	111.639***	0.189***	-0.457***	-24.910***	10.991***
	Evacuation	585.724***	-0.125	-0.572	160.968***	176.748***
word2vec Score	Home Activity	0.667***	-0.0006***	0.002***	0.035***	-0.0272***
	Office Activity	0.577***	0.0005***	0.001***	-0.030***	0.019**
	Other Activity	0.571***	-0.001***	0.011***	0.015***	-0.013***
	Evacuation	0.563***	-0.0005***	0.002***	-0.040***	-0.036*

Note: ~p <0.1 ; **~p <0.05 ; ***~p <0.01 ;

For word2vec evacuation similarity score, home activity has the highest intercept value. It means that if all the independent variables are equal to zero (equivalent to the landfall day and no evacuation order has been issued), users are more likely to post about evacuation from their homes. This is reasonable as users who are not required to evacuate are more likely to stay at home and may post evacuation related tweets. For evacuation activity, a negative coefficient of I^1 (-0.0005) and a positive coefficient of I^3 (0.002) indicate that evacuated users post more about evacuation during a pre-landfall period as time approaches to landfall compared to a post-landfall period. In other words, an evacuated user posts more about evacuation before landfall, probably because they have already evacuated and expressing concerns who are yet to evacuate. These variables (I^1, I^3) have similar effect for home activity and other activities, indicating that in general users are expected to post more about evacuation before the landfall than in a post-landfall period. For home activity and other activity, mandatory evacuation order (I^4) has a positive coefficient and voluntary evacuation order (I^5) has a negative coefficient for word2vec evacuation similarity score. It means that if all other variables remain constant, while staying at home or participating in other activity, compared to a user from no evacuation order zone, a user from a mandatory evacuation order zones is likely to post more about evacuation whereas a user from voluntary evacuation order zones is likely to post less about evacuation. On the other hand, for evacuation activity, variables representing mandatory order zone and voluntary order zone have negative coefficients for word2vec evacuation similarity score. This indicates that evacuated users from a mandatory or voluntary order zone post less about evacuation compared to evacuees from a zone with no evacuation order. This is plausible since evacuated users may have less time to tweet while traveling.

Table 3 shows the coefficients of multinomial logistic regression (MNL) models for the transition models of IO-HMM. Given the current state, there are 4 MNL models to capture the transition among the hidden states (activity types). Here, any positive coefficient means that an increase in the associated variable will increase the probability to make a transition between the corresponding states. We have 80 different coefficients to capture the dynamics of transition between any two states. We mainly focus on interpreting the coefficients associated with evacuation. For instance, if we observe the coefficients of home: evacuation (see Table 3), a negative sign of the variable I^2 (i.e., a post landfall period) represents that if a user's current state is a home activity, in comparison to a pre-landfall period, a post-landfall period decreases the likelihood to evacuate if all other variables remain constant. It is also same for office: evacuation, other: evacuation and evacuation: evacuation transitions (see Table 3). These results are quite expected as individuals are less likely to evacuate after the landfall.

The coefficient of the variable *time difference from landfall* (I^1) is insignificant for the evacuation: evacuation transition; but it is significant and has negative coefficients for other 3 transitions (i.e., home: evacuation, office: evacuation, other: evacuation). This means that if everything remains constant, with increase in time difference from the landfall (as time becomes closer to landfall or away from landfall) these transitions are less likely to occur. The input variable I^4 (*home location under mandatory evacuation order*) has positive coefficients for home: evacuation, office: evacuation, and other: evacuation but it has a negative coefficient for the evacuation: evacuation transition. A plausible explanation is that a user may evacuate directly from

home/office (some user's home and office are same) or may perform some other activities (distance < 200 km) and then evacuate. The positive coefficient of I^4 indicates that compared to the users from zones with no or voluntary evacuation order, users from mandatory evacuation order zones are more likely to evacuate. A negative coefficient of I^4 for evacuation: evacuation transition means that users who evacuate from mandatory evacuation zone are less likely to remain in the evacuation state. It might be because due to their concerns about the damage of their home caused by the hurricane. The input variable I^5 (*home location under voluntary evacuation order*) has positive coefficients for home: evacuation, office: evacuation transition and negative coefficients for other: evacuation and evacuation: evacuation transition. The positive coefficient of (I^5) indicates that compared to the users from no evacuation order, the users from voluntary evacuation order are more likely to evacuate given their current activity is home or office. The negative coefficient of I^5 indicates that given the current state is other or evacuation, compared to the users from zones with no evacuation order, users from voluntary evacuation zone are less likely to evacuate or continue to maintain evacuation state.

TABLE 3 Coefficients of the Transition Models for IO-HMM

From Activity: To Activity	Intercept	Time difference from landfall (hour), I^1	Whether time is post landfall, (I^2)	Home location under mandatory evacuation order, I^4	Home location under voluntary evacuation order, I^5
Home: Home	0.034***	0.253***	0.0001***	0.521***	0.051
Home: Office	0.037***	-0.084	0.002***	-0.032	0.083
Home: Other	0.121***	0.521***	-0.001***	0.088*	0.067
Home: Evacuation	-0.191***	-0.691***	-0.0003	0.401***	0.201***
Office: Home	0.174***	0.061	0.001***	-0.062**	-0.293***
Office: Office	0.393***	0.450***	0.002***	0.588***	0.391***
Office: Other	0.282	0.364***	-0.001***	0.196***	0.069*
Office: Evacuation	-0.242***	-0.875***	-0.003***	-0.722***	0.167***
Other: Home	0.159***	0.023	0.001***	0.228***	0.065
Other: Office	-0.062***	-0.157	0.002***	0.111*	-0.088
Other: Other	0.103***	0.957***	-0.004**	0.187***	0.233***
Other: Evacuation	-0.165***	-0.734***	-0.0002	0.236***	-0.177*
Evacuation: Home	-0.099	0.103	0.010***	-0.072	0.137
Evacuation: Office	-0.132	-0.460	0.0008	0.049	0.237
Evacuation: Other	-0.053	0.025	-0.0009	-0.372***	0.147
Evacuation: Evacuation	0.283***	0.332	-0.010***	-0.394***	-0.523***

Note: ~ $p^ < 0.1$; ~ $p^* < 0.05$; ~ $p^* < 0.01$;

Figure 5 shows the combined effect of different variables contributing to the transition probability from one activity to another. The color of each cell represents the probability of making a transition from the associated row activity to the associated column activity. The sum of each row equals to 1 indicating that from the current state/activity type, it will make transition to any of the four activity types. Figure 5 (a) and 5 (b) show the transition probabilities 100 hours before landfall, whereas Figures 5 (c) and 5 (d) show the transition probabilities 100 hours after the landfall. Overall, before the landfall, given a current state, it has higher probabilities to make a transition to evacuation state compared to the post-landfall period. Figure 5 (a) and 5 (b) show the differences in transition probabilities between voluntary evacuation order and mandatory evacuation order at the home location, 100 hours before the landfall. Given the current state is a home activity, compared to the voluntary evacuation order, a mandatory evacuation order has a slightly higher probability of evacuation (0.25 vs. 0.24) and a lower probability of transitioning to the office (0.16 vs. 0.21). In both cases, we see that given that the current state is a home activity, the probability to participate in other activity is high (0.30 under a voluntary order and 0.26 under a mandatory order). Given the current state represents other activity, the probability to evacuate increases from 0.16 to 0.21 from zones with a voluntary evacuation order to zones with a mandatory evacuation order, respectively. Moreover, given the current state represents an office activity, the probability to evacuate decreases from 0.28 to 0.12 from a voluntary evacuation zone to a mandatory evacuation zone, respectively. Besides, the probabilities that an evacuated user will continue to remain evacuated for voluntary and mandatory evacuation zones are 0.45 and 0.56, respectively.

Similarly, Figure 5 (c) and (d) show the transition probabilities after 100 hours of the landfall for users under voluntary and mandatory evacuation zones, respectively. Given any state, the probabilities to evacuate, 100 hours after the landfall, are very low. For an evacuated user, an individual from a voluntary evacuation zone has a lower probability (0.076) to remain evacuated (see Figure 5c), compared to an individual from a mandatory evacuation zone (0.11) (see Figure 6d). In both Figures 5 (c) and 5(d), the highest probability values are observed for the transition from evacuation to home. However, there is not much difference in the probability of returning to home for users, 100 hours after hurricane landfall, from a voluntary evacuation zone (0.59) or a mandatory evacuation zone (0.60).

		Transition To			
		home	office	other	evacuation
Transition From	home	0.25	0.21	0.3	0.24
	office	0.18	0.28	0.27	0.28
	other	0.24	0.15	0.45	0.16
	evacuation	0.081	0.22	0.25	0.45

(a) 100 hours before landfall, home location under voluntary evacuation order

		Transition To			
		home	office	other	evacuation
Transition From	home	0.33	0.16	0.26	0.25
	office	0.23	0.35	0.31	0.12
	other	0.25	0.16	0.38	0.21
	evacuation	0.072	0.2	0.17	0.56

(b) 100 hours before landfall, home location under mandatory evacuation order

		Transition To			
		home	office	other	evacuation
Transition From	home	0.28	0.25	0.37	0.098
	office	0.18	0.52	0.25	0.05
	other	0.28	0.18	0.48	0.066
	evacuation	0.59	0.14	0.19	0.076

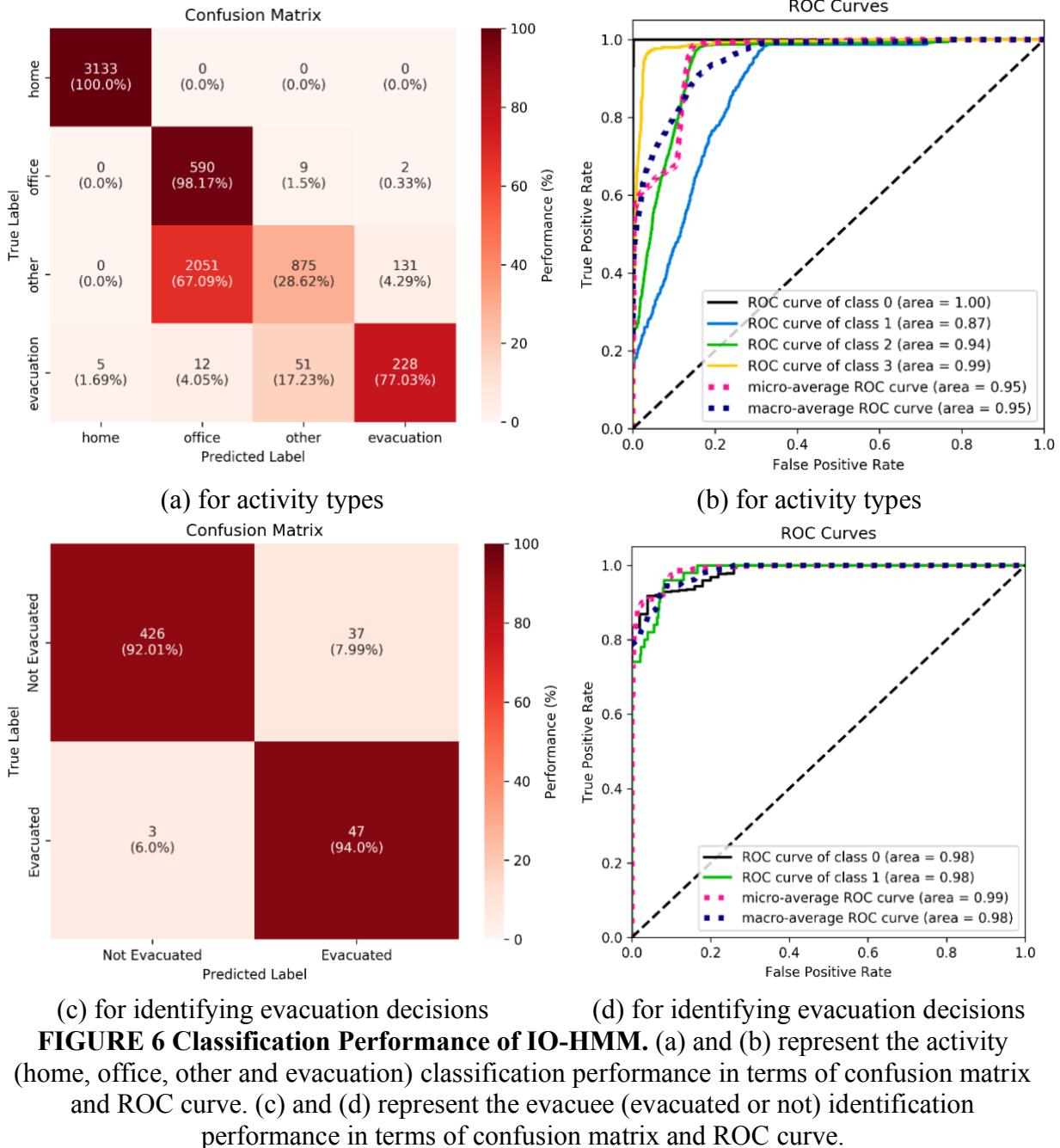
(c) 100 hours after landfall, home location under voluntary evacuation order

		Transition To			
		home	office	other	evacuation
Transition From	home	0.39	0.19	0.32	0.1
	office	0.2	0.54	0.25	0.018
	other	0.3	0.2	0.42	0.09
	evacuation	0.6	0.15	0.14	0.11

(d) 100 hours after landfall, home location under mandatory evacuation order

FIGURE 5 Activity Transition Matrices under different scenarios

Using the trained IO-HMM model, we predict the activity sequences for the test data (20% of the labeled dataset). Figure 6 shows the performance of activity recognition of IO-HMM. The confusion matrix reports the numbers of predicted labels and the ratio of correctly predicted label to actual label. IO-HMM has 100 %, 98.17 %, 28.62% and 77.03% accuracy for recognizing home, office, other, and evacuation activities, respectively (see Figure 6(a)). Using the standard HMM, we obtain 100%, 92.38%, 29.01%, and 62.51% accuracy for home, office, other, and evacuation activity recognition, respectively. Thus, using an IO-HMM structure instead of a standard HMM structure improves accuracy.



We observe that the model has relatively low accuracy in identifying ‘other’ activity types than ‘home’, ‘office’, and ‘evacuation’ activities. This happens because of the similarity between office activity and other activity and the overlap between the learned probability distributions of these two activity types. The result indicates that the model predicts a significant number of ‘other’ activities as ‘office’ activities (see Figure 6a). Although the model has a low performance in identifying ‘other’ activity, it is unlikely to have any consequence in predicting evacuated users (see Figure 6c).

Figure 6(b) shows the ROC curves which plot true positive rates vs. false positives rates under every possible classification threshold. For example, for home activity recognition, a true positive rate answers the question when an actual activity is at home how often the model predicts it as a home activity (true home activity/all home activity). On the other hand, false positive rate for home activity recognition answers the question when actual activity is not at home how often the model predicts it as a home activity (false home activity/ all not home activity). In Figure 6(b) classes 0, 1, 2, and 3 represent home, office, other, and evacuation activities, respectively. Area under the curve or AUC represents the classification performance where AUC is percentage of the whole box which is under the ROC curve (range 0 to 1). If any ROC curve is close the diagonal line or $AUC = 0.5$, the model is not any better than random guessing. We can see that the model has the AUC values of 0.98, 0.87, 0.94, 0.99 for home, office, other, and evacuation activities, respectively.

We also report the performance of the IO-HMM model in identifying evacuation decision (if a user has evacuated or not) at an individual level. Using the test set, for each user, we convert the predicted activity sequence as a binary output by checking if any evacuation state is present in the predicted activity sequence or not. Then we compare the converted evacuation identification result against our labeled data to estimate the model performance using confusion matrix and ROC curve. Figure 6(c) and 6(d) show the confusion matrix, ROC curve, respectively, for individual-level prediction. For identifying individual evacuation decision, the model has 92% and 94% accuracy for non-evacuated and evacuated users, respectively (see Figure 6(c)). The model has the same AUC value of 0.98 for both evacuated and non-evacuated users.

Figure 7 shows evacuation participation rates over time. From the labeled data, we divide evacuations in two categories: evacuations generated from zones under a mandatory evacuation order and evacuations generated from zones under a voluntary or no evacuation order. The evacuations from later zones are also known as shadow evacuation (Sorensen and Vogt, 2006; Zeigler et al., 1981). We find that from our collected samples, around 65% evacuations are generated from the mandatory evacuation order zone and the remaining 35% are from a zone with either a voluntary or no evacuation order. Shadow evacuation causes additional traffic congestion and often hampers the evacuation of the actually threatened population (Murray-Tuite and Wolshon, 2013). Using the trained IO-HMM model, we predict the activity sequences of all the users (including both training and test data). We compare the predicted timing of evacuation state and the number of evacuated users with the labeled data. From Figure 7, we see that on aggregate the model identifies around 62% of total evacuation as mandatory evacuation and around 38% as shadow evacuation. The model captures the overall trend of the evacuation timing and participation numbers.

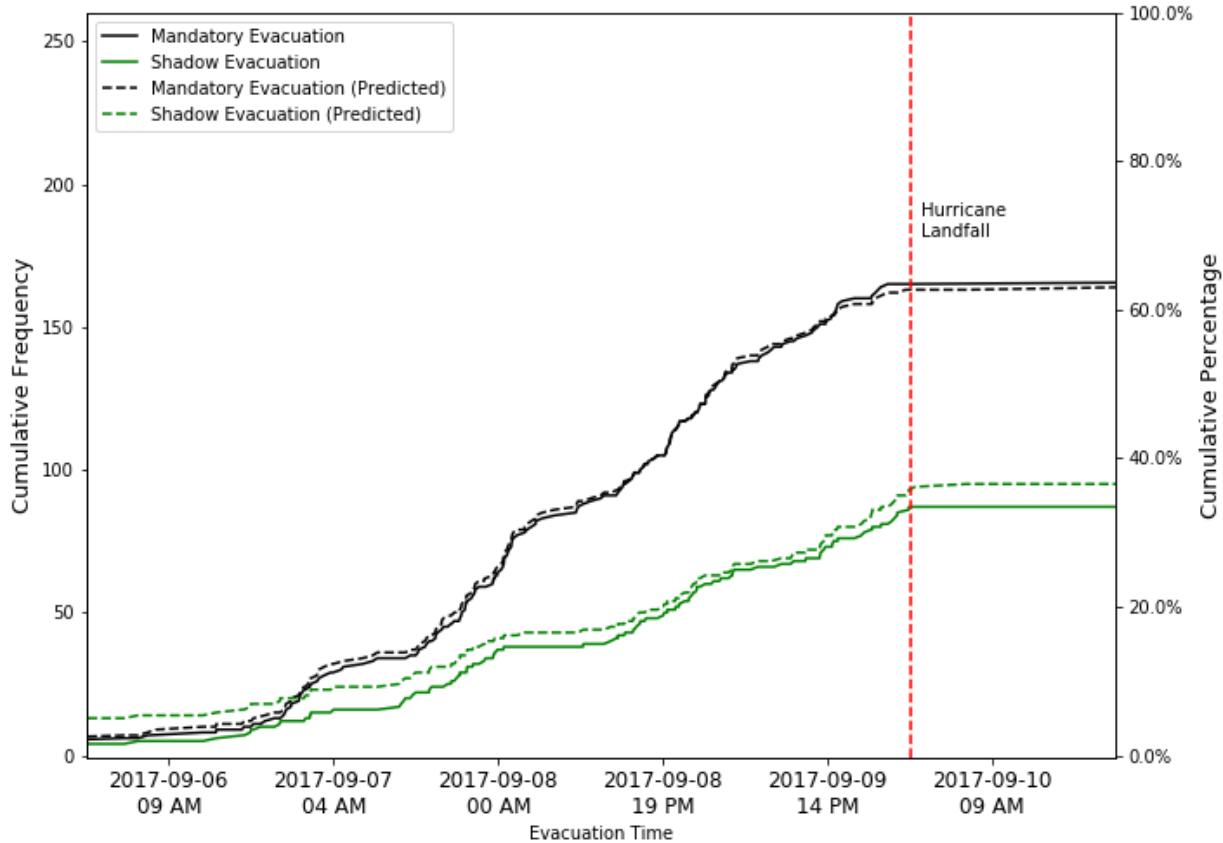


FIGURE 7 Cumulative evacuation frequencies and predicted evacuation frequencies across time.

8. CONCLUSIONS

To better capture the dynamics of individual-level evacuation behavior, longitudinal spatio-temporal data are needed covering both pre- and post-disaster periods. Traditional data collection approaches such as household surveys are static and conducted in a post-disaster period. This limits our ability to capture the dynamics of evacuation decision-making process such as determining the probability of evacuation given the states of the variables (e.g., evacuation order, projected landfall time) changing over time. With longitudinal data collected, we can determine the effects of the changes in variables over time on evacuation decisions. In addition, since the data are collected in real time, we are able to capture the dynamics when the situation is evolving, instead of at a post-disaster period.

In this study, we use Twitter data from Hurricane Irma to develop a model for inferring individual hurricane evacuation dynamics. We have collected evacuation data from Twitter covering all counties of Florida. Based on the tweets of active users during an evacuation period, we develop an input output hidden Markov model to infer what type of activities individuals participate, the locations and timing of those activities, when they evacuate, and where they evacuate to. We model individual participation in four activity types (home activity, office activity, other activity, and evacuation) during a hurricane.

The modeling approach provides rich insights on evacuation and other activity types during a hurricane both spatially and temporally. For instance, we have learned from real-time Twitter data to what extent individual social communication and evacuation distance depend on evacuation order type and time to landfall.

The results associated with the spatial variables (e.g., home location under a mandatory evacuation order, home location under voluntary evacuation order) indicate that if a user's home location is under a mandatory or voluntary evacuation order, he/she is likely to evacuate longer distance compared to the users under no evacuation order. We also find that users from a mandatory evacuation zone are likely to post more about evacuation during home activity and the users from a voluntary evacuation zone are more likely to post about evacuation during an office activity compared to the users from no evacuation order zone. From the activity transition dynamics, we find that given the current activity is a home activity, the probability to evacuate increases for both mandatory and voluntary evacuation order; given the current activity is an office activity, the probability to evacuate increases for mandatory evacuation order and decreases for voluntary evacuation order; and given the current activity is other activity, the probability to evacuate increases for mandatory evacuation order and decreases for voluntary evacuation order.

The results associated with the variables related to temporal dynamics show that evacuation distance is likely to decrease with a decrease in time difference from landfall in a pre-disaster period. The number of evacuation related tweets (representing evacuation context) are likely to increase with decrease in time difference from landfall in a pre-disaster period. We also find that, before the landfall, as the time difference from landfall increases (reaching closer to the landfall), the likelihood of evacuation decreases. And after the landfall, as the time difference from landfall increases, the probability of returning to home increases. Thus, this study can capture the dynamics of evacuation behavior both spatially and temporally within a single modeling framework. Such insights for hurricane evacuation are critical for emergency management. For instance, identifying the evacuated and not evacuated population during a hurricane can make its preparation more effective and dynamic. Another benefit of our modeling framework is that, with the parameters estimated in this study, we can generate the behavior of a synthetic population by simulating their activity dynamics. Such simulated data from the model based on the total population of a region will allow us to determine evacuation demand in real-time.

This study has some limitations such as Twitter may have different penetration in different areas. Twitter users are not equally distributed across different age groups. Consequently, geotagged tweets may not represent the behavior of all population segments. We have assumed 200 km as a threshold distance to identify an evacuation and 20 km to detect a return from an evacuation. Thus, our approach cannot detect shorter distance evacuation such as relocating to higher ground or a better-protected place or a shelter within one's locality. This is due to the lack of granularity in our data since some tweets have a city/county level location instead of a precise GPS location. Other variants of HMM can be applied to get better accuracy. Also, to verify our results data from other sources are not used as they are not currently available.

In spite of the above limitations, this study adds to the growing literature on modeling the dynamics of evacuation behavior. In particular, it investigates the potential of using social media

data for understanding evacuation dynamics. However, future research should focus on how to account for potential biases present in Twitter data. As social media data can be gathered in real time at large scale during a hurricane, our model can make evacuation traffic predictions and provide behavioral insights in real time. Since traditional survey data are costly and often conducted at a post-hurricane period, our method of using social media data can complement the traditional approaches of modeling evacuation behavior.

ACKNOWLEDGEMENTS

This study was supported by the U.S. Department of Transportation University Transportation Centers Program under the project “Disaster Analytics: Disaster Preparedness and Management through Online Social Media” and the U.S. National Science Foundation through the grant CMMI-1917019. However, the authors are solely responsible for the facts and accuracy of the information presented in the paper.

SUPPORTING INFORMATION

Manual Checking of the Labeled Dataset

We created an interactive map to manually check whether a user evacuated or not. For each user, we visualized the home, office, evacuation destination (if any), the visited locations, and the tweets. We checked the tweets if there was any mention that the user was evacuating or leaving home during the evacuation period. As an example, Figure S1 shows the snapshot of our manual checking process for a user. The locations are plotted with a 5-km precision to protect user privacy. The user, shown in Figure S1, had home and office in Lee County, FL and evacuated to Birmingham city, Alabama. While evacuating, the user tweeted from Tampa, FL indicating that he/she was aware of hurricane Irma’s changing path. We checked each user’s home location, evacuation destination (if any), traveled distance, and tweet text to infer whether the user evacuated or not. Using this process, we checked 252 evacuated users and 2,319 non-evacuated users. The manual checking was performed by two individuals.

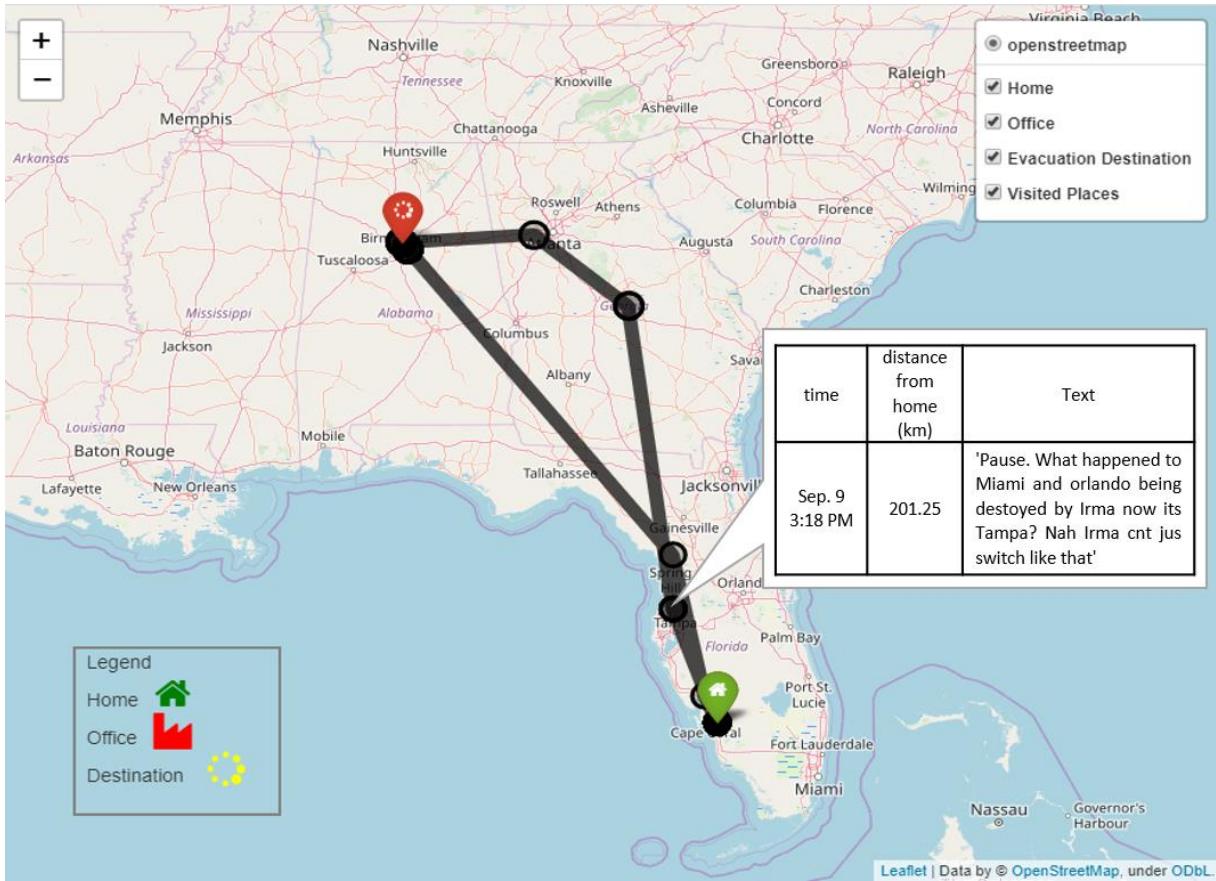


Figure S1: Demonstration of the Manual Checking Process. It shows a snapshot of the interactive visualization a user's home, evacuation destination, and visited places—containing tweet time, tweet text and distance from home.

Word2Vec Model

Word2vec is a predictive model developed by Mikolov et al.(Mikolov et al., 2013b, 2013a). It contains two distinct algorithms: Continuous Bag-of-Words (CBOW) and Skip-Gram. Skip-Gram predicts context word given a target word and CBOW predicts the target word given the context word. Details of word2vec model can be found in refs (Meyer, 2016; Mikolov et al., 2013a, 2013b). It is a very simple, scalable, fast to train model that can learned over billions of words of text that will produce exceedingly good word representations. Word2vec uses the theory of meaning to predict between each word and the context word. Word2Vec contains two distinct algorithms, Continuous Bag-of-Words (CBOW) and Skip-Gram, where Skip-Gram predict context word given the target word and CBOW predict the target word given the context word. Figure S2 shows the CBOW architecture.

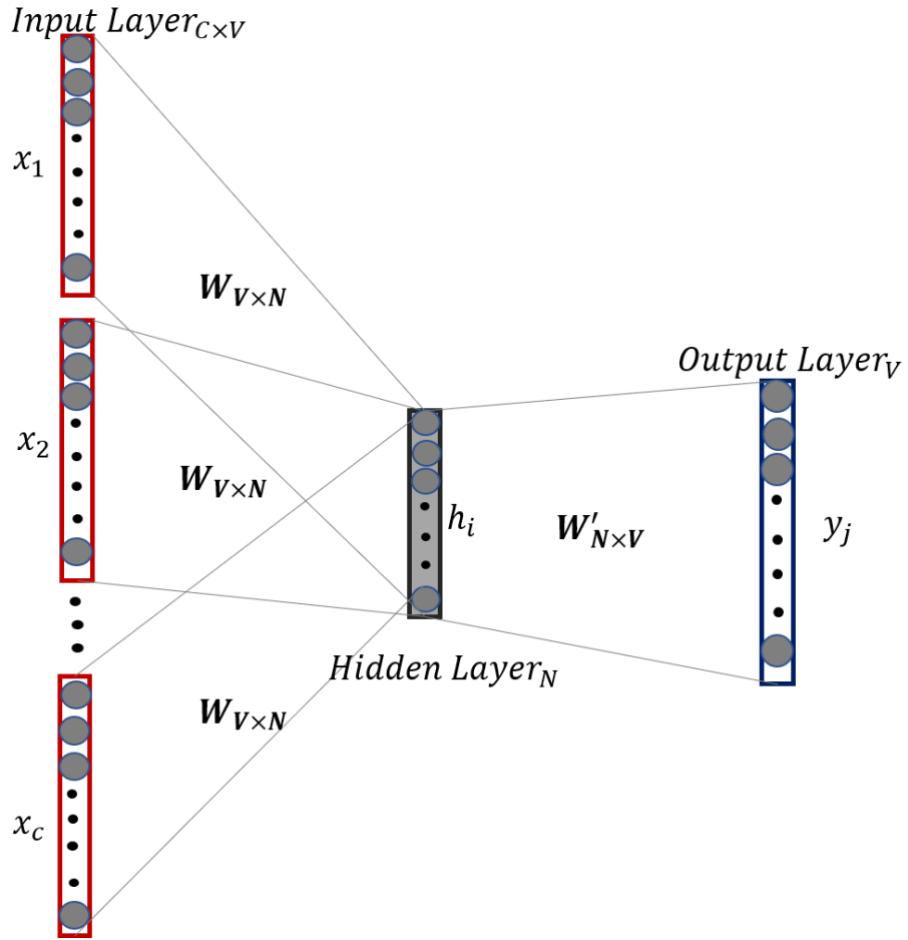


Figure S2: The CBOW architecture predicting the current word based on the context

In CBOW, for a window size C , the inputs are one-hot (size equal to vocabulary size, V) encoded context words $\{x_1, \dots, x_c\}$. The hidden layer h is N -dimensional. The output/target word y_j for the context input words is also one hot encoded of size V . The input layer and hidden layer are connected by weights matrix W of dimension $V \times N$ and the hidden layer and output layer are connected by another weight matrix W' of dimension $N \times V$. The workflow of CBOW can be described in three steps described below.

Forward Propagation

This section describes how the output is computed from the input given that the input and output weight matrixes are known. Hidden layer output is computed first from the input layer and weight matrix W . This is computed as shown in equation (7)

$$h = \frac{1}{C} W \cdot \left(\sum_{i=1}^C x_i \right) \quad (7)$$

which is the weighted average of the input vectors and weight matrix W . Next, the input to each node of output layer is computed by the following

$$u_j = v'_{w_j}^T \cdot h \quad (8)$$

Where v'_{w_j} is the j^{th} column of the output weight matrix W' . Finally, the outputs y_j of the output layer are computed by applying a soft-max function as shown in equation (9).

$$y_j = p(w_{y_j} | w_1, \dots, w_c) = \frac{\exp(u_j)}{\sum_{j=1}^v \exp(u_j)} \quad (8)$$

As the output is computed, the weight matrix W and W' can be learned from by back-propagating the errors. The process is discussed in the next section.

Learning the Weight Matrices

To learn the weight matrices, at first the W and W' are randomly initialized. By feeding the training examples sequentially and observing the predicted output, we get the error which is a function of difference between the actual and predicted output. It is also known as loss function. The objective is to maximize the conditional probability of the output word given the input context, therefore our loss function will be the following:

$$\begin{aligned} E &= -\log p(w_0 | w_I) \\ &= -u_{j^*} - \log \sum_{j'=1}^v \exp(u_{j'}) \\ &= -v_{w_0}^T \cdot h - \log \sum_{j'=1}^v \exp(v_{w_j'}^T \cdot h) \end{aligned} \quad (9)$$

Here j^* is the index of the actual output word. The next step is to update the weight matrices based on the gradient. The gradient of this error is computed with respect to both weight matrices and correct them in the direction of this gradient. This optimization procedure is known as stochastic gradient descent. Details of the optimization procedure can be found here (Bottou, 2010).

Word2Vec Sample Results

We train the model with the corpus of hurricane related tweets collected from 4 hurricanes (Irma, Matthew, Harvey and Sandy). We use minimum word count=3 for preparing the vocabulary. We train the model using a window size, C=32 for context words. Once the model is trained each word in the vocabulary will have a vector representation with its context words. The cosine similarity between two word vectors is computed using the equation 6. Figure S3 shows the top 15 similar words for ‘Evacuation’, ‘Evacuating’ and ‘Sheltering’. For example, similar words to evacuation

Evacuation		Evacuating		Sheltering	
evac	0.906	evacuate	0.853	evacuating	0.784
evacuations	0.902	evacuated	0.848	evacuated	0.772
evacs	0.831	leave	0.787	relocating	0.766
prisoninsider	0.824	sheltering	0.784	shelter	0.758
evacuations....	0.805	flee	0.765	housed	0.747
evacuate	0.777	stranded	0.76	include...	0.738
mandatory	0.775	stuck	0.759	displaced	0.732
emics	0.768	fleeing	0.755	relocated	0.723
evac*	0.764	hunkering	0.737	rehabilitate	0.722
evacuation...	0.762	**some**	0.736	featureso	0.722
curfews	0.759	locked	0.731	prepare	0.72
patrols	0.743	evac	0.731	storypirates	0.715
409-283-2172	0.741	relocated	0.724	hunkering	0.713
lizholmes7	0.74	escape	0.721	hunker	0.713
		trapped	0.72	sos*	0.712

FIGURE S3 Top 15 Word2Vec Cosine Similarity Score of Evacuation, Evacuating, and Sheltering

contains evac, evacuations, evacs, evacuations..., evacuate etc. which may have been used as a short form of evacuation and also evac is emergency service provider name in Volusia county. Other similar words are mandatory, curfews, patrols and a cell number (4092832172) etc. Evacuation is very related with mandatory order for evacuation. And patrol, curfew is also related to evacuation because during state of emergency, state issue curfew and police patrol monitor the situation during hurricane evacuation. This cell number (409-283-2172) is the contact number of Tyler County - Sheriffs' Association of Texas which was very active during Harvey evacuation period. Thus, the result shows very good consistency in finding out the related/similar word

REFERENCES

ABC News, 2017. Hurricane Irma begins to reach Florida as millions of residents evacuate ahead of monster storm [WWW Document]. ABC NEWS. URL <https://www.abc.net.au/news/2017-09-10/hurricane-irma-begins-to-impact-florida-as-residents-evacuate/8889076>

Baker, E.J., 1979. Predicting response to hurricane warnings: A reanalysis of data from four studies. Mass emergencies 4, 9–24.

Balkić, Z., Šoštarić, D., Horvat, G., 2012. GeoHash and UUID identifier for multi-agent systems, in: KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications. pp. 290–298.

Baroni, M., Dinu, G., 2014. Don 't count , predict ! A systematic comparison of context-counting vs . context-predicting semantic vectors. Proc. 52nd Annu. Meet. Assoc. Comput. Linguist. 1.

Beiró, M.G., Panisson, A., Tizzoni, M., Cattuto, C., 2016. Predicting human mobility through the assimilation of social media traces into mobility models. EPJ Data Sci. 5.

Bengio, Y., Frasconi, P., 1995. An Input Output HMM Architecture. Neural Inf. Process. Syst. 427–434. <https://doi.org/10.1093/europace/euq350>

Blanton, B., Dresback, K., Colle, B., Kolar, R., Vergara, H., Hong, Y., Leonardo, N., Davidson, R., Nozick, L., Wachtendorf, T., 2020. An Integrated Scenario Ensemble-Based Framework for Hurricane Evacuation Modeling: Part 2—Hazard Modeling. Risk Anal. 40, 117–133.

Bottou, L., 2010. Large-scale machine learning with stochastic gradient descent, in: Proceedings of COMPSTAT'2010. Springer, pp. 177–186.

Chaniotakis, E., Antoniou, C., Pereira, F.C., 2017. Enhancing resilience to disasters using social media, in: 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS). pp. 699–703. <https://doi.org/10.1109/MTITS.2017.8005602>

Cheng, G., Wilmot, C.G., Baker, E.J., 2008. A destination choice model for hurricane evacuation, in: Proceedings of the 87th Annual Meeting Transportation Research Board, Washington, DC, USA. pp. 13–17.

D. E. Rumelhart, G. E. Hinton, R.J.W., 1986. Learning internal representations by backpropagating errors. Nature 323.

Davidson, R.A., Nozick, L.K., Wachtendorf, T., Blanton, B., Colle, B., Kolar, R.L., DeYoung, S., Dresback, K.M., Yi, W., Yang, K., others, 2020. An Integrated Scenario Ensemble-Based Framework for Hurricane Evacuation Modeling: Part 1—Decision Support System. Risk Anal. 40, 97--116.

Dong, X., Mavroeidis, D., Calabrese, F., Frossard, P., 2015. Multiscale event detection in social media. Data Min. Knowl. Discov. 29, 1374–1405.

Duong, T. V., Bui, H.H., Phung, D.Q., Venkatesh, S., 2005. Activity recognition and abnormality detection with the switching hidden semi-Markov model. Proc. - 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition, CVPR 2005 I, 838–845. <https://doi.org/10.1109/CVPR.2005.61>

Eagle, N., Pentland, A., 2006. Reality mining: Sensing complex social systems. Pers. Ubiquitous Comput. 10, 255–268. <https://doi.org/10.1007/s00779-005-0046-3>

FLKEYSNEWS, 2017. Millions of Floridians who fled Irma are eager to get home. Patience will be necessary [WWW Document].

Fu, H., Wilmot, C., 2004. Sequential logit dynamic travel demand model for hurricane evacuation. *Transp. Res. Rec. J. Transp. Res. Board* 19–26.

Ghahramani, Z., Jordan, M.I., 1996. Factorial hidden Markov models, in: Advances in Neural Information Processing Systems. pp. 472–478.

Gladwin, C.H., Gladwin, H., Peacock, W.G., 2001. Modeling hurricane evacuation decisions with ethnographic methods. *Int. J. Mass Emerg. Disasters* 19, 117–143.

González, A.M., Roque, A.M.S., Garc\'ia-González, J., 2005. Modeling and forecasting electricity prices with input/output hidden Markov models. *IEEE Trans. Power Syst.* 20, 13–24.

Guha-sapir, D., Hoyois, P., Below, R., 2017. Annual Disaster Statistical Review 2016: The numbers and trends. *Cent. Res. Epidemiol. Disasters*. <https://doi.org/10.1093/rof/rfs003>

Han, S.Y., Tsou, M.-H., Knaap, E., Rey, S., Cao, G., 2019. How Do Cities Flow in an Emergency? Tracing Human Mobility Patterns during a Natural Disaster with Big Data and Geospatial Data Science. *Urban Sci.* 3, 51.

Hasan, S., Foliente, G., 2015. Modeling infrastructure system interdependencies and socioeconomic impacts of failure in extreme events: emerging R&D challenges. *Nat. Hazards* 78, 2143–2168.

Hasan, S., Mesa-Arango, R., Ukkusuri, S., 2013. A random-parameter hazard-based model to understand household evacuation timing behavior. *Transp. Res. Part C Emerg. Technol.* 27, 108–116.

Hasan, S., Mesa-Arango, R., Ukkusuri, S., Murray-Tuite, P., 2011a. Transferability of hurricane evacuation choice model: Joint model estimation combining multiple data sources. *J. Transp. Eng.* 138, 548–556.

Hasan, S., Ukkusuri, S., Gladwin, H., Murray-Tuite, P., 2011b. Behavioral Model to Understand Household-Level Hurricane Evacuation Decision Making. *J. Transp. Eng.* 137, 341–348.

Hasan, S., Ukkusuri, S. V., 2017. Reconstructing Activity Location Sequences From Incomplete Check-In Data: A Semi-Markov Continuous-Time Bayesian Network Model. *IEEE Trans. Intell. Transp. Syst.* 1–12. <https://doi.org/10.1109/TITS.2017.2700481>

Huang, A., 2008. Similarity measures for text document clustering. *Proc. Sixth New Zeal.* 49–56.

Kang, J.E., Lindell, M.K., Prater, C.S., 2007. Hurricane Evacuation Expectations and Actual Behavior in Hurricane Lili 1. *J. Appl. Soc. Psychol.* 37, 887–903.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes11Edited by F. Cohen. *J. Mol. Biol.* 305, 567–580. <https://doi.org/10.1006/jmbi.2000.4315>

Kryvasheyeu, Y., Chen, H., 2015. Performance of Social Network Sensors During Hurricane Sandy. *PLoS One* 10, e0117288.

Kryvasheyeu, Y., Chen, H., Obradovich, N., Moro, E., Hentenryck, P. Van, Fowler, J., Cebrian,

M., 2016. Rapid assessment of disaster damage using social media activity. *Sci. Adv.* 2.3 e1500779.

Kumar, D., Ukkusuri, S. V, 2018. Utilizing Geo-tagged Tweets to Understand Evacuation Dynamics during Emergencies: A case study of Hurricane Sandy, in: Companion of the The Web Conference 2018 on The Web Conference 2018. pp. 1613–1620.

Lachlan, K.A., Spence, P.R., Lin, X., Najarian, K., Del Greco, M., 2016. Social media and crisis management: CERC, search strategies, and Twitter content. *Comput. Human Behav.* 54, 647–652.

Lee, K.L., Meyer, R.J., Bradlow, E.T., 2009. Analyzing risk response dynamics on the web: The case of Hurricane Katrina. *Risk Anal. An Int. J.* 29, 1779–1792.

Lin, D.-Y., Eluru, N., Waller, S., Bhat, C., 2009. Evacuation planning using the integrated system of activity-based modeling and dynamic traffic assignment. *Transp. Res. Rec. J. Transp. Res.* Board 69–77.

Lindell, M.K., 2008. EMBLEM2: An empirically based large scale evacuation time estimate model. *Transp. Res. part A policy Pract.* 42, 140–154.

Luz Lazo, L.A., 2017. Airlines scramble and roads fill as residents and visitors rush to get out of Florida ahead of Irma [WWW Document]. Washington Post.

Marcel, S., Bernier, O., Viallet, J.-E., Collobert, D., 2000. Hand gesture recognition using input-output hidden markov models, in: Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference On. pp. 456–461.

Marshal, A., 2017. 4 Maps That Show the Gigantic Hurricane Irma Evacuation [WWW Document]. wired.

Martín, Y., Li, Z., Cutter, S.L., 2017. Leveraging Twitter to gauge evacuation compliance: Spatiotemporal analysis of Hurricane Matthew. *PLoS One* 12, 1–22. <https://doi.org/10.1371/journal.pone.0181701>

McLachlan, G., Krishnan, T., 2007. The EM algorithm and extensions. John Wiley & Sons.

Mesa-arango, R., Hasan, S., Ukkusuri, S. V, Asce, A.M., Murray-tuite, P., 2013. Household- Level Model for Hurricane Evacuation Destination Type Choice Using Hurricane Ivan Data. *Nat. Hazards Rev.* 14, 11–20.

Metaxa-Kakavouli, D., Maas, P., Aldrich, D.P., 2018. How Social Ties Influence Hurricane Evacuation Behavior. *Proc. ACM Human-Computer Interact.* 2, 122.

Meyer, D., 2016. How exactly does word2vec work ? 1–18.

Meyer, R.J., Baker, J., Broad, K., Czajkowski, J., Orlove, B., 2014. The dynamics of hurricane risk perception: Real-time evidence from the 2012 Atlantic hurricane season. *Bull. Am. Meteorol. Soc.* 95, 1389–1404.

Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013a. Distributed Representations of Words and Phrases and their Compositionality. *Adv. Neural Inf. Process. Syst.* 3111--3119.

Mikolov, T., Corrado, G., Chen, K., Dean, J., 2013b. Efficient Estimation of Word Representations in Vector Space Tomas. *arXiv Prepr.* 1–12.

Murray-Tuite, P., Ge, Y.G., Zobel, C., Nateghi, R., Wang, H., 2019. Critical Time, Space, and Decision-Making Agent Considerations in Human-Centered Interdisciplinary Hurricane-Related Research. *Risk Anal.* <https://doi.org/10.1111/risa.13380>

Murray-Tuite, P., Wolshon, B., 2013. Evacuation transportation modeling: An overview of research, development, and practice. *Transp. Res. Part C Emerg. Technol.* 27, 25–45.

Pan, B., Zheng, Y., Wilkie, D., Shahabi, C., 2013. Crowd sensing of traffic anomalies based on human mobility and social media, in: *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.* pp. 344–353.

Pham, E.O., Emrich, C.T., Li, Z., Mitchem, J., Cutter, S.L., 2020. Evacuation Departure Timing during Hurricane Matthew. *Weather. Clim. Soc.* 12, 235–248.

Rabiner, L.R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* 77, 257–286. <https://doi.org/10.1109/5.18626>

Rambha, T., Nozick, L., Davidson, R., 2019. Modeling Departure Time Decisions During Hurricanes Using a Dynamic Discrete Choice Framework, in: *Transportation Research Board 98th Annual Meeting.*

Re, S., 2013. Mind the risk: a global ranking of cities under threat from natural disasters. *Swiss Re.*

Roy, K.C., Cebrian, M., Hasan, S., 2019. Quantifying human mobility resilience to extreme events using geo-located social media data. *EPJ Data Sci.* 8, 18.

Roy, K.C., Hasan, S., Sadri, A.M., Cebrian, M., 2020. Understanding the efficiency of social media based crisis communication during hurricane Sandy. *Int. J. Inf. Manage.* 102060.

Sadri, A.M., Hasan, S., Ukkusuri, S. V., Cebrian, M., 2017a. Understanding Information Spreading in Social Media during Hurricane Sandy: User Activity and Network Properties. *arXiv Prepr. arXiv1706.03019.*

Sadri, A.M., Hasan, S., Ukkusuri, S. V., Cebrian, M., 2017b. Crisis Communication Patterns in Social Media during Hurricane Sandy. *Transp. Res. Rec.* 0361198118773896.

Sadri, A.M., Ukkusuri, S. V., Murray-Tuite, P., Gladwin, H., 2014. Analysis of hurricane evacuee mode choice behavior. *Transp. Res. Part C Emerg. Technol.* 48, 37–46.

Sadri, A.M., Ukkusuri, S. V., Murray-Tuite, P., Gladwin, H., 2015. Hurricane evacuation route choice of major bridges in Miami Beach, Florida. *Transp. Res. Rec. J. Transp. Res. Board* 164–173.

Sahlgren, M., 2008. The distributional hypothesis. *Ital. J. Disabil. Stud.* 1–18.

Sarwar, M.T., Anastasopoulos, P.C., Ukkusuri, S. V., Murray-Tuite, P., Mannering, F.L., 2018. A statistical analysis of the dynamics of household hurricane-evacuation decisions. *Transportation (Amst).* 45, 51–70.

Sorensen, J., Vogt, B., 2006. Interactive emergency evacuation guidebook. *Chem. Stock. Emerg. Prep. Program. Dep. Homel. Secur.* Washington, DC.

Tousignant Lauren, 2017. The cost of natural disasters nearly doubled in 2017. NEWYORK POST.

Ukkusuri, S. V., Hasan, S., Luong, B., Doan, K., Zhan, X., Murray-Tuite, P., Yin, W., 2017. A-RESCUE: An Agent based regional evacuation simulator coupled with user enriched behavior. *Networks Spat. Econ.* 17, 197–223.

Wang, Q., Taylor, J.E., 2014. Quantifying human mobility perturbation and resilience in hurricane sandy. *PLoS One* 9, 1–5.

Whitehead, J.C., Edwards, B., Van Willigen, M., Maiolo, J.R., Wilson, K., Smith, K.T., 2000. Heading for higher ground: factors affecting real and hypothetical hurricane evacuation behavior. *Glob. Environ. Chang. Part B Environ. Hazards* 2, 133–142.

Wong, S., Shaheen, S., Walker, J., 2018. Understanding evacuee behavior: A case study of hurricane Irma. <https://doi.org/10.7922/G2FJ2F00>

Wong, S.D., Pel, A.J., Shaheen, S.A., Chorus, C.G., 2020. Fleeing from hurricane Irma: Empirical analysis of evacuation behavior using discrete choice theory. *Transp. Res. Part D Transp. Environ.* 79, 102227.

Xiao, Y., Huang, Q., Wu, K., 2015. Understanding social media data for disaster management. *Nat. Hazards* 79, 1663–1679. <https://doi.org/10.1007/s11069-015-1918-0>

Xu, K., Davidson, R.A., Nozick, L.K., Wachtendorf, T., DeYoung, S.E., 2016. Hurricane evacuation demand models with a focus on use for prediction in future events. *Transp. Res. Part A Policy Pract.* 87, 90–101. <https://doi.org/10.1016/j.tra.2016.02.012>

Yabe, T., Tsubouchi, K., Shimizu, T., Sekimoto, Y., Ukkusuri, S. V, 2019. Predicting Evacuation Decisions using Representations of Individuals' Pre-Disaster Web Search Behavior, in: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.* pp. 2707–2717.

Yang, K., Davidson, R.A., Blanton, B., Colle, B., Dresback, K., Kolar, R., Nozick, L.K., Trivedi, J., Wachtendorf, T., 2019. Hurricane evacuations in the face of uncertainty: Use of integrated models to support robust, adaptive, and repeated decision-making. *Int. J. Disaster Risk Reduct.* 36, 101093.

Yanjie Duan, Yisheng Lv, Fei-Yue Wang, 2016. Travel time prediction with LSTM neural network. *2016 IEEE 19th Int. Conf. Intell. Transp. Syst.* 1053–1058. <https://doi.org/10.1109/ITSC.2016.7795686>

Ye, J., Zhu, Z., Cheng, H., 2013. What's Your Next Move : User Activity Prediction in Location-based Social Networks. *Sdm* 171–179. <https://doi.org/10.1137/1.9781611972832.19>

Yin, M., Sheehan, M., Feygin, S., Paiement, J.F., Pozdnoukhov, A., 2017. A Generative Model of Urban Activities from Cellular Data. *IEEE Trans. Intell. Transp. Syst.* 1–15. <https://doi.org/10.1109/TITS.2017.2695438>

Yin, W., Murray-Tuite, P., Ukkusuri, S. V, Gladwin, H., 2014. An agent-based modeling system for travel demand simulation for hurricane evacuation. *Transp. Res. part C Emerg. Technol.* 42, 44–59.

Zeigler, D.J., Brunn, S.D., Johnson Jr, J.H., 1981. Evacuation from a nuclear technological disaster. *Geogr. Rev.* 1–16.