Evacuate or not?
Modeling the Decision Making of Individuals in Hurricane Evacuation Zones using Influence Diagrams

by

Adithya Raam Sankar

(Under the Direction of Prashant Doshi)

Abstract

Recent hurricanes in the Atlantic region of southern United States triggered a series of evacuation orders in the coastal cities of Florida, Texas and Georgia. While some of these were voluntary evacuations, most of them were mandatory orders that the residents had to follow. In spite of the government asking people to vacate their homes for their own safety, many did not evacuate. Various reasons motivate individuals to choose to stay or vacate. We aim to understand the factors involved in this decision-making process and model these in a partially observable Markov decision process, which predicts whether a person will evacuate or not given his or her current situation. We consider the features of the particular hurricane, the situation that the individual is experiencing, and demographic factors that influence the decision making of individuals. The model is represented as a dynamic influence diagram and evaluated on data collected through a comprehensive survey of hurricane-impacted individuals. We also propose an improvised method of k-means clustering for tweets that considers the context of the text rather than just the cosine similarity.

Index words: Decision Making, POMDP, Impending Disasters, Evacuation Decision

Evacuate or not?
Modeling the Decision Making of Individuals in Hurricane Evacuation Zones using Influence Diagrams

by

Adithya Raam Sankar

B.Tech., SRM University, 2016

A Thesis Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Master of Science

Athens, Georgia

2019

Evacuate or not?

Modeling the Decision Making of Individuals

in Hurricane Evacuation Zones using Influence Diagrams

by

Adithya Raam Sankar

Major Professor: Prashant Doshi

Committee: Adam Goodie
Budak Arpinar

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2019

Acknowledgments

First, I would like to thank my major professor and principal instructor, Dr. Prashant Doshi, Director of Thinc Lab, for his supervision, advice and guidance. His patience, encouragement and faith has truly been the driving force for me to pursue this project with a higher level of enthusiasm.

I am grateful to Dr. Adam Goodie, the co-principal instructor and the graduate coordinator, for his advice and crucial contributions, making him one of the important contributors for this research.

I express my gratitude to my parents, and all my friends, for supporting me throughout the course of my graduate studies. Without their help, both my coursework as well as research would not have seemed possible.

Last but not the least, I would like to thank my fellow researchers at the Thinc Lab. All of the project related discussions, that we were indulged in, have benefited me more than they actually know.

The research was supported by the National Science Foundation (https://www.nsf.gov/) through the RAPID award #1761549.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Decision making is a complex process that is constantly taking place in our brain. Any choice that we make from a set of options is a decision. It could be as simple as whether to take an umbrella or not and get as complicated as deciding whether to evacuate during a disaster or not. The level of uncertainty in the situation determines the level of difficulty in arriving at a decision. Even if the set of options or actions available is the same, the factors influencing our decision will vary with the changes in the environment and ultimately affect our decision. This kind of context-specific decision making increases in complexity with the decrease in observability of the situation. That is, as we lose knowledge about the situation that we are in, we start taking approximations of our current state and hence the decision becomes difficult. Unfortunately, most of the real world situations are not fully observable. Our brain first gathers information about the environment through the sensors and arrives at a hypothesis of what the state could possibly be. With this assumption, we choose the decision that gives the best expected outcome for the available set of actions.

The decision-making process in a software agent is not very different. Decision theory is composed of both probability theory and utility theory. The uncertainty and observability are modeled as a distribution of probabilities, and the goodness of outcomes is specified by the rewards or utilities. In some situations, it is also important to consider the actions of other agents in the environment. This elevates the environment as an interactive decision-making problem where the agent must consider the possible actions taken by the other agents. For the purposes of this research, we will model the environment as a single-agent decision-making problem with the assumption that there are no other agents whose actions are capable of altering the description of the environment.

Moreover, real-time situations involve time-sensitive data and processing. This splits the problem into multiple time slices with a decision to be taken in each step, that is capable of altering the state and observation of the next step. The agent assumes that it may take another decision in the next step again. The transitions between one step to another are also modeled as probability distributions. The method of solving the decision-making problem varies for each of the above mentioned situations. That is, a fully observable, time insensitive, single-agent problem is solved differently as compared to a partially observable, time-sensitive, multiple-agent environment.

In this chapter, we will gain some basic knowledge about the definition, working, and the constraints of a decision-making agent. After which we will explain the scope of this research and lay down the problem domain. This will help in understanding the situation and our approach to decision making. Finally, we will understand the structure of the rest of this document to aid the flow of the reader.

1.1 Relevance to Artificial Intelligence

Artificial Intelligence is an attempt to build a computer software or an intelligent software agent that has human-level ability to think and act. Once this piece of software is able to think and act like a human, the next step is to make it think and act rationally. A rational agent is the one that can think in the right direction and do the action that results in the best outcome[1]. In a case with uncertainty, the agent chooses to pursue in the direction that provides the best outcome expected for the given condition. These are also called intelligent software agents.

This research aims at building an intelligent software agent that considers the given situation and makes the best possible decision. For convenience, intelligent software agents will be referred to as just agents or intelligent agents in this document, unless stated otherwise.

1.2 Intelligent Agents

A human agent is always in the constant process of thinking. The brain receives information from all the sensors in our body - the eyes, ears, nose, tongue, and skin. It then processes this information and decides what the next action should be. For example, if we feel a thorn in our heel, then the next action should be to remove the thorn. And if we see water flooding our house, the decision will be whether to evacuate or not and subsequently where, when, how, and so on.

On the other hand, an intelligent agent is most likely a software program that is designed to receive inputs, process them and return decisions. If this software is running on a physical agent, like a robot, the input is received from various sensors, like cameras, and the decisions will affect the agent's movement in terms of turning the wheels, moving the arms, etc. But if there are no sensors or actuators, then the software receives input through data, and displays the conclusion on the screen after performing the series of tasks it was designed to do.

A rational agent, be it a human or software, makes a decision that maximizes the benefit for itself, given all the information about the situation/environment and the inferences it draws from this data. These environments are characterized in different ways - fully or partially observable, single-agent or multi-agent, deterministic or stochastic, episodic or sequential, static or dynamic, discrete or continuous, and known or unknown [1]. Environments that are fully observable, deterministic, and static are less common in everyday life while the ones that are partially observable (with uncertainty in observations), and dynamic are more prominent in the real world. A rational agent attempts to maximize the average performance over the specified environment.

1.3 Rational Decision Making

If the action chosen by an agent reconciles with that of its preferences and beliefs, then such a choice is known as a rational decision [2] and the process is called rational decision making.

The preferences between the alternative actions are determined by ranking the states of the environment along with the agent's beliefs of being in a particular state. This is often combined with the benefits of executing a particular action. The agent must also adhere to the axioms of utility as described below (where A, B, and C are different actions):

1. *Orderability*: An action is either preferred over another or the agent is indifferent between two actions. That is, $A \succ B$ or $A \prec B$ or $A \sim B$.
2. *Transitivity*: If $A \succ B$ and $B \succ C$ then $A \succ C$.
3. *Continuity*: If $A \succ B$ and $B \succ C$, then there exists a p such that $[p, A;\ 1-p, C] \sim B$.
4. *Substitutability*: If $A \sim B$, then $[p, A;\ 1-p, C] \sim [p, B;\ 1-p, C]$ for any value of p.
5. *Monotonicity*: If $A \succ B$ then, for any value of p and q,

$$(p \geqslant q \Leftrightarrow [p, A;\ 1-p, B] \succsim [q, A;\ 1-q, B])$$

1.4 Problem Domain

The coastal states of the United States of America encounter a number of tropical cyclones every year during the months of August to October. Depending on the speed of winds, they are called as tropical depressions, tropical storms or hurricanes.

1.4.1 Severity of Hurricanes

The strength of a hurricane is determined by the speed of winds around its eye. The wind speed typically increases when the hurricane is moving over water and reduces as it moves over land. The longer it spends over water, the stronger it is when it reaches land. The speed of wind also diminishes rapidly as it transitions over to land from water. Hurricanes are categorized using the standard Saffir-Simpson Hurricane Wind Scale [3]. The factor for this five-point scale is the peak one minute sustained wind at a height of 10 meters over unobstructed exposure. The wind speeds for each category level is mentioned in table 1.1.

Hurricanes reaching category 3 or higher is known as a major hurricane due to their potential for impact on life and property. If the wind speeds are lower than 119 KM/h, then they are classified as tropical storms and tropical depressions. Though these also result in heavy rain, they are not strong enough to cause extensive damage to people and property. It is also important to note that the scale does not specify the extent of hurricane-related impacts like storm surge, flooding due to rainfall, the possibility of tornadoes, etc.

Table 1.1: Saffir-Simpson Hurricane Wind Scale

Category	**Wind Speed**	**Level of Damage**
1	119-153 km/h	Some
2	154-177 km/h	Extensive
3	178-208 km/h	Devastating
4	209-251 km/h	Catastrophic
5	$>$ 252 km/h	Catastrophic

1.4.2 Evacuation During Hurricanes

Assessing the expected impact on the people and the potential damage, both the federal and state governments take a number of precautionary steps to keep the citizens safe. Depending on the severity of the hurricane, a series of evacuation orders are issued to the areas expected to be hit, in an attempt to keep the residents safe. Most of these orders cater to the areas closer to the beach as the impact is even higher when the eye of the hurricane makes landfall. Once an order is announced, a lot of supplementary preparations become underway in order to ensure the safety of the citizens. These actions include: setting up of evacuation centers, ensuring availability of supplies etc. Despite repeated attempts by the government, the decision for evacuation is made solely by the residents. When inquiring closely, it can be found that there are a host of other factors taken into account by the residents during the process of making this decision.

1.5 Research Contributions

In the previous sections, we saw the basic concepts of decision making. In this section we put forth our claims and contributions through this research:

- The primary focus of this research is to model the decision-making process for evacuation during an impending hurricane.
- This research is expected to open up a new direction in computationally modeling human behavior when under uncertainty.
- The findings of this research would help emergency service personnel to better understand the decisions and assist the citizens appropriately during a future hurricane impact.
- We also introduce the context based similarity as a distance metric in the k-means clustering algorithm for natural language processing.

1.6 Structure of this Work

Being a less explored application of the decision-making domain, the availability of previous research is considerably limited. This demands the need for explaining our methods and approaches more clearly, so as to allow the extension of this work easy and feasible. The report of our research is structured as follows.

The focus of *this chapter* is to give a general idea about the research area and introduce some basic terminologies. It also explains the problem domain and outlines our contributions to the field.

In *chapter 2*, we briefly review some of the concepts that form the building blocks of the different modules of this project. Most importantly, it will help in understanding the rest of the report.

Some of the previous work that closely relates to the modeling of evacuation decisions have been discussed in *chapter 3*. Though they do not significantly impact the course of this research, it is recommended to know the existing methodologies so as to understand the advantages of our work.

Being an application-oriented research project, the underlying data forms the foundation on which the rest of the research is based. *Chapter 4* explains the different sources of our data and also presents the preprocessing methods.

Chapter 5 describes the first computational module of our project, the natural language processing. Due to the nature of our data, there was a need for automatic processing of the texts. This chapter talks about the different techniques implemented and also explains the results obtained.

Based on the variety of information we processed, we present a concise list of variables affecting the decision-making process during hurricane evacuation in *chapter 6*. We also discuss the dependencies between the variables and the proposed decision-making model along with the results obtained.

Finally, in *chapter 7*, we conclude by pointing out the limitations we faced throughout the study along with some possible ways of extending this research.

At the end of the report, in the appendix, we have included the survey instrument, some supplementary results for natural language processing and additional information about the decision-making model.

Chapter 2

Background

In this chapter, we will briefly go over various methods and techniques that form the building blocks of this research project. These concepts will help in understanding the later chapters. We will understand the basics of web crawling, a popular way of gathering data. Then, we will see the working of K-Means clustering as a method for unsupervised learning. After that, the ways of computing the similarity of two given texts. And finally, we will understand the different decision-making models.

2.1 Web Crawling

One of the major sources of data for any kind of research is the internet. The availability of information and uniformity of representation makes it easy for a software program to read and save the required content without much human interference. There are a number of packages and Application Programming Interface (API) developed for many of the modern programming languages that are designed to read and understand the contents of different web pages. There are also packages that are capable of parsing a specific type of web page, e.g. news articles. Thus, these packages make it easy to write a software program that reads, analyzes and saves the required web pages for further processing. This kind of software program is often referred to as a spider as it is "crawling the web".

The basic working of a web crawler is to take in a seed Uniform Resource Locator (URL), and process all web pages that are accessible from there. First, all the links listed on the given URL is added to a list of pages to be crawled. After that, the required information is searched for within this page. This can be done in various ways. Since most of the websites are written using Hypertext Markup Language (HTML) tags, we can look for a specific sequence of nested tags to search for the required content. Another common method is to search for a class name attached to any tag. The major limitation in both of these methods is that the structure of the web pages need to be known beforehand in order to develop the crawlers. This can be overcome by the use of domain-specific packages that already know how to look for relevant content in a given URL.

In recent times, any web page designed with the modern set of principles and guidelines contain a large number of links. This means that the number of links to be processed by our crawler would increase exponentially with each page accessed. But, following some basic filtering techniques can restrict the crawler from going off the domain or topic. It is also common practice to give an exit criterion to stop the crawler from execution. This could be a cap on the number of pages processed or the number of required pages found. It is also good to use a combination of these conditions so as to reduce the time spent on processing irrelevant web domains.

Another factor to be considered while designing these crawlers is the type of links that they encounter. The web is filled with a lot of files that are not in the HTML format, e.g. Portable Document Format (PDF) files, Portable Network Graphics (PNG) images etc. As these files do not contain HTML tags, the program would possibly crash while processing them. The crawler should either be capable of processing these links differently or be able to avoid them. To add to the complexity of the crawler, some e-mail IDs and telephone numbers also appear as hyperlinks in a HTML document. These do not have a separate page and will throw exceptions if not handled properly.

2.2 K-Means Clustering

K-means clustering is one of the simplest and most popular unsupervised learning algorithms. The procedure follows an easy way to classify a given set of data points into a certain number of clusters, k, fixed apriori. The algorithm is executed in the following order [4]:

1. The main idea is to start with k centers, usually initialized randomly, one for each cluster. These centers should be chosen in an intelligent way as different starting points yield different results. So, the better choice is to place them as far away, from each other, as possible.

2. Each point in the dataset is then associated to the nearest center calculated using Euclidian distance.

3. Once every point is mapped to a center, a new set of centers are calculated from the clusters obtained.

4. Steps 2 and 3 are repeated until there are no changes to the centers in consecutive iterations.

Some of the advantages of the K-Means algorithm is that it is fast, robust and easy to understand. It is considerably efficient with a time complexity of $O(tknd)$, where n is the number of data points, k is the number of clusters, d is the number of dimensions of each point, and t is the number of iterations. It is also known to give great results when the data points are distinct and well separated from each other. This allows the algorithm to draw clear boundaries between the clusters.

2.3 Markov Decision Processes

A particular type of problem where the environment is fully observable, stochastic and sequential in which the transitions and rewards only depend on the current state and the current action, is called a Markov Decision Process (MDP) [5]. An MDP is defined by $< S, A, T, R >$ [1] where S is the set of all possible states and A is the set of all actions for each state. T is the transition matrix that holds the probability of the change from s, the current state, to s', the next state when taking the action a. It is represented by $P(s'|s, a)$. R is the reward function that holds the expected rewards for each state s.

Being a stochastic process, there cannot be a single fixed action sequence that will solve the problem. A solution that specifies the optimal action from each state of the environment is called a policy. It is usually denoted by π and $\pi(s)$ is the action to be taken from state s when following the policy π. Having a complete policy, the agent always knows the next optimal action irrespective of the outcome of the previous action [1].

2.4 Partially Observable MDPs

The description of MDP assumes that the environment is fully observable, which means that the agent always knows it's current state. In real-world situations, the environment is usually not completely observable. The agent receives a set of observations through its sensors and arrives at an estimate as to where it would most-likely be. The estimate is usually a distribution of probabilities associated with each state. This generalized problem is called a Partially Observable Markov Decision Process (POMDP).

Since POMDP is an extension of MDP, it's definition consists of all the elements of an MDP along with the observation function or the sensor model. Thus, it is defined as $< S, A, \Omega, T, O, R >$, where the elements S, the set of all states, A, the set of all actions, T, the transition model, and R, the reward function, are the same as that for an MDP. Ω is the set of all possible observations and O is the observation function represented as $P(o|s, a)$ which is the probability of getting the observation o when executing action a from the state s.

2.5 Influence Diagrams

An influence diagram (also known as a decision network) is a compact graphical and mathematical representation of a decision problem. It is a generalization of Bayesian Networks that can model and solve probabilistic as well as decision-making problems. It accurately depicts all the factors involved in the decision-making process along with the dependencies between the variables. A complete influence diagram describing a decision-making problem has three kinds of nodes, namely:

1. *Chance Node*: The nodes drawn as ovals or rounded rectangles represent the random variables affecting the network. In an MDP or a POMDP, these are the states and observations. Each node is associated with a conditional probability table, or CPT, that is indexed by the values of its parents.

2. *Decision Node*: The nodes drawn as rectangles represent the set of actions available to the decision-making agent at each point in the process.

3. *Utility Node*: The nodes drawn as diamonds or hexagons represent the rewards received by the agent for any possible combination of the values of its parents. This node carries the reward function in a MDP or a POMDP.

2.6 Dynamic Influence Diagrams

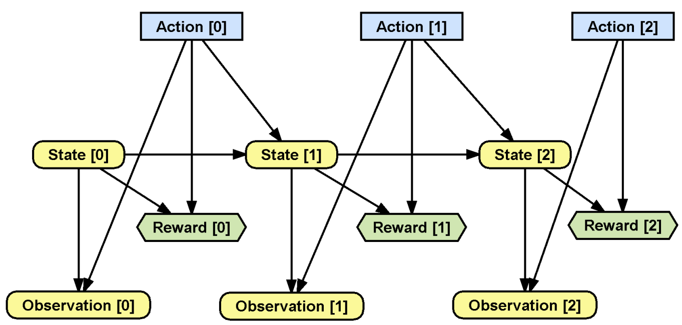

Figure 2.1: A sample dynamic influence diagram showing three time steps

As we have seen earlier, MDPs and POMDPs are sequential decision-making problems spanning over multiple time steps. The dynamic influence diagrams extend the concept of influence diagram to multiple time-slice situations. This means, there can also variables that affect other variables across time steps. This allows the representation of transition functions that are part of the definition of the above-mentioned decision-making models. A sample dynamic influence diagram is shown in Fig. 2.1.

Chapter 3

Related Research

Being a fairly unexplored area of research, there has not been significant attempts in computationally modeling human behavior during the evacuation of hurricanes. This chapter discusses some papers related to the methods or the model presented later.

3.1 Evacuation During Hurricane Sandy: Data from a Rapid Community Assessment

Researchers from the New York City Department of Health and Mental Hygiene analyzed data from a mental health needs assessment survey conducted among residents of New York City soon after Hurricane Sandy [6]. A subset of questions relating to the factors affecting the evacuation and evacuation timing after mandatory evacuation orders were issued to these residents were analyzed statistically.

While 49% of the 420 respondents had evacuated, the others had not. It was found that those who had witnessed the sufferings of others during the World Trade Center attacks had evacuated more than those who did not. More residents having their houses on lower floors had evacuated before the storm. It was also identified that those who reported significant damages to their household, after Hurricane Sandy, showed an increased rate of evacuation than those with fewer damages. Overall, the evacuation rate among the residents of the targeted regions of New York was less than optimal considering the warnings to evacuate was issued before the hurricane made landfall.

In an attempt to verify whether the above-mentioned set of factors are also significant predictors of evacuation during Hurricane Harvey and Hurricane Irma, similar questions were included in our survey. The results revealed that many of these factors were important decision modifiers in the states of Florida, Georgia, and Texas as well.

3.2 Modeling Hurricane Evacuation Decisions with Ethnographic Methods

A group of researchers from the University of Florida and the Florida International University interviewed residents from south Florida who experienced Hurricane Andrew (1992) and Hurricane Erin (1995) [7]. Based on the responses of the participants, a decision model was generated in the form of a tree composed of binary questions. Later, a general decision model was derived from these individual models.

Although the model was able to achieve high accuracy of prediction, it is important to note that there is no room for ambiguity or an option to skip a question. In a situation where there is no answer to a particular question, the model fails altogether. Moreover, the process of gathering more information before taking a decision is not facilitated, though it is mentioned that it is the first step. Also, it does not consider the actual situation that the person is in, but only takes their observations into consideration.

3.3 Evacuation Decision Making and Behavioral Responses: Individual and Household

A survey of the history of literature on decision-making during hurricanes reveals a lot of information about how the trends in research changed over the years [8]. Initially, it was believed that decisions are solely based on different kinds of warnings. During this time, it was also identified that different people react differently to the same announcement of disaster warning [9]. Although warnings play an important role, the perception of risk also has a significant part in influencing the decision to evacuate or not. In fact, warnings do not directly influence evacuation, rather induce the perception of higher risk which in turn motivates evacuation. A variety of other reasons like traffic and roadblocks were also identified as important factors influencing the evacuation. Demographic features of the decision maker like age, the presence of dependents, etc are also major contributors to the evacuation outcome. Finally, the researchers suggest three potential objectives for the future including a better prediction model for evacuation behavior and a geographically specific understanding of evacuations in order to help the officials focus on areas where the residents do not follow the orders and put themselves in danger.

3.4 Heading for higher ground: factors affecting real and hypothetical hurricane evacuation behavior

While assessing the determinants of hurricane evacuation in coastal regions of North Carolina, it is identified that the presence of evacuation orders and the perception of risk from flooding are important [10]. A logit model is used to examine the magnitude and direction of the independent variables. The factors related to risk like living in a mobile home are found to be significant predictors of evacuation decisions rather than household income. Also, the difference in evacuation influencing factors between real and hypothetical situations are highlighted. This shows that the use of mock scenarios for finding the decision variables would be misleading.

3.5 Behavioral Model to Understand Household-Level Hurricane Evacuation Decision Making

This research proposes a mixed-logit model composed of numerous factors and variables [11]. Some of these are state/location, home ownership, prior hurricane experience, education level, dependents, and income. Also included were some random parameters in order to circumvent the assumptions made by the discrete-outcome models. The resulting model was tested on the data from a survey conducted after Hurricane Ivan.

Most of the variables specified in their model were included in our survey questionnaire. While some of them were found to be significant predictors of evacuation and included in our model, those that were not found to be influencing the decision were dropped from subsequent analyses.

3.6 Summary

A long history of research on the evacuation behavior of residents in areas frequently affected by hurricanes is present. Most of these focus on the psychological, behavioral and statistical modeling of the decision-making process rather than a computational decision-making process. That way, this research will serve as a step ahead in the domain and open a new direction in the modeling of decision making during hurricanes.

Chapter 4

Data Gathering

For the purposes of this research project, data from a variety of sources was required. In order to understand the correlation between the situation of the citizens and the way they reacted to it, a directed survey was conducted. Another platform where the people publicly post their decisions, sometimes with the reasons, is the online social media. And lastly, the reports from news agencies who interacted with the citizens from affected and to be affected areas are also expected to be useful. The rest of the chapter explains how we gathered data from these sources and also some constraints and characteristics of the retrieved data. Finally, we also talk about how this data was processed and stored.

4.1 Directed Survey

In order to achieve the target of this research, it was important to understand how the real people responded to different situations. This will also serve as the ground truth that the model will have to behave like as it depicts the actual actions of the people along with the reasons behind the same. We administered a survey over the course of one week in December 2017 to participants located in the evacuation zones of hurricanes Harvey and Irma to collect data for understanding and modeling their evacuation decision making.

4.1.1 Survey Questions

The survey consisted of 39 questions which dealt with a combination of experiential factors, psychological biases as well as demographic information. Most of these questions had a list of options from which the participant could choose. Many of the questions also included a

"Prefer not to answer" option to protect the privacy of the participant in case they were not comfortable answering a question. It also consisted of an open text question where the participant was able to enter the reason for their decision depending on the answer to the question which asked whether they evacuated or not. The participants also had to self report their state, county and zip code in order to reassure that they actually received an order. The full list of questions is attached in appendix A.

4.1.2 Participants and Responses

The participants of the survey were recruited from areas that received an evacuation order in Texas, Florida and Georgia. While 12 counties in Texas and 23 counties in Florida received evacuation orders, Georgia received orders by zip codes. Some of these were mandatory evacuation instruction whereas some others were voluntary. Also, some of the orders were changed from voluntary to mandatory depending on the changes in path and expected impact of the respective hurricanes. The time between the announcement of the order and the expected impact was not taken into consideration in the analysis as the citizens are expected to vacate as soon as the order was given out irrespective of the time left. The targeted areas are shown by a red shade in figure 4.1.

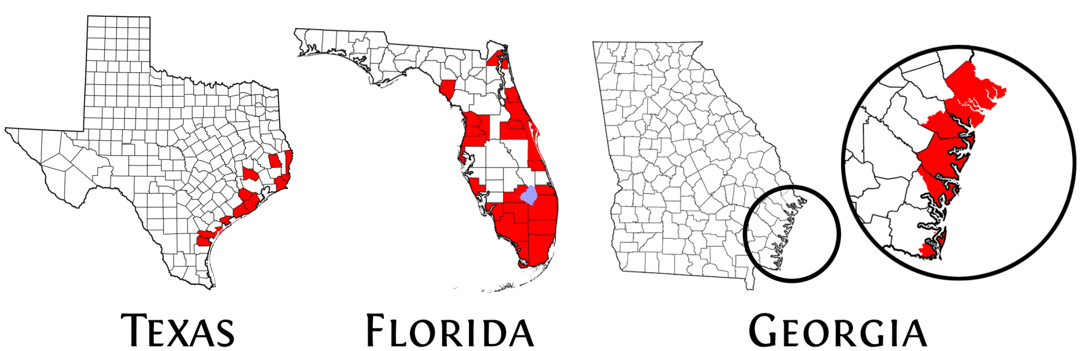

Figure 4.1: Maps of the target states with county boundaries showing the surveyed regions in red

The survey received a total of 825 responses, from 54 pilot participants recruited between 14-15 December 2017, and the remainder recruited between 18-21 December 2017. This included 3 duplicate participants who were identified and eliminated, based on their IP address and demographic factors. Out of the remaining 822 participants, 330 self-reported evacuations during the hurricane, while 492 said that they did not.

4.2 Social Media Posts

Throughout the duration of the disasters under consideration, a large number of the public is expected to share their views, opinions, and situations on various social media platforms, such as Twitter and Facebook.

4.2.1 Twitter

Twitter is a micro-blogging service where the users can publicly share their thoughts. The posts, known as tweets, include texts, images, videos, short animations, etc. Twitter sells historic tweets for research as well as commercial purposes. A total of 778863 tweets and retweets from 243762 different handles posted between August 16, 2017, to September 16, 2017, was retrieved. The filtering was based on location, the presence of selected keywords and tweets from a set of selected handles. There were also tweets from news agencies, government departments and officials that spoke about evacuation orders, weather conditions, and other official announcements.

4.2.2 Facebook

Unlike Twitter, Facebook is a social networking service that allows its users to post content of any length. All the posts shared on Facebook is visible only to the user's friends unless specified otherwise. Thus, these posts are not available for access, neither for research nor commercial purposes. However, there is a feature called "Pages" that are created by the users for interaction between people with common interests. Some of these are public where the posts can be viewed by everyone. Such posts are accessible through the Facebook API. A total of 1116 posts were retrieved from selected "Pages" from August 16, 2017, to September

16, 2017. These pages were selected through a manual search on Facebook using the names of the hurricanes. An additional set of pages managed by national and local news agencies as well as government organizations that track and report about hurricanes were also selected.

4.3 News Articles

The primary task of mass media, be it print or televised, is to deliver the latest information to the public in a timely manner. During a natural disaster, like a hurricane, many news agencies attempt to capture and publish the situation from the areas expected to be affected. In this attempt, they tend to interact with the citizens in these regions and, many times, even quote them. A part of this interaction deal with their preparations, past experiences, evacuation decisions and, the reasons also. Although the probability of articles with such concentrated information is low, it could reveal some important aspects of human behavior.

Expecting the same during the hurricane season as well, a set of local and national news reporting websites were searched to retrieve the articles published online. A total of 3283 articles, that passed through a simple keyword matching filter, were saved. A small set of meta-data, such as date published, media links, prominent words in the article, were included along with the complete text. Since the web scrapping was performed after the target hurricanes hit, many of the articles were talking about the after effects and the rescue operations that took place once the hurricane had passed.

4.4 Data Processing and Storage

Anyone posting a tweet is bound to use a combination of text as well as emoticons. One simple way to process them is to convert it to text which also allows for easy storage in a relational database. Although emoticons reflect the mood of the user, assisting in identifying the sentiment of the texts, they are expected to interfere in the comparison of the context of the texts. For example, consider two tweets talking about completely different things that have the same set of emoticons. Since the emoticons have been converted to text, these two tweets would possibly be grouped together. In order to avoid this, the emoticons were removed from the tweets before analysis. Similarly, it was also found that mentions to specific people were only adding noise to the tweet as these do not provide any context or have any meaning. Thus, these were also removed from the texts before analysis. However, hashtags, another popular addition to simple text in tweets, would be helpful in grouping similar tweets. Thus, these were allowed to be considered in the similarity analysis.

On the other hand, news agencies tend to add general information about the context of an article before going into the main content. In order to eliminate the interference of the header content of an article in the analysis, only the part of the article from the sentence with the first occurrence of the word "*hurricane*" until the end was taken into consideration, while the initial portion was trimmed.

In order to facilitate easy readability and uniform access from multiple machines for data processing, the data from all the aforementioned sources were stored in a relational database. This database was hosted on a local server setup in one of the lab machines allowing access to the data within the university network. This also facilitated some basic analysis and filtering to be performed with Structured Query Language (SQL) queries.

Chapter 5

Textual Data Processing

As we have seen in chapter 4, the underlying data in this research involves a lot of unstructured and uncategorized text. Processing such data programmatically involves a number of challenges. The first step is to train a model to understand the given texts. This becomes more complex when the texts are from the general public as there is an innumerable number of ways to express the same thing. Next, we should be able to use the model to categorize the text based on what they mean. Once categorized and tagged with a set of target keywords, it will be easier to work with. The process of categorization involves the computation of probabilities of the text being in each of the categories and then selecting the category with the highest probability. This probability can be computed using Bayes' rule as:

$$\underset{c}{\operatorname{argmax}}\, P(c \,|\, text) = \underset{c}{\operatorname{argmax}}\, P(text \,|\, c)\, P(c) \tag{5.1}$$

where the $P(text \,|\, c)$ comes from the model and $P(c)$ is the ratio between the number of texts in this category to the total number of texts [1]. For the purposes of this research, we have used a slightly different approach in arriving at the $P(text \,|\, c)$ by using the contextual similarity between the *text* and a representative text from category c.

In this chapter, we will first see about the common ways in which two texts can be compared. We will also introduce the concept of cosine similarity and put forth two packages for the Python language that extend this to include the context in computing the similarity. After that, we will explain a modified clustering algorithm that implements these packages and proved to be yielding relevant clusters. Finally, we will see how the raw data was classified and how it will be used in the upcoming stages of the research. Throughout this process, we will also talk about some of the difficulties that we faced and the techniques that we followed to overcome them. It is also important to note that this methodology can be extended to any text corpus, a set of texts, and the experiments can be repeated to achieve similar results as long as there are similarities in the texts that are recognizable by a computer program.

5.1 Text Similarity

A comparison between any two entities is usually represented as a percentage that describes the extent to which they are similar. This can also be converted to a probability of the first text being similar to the second one. The result of this computation is symmetrical, i.e. the probability of the first text being similar to the second one is equal to the probability of the second text being similar to the first one.

5.1.1 Cosine Similarity

One of the common ways to compute the similarity between texts is to represent them on a vector space and then calculating the cosine of the angle between them. The reason for choosing the cosine is that the cosine of identical vectors results in a 1 and that of orthogonal vectors results in a 0 [12]. Moreover, the result of a cosine of diametrically opposite vectors is a -1, making the range of the similarity function to be -1 to 1. This can be normalized later in order to arrive at the probability of text being similar to another one.

In order to make sure that the texts acquire unique positions on the vector space, the *terms* in the text itself is used to arrive at the vectorized representation of the text. A *term* may be defined as a word or a phrase in the text. Each *term* has its own dimension in the vector space. When a text has this particular term, then the texts gets a positive value in that dimension. Though the number of dimensions can be high in number, the texts have a limited number of terms. Thus, the number of non-zero values for each text is low. This can be made even more unique by taking the number of words in the text also into account.

This method of vectorization is known as TFIDF or Term Frequency – Inverse Document Frequency.

5.1.2 spaCy

spaCy is a free and open source industrial-strength natural language processing library developed in Python and Cython [13]. It is capable of tagging parts of speech to the words in a text, recognizing named entities and parsing dependencies using convolutional neural networks. It features one of the fastest syntactic parsers in the world. Some of its other capabilities include tokenization, lemmatization, sentence boundary detection, and text classification. It also has a built-in similarity analysis function that computes the extent of similarity between any pair of texts, words, phrases or paragraphs, whose underlying principle is cosine similarity.

Although spaCy supports more than 28 different languages, it ships with pre-trained models for a selected few of them. These models include word vectors, binary weights, lexical entries, and some configuration options. The word vectors are represented in a 300-dimensional space corresponding to the meaning of the words. The vectors for phrases and sentences are computed by averaging the vectors of words in them. These vectors are used in arriving at the similarity of a pair of entities. The similarity function returns a scalar floating point score that ranges from 0 to 1. A sample set of similarity scores obtained from spaCy is illustrated in table 5.1. As we can see, the output is a symmetric table indicating that the order of the entities does not matter while computing the similarity. The terms *parents* and *family*, being closely related, have a high score of 0.69. While *hurricane*, being considerably less relevant to *parents* and *family* yields a low score of 0.17 and 0.19 respectively.

Table 5.1: Sample similarity scores from spaCy

	Parents	Family	Hurricane
Parents	1.0	0.69	0.17
Family	0.69	1.0	0.19
Hurricane	0.17	0.19	1.0

5.1.3 GenSim

GenSim[14] is one of the robust and efficient software for unsupervised semantic modeling of plain text. Its python API features a number of different modules that allow various natural language processing methods to understand and work with textual data. In particular, the *docsim* module in the *similarities* package enables the comparison of texts once the complete dataset is parsed using Latent Semantic Indexing (LSI) [15]. After this, the text can be compared with each other with the help of $MatrixSimilarity$ function. This returns a two-dimensional table, where the rows and columns represent each of the texts in the dataset and each cell stores the extent of similarity between the row and the column texts. Since the dictionary for LSI is created by just using the words in the dataset, the algorithm does not consider the actual English dictionary meaning of the words in the texts.

5.2 Motivation for Unsupervised Clustering

The responses from the survey revealed a large number of reasons with varying counts. However, since the population size was small, it would be unfair to believe that it is an exhaustive list. Thus, the primary motive of the clustering step was to identify new reasons from other sources of data. For this, the algorithm was to cluster similar texts together, and then deduce the common theme in each cluster.

When we use supervised clustering, the objective is to match the texts to a specific set of cluster themes. In this case, every data point to be clustered is forced to be put into one of these clusters. This process assumes that all the data points definitely belong to at least one of the clusters. Whereas, the data available to us does not follow this condition. Moreover, this will not facilitate the identification of new reasons either.

Specifically, the tweets consisted of a large amount of noise including those that dealt with evacuations not related to hurricanes. In order to filter these, it would be required to set a threshold value that will decide the candidacy of a tweet in a cluster. The specification of this threshold value is subjective and time consuming as we do not have the ground truth of clusters.

Thus, the unsupervised clustering method was chosen to group the data points based on a theme. Additionally, due to the volume of data and the amount of noise, it was necessary to perform this clustering hierarchically to narrow down on the reasons.

5.3 Hierarchical Context-Based Clustering

Now that we can get the similarity between any pair of texts, we can proceed to use this as a distance measure to perform clustering. The K-Means clustering algorithm presented in chapter 2 is not designed to work with textual inputs. So far, to cluster texts, the algorithm is given just the vectorized inputs and then the vectors are remapped to the texts once the clustering is completed. In such a case, the algorithm treats the problem like a usual numerical clustering on a multi-dimensional space. This does not consider the context of the texts. Another important factor to note is that the computed centers, using averaging of the vectors in a particular cluster, might not correspond to a text in the input dataset. This means the clusters are centered around a point that is not a representative of the texts in the cluster. In order to address these concerns, we modified the K-Means clustering algorithm to work with the raw texts during the distance computation phase but use the vector representation of the texts while updating the centers. Once the average vector is found, a vector corresponding to the texts that is closest to the average vector is selected and the text that it represents is used in the next iteration as the new center. The algorithm proceeds until there is no change in the centers of consecutive iterations. This process is illustrated in algorithm 1. Each tweet was vectorized using SpaCy, as explained in the previous section. A second level of clustering was performed on those clusters that contained many tweets which identified distinct reasons for evacuation or not evacuation. The clustering method was similar to the method used in the first level of clustering.

Algorithm 1: Tweet Clustering

```
input: k: Number of Clusters
input: tweets: List of Tweets
input: initCenters: List of Tweets to be used as initial centers
text2vecnlp_map ← {};
foreach tweet in tweets do
    text2vecnlp_map[tweet_text] ← [tweet_nlp, tweet_vector, tweet_BoW]
end
centers ← initCenters;
clusters ← [[ ]*len(centers)];
numIterations ← 0;
repeat
    // compute the clusters
    Find the most similar center for each tweet;
    foreach tweet in tweets do
        similarities ← [ ];
        foreach center in centers do
            similarities.append(tweet.similarity(center));  // computed using SpaCy
        end
        clusters[similarities.index(max(similarities))].append(tweet);
    end
    // compute new centers
    new_centers ← [ ];
    foreach cluster in clusters do
        center_vector ← average([tweet_vector for tweet_text in cluster]);
        centerTweet_vector ← nearestTweet(center_vector);
        centerTweet_text ← getText(centerTweet_vector);
        new_centers.append(centerTweet_text);
    end
    if centers == new_centers then
        break;
    else
        centers ← new_centers;
        numIterations ← numIterations + 1;
    end
until numIterations < 10;
return clusters, centers
```

5.4 Experiments and Results

Initially, the clustering algorithm was implemented on the complete corpus of tweets. This resulted in very generic clusters as the tweets included news reports, weather updates, public announcements, etc. This kind of high-level clustering did not isolate the tweets mentioning reasons. Thus, the tweets were filtered using a simple keyword matching technique and those tweets with the letters "*evac*", occurring in that specific order, were extracted for further analysis. This resulted in a subset of 29213 tweets that included all the different forms of the word "*evacuate*" along with the ones that had any popular hashtags used during the evacuation for Hurricane Harvey and Hurricane Irma. Also, after a few different experiments and exploration, it was found that GenSim is not reflecting the true similarity between data points due to the way in which the dictionary is created for its similarity computation. Hence, it was discarded as a similarity measure and SpaCy was only used during the clustering process.

5.4.1 First Level of Clustering

The extracted subset of tweets was divided into 25 different clusters. For this, 20 of the initial centers were seeded by choosing the tweets that mentioned some distinct reasons identified in the survey responses. The remaining 5 were selected at random in order to allow the finding of any new reasons. After convergence, the topic of the clusters, if any, were identified by reading a random set of tweets from the cluster. While few of the clusters did not reveal any specific topic, others talked about: news announcements, evacuation orders, preparations, and questions to family members. The clustering also helped to filter out the tweets that were irrelevant to the hurricane domain as they mentioned evacuations due to other reasons such as bomb explosions elsewhere. Three of the clusters, with counts 4,984, 1,443, and 4,458, revealed topics that were talking about reasons for evacuation-related actions due to the hurricanes: What decisions were taken; Location and pets being the reason; Traffic and family-related reasons, respectively.

5.4.2 Second Level of Clustering

Tweets in each of the three identified clusters were further divided into 10, 5 and 10 clusters respectively, thereby performing hierarchical clustering. Once again, some of the centers in each cluster were initialized based on the target clusters and the rest were chosen randomly. After convergence, tweets that were at least 95% similar to the center of the cluster were counted as belonging to the cluster, and the rest were dropped out. This was implemented in order to account for those tweets that did not belong to any of the clusters. While our hierarchical clustering method did not uncover all the reasons mentioned in the tweets, it successfully yielded those that were prominent. For example, the true counts (from a manual tagging effort) indicated that flooding was the top reason for evacuation and the counts from the "Location and Pets" cluster agreed with this (shown in figure 5.3). Some additional results that supplement the reasoning of the minimum similarity score are explained in appendix B.

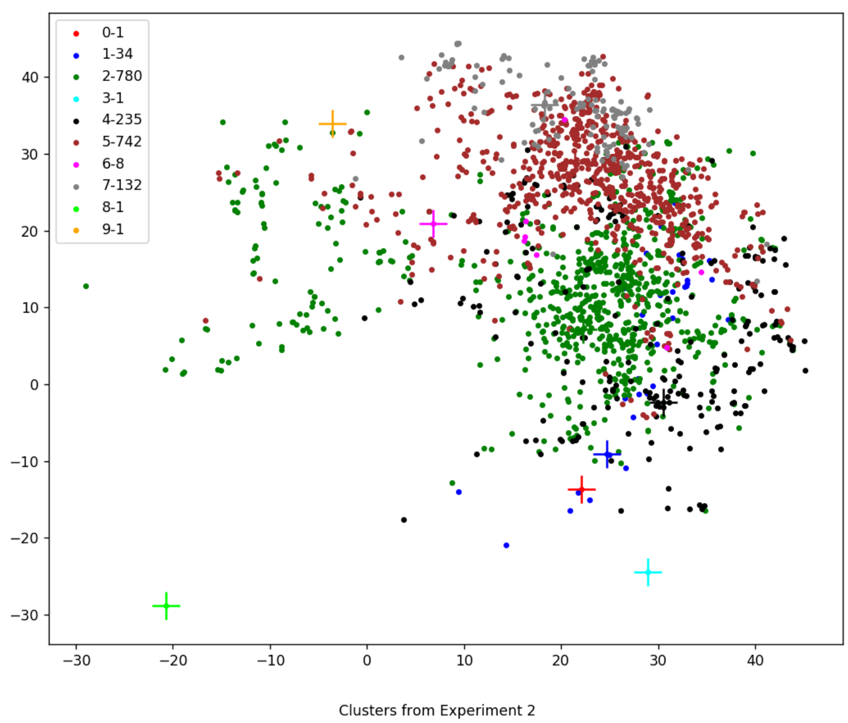

Figure 5.1: Visualization of second level of clustering on "Evacuation Decisions" cluster with minimum similarity score of 95%

5.4.3 Visualizing the Clusters

The results of the second level of clustering are visualized as scatter plots. Since the actual clustering was performed with 300 dimensions, it has been reduced to two dimensions using the t-SNE package [16] in Python. Each point on the scatter plot represents a tweet. The tweets belonging to a single cluster are represented by the same color.

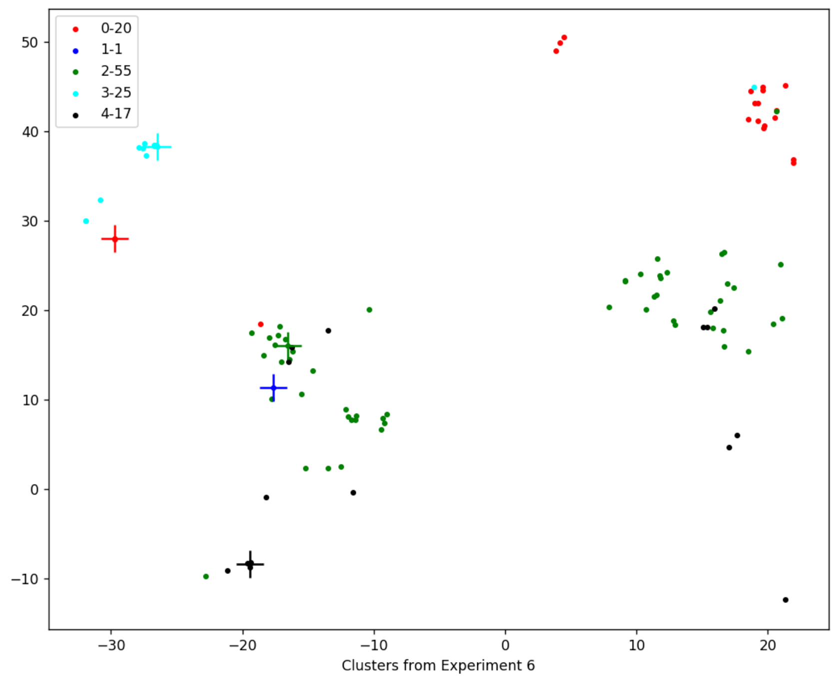

Figure 5.2: Visualization of second level of clustering on "Location and Pets" cluster with minimum similarity score of 95%

5.5 Limitations of Clustering Algorithm

One of the major obstacle in achieving the desired result is the unavailability of robust natural language processing algorithms. Most of the currently available algorithms work exceptionally well on structured and formal text. But when it comes to text posted by general public, they fail to accurately capture the context. Adding to the complexity is the character limit imposed by Twitter. This induces the people to use shorted versions of words with many characters. Most of the time, there is no standard way of shortening a word.

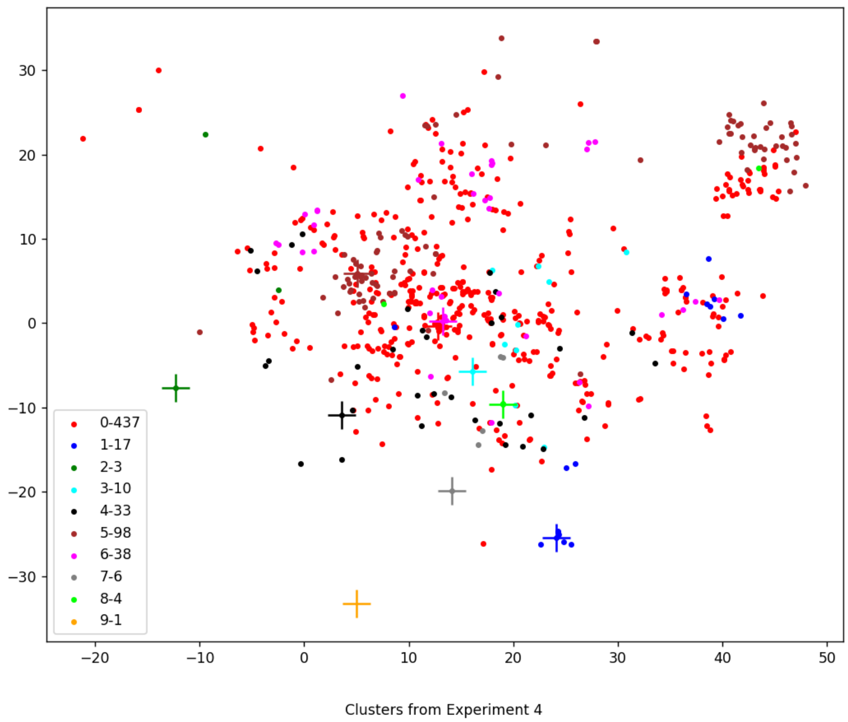

Figure 5.3: Visualization of second level of clustering on “Traffic and Family” cluster with minimum similarity score of 95%

Also, the same short form might be used to represent different words by different people. This makes it difficult for the program to understand the context accurately.

Chapter 6

Decision Model

The goal of this research is to build a model that can predict the actions of an individual during a hurricane evacuation. In this chapter we describe the structure of the POMDP model and explain the reason behind the choice of variables and its probability distributions. We will look at the experimental setup along with the results for the training and testing sets of data. Finally, we will understand the limitations of the research that has constrained the generalized implementation of the model. While it is possible to circumvent a few of them in future research attempts, there are also some that cannot be avoided at all.

6.1 Statistical Analysis

The primary source of information about the factors affecting the process of making the decision to evacuate was the directed survey conducted among the residents of the affected areas as explained in chapter 4. Initially, the responses from this survey were analyzed using statistical methods in order to understand the correlation between the question about the evacuation decision and the others.

Logistic regression was conducted to predict self-reported evacuation (a binary, yes/no variable), from those variables demonstrating statistically significant ($p < .05$) correlations with evacuation. The logistic regression was conducted in two blocks: first demographic variables and then experienced based variables. The logistic regression predicted evacuation from the significant demographic predictors (age and number of children) in Block 1, and from the significant experience based predictors (times asked to evacuate, assessed risk, preparation level, prior trauma) in Block 2. Note that 301 participants (37.2%) did not respond to at least one item in the regression, and were excluded. The evacuation choices of 75.8% of the remaining participants were predicted correctly using the full model with both blocks of variables.

It was found that the number of times people were asked to evacuate, the perception of neighbors having evacuated, and the assessment of the level of risk they experienced, correlated with each other. Moreover, these were also significant factors in predicting the decision to evacuate. While prior trauma was a less significant predictor, prior experience with hurricanes did not correlate with the decision to evacuate at all. Among the demographic variables, sex, race, relationship status, education level, number of elder dependents in the home, the level on which one lived (1st floor, 2nd floor, etc.), and whether the participant owned or rented his or her home, were all unrelated to evacuation.

The most common reasons provided for not evacuating were that respondents did not receive an order to evacuate or did not believe they were in the evacuation zone (118 out of 492); assessed risk had been mitigated through other steps that had been taken (104 of 492); did not believe they were at high risk in general (70 of 492); or did not believe risk was significant in their particular location (65 of 492). No other category of reasons were provided by more than 35 respondents. Among those who did evacuate, top reasons for their decisions included warnings from government and media (70 out of 330); concern about the severity of the storm (56 of 330); general references to safety or comfort (51 of 330); and concern for the well-being of family members or other close associates (42 of 330). These observations are highly consistent with the results from the statistical analysis.

6.2 Motivation for Probabilistic Graphical Model

As mentioned earlier, most of the prior research attempts in predicting the evacuation behavior took the deterministic route. The major limitation of this approach is that the model can only work on completely known situations. In a case where any of the variables/parameters are unknown, then the model becomes invalid or one of the options need to be chosen randomly. This can be overcome by implementing a probabilistic model where there is an option of arriving at an inference based on other factors. Here, even when not all the variables are known, the values can be inferred based on the values of the parents or children.

After the initial statistical analysis and the identification of basic parameters, it was found that not all the variables are dependent on every other variable. This demanded a graphical representation of the model that explicitly defines the influences between the variables. This would also help in reducing the dimentionality of the problem as it reduces the state space.

As for the hurricane domain, the decision of whether to evacuate or not is taken at different time steps based on the availability of information. For example, some people choose to evacuate immediately after an order is announced. While others tend to wait and understand the situation better before making a decision. In order to capture this nature of the problem, the model was split into four time steps and represented as a dynamic influence diagram. The number of time steps was decided based on the number of major changes in path or strength of the hurricanes Harvey and Irma in 2017.

6.3 Decision Variables

As for any POMDP, this model is also composed of multiple variables that influence the decision. These variables not only depict the situation but also capture the thought process of the people in those situations.

6.3.1 State Variables

Any decision taken in our everyday life takes the situation at hand into consideration. The state variables help in defining the situation of the problem as detailed as possible. With respect to the hurricane domain, it reflects the information about the hurricane and also the common factors and precautionary measures that follow it. These variables are more of public experiences and not individualistic. Also, these variables depend on their state in the previous time step during the transition.

1. *Hurricane Intensity*: As explained in chapter 1, a hurricane is divided into five categories based on wind speeds around the eye. In addition to the five categories, a hurricane could become weaker and turn into a tropical storm or a depression. This was represented in our network model as a category 0, indicating that the wind speeds are lower than a category 1. The prior distribution for this variable is the distribution of hurricanes that occurred in 2017 [17].

2. *Hurricane Path*: The possible path of a hurricane is predicted, by the National Hurricane Center (NHC), up to five days in advance and represented as a probability cone. As far as a resident in the potential path is concerned, the only factor that they will care about is if the hurricane is heading towards or moving away from them. Therefore, this variable takes one of these two values with a prior distribution $\langle .33, .66 \rangle$ obtained from the probable track of the center of hurricanes historically.

3. *Evacuation Orders*: As mentioned in chapter 1, the government issues a series of evacuation orders over multiple days, either mandatory or voluntary, depending on the severity of the hurricane. The prior for this variable assigns a probability of 0.42 to the true value of this Boolean variable. As this variable is influenced by the hurricane level, we obtain the probability from the prior over various hurricane levels. We seek to model the decision making of individuals under evacuation orders, and this variable is set to true.

4. *Rain*: As an immediate effect of the high winds, rain is naturally bound to occur. However, this is not directly related to the level of the hurricane. This random variable can assume a value from this set, $\{$Light, Medium, Heavy$\}$, and the prior over this set is $\langle .5, .4, .1 \rangle$.

5. *Preparation to Stay*: One of the common reasons to not evacuate is the precautionary measures taken by the residents. This could include stocking up supplies, boarding up windows and doors and safeguarding vehicles or other property. This random variable assumes a value {None, Somewhat prepared, Very prepared}, and the prior is $\langle .11, .35, .54 \rangle$ obtained from the responses in the survey.

6. *Decision of Neighbors*: In a case where the person is not experienced with hurricanes or if they are new to the area, they tend to look at the neighbors to assess the risk. Most of the times, the reason behind their decisions are not taken into account which might lead to incorrect decisions. This is a Boolean variable whose prior $\langle .4, .6 \rangle$ is obtained by averaging the percentage of other people in the same zip code as the responder who evacuated.

7. *Traffic*: The most common mode of transportation during an evacuation is by road due to the low cost. This subsequently crowds all the major interstate highways and eventually blocks the smaller roads too. Some people often stay put as they do not want to deal with this extreme traffic or sit in their car for extended hours.

In addition to these experiential variables, demographic variables such as age, health condition, prior trauma, number of dependents, availability of transportation, employment status, and the safety of the house also compose the state of the decision-making problem. Age takes one of {$\leq$20, 21-40, 41-60, $\geq$61} values and its prior is $\langle .05, .35, .28, .32 \rangle$ based on the distribution in the survey. Health condition takes one of {Good, Need Support} values and its prior is uniform. Prior trauma is a binary variable and is true with probability 0.35 based on the survey responses. The dependents variable includes younger and older dependents as well as any pets. It takes a true value with a probability of .26 obtained from our survey data. The availability of transportation is a binary variable which is true 90% of the time. Employment is taken as true with 60% probability based on the 2017 US population statistics. The safety of the house is a binary variable with a uniform distribution. Since the value of these variables do not change over course of the hurricane, they are time invariant.

6.3.2 Observation Variables

Apart from capturing the situation, it is also important to understand what each person is experiencing in their immediate surroundings. For example, there might be torrential rain in the area but a few houses might not flood due to a variety of reasons. Similarly, there could also be some false conceptions based on previous experiences that deviate from the actual situation. These differences play a significant role in the decision-making process of any individual. Unlike state variables, the observations do not depend on those from prior time steps. Moreover, observations occur only after the first action has been taken. Thus, there are no observations possible at time step 0. Variables prefixed with 'O' are the observation variables.

1. *Flooding and Power Failure*: The most commonly observed effects of a hurricane are flooding and power failures. While extreme rain causes the flood, failure of power could be due to high winds and tree falls.

2. *Times Asked to Evacuate*: An order of evacuation is often followed up by repeated announcements on television, radio, emergency alert systems, and door-to-door canvassing. The more one is asked to leave, the more dangerous the effect is going to be. This observation variable depends on whether an evacuation order is issued and takes one of $\{0, 1\text{-}2, 3\text{-}4, \geq 5\}$ values.

3. *Rain and Wind*: While the hurricane level and rain as state variables reflect the actual situations, there could be differences in each person's perception of their severity. For example, if someone has never seen torrential rain, then they might perceive a normal rain itself to be high.

4. *Traffic Perception*: Sometimes, even when the traffic is not significant, residents may perceive it to be high based on prior experiences. This false assumption eventually affects their decision.

5. *Neighbor Evacuations*: From the survey results, it was found that people often misjudge the decisions made by their neighbors. Thus, it was important to include the difference between actual decisions and perceived decisions.

6. *Tweets from Peers*: Social media allows families and friends to remotely keep track of the disaster situation and any evacuation orders. It allows them to quickly convey their advice and worries about the individual's decision so far. We collected several such tweets issued during the time period hurricanes Harvey and Irma were active. This variable is influenced by evacuation order and takes a Boolean value.

7. *Call from Work*: Although a lot of people would be employed, there is a high possibility that their employers do not require them to work during these tense situations.

6.3.3 Hidden Variables

In order to reduce the dimensionality of the problem at each step, we introduce some hidden variables that collectively reflect the values of its parents. These nodes help in combining related variables and avoiding combinatorial explosion due to multiple variables affecting a single node. For clarity, we prefix hidden variables with 'HN'.

1. *Weather Perception*: Wind and rain together indicate the current status of the weather. This hidden variable may assume a value in the set { Good, Bad } and its CPT is shown below.

2. *Safety Perception*: Flood, power failure and safety of the house can be combined to understand how safe the person/area is. The conditional probability table of this hidden variable is shown below.

O Rain	**O Wind**	**Good**	**Bad**
None	Normal	.95	.05
None	High	.7	.3
None	Extreme	.1	.9
Normal	Normal	.8	.2
Normal	High	.5	.5
Normal	Extreme	.05	.95
Heavy	Normal	.5	.5
Heavy	High	.1	.9
Heavy	Extreme	.01	.99

Table 6.1: CPT for hidden variable representing *weather threat perception*

House Safety	**O Power Failure**	**O Flooding**	**Safe**	**Unsafe**
Safe	Not Possible	Not Possible	.99	.01
Safe	Not Possible	Possible	.9	.1
Safe	Not Possible	Flooded	.7	.3
Safe	Possible	Not Possible	.8	.2
Safe	Possible	Possible	.7	.3
Safe	Possible	Flooded	.6	.4
Safe	Failed	Not Possible	.65	.35
Safe	Failed	Possible	.6	.4
Safe	Failed	Flooded	.5	.5
Unsafe	*	*	0	1

Table 6.2: CPT for the hidden variable representing *safety threat perception*.

3. *Mobility*: Demographic factors like age, health, and transportation influence the mobility of a person. Due to the nature of its parents, this node is also time insensitive. An extract from the full Dynamic Influence Diagram (DID) depicting this node and its several parents is shown in Figure 6.1.

A fourth variable that somewhat falls in this category is *assessed risk* posed by the approaching hurricane to the individual's own safety and her property. In contrast to the previous hidden variables, our survey specifically asked the participants to assess this risk that they faced on a 7-point Likert scale with 1 being the lowest risk.

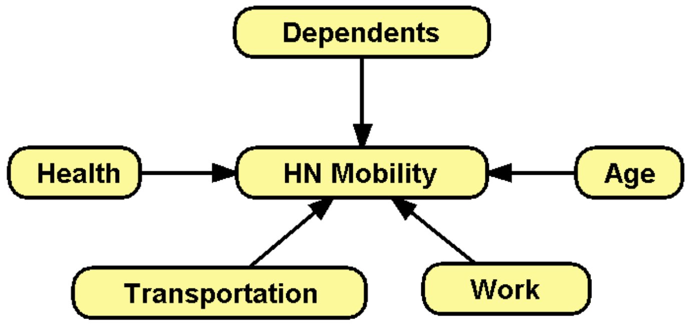

Figure 6.1: Factors influencing the hidden variable *mobility*.

We modeled the individual's assessed risk as shown in Figure 6.2. While hidden variables *weather threat perception* and *safety threat perception* obviously influence the risk, an order to evacuate raises the risk considerably. On the other hand, the individual's preparation to stay put and ride out the storm mitigates it.

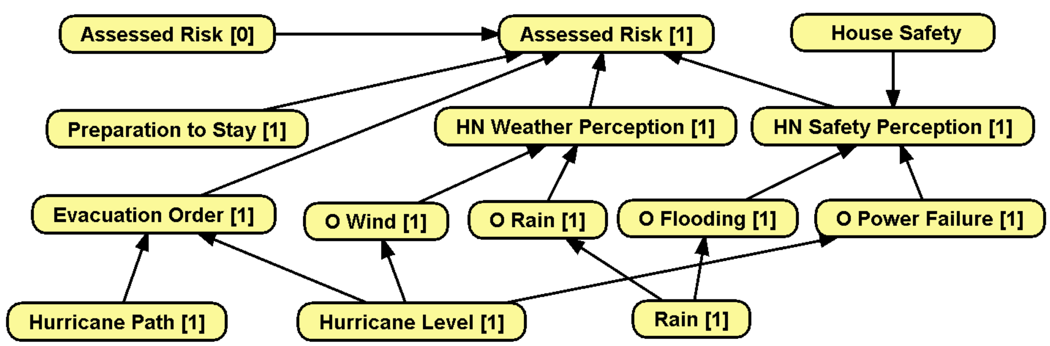

Figure 6.2: Factors influencing an individual's assessed risk.

6.3.4 Actions

Individuals in impending disaster areas face the choice of evacuating or not evacuating. They may also choose to make a decision about evacuating after collecting more definite information on the observed variables. We label this third choice as *Get info*. As such, an individual can choose between three actions at each time step. Figure 6.3 is an extract from the full DID that shows the action decision node. Variables that have outgoing edges to the decision node form the information set of action.

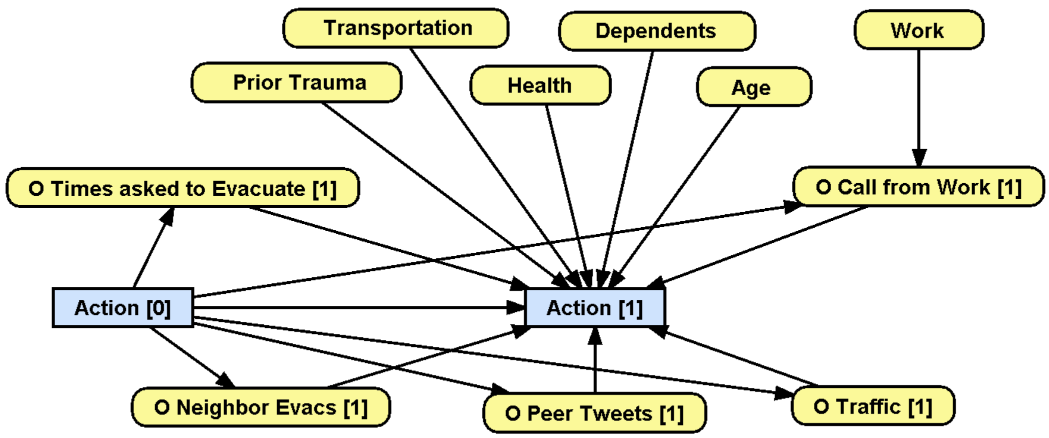

Figure 6.3: The Action decision node is shaded in blue.

6.3.5 Rewards

Based on the action taken and a few other variables (shown in figure 6.4), there is a specific reward assigned. The variables were chosen based on the results from the statistical analysis of the survey responses. Each of these variables are given a specific weight based on their importance in the decision making process. "Mobility" and "Assessed Risk" are given the weights of 4 and 3 respectively as they strongly correlated with evacuation decision. The other variables namely "Prior Trauma", "Traffic" and "Neighbors" were given a lower and equal weight of 1 as these were less significant in predicting the evacuation behavior.

The reward is computed based on a conditional equation. In the case where a person is not able to move for any reason, then there is a high cost for "Evacuate". There is also a standard cost for "Get Info" and a small cost for "Not Evacuate" depending on the risk involved. Otherwise, it is calculated as a weighted sum for each of the possible combinations of other variables. The weights mentioned above is multiplied by a predetermined reward for each value of the variable, given the action. The result for each variable is added up to get the cumulative reward/cost for a given action and a given combination of variable values.

The model attempts to maximize the overall reward at the end of all time steps while solving the network. The costs are represented as negative rewards.

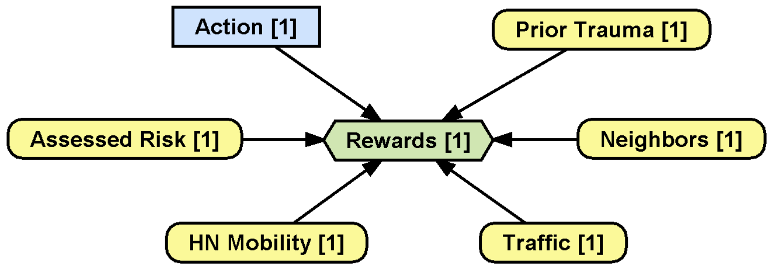

Figure 6.4: Factors influencing reward function.

6.3.6 Representation and Time Steps

The network is represented as a dynamic influence diagram that is split into four time steps. The transition between consecutive time steps occurs due to a major change in either the intensity or the direction of the hurricane in consideration. The number of time steps was decided using this criterion in the case of Hurricane Harvey and Hurricane Irma. The number in square parenthesis denotes the index of the time slice in which that node is located.

Being a parameterized description of the problem, not all variables influence every other variable. Thus, understanding the selective dependencies of each variable will help in reducing the dimensionality of the problem. Due to the lack of any prior efforts in the hurricane domain to discover the dependencies between the variables, all the influences were a result of our intuitive understanding of the domain. A two-timeslice representation of the network is illustrated in figure C.1.

6.4 Conditional Probability Tables

Each node is associated with a conditional probability table that indicates what possibility of occurrence of each of the values of a variable is, given the values of its parents. In this section, we explain how these tables were populated.

6.4.1 Transition Functions

Transition functions are those that show the change of states across the time steps. However, the change in the strength of hurricanes was the only factor that was modeled as a Markov process previously [18]. While the probability of transition from one category level to another was explained clearly, the specifics about the change in time steps was not elaborate.

Participants' responses to the question on assessed risk and their responses to its parents were utilized as part of expectation-maximization [19] to learn the CPT of *assessed risk*. During this process, a subset of the survey responses were given as evidence to the network in order to learn the CPT. With each case loaded as evidence, the probability distribution is updated based on the value for the *assessed risk* node and the number of times this combination of states of the parents was encountered earlier (experience). The CPT is initialized with an intuitive Gaussian distribution and the experience values are initialized with a uniform distribution.

Preparation to stay back is significant only when the previous decision is to get information. Otherwise, the distribution is uniform. Also, there is no possibility for the individual to become less prepared than before. A neighbor who has evacuated is generally not expected to return until the hurricane has passed. However, if they had not evacuated then the probability of her evacuating in the next tiem step is same as the prior.

If it has started raining, then it continues with the same intensity with a probability of .85 whereas there is a 40% chance for it to start raining and a 10% chance of it being heavy. There is a high chance of traffic increasing when the previous decision is to evacuate. Otherwise, the probability of traffic depends on whether an evacuation order is present or not.

6.4.2 Observation Functions

The observation functions are designed to reflect the uncertainty in the perception of the environment. Thus, all of these were populated based on the survey responses. For example, we show the CPT for the observation node *times asked to evacuate* in Table 6.3. Given an evacuation order, we utilized the distribution of answers to our survey question to arrive at the probabilities. Otherwise, an individual most likely did not hear it.

Action (t-1)	**Evacuation Order**	**0**	**1-2**	**3-4**	**$\geq$ 5**
Get info	Announced	.45	.42	.08	.05
Get info	Not Announced	.95	.04	.01	0
*	*	.25	.25	.25	.25

Table 6.3: CPT for observation variable representing *times asked to evacuate*

6.5 Sensitivity Analysis

A sensitivity analysis allows us to understand the impact of a change in the conditional probabilities of a hypothesis variable on a target random variable. As our model is strongly driven by data, the analysis serves as a tool to partly verify whether the variables and their CPTs are reflecting the influences intuitively. We deploy the sensitivity analysis to answer a series of questions, which clarify how the model works.

While sensitivity analysis can be performed in multiple ways [20], we rely on the method used by the Hugin Expert system. First, select a cell in the CPT of the hypothesis variable. For various hypothetical probabilities in this cell, we may obtain the corresponding inferred probability of the values of the target random variable. This sensitivity function is shown as a line graph – one line for each value of the target variable. A sensitivity value for each state of the target variable is then simply the derivative of the sensitive function (the slope of the line).

Is assessed risk influenced by the strength and path of the hurricane? Figure 6.5 shows the sensitivity analysis with *assessed risk* as the target variable. Observe that the risk is indeed influenced by the strength of the hurricane: as probability of the hurricane initially making landfall as category 5 increases, risk levels ≥ 5 exhibit a positive slope and for other levels the slope is negative. However, the path of the hurricane did not impact the risk. We

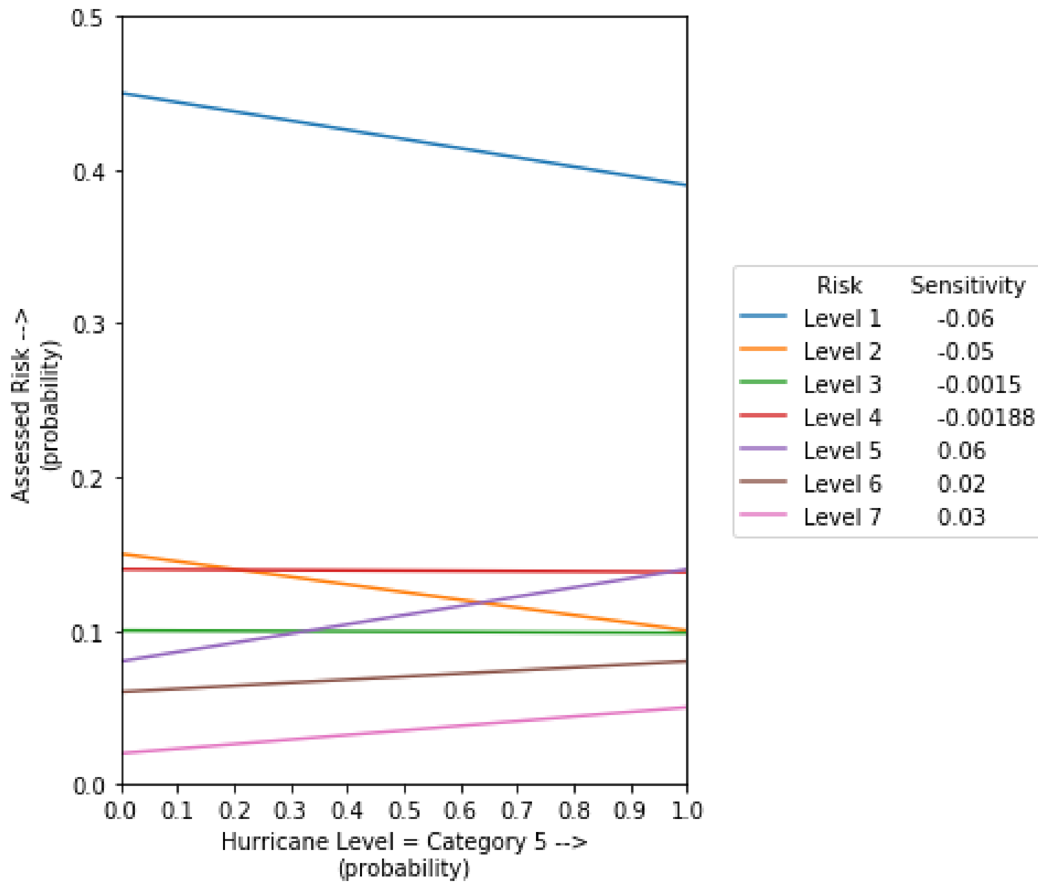

Figure 6.5: Sensitivity functions showing how the probabilities of various *assessed risk* levels changes with the probability of the hurricane initially being a category 5 at landfall.

believe this is because the network has only seen cases where the hurricane is headed toward the individual, and none where the hurricane veers away.

Does the safety of the house significantly impact risk assessment? Observe from Fig. 6.6 that assessed low risk levels of 1 and 2 exhibit positive slopes as the probability of house being safe increases. Indeed, risk level 2 probability increases steeply. Furthermore, the probabilities of higher risk levels reduce as we may expect.

Does the presence of transportation most affect the individual's mobility? What are the other factors influencing mobility? Figure 6.7 reveals that mobility is highly sensitive to changes in the probability of having transportation (sensitivity value is .58). This was closely

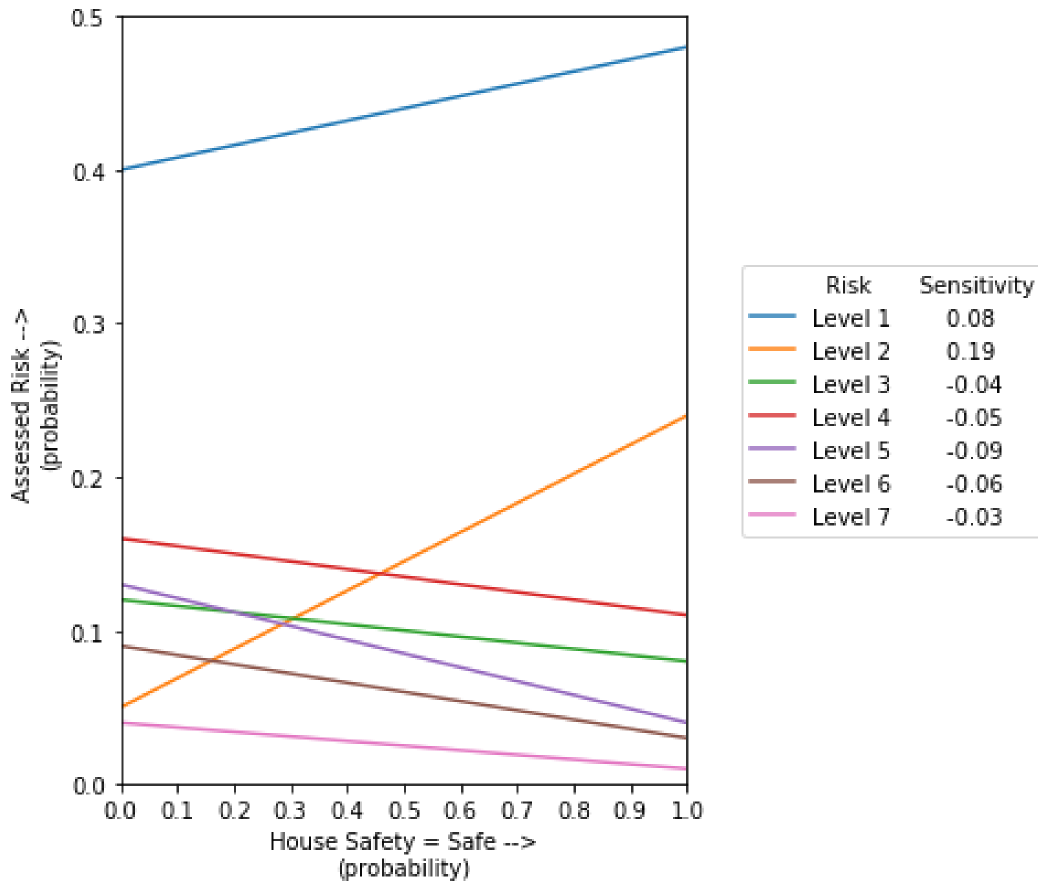

Figure 6.6: Sensitivity functions showing how the probabilities of assessed risk levels are affected by the changes in probability of individual's house being safe.

followed by the individual's employment situation with a sensitivity value of .32 – employment reduces the ability to move.

6.6 Experimental Setup

The network was converted into machine readable format to be processed using Hugin Expert [21] and saved as a *.net* file. This was then used by the Hugin Java API to process and test the prediction accuracy of the network.

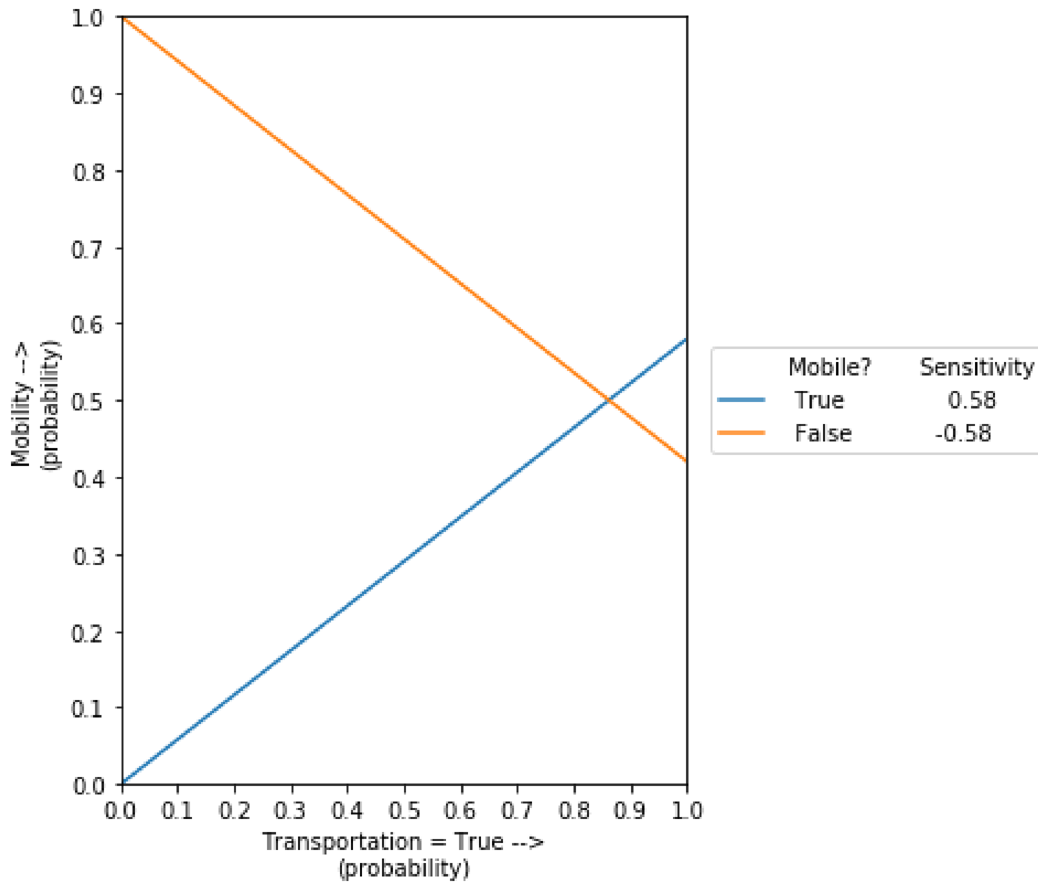

Figure 6.7: Sensitivity functions showing how the probabilities of mobility values change with the probability of the presence of transportation.

6.6.1 Training and Testing Datasets

The responses from the survey was randomly divided into training and testing sets in the ratio 2 : 1. A Python program was designed to access the database, get the required columns, convert to key-value pairs accepted by Hugin and save it as a file. At time step 0, the variables that are indifferent across time steps and the variables that define the environmental situation are entered. All the other known variables are entered in time step 1 after the decision at time step 0 is taken. A sample list of evidence values are shown below. Moreover, during the evaluation of the network, the decision at time step 0 is manually set to $GetInfo$ irrespective of what the decision is. This was to facilitate the confirmation of the evacuation orders and the correct understanding of the situation before taking a decision.

Dependents: "Absent"	Preparation_to_Stay_1: "Very"
Age: "a61_above"	O_Flooding_1: "Possible"
Prior_Trauma: "Absent"	O_Power_Failure_1: "Failed"
Hurricane_Level_0: "Cat_5"	O_Times_asked_to_evacuate_1: "t1_2"
Evacuation_Order_0: "Announced"	O_Neighbor_Evacuations_1: "Not_Evacuated"

In order to learn the CPT for the "$AssessedRisk''$ node at time step 1, a separate file with the values from all the training cases was generated using the same Python program. Along with the variables listed above, this file also had the value for "$AssessedRisk1''$ node. This file was parsed by the Hugin GUI during the EM-Learning process. Although all evidences were in the same file, the algorithm made sure any previous evidence was removed before loading a new set of evidence.

6.6.2 Evaluation Phase

In order to evaluate the performance of the network, each evidence case was propagated individually using a Java program implementing the Hugin API. At each time step, the decision made was captured. This decision was explicitly set to the action node and the evidence was re-propagated before capturing the decision in the next time step. The same process was repeated until all the decisions have been recorded. These decisions are processed by another Python program to arrive at the accuracy of predictions. If required, the inferred value of individual variables were also analyzed.

6.6.3 Cross Validation

To verify that the network has not been overfit for the initial split of training and testing data, a three-fold cross validation was implemented. While care was taken to not have any overlap between the test sets, it was found that seven of the survey responders had not mentioned the risk that they assessed. Hence, these cases were always placed in the test set and were replaced by random cases in the training set.

6.7 Result Analysis

The prior distributions and the rewards were set in such a way that without any observations, the default action is to either get more information or to evacuate. Unless any of the observations explicitly warrant staying back, the decision should be to evacuate. The decision from the network at time step 0 was used to validate the same. As mentioned earlier, this decision was overridden by $GetInfo$ before moving to subsequent time steps. Although the decision at every time step was captured, it was only processed if the decision at the previous time step was $GetInfo$. Once the network predicts *Evacuate* or *NotEvacuate*, the subsequent decisions become irrelevant.

A standard confusion matrix depicts the total counts of true and false calculations. The evacuation decision captured through the survey had only two options (Evacuate and Not Evacuate). Whereas, our model has an additional option to get more information and wait for taking the decision in the next time step. In case there is no decision taken at the end of time step 3, then it is counted as $GetInfo$. If this situation occurs, it means that the model did not have enough time steps to make an evacuation decision. Another two-dimensional matrix was used to depict the decision taken at each time step. Here, the decision at time step 0 shows the actual decision from the model. This was overridden with $GetInfo$. In subsequent time steps, the counts will match the number of decisions to $GetInfo$ in the previous time step.

Tables 6.4, 6.5, and 6.6 show the results for each of the three cross validation runs. We show both the confusion matrix and the predicted decisions at various time steps for the test fold in each run. Recall that the decision at time step 0 is set as *Get info* and we do not show it. Once the network recommends evacuate or not, subsequent decisions become irrelevant. The average accuracy across all three runs is 71.77% with a standard deviation of 1.89%

Table 6.4: Results for Cross Validation 0

(a) Training Time Matrix

t	0	1	2	3
GI	44	151	114	1
NE		166	33	79
EV	504	231	4	34

(b) Testing Time Matrix

t	0	1	2	3
GI	10	12	1	
NE		116	11	1
EV	264	146		

(c) Training Confusion Matrix

	NE	EV
GI	1	
NE	216	62
EV	112	157

Accuracy: 68.07%

(d) Testing Confusion Matrix

	NE	EV
GI		
NE	105	23
EV	58	88

Accuracy: 70.44%

and a high of 74.55%. Our model predicts a decision to evacuate or not at time step 1 for a majority of the participants. For example, in run 1 it rendered a decision that the participant will evacuate or not in time step 1 for 95% of the respondents. Indeed, coastal residents in Georgia began evacuating immediately on receiving the evacuation order. The DID takes about 45 seconds on an Intel Core i7, 64GB RAM, Ubuntu PC to render a decision.

Among the mispredictions, our model incorrectly predicts evacuate for more respondents than an incorrect prediction of not evacuate. A deeper analysis of the data reveals that the false predictions in the confusion matrices often reflect behavior that is hard to understand. We found that 62 respondents had decided to not evacuate despite assessing the risk level to be very high – at 6 or 7 – and they were mobile. Additionally, more than 100 respondents who did not evacuate falsely believed that they were not in an evacuation zone although an evacuation order was issued for their place of residence. This induced them to not evacuate. For these reasons, the number of respondents for whom the model mispredicted evacuate is higher.

Table 6.5: Results for Cross Validation 1

(a) Training Time Matrix

t	0	1	2	3
GI	35	129	105	
NE		184	22	75
EV	513	235	2	30

(b) Testing Time Matrix

t	0	1	2	3
GI	26	21		
NE		130	21	
EV	248	123		

(c) Training Confusion Matrix

	NE	EV
GI		
NE	204	77
EV	120	147

Accuracy: 64.05%

(d) Testing Confusion Matrix

	NE	EV
GI		
NE	124	27
EV	43	80

Accuracy: 74.45%

6.8 Limitations

The proposed model had been constrained both in terms of availability of data as well as computational resources. The model is strongly based on the survey conducted after the occurrence of a hurricane which limits our knowledge about the timeline of events. For example, one participant might have evacuated immediately after an evacuation order was announced but another participant might have waited until the situation got worse and then evacuated. These two cases are not differentiated by our survey, nor are they accurately represented in the model.

Human decision making is known to suffer from cognitive biases and these may affect evacuation decision making as well. Indeed, there is preliminary research [10, 8] on the evacuation behavior of residents in areas frequently affected by hurricanes with a focus on the psychological aspects of the decision-making process (but lacking a computational model). Our administered survey also tested for some cognitive biases in individuals. In particular, it addressed three well-established and potentially relevant biases generally evident across

Table 6.6: Results for Cross Validation 2

(a) Training Time Matrix

t	0	1	2	3
GI	46	148	116	
NE		169	27	90
EV	502	231	5	26

(b) Testing Time Matrix

t	0	1	2	3
GI	19	21	3	
NE		105	18	3
EV	255	148		

(c) Training Confusion Matrix

	NE	EV
GI		
NE	210	76
EV	117	145

Accuracy: 64.78%

(d) Testing Confusion Matrix

	NE	EV
GI		
NE	105	21
EV	60	88

Accuracy: 70.44%

populations: gambler's fallacy associated with overconfidence in predictive ability [22]; the illusion of control which occurs when people confuse chance with skill and thus behave as though their actions control the outcome of a random event [23], and the confirmation bias [24]. An ongoing analyses will reveal whether any of these biases were significantly evident in the surveyed population. However, we caution that mathematical models of these biases do not exist, which makes their inclusion in our decision-making model not trivial.

Chapter 7

Conclusion

In this chapter, we summarize the overall research and discuss the basic methodologies along with some results. Throughout the experimentation phase of the project, we encountered a variety of ideas that could have been implemented at different sections. Some of these are detailed in the later part of this chapter.

7.1 Discussion

This project is an attempt to model the variables influencing the decision to evacuate or not before the occurrence of a hurricane. This domain specially increases the uncertainty factor as it involves both a natural phenomenon as well as the human thought process.

The responses of a directed survey received from the residents of Georgia, Texas and Florida who were asked to evacuate during Hurricane Harvey and Hurricane Irma formed an important part of the research. Data from other sources such as social media and news articles were used to support the results and findings.

During the course of the research, we had to modify the K-Means Clustering algorithm to incorporate the context of texts rather than just their vector representations. This is implemented in processing the tweets posted over the duration of Hurricane Harvey and Hurricane Irma. The results reflected the findings from the survey and helped in confirming the significant variables identified.

We built a POMDP model with four time steps, represented as a dynamic influence diagram, to depict the factors that play an important role in the decision-making of an individual residing in an area that is expected to be affected. The interrelations between the variables and the conditional probability tables helped in depicting the situation and the differential importance of the factors in the decision-making process. The data from the survey responses were used to enter the evidences for the variables in the model and arrive at the decision. The model resulted in an average testing accuracy of 71.77% over a three-fold cross validation experiment.

7.2 Future Implementations

The proposed POMDP model can be extended into an Interactive Partially Observable Markov Decision Process (I-POMDP) [25] by treating the hurricane itself as an agent. In such a situation, the actions of the hurricane could be a combination of intensity (Category 1 through 5) and direction (heading towards and moving away). The other agent would be the inhabitant of the area who must take the decision to evacuate or not. This makes the problem more generalized and, at the same time, more complex. Such a model could be implemented to any population irrespective of geographic location.

The survey included questions that were intended to capture the cognitive biases that might have played a role in the decision-making of the responders. However, these factors are yet to be modelled computationally. Thus, these biases could not be represented in the model explicitly. Implementing these biases would help in understanding the human behaviour better as well as in predicting their actions more accurately.

Bibliography

[1] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall Press, Upper Saddle River, NJ, USA, 3rd edition, 2009.

[2] Jon Doyle. Rational decision making. *MIT encyclopedia of the cognitive sciences*, pages 701–703, 1999.

[3] Timothy Schott, Chris Landsea, Gene Hafele, Jeffrey Lorens, Arthur Taylor, Harvey Thurm, Bill Ward, Mark Willis, and Walt Zaleski. The saffir-simpson hurricane wind scale. *NOAA/National Weather Service [Internet]*, pages 1–4, 2012.

[4] Rakesh Agrawal and Ramakrishnan Srikant. Fast algorithms for mining association rules. In *International Conference on Very Large Data Bases*, pages 487 – 499, 1994.

[5] Martin L Puterman. Chapter 8 markov decision processes. In *Stochastic Models*, volume 2 of *Handbooks in Operations Research and Management Science*, pages 331–434. Elsevier, 1990.

[6] Shakara Brown, Hilary Parton, Cynthia Driver, and Christina Norman. Evacuation During Hurricane Sandy: Data from a Rapid Community Assessment. In *PLOS Currents: Disasters*, pages 45–50, January 2016.

[7] CH Gladwin, Hugh Gladwin, and Walter Gillis Peacock. Modeling hurricane evacuation decisions with ethnographic methods. *International journal of mass emergencies and disasters*, 19(2):117–143, 2001.

[8] Nicole Dash and Hugh Gladwin. Evacuation decision making and behavioral responses: Individual and household. *Natural Hazards Review*, 8(3):69–77, 2007.

[9] Dennis S. Mileti and E.M. Beck. Communication in crisis: Explaining evacuation symbolically. *Communication Research*, 2(1):24–49, 1975.

[10] John C. Whitehead, Bob Edwards, Marieke Van Willigen, John R. Maiolo, Kenneth Wilson, and Kevin T. Smith. Heading for higher ground: factors affecting real and hypothetical hurricane evacuation behavior. *Global Environmental Change Part B: Environmental Hazards*, 2(4):133–142, 2000.

[11] Samiul Hasan, Satish Ukkusuri, Hugh Gladwin, and Pamela Murray-Tuite. Behavioral model to understand household-level hurricane evacuation decision making. *Journal of Transportation Engineering*, 137(5):341–348, 2011.

[12] Amit Singhal et al. Modern information retrieval: A brief overview. *IEEE Data Eng. Bull.*, 24(4):35–43, 2001.

[13] Matthew Honnibal and Ines Montani. spacy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing. In *To appear*, 2018.

[14] Radim Řehůřek and Petr Sojka. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta, May 2010. ELRA.

[15] Scott Deerwester, Susan T. Dumais, George Furnas, Thomas Landauer, and Richard Harshman. Indexing by latent semantic analysis. In *Journal of the American Society for Information Science*, volume 41, pages 391–407, 09 1990.

[16] L.J.P.V.D. Maaten and GE Hinton. Visualizing high-dimensional data using t-sne. *Journal of Machine Learning Research*, 9:2579–2605, 01 2008.

[17] National Centers for Environmental Information. State of the Climate: Hurricanes and Tropical Storms for Annual 2017. `https://www.ncdc.noaa.gov/sotc/tropical-cyclones/201713/`, 2018. [retrieved March 7, 2019].

[18] Rebecca Wooten and Chris Tsokos. A markovian analysis of hurricane transitions. *Neural, Parallel and Scientific Computations*, 16:1–16, 03 2008.

[19] Steffen L. Lauritzen. The em algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis*, 19(2):191 – 201, 1995.

[20] D. M. Hamby. A review of techniques for parameter sensitivity analysis of environmental models. *Environmental monitoring and assessment*, 32:135–154, 09 1994.

[21] Hugin Expert. White Paper. `https://hugin.com/wp-content/uploads/2016/05/HUGIN-WHITE-PAPER-NEW-AND-REVISED_2013.pdf`, 2013. [Online; accessed 06-January-2019].

[22] Peter Ayton and Ilan Fischer. The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition*, 32(8):1369–1378, Dec 2004.

[23] Paul K. Presson and Victor A. Benassi. Illusion of control: A meta-analytic review. *Journal of Social Behavior & Personality*, 11(3):493 – 510, 1996.

[24] Raymond S. Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2):175–220, 1998.

[25] Piotr J. Gmytrasiewicz and Prashant Doshi. A framework for sequential planning in multi-agent settings. In *Journal of Artificial Intelligence Research*, volume 24, pages 49–79, USA, July 2005.

Appendix A

Survey Questions

A directed survey comprising of a total of 39 questions was circulated among selected residents in the areas of Florida, Georgia and Texas that received a voluntary or a mandatory evacuation order. The list of questions are given in table A.1.

Table A.1: Survey Questions

Question	Text	Response Options
1	(Introduction Text)	
2	What is your current state of residence?	(Florida, Georgia, Texas, Other State)
3	If Texas to Q2 : What is your county of residence?	(Target counties in Texas + Other)
4	If Florida to Q2 : What is your county of residence?	(Target counties in Florida + Other)
5	If Georgia to Q2 : What is your residential zip code?	(Target zip codes in Georgia + Other)
6	How many times did government officials ask you to evacuate?	(0, 1-2, 3-4, 5 or more, Unsure, Prefer not to answer)
7	Did you believe government officials were accurately reflecting the true level of risk?	(Yes Completely, Yes to some extent, No, Don't know, Prefer not to answer)
8	Prior to the hurricane impacting your area, did you evacuate from your home?	(Yes, No)
9	If Yes to Q8 : Did you evacuate because of the warnings you received from government officials?	(Yes Completely, Yes to some extent, No, Prefer not to answer)
10	If Yes to Q8 : When you evacuated, where did you go?	(Friend of Family's house in town, Hotel in town, Left town, Public shelter, Other(specify), Prefer not to answer)
11	If Yes to Q8 : What convinced you to leave your home to go to someplace safe?	Open-ended, text entered

Continued on next page

Question	Text	Response Options
12	If No to Q8 : Why did you not evacuate?	Open-ended, text entered
13	How many of the residents of your affected area would you estimate evacuated, based on what you know of their actions?	(Less than 10%; 10-45%; 45-55%; or about half; 55-90%; More than 90%; Don't know/Decline to answer)
14	Thinking about the recent hurricane in your area, how prepared were you with adequate food, water and other necessities to survive on your own for three days or more after the storm hit?	(Very prepared, Somewhat prepared, Somewhat unprepared, Very unprepared, Prefer not to answer)
15	Have you experienced a hurricane previously (prior to most recent one in your area)?	(Yes, No, Prefer not to answer)
16	If Yes to Q15 : Did your previous experience make you more or less likely to evacuate during the recent hurricane?	(Much more likely, Somewhat more likely, No effect, Somewhat less likely, Much less likely, Prefer not to answer)
17	Were you personally affected by the hurricane in your area? If yes, how were you affected? (Please mark all that apply)	(Lost power or other services for less than one week, Lost power or other services for more than one week, Home was damaged, Had to take extended time off from job, Car damaged, Other (Specify), Not affected, Prefer not to answer)
18	On a scale from 1 to 7, where 1 means no risk at all, and 7 means the greatest risk possible, as the hurricane approached, how much risk did you think the hurricane posed to your property or safety?	(1-7)
19	Prior to this year's hurricane, has something terrible ever happened to you, similar to the impact of a hurricane?	(Yes, No, Prefer not to answer)
20	Next, we will ask a few questions about your general approach to decision making and risk taking. Some of these might sound like questions from a statistics class, but please try to do your best.	

Continued on next page

Question	Text	Response Options
21	Imagine that 10% or 1 out of 10 hurricanes that develop near your state are known to seriously impact your area over the long run. After several years, nine hurricanes have developed near your state, and none of them seriously impacted your area. Which of the following statements about the next hurricane is most correct?	(It is less than 10% likely to seriously impact your area, It is 10% likely to seriously impact your area, It is more than 10% likely to seriously impact your area, Prefer not to answer)
22	Imagine you are playing a random game like flipping a coin. After losing many times in a row, you are more likely to win. Do you. . .	(Strongly agree, Agree, Disagree, Strongly disagree, Prefer not to answer)
23	True or false, supposedly random events can usually be predicted if you know the system?	(True, False, Prefer not to answer)
24	Now imagine that you have two lottery tickets. Ticket A was selected with your personal lucky numbers. Ticket B was selected by a computer at random. If you needed to keep one and give one away, which ticket would you be most likely to keep?	(Ticket A, Ticket B, No preference, Prefer not to answer)
25	Would you exchange Ticket A (with your lucky numbers) for two Ticket B's (which are computer generated)?	(Yes, No, Prefer not to answer)
26	True or false, I usually end up doing better at supposedly random tasks than most other people.	(True, False Prefer not to answer)
27	You are shown a set of four cards placed on a table, each of which has a number on one side and a colored patch on the other side. The visible faces of the cards show 3, 8, red and brown. Which card(s) must you turn over in order to test the proposition that if a card shows an even number on one face, then its opposite face is red? (Please check all that apply)	(3, 8, Red, Brown, Prefer not to answer)

Continued on next page

Question	Text	Response Options
28	Now we would like to ask you a few final questions for statistical purposes only.	
29	What was your zip code of residence at the time the hurricane impacted your area?	(Zip code (Textbox), Prefer not to answer)
30	Sex	(Male, Female, Prefer not to answer)
31	Please record your age:	text entered
32	Which race do you most identify with?	(Caucasian/White (Non-Hispanic), African American/Black, Hispanic/Latin American, Native American/Alaskan Native, Asian or Pacific Islander (including Hawaii and Philippines), Arabic, Mixed Race, Other, Prefer not to answer)
33	What is your relationship status?	(Single, Married, Long Term Relationship, Separated/Divorced, Widowed, Other, Prefer not to answer)
34	What is the highest level of education you have attained?	(Some high school, High school graduate, Some college, Associate's degree, Bachelor's degree, Master's degree, Doctorate Degree, Other, Prefer not to answer)
35	How many children (less than 18 years old) live in your home?	(0, 1-2, 3 or more, Prefer not to answer)
36	If 1-2/3 or more to Q35 : What is the age range your children fall within?	(0 to 5 years of age, 6 to 11 years of age, 12 to 17 years of age)
37	How many elder dependents live in your home?	(0, 1-2, 3 or more, Prefer not to answer)
38	What level is your home or apartment on?	(1st floor, 2nd floor, 3rd to 5th floor, 6th floor or higher, Prefer not to answer)
39	Did you rent or own your home at the time of the hurricane?	(Own, Rent, Other, Prefer not to answer)

Appendix B

Supplementary Results for Text Clustering

Additionally, in order to verify that all the relevant tweets were dropped, a few more experiments were performed with the minimum similarity score at 96%, i.e. all the tweets that have a score less than 0.96 were dropped after convergence. This resulted in multiple clusters with less than 10 tweets. This means that the number of clusters are very high for the number of data points. Since it was known, from the ground truth counts received through a manual clustering attempt, that the number of clusters could not be lower, it was confirmed that the threshold being set at 95% is the best option. The results from the additional experiments are shown below.

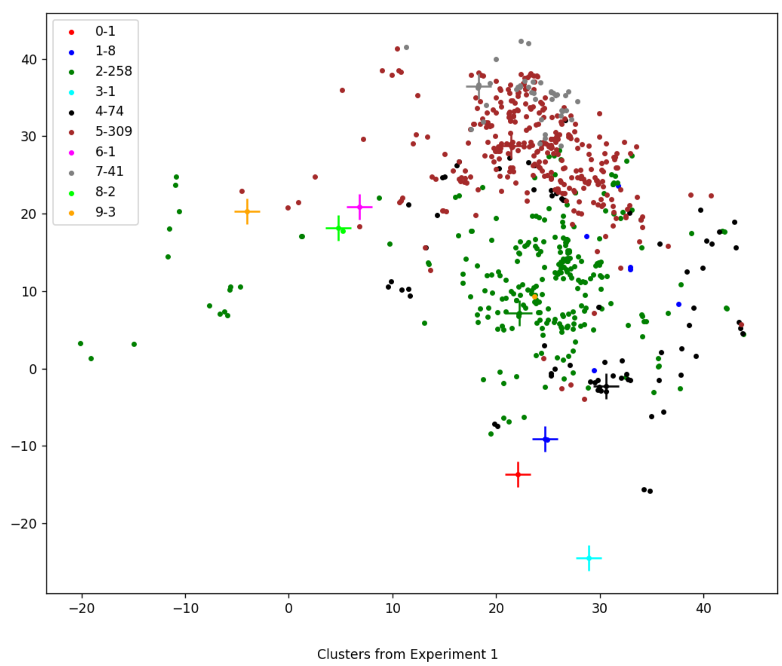

Figure B.1: Visualization of Clustering Experiment 1

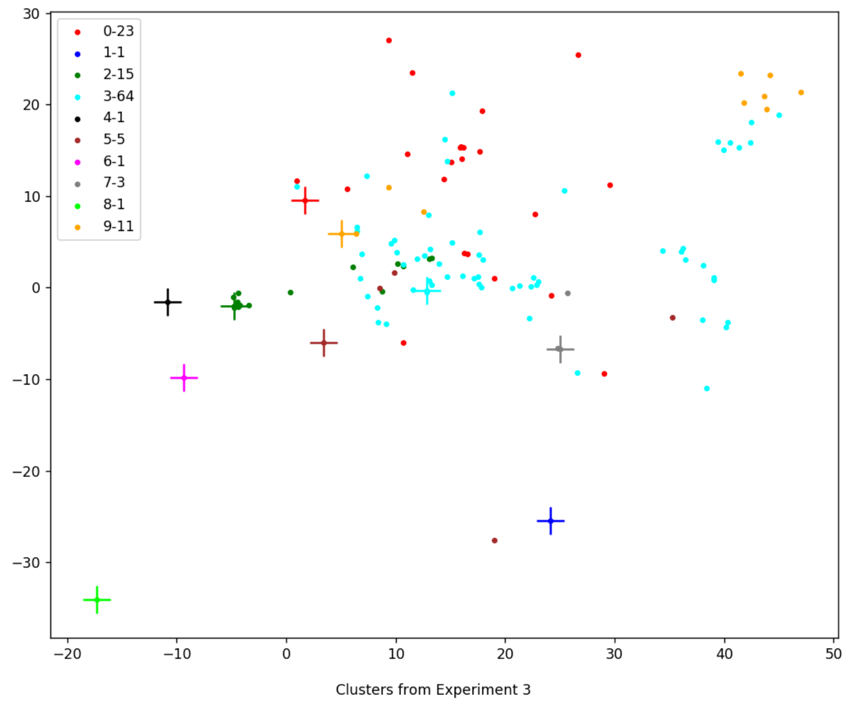

Figure B.2: Visualization of Clustering Experiment 3

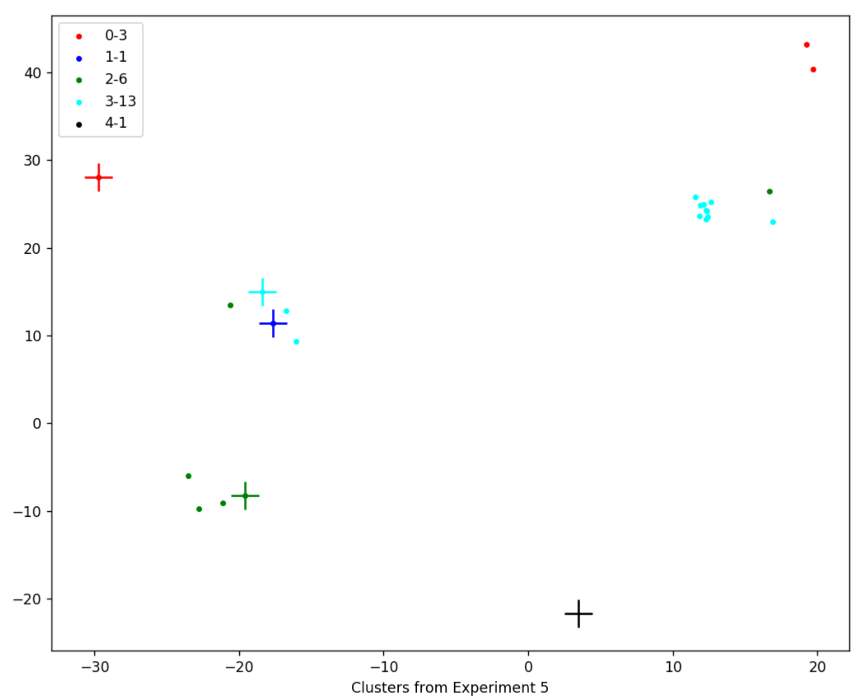

Figure B.3: Visualization of Clustering Experiment 5

Appendix C

Two time slice representation of DID

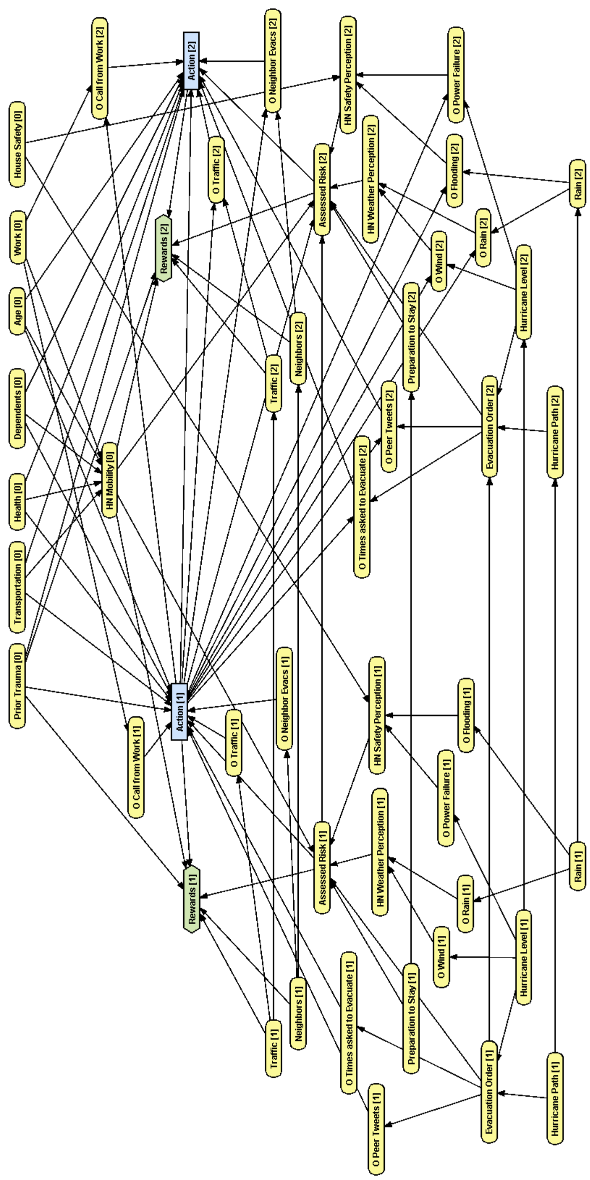

Figure C.1: A two-timeslice representation of the network