SafeR Mapping for Safe Refugee and IDP Relocation

Ishaan Shaurya Chamoli

Ridge Valley School, Gurugram, Haryana, India

Ishaan Shaurya Chamoli, email@ishaanchamoli.com

ABSTRACT

The study explores safe routing based on geographic danger scoring/mapping, specifically for use in conflict regions globally, producing internally displaces person's and refugees. This study takes inspiration from prior research on safe routes for cities, based on the extent of data available for such a use case. Estimating lower data availability, this research aims to introduce the creation of a new algorithm tailored specifically to make optimal use of the kind of data available.

Research has been conducted by analysing data in New Delhi, India, and New York City, USA, while aiming for a predictive estimation of danger rather than one based on the overall accumulation of crime.

The modern era faces multiple conflicts across the world, and many people are forcefully displaced. They have to relocate within borders as Internally ADisplaced Persons (IDPs) or escape their countries as Refugees. Some examples include the Syrian Civil War, Myanmar's Rohingya Crisis, and the recent Russia-Ukraine War. While the algorithm may not be easily accessible for most individuals in such regions, with further development, it may be of especially great benefit to aid camps and organisations working in the region such as the United Nations, Doctors Without Borders, Red Cross, etc. who can directly be provided with the application.

They may use the application to suggest the routes to those migrating and also for their own safety such as during the movement of personnel/resources, and even to decide the best locations for the establishment of camps.

INTRODUCTION

The modern era faces multiple conflicts across the world, and many people are forcefully displaced. They have to relocate within borders as Internally Displaced Persons (IDPs) or escape their countries as Refugees. According to the Office of the United Nations High Commissioner for Refugees (UNHCR), there exist over 89.3 million forcibly displaced people across the globe, over half of whom are under the age of 18.

Some conflicts resulting in forcible displacement include the Syrian Civil War, Myanmar's Rohingya Crisis, and the recent Russia-Ukraine War.

During this migration, people often have to trust unreliable sources to obtain information and aid. They face multiple dangers amidst the conflict, and a lack of medical assistance and food supply. Over 4000 asylum seekers lose their lives annualy, during this process. There is a need to guide them in this process and ensure contact with existing aid camps in the area.

Though largely limited, there has been some prior research on the detection of safe routes [1]. However, this research has been conducted primarily for cities, with the assumption of high levels of constantly updating data. There is great opportunity for such a mapping system to be used in conflict locations where even the slightest of differences in danger levels can be a matter of life and death. Such conflict areas will in most cases not have data as readily available or updated as cities, due to other governmental priorities or instabilities. Though this limits the accuracy of danger scoring for safe route prediction, an algorithm can still be created to make optimal use of the data that may be available. In this paper, this objective is pursued by making a prediction of the danger score based on past annual danger trends rather than an overall accumulation of all crimes in a set time period. The region is divided into smaller segments and given individual scores based on the weighting for crimes as well as how recently they occurred. The regional scores are used to find total route danger levels and pick the safest of all available routes. This serves as a mapping application for refugees that takes danger into consideration rather than just distance and estimated time.

This paper is created so that it can be used as an introduction to the use of safe mapping technology in conflict locations. Though a limited number of refugees with devices would be able to access the maps on their own, the project would be of especially great benefit to aid camps and organisations working in the region such as the United Nations, Doctors Without Borders, Red Cross, etc. who can directly be provided with the application.

They may use the application to suggest the routes to those migrating and also for their own safety such as during the movement of personnel/resources, and even to decide the best locations for the establishment of camps. There is also further scope for development of the algorithm that

takes aid camp locations, and international aid organization or NGO reports into account for more relevant routing.
(Will then add findings and novelty, once done with the results)

METHODOLOGY

Aim of the study

To create an algorithm that makes the best use of available data, to calculate danger scores for regions in a map and considers them to find the safest route.

Research Design

Building and explaining a new algorithm

Consent and Ethical Issues

All ethical considerations were followed for the current study. All data utilised was taken datasets readily available in the public domain, specifically for such research. All crime datapoints include only the type of crime and approximate location, without procuring any information of the victims of the crime.

Sample

A sample of 3361 crimes in various parts of Delhi, India [2] have been used for building the algorithm. The data has been cleaned such that each datapoint include the coordinates for the location where the crime occurred, the type of crime (divided into murder, rape, robbery, and theft), the quantity of each such crime, and the year it occurred.
Static, limited data has been used to replicate a comparatively low availability of data in conflict locations, so the algorithm can be readily repurposed.

Algorithm

a. **Region Classifier**
   Data is divided based on when each crime took place and its location.
   Based on the coordinates available, the dataset is divided into small regions of area 1 km^2 each. Each region is given an annual danger score based on the process that follows.

b. Danger Scoring

The annual crime data points that lie within each region are weighted by intensity of crime considered to reach the annual danger score $D_{[lat,\ long]}$

$D_{[lat,\ long]} = \Sigma\ [K_i * n(C_i)]$

K_i = Crime Weight (Based on prison terms for relative intensity of each crime)

C_i = Each type of crime lying within [lat, long]

Crime	Crime Weight
Murder	14
Rape	7
Robbery	5
Theft	0.5

*A few such examples are visible in the above table. These have been assumed with thorough analysis based on minimum sentences for each crime type according to the Indian Penal Code and can always be refined further with additional research.

c. Utilising Annual Scores

Once we have an annual score for each region, they are all considered in finding a final score that will be used to choose the safest route.

This is implemented by providing a weight to each year's score and taking their sum. For the dataset we have worked on, the following years and weights have been used.

Year	Year Weight
2017	1
2018	2
2019	3
2020	4
2021	5

More recent years are given a higher weight, to reach a more accurate current crime score.

An older crime may not be as indicative of the crime levels in an area as more recent crimes.

d. Route Detection

All possible routes on the google maps API from one location to another are considered and each route is given a total score R_i.

This is done by considering each region touched/crossed by the route. Each individual region's final score is multiplied with the meter distance travelled within it.

The value all such regions are summed up for a route score.

$R_i = \Sigma$ (region crossed * distance of route within region)

The Route with the lowest score is displayed.

RESULTS

Region	Bottom-Left Coordinates		Top-Right Coordinates		Final Danger Score
	Latitude	Longitude	Latitude	Longitude	
lat14_long0	28.592102042613995	76.91555	28.601145759943567	76.92343256063715	5685
lat18_long6	28.62827691193228	76.96284536382292	28.63732062926185	76.97072792446008	9835
lat23_long7	28.673495498580134	76.97072792446008	28.682539215909706	76.97861048509724	12255
lat8_long34	28.537839738636567	77.18355706166328	28.54688345596614	77.19143962230044	17185
lat41_long39 (e.g. high danger zone)	28.836282410512418	77.22296986484905	28.845326127841986	77.2308524254862	418435
…					

The table represents examples of the regions that delhi has been divided into, marked by the coordinates of their bottom-left and top-right corners. Each region has been given a score using the explained algorithm that adds weightage to the type and year of the crime committed. For example, the region labelled lat41_long39 has a much higher danger score and can statistcially be considered more dangerous than the other displayed regions.

DISCUSSION

The utilised data has thus been used to calculate smaller regional scores for a conflict location which can be utilised for safer route detection when data is limited.

The regional scoring can now provide scores to the available routes from point A to B, for final route detection.

CONCLUSION

A highlight on using limited data for safe route detection in conflict locations, where it's most impactful, also makes way for the future use of data extracted from online media and reports in the algorithm. While the accuracy of the recommendation is directly proportional to the amount of data available, this paper serves as a stepping stone to highlighting the need for greater research with a focus on danger levels in conflict locations, such that the comparatively limited data is used in the best manner possible.

LIMITATIONS

(i) In the current study, multiple square kilometre regions remain unscored due to a lack of available data, which is in progress of being addressed through communication and crime statistics request to the Delhi Government based on the Right to Information Act, 2005. (ii) At large, the mapping of danger zones for safe route recommendations reveals the possibility for bias or the marking of specific communities as "high danger zones". (iii) There is also scope for improvement in the coefficients or "weights" for the year in which each crime occured, to output more accurate danger zones, which take better account of how recent each crime is.

ACKNOWLEDGEMENT

I would like to acknowledge the continued insights of **Dr. Debayan Gupta**, Faculty at the Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, **Dr. Benjamin Larsen**, Artificial Intelligence and Machine Learning Lead at the World Economic Forum, the **Intel AI4Youth team** for their training and mentorship, and the entire **SafeR Mapping Team** namely - Pranav Kulkarni, Trinity Power, Aryan Reddy, Karan Chawla, Evan Osgood, and Gaurav Shah.

REFERENCES

[1] National Crime Records Bureau; https://ncrb.gov.in

[2] Soni, Shivangi & Gauri Shankar, Venkatesh & Sandeep, Chaurasia. (2019). Route-The Safe: A Robust Model for Safest Route Prediction Using Crime and Accidental Data. International Journal of Advanced Science and Technology, Vol 28. 1415-1428