ALGORITHM FOR SAFETY AND DISTANCE OPTIMIZATION ON ROUTES IN MEDELLÍN

Andrés Parra
Universidad Eafit
Colombia
afparram@eafit.edu.co

Sharif Velasquez
Universidad Eafit
Colombia
Svelasqu22@eafit.edu.co

Andrea Serna
Universidad Eafit
Colombia
asernac1@eafit.edu.co

Mauricio Toro
Universidad Eafit
Colombia
mtorobe@eafit.edu.co

ABSTRACT

The objective of this report/article is to analyze the growing problem of sexual street harassment in the city of Medellín. It is intended to develop a method, which, through various algorithms and data structures, allows to find the most optimal path between two points in the city of Medellin. This most optimal path will be calculated in reference to the distance required to travel it, and the insecurity involved.

We consider this problem to be extremely important, since a solution to it would involve the reduction of gender gaps, especially towards the female gender, and would allow a free development of identity, since it would reduce the social obstacles to achieve a boom in the self-determination of one's own tastes (there would be independence on the type of clothing to use, the way to act, and evidently, life would be safeguarded).

Key words

Shortest route, street sexual harassment, identification of safe routes, crime prevention

1. INTRODUCTION

Sexual harassment is one of modern's society biggest concerns, mainly, because it represents a threat to the development of self-identity, as well as it increases gender inequality. Its consequences can be experienced by everyone, nonetheless, women are more propense to be the objective of considerable attacks involving this problematic.

In Medellín, sexual harassment has increased in the last few years, and therefore, has become an issue worthy of deep analysis, not only for mayoralties or public authorities, but for scholars, researchers, and data scientists.

That is why we have decided to contribute to the reduction of consequences of sexual harassment by creating an algorithm to find the shortest and most convenient path for pedestrians, by finding a balance between distance and risk of sexual street harassment.

1.1. The problem

Through the analysis of algorithms and use of data structures, create and calculate the best three paths that reduce both distance and the risk of facing sexual harassment. Also, find a correlation between both parameters that is useful to measure the level of risk that will be used in the process of creating and implementing the algorithm.

1.2 Structure of the article

Next, in Section 2, we present work related to the problem. Then, in Section 3, we present the datasets and methods used in this research. In Section 4, we present the algorithm design. Then, in Section 5, we present the results. Finally, in Section 6, we discuss the results and propose some directions for future work.

2. RELATED WORK

Below, we explain four works related to finding ways to prevent street sexual harassment and crime in general.

2.1 Preventing Sexual Harassment Through a PathFinding Algorithm Using Nearby Search

The author describes how his fellow teams and he built a pathfinding algorithm in combination with heatmaps to identify ‘safe spots’ relative to a user’s coordinates and directions. Employing the Bresenham's line algorithm to search possible obstacles on the route. Using nearby search and Directions API that prioritize in first place the overall route safety-ness and later choosing the closer destination.

2.2 Incorporating a Safety Index into Pathfinding

The authors incorporate the shortest and safest path using the minimum and maximum deacceleration ratio of a vehicle as well as the probability of a collision on the road (the lower the value, the safer it is). Factors like road traffic density data, average highway speed and user's vehicle specifications are included in the safety index model.

2.3 Data Integration and Analysis System for Safe Route Planning

This is a study focused on finding the safest route for visitors and/or female inhabitants in a city in India, given that this is a country with a high rate of sexual street harassment. The authors propose parameters such as the potential danger that an area may have in the future, or various outliers (irregular data) as determinants in the calculation of an insecurity index.
After methodologically prioritizing the parameters to be taken into account, they perform a shortest path analysis using Dijkstra's algorithm, and make some modifications that include the implementation of a safety index (for each street). This is how the authors end up concluding that a function that will illustrate very well the cost of travel between nodes as a function of safety, would be: p = p - (2p/350) * Ys. Where p is the street distance cost.

2.4 Safety Routing for Motorized Tourists Based on Open Data and VGI

Design for vehicle routing application in LA, this algoritm uses information about street lights, police stations, highways and uniform crime reports recently to generate a cost surface. The further steps are connecting the cost surface with arcs of a road network and the derivation of obstacles based on crime hot spots. Using these two types of elements, calculates the relative least dangerous path.

3. MATERIALS AND METHODS

In this section, we explain how the data were collected and processed, and then different alternative path algorithms that reduce both the distance and the risk of sexual street harassment.

3.1 Data collection and processing

The map of Medellín was obtained from *Open Street Maps* (OSM)[1] and downloaded using the Python API[2] OSMnx. The map includes (1) the length of each segment, in meters; (2) the indication of whether the segment is one-way or not, and (3) the known binary representations of the geometries obtained from the metadata provided by OSM.

For this project, a linear combination (LC) was calculated that captures the maximum variance between (i) the fraction of households that feel insecure and (ii) the fraction of households with incomes below one minimum wage. These data were obtained from the 2017 Medellín quality of life survey. The CL was normalized, using the

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

[2] https://osmnx.readthedocs.io/

maximum and minimum, to obtain values between 0 and 1. The CL was obtained using principal components analysis. The risk of harassment is defined as one minus the normalized CL. Figure 1 presents the calculated risk of bullying. The map is available on GitHub[3] .

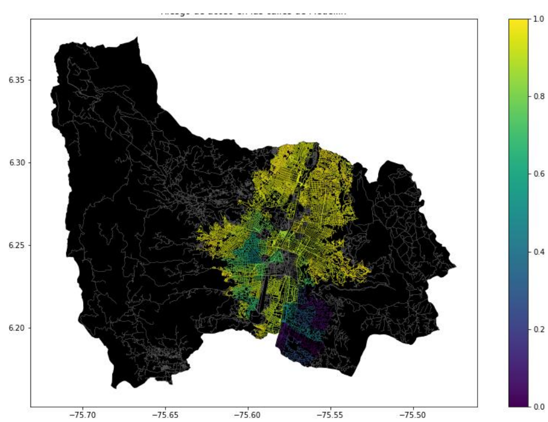
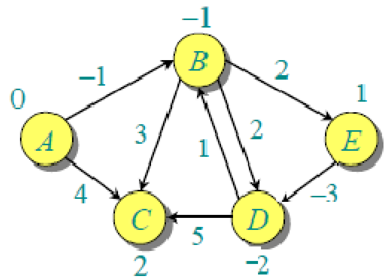


Figure 1. Risk of sexual harassment calculated as a linear combination of the fraction of households that feel unsafe and the fraction of households with income below one minimum wage, obtained from the 2017 Medellín Quality of Life Survey.

3.2 Algorithmic alternatives that reduce the risk of sexual street harassment and distance

In the following, we present different algorithms used for a path that reduces both street sexual harassment and distance.

3.2.1 Bellman-Ford



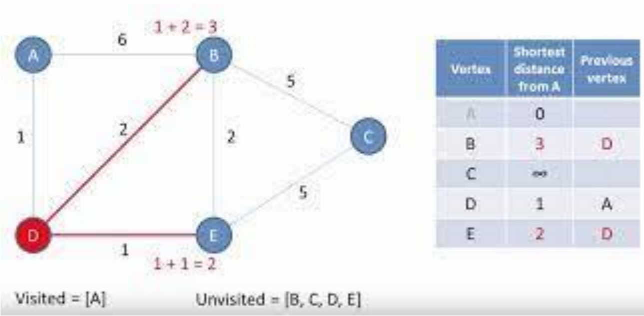
A	B	C	D	E
0	∞	∞	∞	∞
0	−1	∞	∞	∞
0	−1	4	∞	∞
0	−1	2	∞	∞
0	−1	2	∞	1
0	−1	2	1	1
0	−1	2	−2	1

This algorithm was created by computer scientists Richard Bellman, Samuel End and Lester Ford in 1956.

Its function is to compute the shortest path from a single source node to all other nodes in a graph. It works by setting to infinity the cost required to get from the source to each node in the network (given that the distances are unknown). Subsequently, several iterations are carried out in which the distance cost of each node is modified by one of its neighboring nodes. The process followed in each iteration is the comparison of whether the current node represents a shorter distance cost than the one temporarily stored by its neighbour node. The algorithm then establishes the minimum distance costs required to reach each node of the network from the source node.

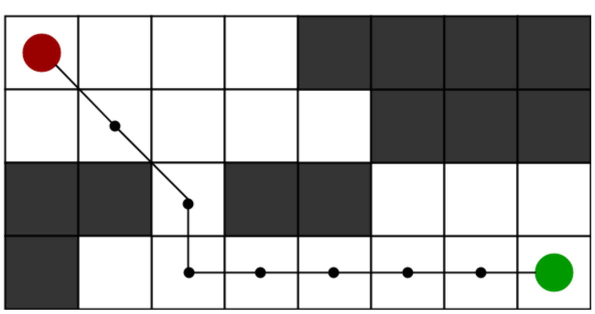
[3]https://github.com/mauriciotoro/ST0245Eafit/tree/master/proyecto/Datasets

3.2.2 Dijkstra



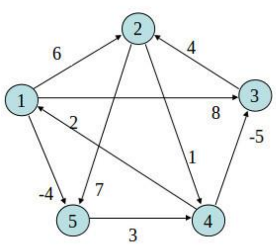
Like the Bellman-Ford algorithm, the Dijkstra's algorithm was created in 1956. Its function is also to find the shortest path to each node of a graph from a source node. It works by storing "visited" and "unvisited" nodes. Initially, the shortest path from the source node to each of its unvisited neighbouring nodes is obtained, and once this goal is achieved, the source node is added to the list of visited nodes, so that no unnecessary iterations are made to re-evaluate its already known distance cost. This approach is applied to the node with the list temporarily calculated distance, thus making this algorithm a "greedy" algorithm. After completion, the algorithm finishes by determining the minimum distances required to reach each node in the network.

3.2.3 A* Algorithm (A-Star)



A* focuses on finding the shortest path (as do all other algorithms). This is truly an addition to Dijkstra's algorithm, in which data external to the graph is used to determine the best path. This data is usually the Euclidean distance from a node to the destination node. Its logic flow is exactly the same as Dijkstra's, but the function that determines the value for setting the priority value in the priority queue changes, and is defined by the distance cost plus a heuristic function (determined by us depending on the situation and context). This algorithm is usually better than Dijkstra, especially when working with grids and/or maps. It is a Best-First-Search (BFS) algorithm and was created in 1968.

3.2.4 Floyd-Warshall algorithm



	1	2	3	4	5
1	0	6	8	∞	-4
2	∞	0	∞	1	7
3	∞	4	0	∞	∞
4	2	∞	-5	0	∞
5	∞	∞	∞	3	0

As is evident, this algorithm also focuses on finding the shortest path. However, it does not return the shortest path from a single source node to all nodes in the network, but instead stores the shortest path between every possible pair of nodes. The algorithm works by handling an array in which the shortest paths between two given nodes are stored. To find these values, three "for" cycles are used, as well as distance cost comparisons, following a methodology very similar to the previous algorithms.

REFERENCES

1. Ma, D. Preventing Sexual Harassment Through a Path Finding Algorithm Using Nearby Search. *Omdena*. Retrieved August 23, 2022, from Omdena: https://omdena.com/blog/path-finding-algorithm/

2. He, Z., Qin, X. Incorporating a Safety Index into Pathfinding. *SAGE Journals*, 2659 (1), 63-70.

3. Guptaa, A., Bhavye, K. A Data Integration and Analysis System for Safe Route Planning. *International Journal of Science and Research*, 9 (10), 1034-1042.

4. Keler, A., Mazimpaka, D. J. afety-aware routing for motorised tourists based on open data and VGI. *Journal of Location Based Services*, 10 (1), 64-77.

ATTACHMENTS

Video Presentación:

https://eafit-my.sharepoint.com/:v:/g/personal/svelasqu22_eafit_edu_co/EawXczIIptRNgFrI8UtwsiMB9uGSQnaeKxAouq6m6LgioA?e=ShqibX

Group Presentation (PowerPoint): https://github.com/SharifVelasquezAlzate/SafetyAndDistanceOptimization/blob/master/Group%20Presentation.pptx

Code for the obtention of Medellín's data (.csv): https://github.com/SharifVelasquezAlzate/SafetyAndDistanceOptimization/blob/master/main.py