FINDING A SAFE PATH TO A DESTINATION

Carlos Esteban Villa
Universidad Eafit
Colombia
Cevillaj@eafit.edu.co

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

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

ABSTRACT
In the following work I will try to solve an issue that has come to the light in recent years thanks to the constant grow of inequality between sexes and the constant search to make the balance even for all of them: sexual harassment. Given how much people have experienced such a horrible experience, the look for ways to evade it become necessary in our day-to-day life, which carries multiple issues such as: when will people feel really safe moving from one point to another in their own city? is it really enough with the work I will present? how can we tell if it's really the best way to approach such problem in modern society?

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

1. INTRODUCTION

As we have seen recently more and more in Medellin has become a constant issue of report sexual assault/harassment in the streets has increased more and more as the days go by, the project is born from the necessity of such things to stop happening or at the very least reduce the number of times it happens, you just have to go and google "sexual harassment Medellin" in the news section and you come across dozens of reports of instances where it happens, even in places where people is supposed to be "safe" like their college or even schools, where recently in the UdeA there were a number of reports of some teachers making inappropriate advances on students.

1.1 The problem

For this research the chosen topic is sexual harassment, which as mentioned in the abstract it's an issue that has come to be substantially more recurrent than what one could think, and cases being mayorly women (which happen to be half of humanity), there's a clear need to a solution to come forwards, let it be preventive (which will be the case of this research) of reactionary, that could be the case for the law to take the issue at hand

1.3 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 Route Safety

Although our issue to solve as of now is that of giving a route that would result in the user going to the route that has no probability of sexual harassment, there's also the issue of road safety, as described by Zhaoxiang He and Xiao Qin in their article [5], in which they mention how we should be capable of solving said issue. In their article they explain that, for rear-end clashes it's the sudden stop of the leading car, the in-depth explanation for the algorithm is in their article but the end-result of their investigation ended with the following formulas:

The expression of the safety hazard index I_{link} for any link is

$$I_{\text{link}} = \max\left\{\frac{\text{DRAC}_A}{\text{MADR}_A}, \frac{\text{DRAC}_B}{\text{MADR}_B}\right\} = \max\left\{\frac{\frac{v^2}{2(D_L - v\Delta t)}}{\text{MADR}_A}, \frac{\frac{v^2}{2(D_L - v\Delta t)}}{\text{MADR}_B}\right\} \quad (6)$$

where

$$D_L = \begin{cases} \frac{1}{d}\left(1 - e^{-dl} - dle^{-dl}\right) & \text{for sparse traffic conditions} \\ \frac{1}{d}\left(1 - \frac{4l^2d^2}{\pi}e^{-\frac{4l^2d^2}{\pi}} - e^{-\frac{4l^2d^2}{\pi}}\right) & \text{for dense traffic conditions} \end{cases}$$

2.2 Nearby Search

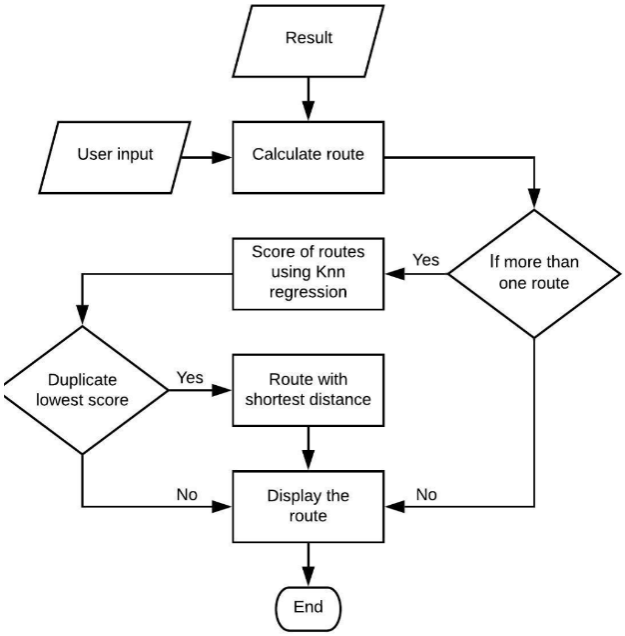
With previous work done on heatmaps where they predict places with high risk of sexual harassment incidents, using this one can create a grid an associate a risk to each point given how high is the risk. Now knowing this one can have a few different solutions, one could be to calculate risk score based on the destination which could be determined by the Euclidean distance and summing the risk associated with each grid-point. The formula would be as follows, for the Euclidean distance:

$$d(a,b) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + ... + (a_n - b_n)^2}$$

Figure 5: Euclidean distance between points a and b

2.3 Balance between routes

First one would generate the K paths and rank them from shortest to longest and then, given the second objective, being this one the safest, that being said, one could also give another second objective, like the one with less traffic. [7]



2.4 Cost Surface [8]

Consist on merging three data points already pre-determined previously such as Police Stations, Street Lights and highways and crime reports in the area to map (Los Angeles, for this investigation) which are combined in one point dataset. Then using the OPTICS algorithm, it takes a distance function to compute the relative distance between two points then a "weight" is given to each area related to the safety index and the crime reports in the last month are interpolated to the route as restrictions.

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 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] .

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

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

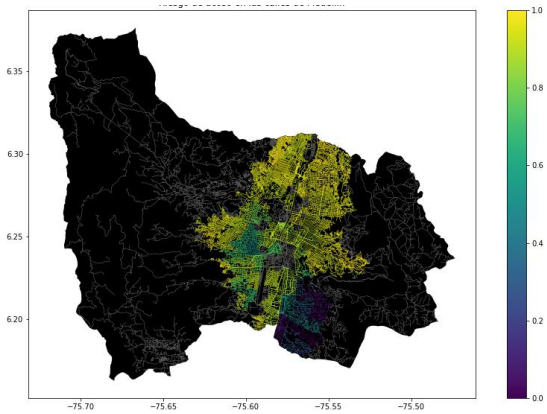


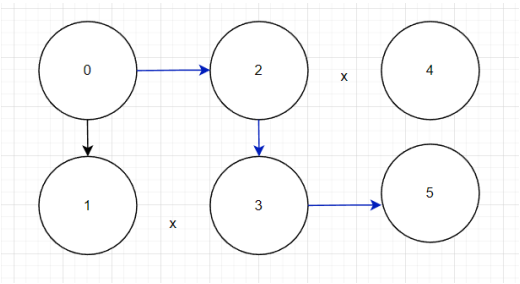
Figure 1. Risk of sexual harassment calculated as a linear combination of the fraction of households that feel unsafe and the fraction of households with income below one minimum wage, obtained from the 2017 Medellín Quality of Life Survey.

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.

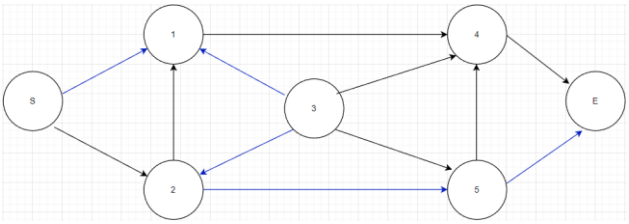
3.2.1 k-Shortest Path Method [1]

It computes all the path between the starting nod and the finish nod and returns all the paths starting with the shortest first.



3.2.2 The Pulse Algorithm [2]

Sends pulse through a network from the starting point all the way to the end and builds a partial path of all the nodes it has gone through, the pulse replicate itself in all the nodes it touches and all of those pulses that got to the end will send back the path they went through, so that the optimal path can be calculated/compared.



3.2.3 Backtracking [3]

Consist on going through only a valid path until there's a dead end, then backtrack to the last traversed node and continuing your path through that one

Start	1	1	1	0	0	0
	0	1	0	1	1	0
	1	1	1	1	0	0
	1	1	0	1	0	0
	0	0	0	1	0	0
	0	0	0	1	1	0
	0	0	0	0	1	0
	0	0	0	0	1	0
	0	0	0	0	1	0
	0	0	0	0	1	0
	0	0	0	0	1	1
	0	0	0	0	0	1
	0	0	0	0	0	1
	0	0	0	0	0	1 End

3.2.4 The Lee Algorithm [4]

It's an algorithm where you create multiple "queues", you start by queueing the starting point of your matrix and then add all the valid nodes next to your starting point, then repeat the same process with all the queued nodes until you reach your endpoint and the queue it's empty

Start	1	1	1	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	0	1	1	0
0	0	0	0	1	0
0	0	0	1	1	0
0	0	0	1	0	0
0	0	0	1	0	0
0	0	0	1	1	0
0	0	0	0	1	0
0	0	0	0	1	0
0	0	0	1	1	0
0	0	0	1	0	0
0	0	0	1	1	End

4. ALGORITHM DESIGN AND IMPLEMENTATION

In the following, we explain the data structures and algorithms used in this work. The implementations of the data structures and algorithms are available on Github[4] .

4.1 Data Structures

Explain the data structure that was used to represent the map of the city of Medellín. Make a figure that explains it. Do not use figures from the Internet. *(In this semester, examples of data structures are adjacency matrix, adjacency list, adjacency list using a dictionary).* The data structure is presented in Figure 2.

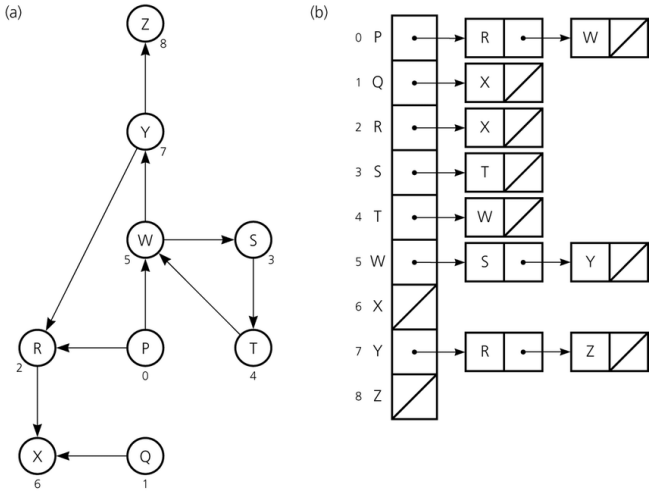


Figure 2: An example street map is presented in (a) and its representation as an adjacency list in (b). (*Please feel free to change this graph if you use a different data structure*).

4.2 Algorithms

In this paper, we propose an algorithm for a path that minimizes both the distance and the risk of street sexual harassment.

4.2.1 Algorithm for a pedestrian path that reduces both distance and risk of sexual street harassment

Explain the design of the algorithm for calculating a path that reduces both distance and risk of harassment and make your own graph. Do not use graphs from the Internet, make your own. *(In this semester, the algorithm could be DFS, BFS, Dijkstra, A*, Bellman, Floyd among others).* The algorithm is exemplified in Figure 3.

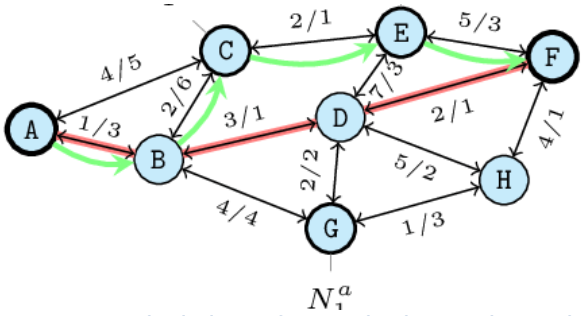


Figure 3: Calculation of a path that reduces both distance and risk of harassment (please feel free to change this figure if you use a different algorithm).

4.2.2 Calculation of two other paths to reduce both the distance and the risk of sexual street harassment

Explain the other two paths that reduce both distance and risk of street sexual harassment and make your own graph. Do not use graphs from the Internet, make your own. *(In this semester, the algorithm could be DFS, BFS, Dijkstra, A*, among others).*) The algorithm is exemplified in Figure 4.

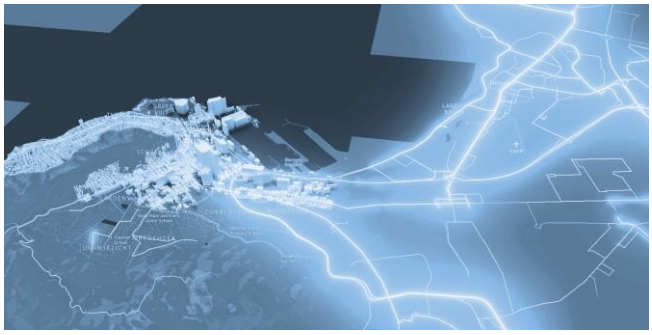


Figure 4: Map of the city of Medellín showing three pedestrian paths that reduce both the risk of sexual harassment and the distance in meters between the EAFIT University and the National University.

4.3 Algorithm complexity analysis

Explain, in your own words, the analysis, for the worst case, using the notation O. How did you calculate these complexities? Explain briefly.

Algorithm	Time complexity
Algorithm name	$O(V^2 * E^2)$
Name of the second algorithm (in case you	$O(E^3 * V * 2^V)$

[4] http://www.github.com/ ????????? /.../project/

have tried two)	

Table 1: Time complexity of the name of your algorithm, where V is.... E is... *(Please explain what V and E mean in this problem). No, do not use 'n'.*

Data Structure	Complexity of memory
Name of the data structure	$O(V*E*2^E)$
Name of the second data structure (in case you have tried two)	$O(2^{E*} 2^V)$

Table 2: Memory complexity of the data structure name used by your algorithm, where V is.... E is... *(Please explain what V and E mean in this problem). No, don't use 'n'. That is, don't use 'n'. Not 'n'.*

4.4 Algorithm design criteria

Explain why the algorithm was designed that way. Use objective criteria. Objective criteria are based on efficiency, which is measured in terms of time and memory. Examples of NON-objective criteria are: "I was sick", "it was the first data structure I found on the Internet", "I did it the last day before the deadline", "it's easier", etc. Remember: This is 40% of the project grade.

5. RESULTS

In this section, we present some quantitative results on the three pathways that reduce both the distance and the risk of sexual street harassment.

5.1 Results of the paths that reduces both distance and risk of sexual street harassment

Next, we present the results obtained from *three paths that reduce both distance and harassment,* in Table 3.

Origin	Destination	Distance	Risk
Eafit	Unal	??	??
Eafit	Unal	???	??
Eafit	Unal	??	??

Distance in meters and risk of sexual street harassment (between 0 and 1) to walk from EAFIT University to the National University.

5.2 Algorithm execution times

In Table 4, we explain the ratio of the average execution times of the queries presented in Table 3.

Calculate the execution time for the queries presented in Table 3.

Calculation of v	Average run times (s)
v = ??	100000.2 s
v = ??	800000.1 s
v = ??	8450000 s

Table 4: *Algorithm* name execution times *(Please write the name of the algorithm, e.g. DFS, BFS, A*)* for each of the three calculator paths between EAFIT and Universidad Nacional.

6. CONCLUSIONS

Explain the results obtained. Are the paths significantly different? How useful is this for the city? Are the runtimes reasonable to use this implementation in a real situation? Which path would you recommend for a mobile or web application?

6.1 Future work

Answer, what would you like to improve in the future? How would you like to improve your algorithm and its application? Will you continue this project working on optimization? Statistics? Web development? Machine learning? Virtual reality? How?

ACKNOWLEDGEMENTS

Identify the type of thank you you wish to write: to a person or to an institution. Keep the following guidelines in mind: 1. The professor's name is not mentioned because he or she is an author. 2. You should not mention the authors of articles that you have not contacted. 3. You should mention students, teachers of other courses who have helped you.

By way of example: This research has been supported/partially supported by [Name of Foundation, Donor].

We are grateful for help with [particular technique, methodology] to [First name Last name, position, name of institution] for comments that greatly improved this manuscript.

The authors thank Professor Juan Carlos Duque, Universidad EAFIT, for providing the data from the 2017 Medellín Quality of Life Survey, processed in a *Shapefile*.

REFERENCES

1. Santos, L., Coutinho-Rodrigues, J. and Current, J., 2006. *An improved solution algorithm for the constrained shortest path problem*. [online] ScienceDirect, Coimbra, POR.
2. Lozano, L. and Medaglia, A., 2012. *On an exact method for the constrained shortest path problem.* Elsevier, Bogota, COL.
3. Pencil Programmer. n.d. *Shortest Path in Maze using Backtracking – Pencil Programmer*. [online] Available at: <https://pencilprogrammer.com/algorithms/shortest-path-in-maze-using-backtracking/> [Accessed 13 August 2022].
4. Techie Delight. n.d. *Shortest path in a maze – Lee Algorithm | Techie Delight*. [online] Available at: <https://www.techiedelight.com/lee-algorithm-shortest-path-in-a-maze/> [Accessed 13 August 2022].
5. He, Z and Qin, X. 2017. Incorporating a Safety Index into Pathfinding, The Safety Section.
6. Daniel Ma. 2021. Preventing sexual harassment through a path finding algorithm using nearby search. (September 2021). Retrieved August 15, 2022 from https://omdena.com/blog/path-finding-algorithm/
7. Shivangi Soni, Chaurasia Sandeep, and Venkatesh Gauri Shankar. 2019. Route-the safe: A robust model for safest route prediction using crime ... (December 2019). Retrieved August 15, 2022 from https://www.researchgate.net/publication/338096313_Route-The_Safe_A_Robust_Model_for_Safest_Route_Prediction_Using_Crime_and_Accidental_Data
8. Andreas Keler and Jean Demascene Mazimpaka. 2016. Safety-aware routing for motorised tourists based on open data and VGI. (April 2016). Retrieved August 16, 2022 from https://www.researchgate.net/publication/301566519_Safety-aware_routing_for_motorised_tourists_based_on_open_data_and_VGI