**REDUCING STREET HARASSMENT THROUGH ALGORITHMS**

|  |  |  |  |
| --- | --- | --- | --- |
| Juan Felipe Restrepo Buitrago  Universidad Eafit  Colombia  jfrestrepb@eafit.edu.co | Sara Valentina Cortes Manrique  Universidad Eafit  Colombia  svcortesm@eafit.edu.co | Andrea Serna Universidad Eafit Colombia asernac1@eafit.edu.co | Mauricio Toro  Universidad Eafit  Colombia  mtorobe@eafit.edu.co |

**Text in black =** Contribution of Andrea and Mauricio

**Text in green** = To complete the first delivery

# **ABSTRACT**

To write an abstract, you must answer the following questions in a single paragraph: What is the problem? Why is the problem important? What are the related problems? What is the algorithm you have proposed to solve the problem? What quantitative results have you obtained? What are the conclusions of this work? The abstract should be **at most 200 words**. (*In this semester, you should summarize here the execution times, and the results obtained with the three paths*).

## **Key words**

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

# **1. INTRODUCTION**

Explain the motivation, in the real world, that leads to the problem. Include some background on this problem. *(In this semester, the motivation is why and why we need to calculate a path that reduces both the distance and the risk of sexual street harassment.)*

# **1.1. The problem**

In a few words, explain the problem, the impact this problem has on society, and why it is useful to solve it. *(In this semester, the problem is to calculate three different paths that reduce both the distance and the risk of street sexual harassment).*

**1.2 Solution**

Explain, briefly, your solution to the problem *(In this semester, the solution is a pedestrian algorithm to reduce both the distance and the risk of harassment. Which algorithms did you choose? Why?)*

**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.

## Explain four (4) articles related to the problem described in 1.1. You can find the related problems in scientific journals. Consider Google Scholar for your search. *(In this semester, the related work is the search for ways to prevent street sexual harassment and crime in general).*

## **2.1 Write a title for the first related problem**

You must mention the problem you solved, the algorithm you used, the results you obtained and the citation in the ACM format.

## **2.2 Write a title for the second related problem**

You must mention the problem you solved, the algorithm you used, the results you obtained and the citation in the ACM format.

## **2.3 Write a title for the third related problem**

You must mention the problem you solved, the algorithm you used, the results you obtained and the citation in the ACM format.

## **2.4 Write a title for the fourth related problem**

You must mention the problem you solved, the algorithm you used, the results you obtained and the citation in the ACM format.

## **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]](#footnote-1)  and downloaded using the Python API[[2]](#footnote-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]](#footnote-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. *(In this semester, examples of such algorithms are DFS, BFS, Dijkstra, A\*, Bellman, Floyd, among others).*

**3.2.1 First Algorithm Name**

Please explain the algorithm, its complexity and include your own vector figure designed at https://www.lucidchart.com/ or equivalent.

**3.2.2 Name of the second algorithm**

Please explain the algorithm, its complexity and include your own vector figure designed at https://www.lucidchart.com/ or equivalent.

**3.2.3 Third Algorithm Name**

Please explain the algorithm, its complexity and include your own vector figure designed at https://www.lucidchart.com/ or equivalent.

**3.2.4 Name of the fourth algorithm**

Please explain the algorithm, its complexity and include your own vector figure designed at https://www.lucidchart.com/ or equivalent.

## **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]](#footnote-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(V2 \*E2 ) |
| Name of the second algorithm (in case you have tried two) | O(E3 \*V\*2V ) |

**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\*2E  ) |
| Name of the second data structure (in case you have tried two) | O(2E\* 2V ) |

**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**

References are obtained using the ACM reference format. Read the ACM guidelines at [http://bit.ly/2pZnE5g.](http://bit.ly/2pZnE5g)

As an example, consider these two references:

Adobe Acrobat Reader 7, Make sure text in reference sections is Ragged Right, Not Justified. http://www.adobe.com/products/acrobat/.

2. Fischer, G. and Nakakoji, K. Amplifying designers' creativity with domain oriented design environments. in Dartnall, T. ed. Artificial Intelligence and Creativity: An Interdisciplinary Approach, Kluwer Academic Publishers, Dordrecht, 1994, 343-364.

Please remove the above references, they are just an example.

1. <https://www.openstreetmap.org/> [↑](#footnote-ref-1)
2. https://osmnx.readthedocs.io/ [↑](#footnote-ref-2)
3. https://github.com/mauriciotoro/ST0245Eafit/tree/master/proyecto/Datasets [↑](#footnote-ref-3)
4. http://www.github.com/ ????????? /.../project/ [↑](#footnote-ref-4)