

## Algorithm for Avoiding Street Harassment by Using Secure Routes

Mariana García  
Universidad Eafit  
Colombia  
mgarciaz2@eafit.edu.co

Isabella Márquez  
Universidad Eafit  
Colombia  
imarquezc@eafit.edu.co

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

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

María Camila Raigosa  
Universidad Eafit  
Colombia  
Mcraigosas@eafit.edu.co

### SUMMARY

Street harassment consists in unwanted comments, gestures, and actions forced on a stranger (mostly females) in a public place without their consent. This Gender-based problem is important since everyone deserves to be treated with dignity and respect and to feel safe in public spaces regardless of gender. Many external problems that affect the victims arise, such as making the person worried about its physical safety and creating an environment of fear and intimidation from the abuser, causing the person to be afraid that when she goes out on the street this could happen again which makes it harder to do things in everyday life.

### KEY WORDS

shortest path, street harassment, identification, secure routes, prevention, crime

### 1. INTRODUCTION

The risk of street harassment is extremely high, especially for women. Paths with less distance that reduce sexual harassment are part of the solution to eradicate this problem, we need them since every day someone is affected either in a small or large magnitude. We demand that this does not continue to grow, to make this reality visible as a problem that requires an urgent solution.

#### 1.1. Problem

The impact of this problem on society consists of several points. Violence can be related to street sexual harassment since both come from people with whom they have never met. Another point of street sexual harassment is the harm that the aggressors cause, which can be physical or sexual, although it also involves verbal harassment. This occurs mostly in women because of all the stigmatization they have in society in relation to men. For this reason, the

latter seek to highlight this superiority that they have been given over women by exercising this type of violence in public spaces.

#### 1.3 Structure of the Article

Following in section 2, we will present the related works with the problem presented on section 1.1. Subsequently, on section 3, we will show the gathering of data and methods used on the investigation. On section 4 exhibit the design of our algorithm. On section 5 there will be the display of the results. Finally on section 6, we will discuss said results and propose some directions to take on future works.

### 2. RELATED WORKS

Now, we will analyze and describe four related works with the search of pathways to prevent street harassment and delinquency in general.

#### 2.1 Un sistema de integración y análisis de datos para una planificación de rutas

La manera de resolver la problemática del acoso callejero y robos que se presentan en el texto es a través del diseño de una aplicación para rutas seguras que puede ser utilizada para determinar la ruta más segura y más rápida en lugar de solo la ruta más rápida.

El algoritmo que utilizaron fue averiguar las condiciones en las que los casos son propensos a producirse: agrupar los lugares y luego evitar los grupos de posibles personas peligrosas para saber cuán segura es la ubicación, teniendo en cuenta los índices de acoso de la zona con el valor de seguridad de la región, teniendo en cuenta las condiciones con las que las mujeres se sentirían inseguras para comparar esto en una base de datos.

Para determinar la seguridad de un lugar hay que tener en cuenta factores como la gente de la zona, hora del día, iluminación, situación monetaria, cámaras de la zona,

etc. Pero como se puede ver es imposible identificar la correspondencia de los factores. Consecuente a esto hay que definir una función de seguridad, que tomará entradas en base a los factores anteriormente mencionados para combinarlos y determinar la seguridad relativa de la región en relación con el lugar que se espera que sea (que las probabilidades de un caso sea mínima) y hacer más prácticas las planificaciones de las rutas cambiando las condiciones según el valor de seguridad. En base a esto se vio como resultado que si sólo se basan en seguridad la ruta sería demasiado larga, ya que los nodos estarían muy alejados entre sí. pero en algunas partes no hay buena iluminación o poca gente en donde los nodos no están tan lejos pero sin seguridad, así se daría el paso a la combinación de distancia y seguridad mediante la influencia máxima que puede tener la seguridad en el coste de la ruta. Así que a los usuarios se les proporcionará un conjunto de preguntas para elegir una ruta, y el algoritmo de toma de decisiones analizará las respuestas y decidirá qué ruta es la mejor para el usuario.

## **2.2 Previendo el Acoso Sexual a Través de un Camino Encontrando Algoritmos Usando Búsquedas Locales**

El problema que se resolvió dentro de este proyecto fue, principalmente, el tener la capacidad de predecir de manera efectiva, los lugares que contengan un contenido de riesgo alto en torno a problemas de acoso sexual a través de “puntos calientes y “zonas seguras”, para que así el usuario pueda identificar tanto las zonas por las cuales desplazarse y las zonas que debe de evitar.

El resultado obtenido de la investigación fue una AI implementada dentro de una cuadrícula “heatmap” (la cual fue mencionada anteriormente con sus “puntos calientes” y sus “zonas seguras”) con un sistema del 0-4, siendo 0 su zona más segura y 4 la más peligrosa, teniendo una distancia entre punto y punto de aproximadamente 1.74 Km<sup>2</sup>.

Gracias a la pregunta “¿Cómo determinar la seguridad de una ruta? la cual hacen los investigadores, dentro de su solución, se llegan a 3 conclusiones, “determinar en general la seguridad de cierta ruta sólo utilizando los valores asociados con el destino” “calcular los factores de riesgo promedio utilizando la cuadrícula en línea recta hasta el destino” y, finalmente “calcular los factores de riesgo promedio utilizando la cuadrícula paso por paso en la ruta hasta el destino”. Estos dos últimos implican un factor no considerado anteriormente, la línea de visión, por lo cual su solución para los obstáculos fue el algoritmo llamado “Línea de Bresenham” Que, en resumen, consiste en que al tener un punto de terminación dentro de una línea de

segmento de la cuadrícula, el siguiente punto dentro de esta misma determina si se puede llegar a ella por medio de la evaluación de dónde la línea de segmento cae más cercana al punto medio, si arriba o abajo. Y con lo anteriormente explicado, aquella “probabilidad de seguridad” se aplica dentro del “heatmap” explicado

## **2.3 Rutas seguras para turistas motorizados basadas en datos abiertos y VGI**

El problema presentado consiste en evitar las áreas peligrosas de una ciudad, en donde mediante enrutamientos estas puedan evitarse, teniendo en cuenta la seguridad de los turistas ya que algunos lugareños las eluden y solo las conocen a nivel local. Se resalta la importancia de la delincuencia en los medios de transporte, y que se quiere detectar sus trayectorias dentro de una solución de navegación que impida que estos entren en zonas peligrosas.

En consecuente a esto, incluyeron el uso de información libremente disponible en forma de datos abiertos y VGI, para proporcionar conocimientos locales sobre los lugares y ser una fuente de información para los turistas. También se incluyeron datos abiertos gubernamentales como la policía local.

Por otro lado, mediante la obtención de datos quisieron detectar los puntos de la delincuencia mediante técnicas de agrupación e información de libre acceso sobre la red de carreteras y en síntesis, combinar los datos de libre acceso, Sus datos de prueba fueron: el alumbrado, comisarías de policía, carreteras e informes de delito en el último mes. Con lo anterior se diseñó una aplicación de enrutamiento consciente de seguridad, teniendo en cuenta los costes, los obstáculos de enrutamiento, y las pruebas que se hicieron con los datos.

## **2.4 Modelo Ruta Segura**

Lo que busca resolver este estudio es los problemas de acoso sexual a mujeres identificados en los últimos años, entendiendo que incluso con la ayuda de herramientas ampliamente implementadas como Google Maps o Waze, estos no tienen la capacidad de identificar zonas peligrosas o de riesgo alto. Este estudio está específicamente enfocado en mujeres turistas y la ciudad de Nueva York, entendiendo que la implementación de una herramienta la cual provea seguridad, especialmente en una ciudad metropolitana tan concurrida por extranjeros, es un concepto que necesita ser implementado lo antes posible.

Su solución se centra principalmente en 3 aspectos: Primero se busca una ruta segura con ayuda de los datos

criminalísticos de cada ciudad mientras se usa la distancia más corta entre el punto de salida y de llegada; segundo, se divide Nueva York en regiones más pequeñas teniendo en cuenta su nivel de peligro; y tercero, calcular el valor de riesgo de las rutas basándose en los valores de los clústeres cercanos.

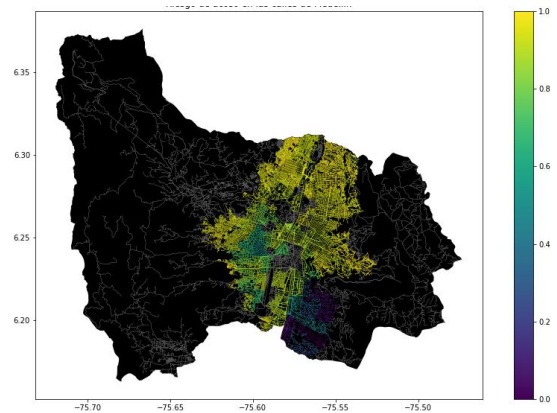
En este proyecto se utiliza el siguiente algoritmo; el dataset de crímenes y accidentes fue extraído directamente de las bases de la NYPD, dejando de lado las columnas que no fueran de utilidad para la investigación, dándoles el valor de null, así mismo renombrando las columnas para que fueran más entendibles. Las dataframes que crearon consistían en un número de clústeres, latitud y longitud de los clústeres y de “C” y “A” (dataset de crimen y dataset de accidentes respectivamente). La última parte consistió en que, ayudándose de Google Maps, encontraran los caminos indicados por medio de la creación de varios comandos dentro de la función de direcciones de la aplicación. Por lo cual el algoritmo se divide en 3; Procesador de datos, “K-mean” (la latitud y longitud) y finalmente el vecino más cercano o “K nearest neighbor”

### 3. MATERIALES Y MÉTODOS

En esta sección, explicamos cómo se recogieron y procesaron los datos y, después, diferentes alternativas de algoritmos de caminos que reducen tanto la distancia como el riesgo de acoso sexual callejero.

#### 3.1 Recogida y tratamiento de datos

El mapa de Medellín se obtuvo de *Open Street Maps* (OSM)<sup>1</sup> y se descargó utilizando la API<sup>2</sup> OSMnx de Python. El mapa incluye (1) la longitud de cada segmento, en metros; (2) la indicación de si el segmento es de un solo sentido o no, y (3) las representaciones binarias conocidas de las geometrías obtenidas de los metadatos proporcionados por OSM. Para este proyecto, se calculó una combinación lineal (CL) que captura la máxima varianza entre (i) la fracción de hogares que se sienten inseguros y (ii) la fracción de hogares con ingresos inferiores a un salario mínimo. Estos datos se obtuvieron de la encuesta de calidad de vida de Medellín, de 2017. La CL se normalizó, utilizando el máximo y el mínimo, para obtener valores entre 0 y 1. La CL se obtuvo mediante el análisis de componentes principales. El riesgo de acoso se define como uno menos la CL normalizada. La Figura 1 presenta el riesgo de acoso calculado. El mapa está disponible en GitHub<sup>3</sup>.



**Figura 1.** Riesgo de acoso sexual calculado como una combinación lineal de la fracción de hogares que se sienten inseguros y la fracción de hogares con ingresos inferiores a un salario mínimo, obtenidas de la Encuesta de Calidad de Vida de Medellín, de 2017.

#### 3.2 Alternatives of routes that reduce the risk of street harassment

Now, we will present different algorithms used to find a pathway that will reduce street harassment as well as the distances from point A to point B .

According to Dijkstra, an initial network F is constructed consisting only of the origin node. It must be ensured that all nodes that can be reached directly from F are added.

You must stop when the vertex to be included in F is the destination node. At each iteration, when you include a node to F you do so using the smallest path from the origin to the node. If another path exists, it must go through one of the other nodes that we can reach directly from F, which are greater or equal distance from the origin than node v, by construction. Since the weights are non-negative, any path that passes through any of these nodes to reach v cannot have a shorter distance than the one we were initially considering.

According to Floyd-Warshall, we must calculate the minimum distance between all pairs of vertices of a weighted network. It can be done in  $O(|V| \cdot (|E| + |V| \cdot \log(|V|)))$  by performing a Dijkstra for each node. But if the weights are negative, this method cannot be used, it must be done in  $O(|E| \cdot |V|^2)$  by performing n times the Bellman-Ford algorithm, one for each vertex. We perform n iterations; in each iteration we consider the optimal paths that have as intermediate nodes those with index less than i. So, after n

<sup>1</sup> <https://www.openstreetmap.org/>

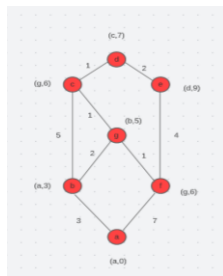
<sup>2</sup> <https://osmnx.readthedocs.io/>

<sup>3</sup> <https://github.com/mauriciotoro/ST0245Eafit/tree/master/proyecto/Datasets/>

iterations we consider the optimal paths that include any node, i.e., the optimal paths in general. There is another matrix (dist) of n rows and n columns such that the minimum distance from node i to j is stored in the box dist[i][j]. Initially the value of all distances is infinite, except for those where there is a direct arc from i to j, in which case the value of dist[i][j] will be the weight of the arc. Also, by convention dist[i][i] = 0. N iterations are performed, in each of which we will consider all pairs of nodes, so that at the k-th iteration if passing through vertex k on the path from i to j is optimal, we will update the value of the distance from i to j.

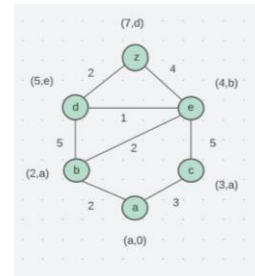
### 3.2.1 First Algorithm

In this algorithm, the circles with letters represent the nodes and lines the beard which represent the connections that exist (what they can run) and the distance is represented by the numbers that are above them. We will find the nearest route by starting from “a” which will be the initial node, by the parenthesis we will put the node from the left and on the right, the accumulated distance from the node of the past node to the next. It begins from “a” to 0 then it will move to either of the next beards having into account the distance. In this case it moved until b, so we are left with “a, 3” (we will need to keep adding). We need to do the same on the other side and it runs the one which is less heavy (b in this case) and it follows this pattern until the end.



### 3.2.2 Second Algorithm

In this algorithm happens the same as the last one, only that the values we found on the parenthesis are crossed out in terms of the last one, but this will not affect the result. It is very similar to the last one mentioned, and it would be the same pattern, starting from “a” due to it being the initial node.



### 3.2.3 Third Algorithm

### 3.2.4 Fourth Algorithm

## BIBLIOGRAPHY

Shankar, V. and Sandeep, C., 2019. *Route-The Safe: A Robust Model for Safest Route Prediction Using Crime and Accidental Data*. [online] ResearchGate. Available at: <[https://www.researchgate.net/publication/338096313\\_Route-The\\_Safe\\_A\\_Robust\\_Model\\_for\\_Safest\\_Route\\_Prediction\\_Using\\_Crime\\_and\\_Accidental\\_Data](https://www.researchgate.net/publication/338096313_Route-The_Safe_A_Robust_Model_for_Safest_Route_Prediction_Using_Crime_and_Accidental_Data)> [Accessed 21 August 2022].

Ma, D., 2022. *Preventing Sexual Harassment Through a Path Finding Algorithm Using Nearby Search*. [online] Omdena | Building AI Solutions for Real-World Problems. Available at: <<https://omdena.com/blog/path-finding-algorithm/>>

Keler, A. and Mazimpaka, J., 2016. *Safety-aware routing for motorised tourists based on open data and VGI*. [online] ResearchGate. Available at: <[https://www.researchgate.net/publication/301566519\\_Safety-aware\\_routing\\_for\\_motorised\\_tourists\\_based\\_on\\_open\\_data\\_and\\_VGI](https://www.researchgate.net/publication/301566519_Safety-aware_routing_for_motorised_tourists_based_on_open_data_and_VGI)>

Sarraf, R., 2016. *A Data Integration and Analysis System for Safe Route Planning*. [online] ResearchGate. Available at: <[https://www.researchgate.net/publication/327390893\\_A\\_Data\\_Integration\\_and\\_Analysis\\_System\\_for\\_Safe\\_Route\\_Planning](https://www.researchgate.net/publication/327390893_A_Data_Integration_and_Analysis_System_for_Safe_Route_Planning)>

Aprende.olimpiada-informatica.org. 2019. *Camino más corto entre nodos*. [online] Available at:

<<https://aprende.olimpiada-informatica.org/algoritmia-dijkstra-bellman-ford-floyd-warshall>>