TÍTULO (UNA BREVE DESCRIPCIÓN DEL PROYECTO, DE ENTRE 8 Y 12 PALABRAS)

Moisés Arrieta Universidad Eafit Colombia mdarrietah@eafit.edu.co smcardonav@eafit.edu.co

Sara Cardona Universidad Eafit Colombia

Juan López Universidad Eafit Colombia imlopezs@eafit.edu.co

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

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

Texto en rojo = Comentarios

Texto en negro = Contribución de Andrea y Mauricio

Texto en verde = Para completar la primera entrega

Texto azul = A completar para el 2º entregable

Texto en color violeta = A completar para el tercer entregable

RESUMEN

En ciertas zonas de la ciudad de Medellín se presenta una alta probabilidad de acoso sexual y delincuencia hacia las mujeres, debido a esto su resolución es muy importante para garantizar a las ciudadanas, que a la hora de desplazarse por la ciudad estas se sienta seguras y tranquilas cuando se dirige a su lugar destino.

Algunos problemas relacionados con este son la inseguridad vial, agresión física, las amenazas e intimidación y otros tantos que se analizaran en este informe.

¿Cuál es el algoritmo que has propuesto para resolver el problema? ¿Qué resultados cuantitativos has obtenido? ¿Cuáles son las conclusiones de este trabajo? El resumen debe tener como máximo 200 palabras. (En este semestre, debes resumir aquí los tiempos de ejecución, y los resultados del camino de menor riesgo y del camino más corto).

Palabras clave

Camino más corto restringido, acoso sexual callejero, identificación de rutas seguras, prevención del crimen.

1. INTRODUCCIÓN

Cada día es más común que una mujer cuando se desplaza de un lugar a otro sea víctima de algún tipo de acoso callejero tales como palabras soeces e incomodas, miradas persecutorias, insinuaciones sexuales entre otros, un estudio del Centro de Atención Integral a Víctimas de Abuso Sexual de la Fiscalía General de la Nación, indica que, en el año 2020, en la capital antioqueña 1.647 mujeres fueron víctimas de delitos sexuales.

Por lo anterior el objetivo de este proyecto es encontrar las rutas más seguras de desplazamiento dentro de la ciudad de Medellín teniendo en cuenta un ponderado de acoso, lo cual mitiga que una mujer pueda vivir situaciones tan deplorables.

1.1. Problema

El problema consiste en encontrar una estructura de datos que permita calcular el camino más corto sin superar un riesgo de acoso entre dos distancias.

Resolver este problema aportaría a disminuir índices de inseguridad con respecto a delitos sexuales y delincuencia los cuales día a día es un problema en común para toda la sociedad.

1.2 Solución

Explica, brevemente, tu solución al problema (En este semestre, la solución son algoritmos para caminos más cortos restringidos. ¿Qué algoritmos has elegido? ¿Por qué?)

1.3 Estructura del artículo

A continuación, en la Sección 2, presentamos trabajos relacionados con el problema. Posteriormente, en la Sección 3, presentamos los conjuntos de datos y los métodos utilizados en esta investigación. En la Sección 4, presentamos el diseño del algoritmo. Después, en la Sección 5, presentamos los resultados. Finalmente, en la Sección 6, discutimos los resultados y proponemos algunas direcciones de trabajo futuro.

2. TRABAJOS RELACIONADOS

A continuación, explicamos cuatro trabajos relacionados con la búsqueda de caminos para prevenir el acoso sexual callejero y la delincuencia en general.

3.1 Algoritmo para determinar la ruta más segura

Los incidentes criminales están en aumento en todo el mundo. Con el avance de la tecnología y los datos fácilmente disponibles relacionados con el crimen, es posible analizar estos datos para aumentar la conciencia y la seguridad de los ciudadanos. Usamos el conjunto de datos de San Francisco para proponer un modelo para determinar el camino más seguro de un lugar a otro.

El conjunto de datos consta de todos los delitos que han ocurrido en los últimos 12 años. Estos delitos van desde infracciones menores de tránsito hasta asalto e intento de asesinato y cada uno tiene una clasificación de seguridad "SI" o "NO"

celle	Aselta	seasi oferees	Conducir bajo el refuendo	Sesaestra	Fobs	Arbitration No. 10	50.00
1.5	ALTO:	AL10	89.10.	:5400	TIAKO	ALTO	HO
2	ngn	AUTO	AFTG	FAXT	ALDO .	DARL	WO.
3	P40	BAO	840	4 TC	6/70	PAIN	st.
4	ALTO	840	840	ALTO	ALTO	WLTC	3
5	ALTO	0.60	hain	470	DAKE.	neio	si
6	840	ALTO	840	E4/0	ALTO	ALTO	SE

Se crean perfiles de calle para cada calle. Un perfil de calle contiene el nombre de la calle y el recuento de las ocurrencias de cada categoría individual de delito cometido en una dirección que contiene la calle correspondiente

(i) Colombia on of each morph of about,

Bask weight of select
$$j = \sum_{i=1}^{\log n} (\operatorname{const}_{i,k} - \operatorname{rink} (\operatorname{actor}_{i,k}))$$

Where,

Total - the number of types of crimes.

Count - the count of command moidents for come,

is shown.

Christofactor - the most my of the strine.

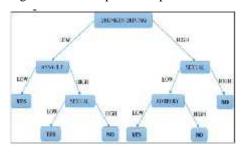
2) - Colombia on of riok of a route.

Blisk of route - $\sum_{i=1}^{\log n} (\operatorname{Bisk} \operatorname{widget} of street_{i,k} + \operatorname{elistation}_{i,k})$

Where,

to - weal number of unsafe streets on that - route.

El árbol de decisiones se utiliza para encontrar la ruta más segura entre múltiples rutas posibles.



3.2 Detección de la ruta más segura a través del cálculo del índice de peligro y la agrupación de K-Means

Formular una solución para identificar la ruta más segura entre dos ubicaciones geográficas ingresadas. Usando el conjunto de datos de la ciudad de Nueva York, que nos proporciona estadísticas de delitos con etiquetas de ubicación; estamos implementando diferentes algoritmos de agrupamiento y analizamos los resultados comparativamente para descubrir el más adecuado.

El conjunto de datos utilizado para el trabajo propuesto es de la ciudad de Nueva York tomado de los recursos en línea NYC Open-Data.

El cálculo de la puntuación de accidentes para cada coordenada utilizando la información de los conjuntos de datos se puede estimar como:

Accident Score (As) = Pk * 2 + Ck * 2 + Mk * 2 + Pi + Ci + Mi (1) where, Pk = Pedestrians Killed, Ck = Cyclists Killed,

Mk = Motorcyclists Killed, Pi = Pedestrians Injured, Ci = Cyclists Injured and Mi = Motorcyclists Injured

Asimismo, para los datos de delincuencia, le hemos dado a cada delito una calificación para considerar su ponderación.

Diman	gonzaradón	Certain	pondencia
Yulia as	15	armas peligrosas	7
Opposite light sequines	15	Florida con service control	6
receive market	14	parto grave no destructo	6
Prostinución y del tos relacionados Delitos	13	Para otras autondades	6
relacionações com refleis Seculações y	13	Robo	5
deblos relacionados Altandorio Maha de	12	Percentage selector	5
manufestion infantil biegos de azar	11	Contaction on a financial propagation are to se	4
- HARMACI NO COLO MONTRO CONTA	10	American Propheroria	4
Conducts deser denads	10	énfraçores de may manes	3
robe de servidos	9	Leves de tráfico y veniculos	3
Worms/Canada'	η	Fastficectón/acosó fraudidento	7
Drigs keep estatulos	.0	Harto mayor	1

Hemos implementado la agrupación en dos capas, la primera en función del índice de peligro y la segunda en función de las coordenadas para aumentar aún más nuestra precisión y permitir que las rutas se marquen con mayor precisión con el índice de peligro.

El resultado mostrado comprende la distancia, el tiempo de duración requerido para recorrer esa ruta y el índice de peligro calculado por el algoritmo propuesto. De esta manera, brinda al usuario una visión general de la seguridad en el camino que está a punto de recorrer.

3.3 Recomendación de ruta segura específica del usuario sistema

Las actividades delictivas han alcanzado niveles sin precedentes en casi todas las partes del mundo. Momentos desesperados como estos requieren medidas desesperadas para garantizar la seguridad de las personas, especialmente de las personas que necesitan viajar diariamente a lugares, conocidos y desconocidos.

Utilizamos la base de datos UCR [17] del FBI, que contiene registros bien equilibrados y seleccionados de víctimas relacionadas con delitos penales. Esta base de datos abarca un período de tiempo de más de 15 años.

Para la siguiente fase del algoritmo, es decir, el perfilado de rutas, utilizamos el conjunto de datos de San Francisco que contiene incidentes delictivos ocurridos durante 12 años.

1) Red de decisión P (Risk) = P(Age) * P(Sex) * P(Time) (1)

2) Cálculo del índice de riesgo y puntuación delictiva

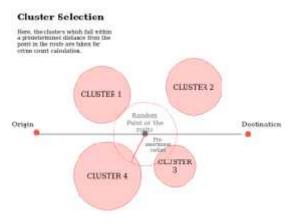
Risk Index = P(Risk) * Crime Score (2)

3) Información geoespacial reducida

Almacenamos los datos de conteo de delitos del grupo con el centroide del grupo correspondiente, obtenido del agrupamiento de K-Means como un atributo de índice.

4) Perfilado de ruta óptimo

El sistema toma la ubicación del origen y el destino como entrada del usuario para la recomendación de ruta. Luego, utilizando la API de Google Maps, se distinguen las tres rutas más cortas desde el origen hasta el destino. El sistema propone la ruta óptima en función del índice de seguridad y la longitud de la ruta.



3.4 Be-Safe Travel, una aplicación geográfica basada en la web para explorar una ruta segura en un área

Esta investigación propone BeSafe Travel, una aplicación basada en la web que utiliza la API de Google a la que pueden acceder las personas a las que les gusta conducir en un área, pero que aún no conocen los caminos que están a salvo de la delincuencia.

Los datos necesarios para la aplicación Be-Safe Travel son la ubicación de los delitos y los tipos de delitos

Para calcular la distancia entre el lugar de origen y el lugar de destino, usamos la fórmula de Haversine

$$a = \sin^2 \frac{\Delta \Phi}{2} + \cos \Phi_1 * \cos \Phi_2 * \sin^2 \frac{\Delta \Phi}{2}$$
 (1)

$$c = 2 * \arcsin 2(\sqrt{a}, \sqrt{1-a})$$
(2)

$$distance = R * c$$
 (3)

El programa comienza con la inicialización de mapas que define un contenedor en una etiqueta div HTML. Este contenedor mostrará Google Maps para la ruta deseada. La aplicación presentada utiliza DirectionService, que es la clase de la API de Google Maps para mostrar rutas.

El siguiente paso es la inicialización de puntos peligrosos y radio, que contiene dos subprocesos.

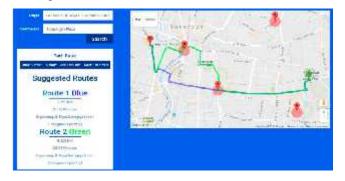
El primer subproceso es conectarse a la base de datos. La aplicación se conecta a la base de datos MySQL y hay dos tipos de consultas. Primero, una consulta para recuperar el

nombre y las coordenadas de ubicación expresadas por latitud y longitud. En segundo lugar, una consulta para obtener el waypoint definido como puntos peligrosos en los mapas.

El segundo subproceso creará el radio peligroso que marca la ubicación más cercana con los puntos del crimen como área peligrosa.

En primer lugar, el procedimiento contará cuántos puntos de delincuencia atraviesan las rutas posibles

Luego, el procedimiento crea una matriz en las secuencias de visualización de direcciones y variables incluidas en la matriz para determinar su secuencia.



3. MATERIALES Y MÉTODOS

En esta sección, explicamos cómo se recogieron y procesaron los datos y, después, diferentes alternativas de algoritmos del camino más corto restringido para abordar el acoso sexual callejero.

3.1 Recogida y tratamiento de datos

El mapa de Medellín se obtuvo de Open Street Maps (OSM) ¹y se descargó utilizando la API² OSMnx de Python. La (i) 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 se obtuvieron de los metadatos proporcionados por OSM.

Para este proyecto, se calculó la 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

² https://osmnx.readthedocs.io/

¹ https://www.openstreetmap.org/

Figura 1 presenta el riesgo de acoso calculado. El mapa está disponible en GitHub³.

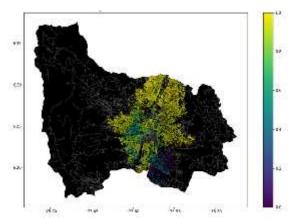


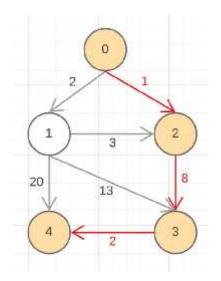
Figura 1. Riesgo de acoso sexual calculado como una combinación lineal de la fracción de hogares que se sienten inseguros y la fracción de hogares con ingresos inferiores a un salario mínimo, obtenida de la Encuesta de Calidad de Vida de Medellín, de 2017.

3.2 Alternativas de camino más corto con restricciones

A continuación, presentamos diferentes algoritmos utilizados para el camino más corto restringido. (En este semestre, ejemplos de dichos algoritmos son DFS, BFS, una versión modificada de Dijkstra, una versión modificada de A*, entre otros).

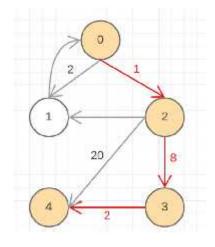
3.2.1 Dijkstra algorithm

El algoritmo de Dijkstra puede encontrar la ruta más corta desde un nodo llamado origen a todos los demás nodos de un grafo, produciendo un árbol de ruta más corta, este básicamente comienza en el nodo origen, luego analiza la ruta más corta desde ese nodo a todos los demás nodos del grafo, después ejecuta un seguimiento de la distancia más corta conocida actualmente desde cada nodo hasta el nodo principal u origen, y actualiza estos datos cada vez que encuentre una ruta más corta, prontamente de que el algoritmo haya encontrado la ruta más corta entre el nodo origen y otro nodo, este nodo se marca como visitado por lo cual se agrega a la ruta, este proceso continua hasta que todos los nodos del grafo se hayan agregado a la ruta. De esta forma tenemos una ruta que conecta el nodo de origen con los demás, siguiendo la ruta más corta posible, para llegar a cada uno de los nodos.



3.2.2 Back tracking algorithm

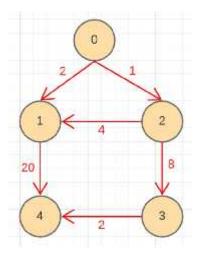
Es un método de búsqueda de soluciones sobre grafos, el cual se acelera mediante la poda de ramas poco prometedoras, normalmente se suele implementar como un procedimiento recursivo, también se generan funciones de cota, de forma que no se generan algunos estados si no conducen a ninguna solución, de esta forma se ahorra espacio en memoria y tiempo de ejecución, para diseñar este tipo de algoritmos se tienen en cuenta los siguientes elementos: una función de poda para eliminar partes del árbol de búsqueda, construir la solución del problema en distintas etapas, en cada caso se elige un posible candidato y se añade a la solución parcial, y si no quedan más valores por probar, se retrocede al candidato anterior, se desecha y se selecciona otro candidato.



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

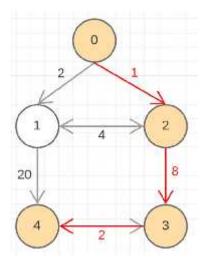
3.2.3 Algoritmo de búsqueda profunda aleatorizada

Este algoritmo busca aleatoriamente una ruta para llegar a una casilla final y verifica en cada caso si hay más de un camino que se pueda recorrer, si ese es el caso este guarda la posición para que cuando encuentre un camino que no tenga salida, regrese a la casilla anterior y continue la búsqueda desde esa posición, este proceso se repite hasta encontrar el camino que mejor se ajuste a la necesidad.



3.2.4 Bellman Ford Algorithm

Este algoritmo genera el camino más corto desde un nodo origen al resto de nodos de un grafo como lo hace el algoritmo de Dijkstra, sin embargo, la diferencia es que este algoritmo soluciona el problema de la ruta mas corta, de una forma más general que Dijkstra, ya que permite valores negativos, en pocas palabras este algoritmo se utiliza únicamente cunado hay aristas negativas presentes en el grafo, lo que lo hace un poco mejor ya que admite longitudes negativas en los arcos.



4. DISEÑO E IMPLEMENTACIÓN DEL ALGORITMO

A continuación, explicamos las estructuras de datos y los algoritmos utilizados en este trabajo. Las implementaciones de las estructuras de datos y los algoritmos están disponibles en Github⁴.

4.1 Estructuras de datos

Explica la estructura de datos que se utilizó para implementar el algoritmo del camino más corto restringido y haz una figura que lo explique. No utilice figuras de Internet. (En este semestre, los ejemplos de las estructuras de datos son la matriz de adyacencia, la lista de adyacencia, la lista de adyacencia utilizando un diccionario). La estructura de los datos se presenta en la Figura 2.

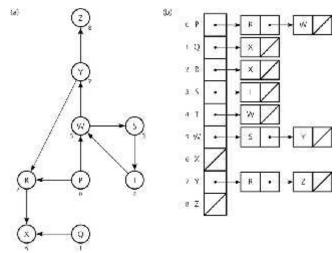


Figura 2: Un ejemplo de mapa de calles se presenta en (a) y su representación como lista de adyacencia en (b). (Por

⁴ http://www.github.com/ ???????? /.../proyecto/

favor, siéntase libre de cambiar esta gráfica si utiliza una estructura de datos diferente).

4.2 Algoritmos

En este trabajo, proponemos algoritmos para el problema del camino más corto restringido. El primer algoritmo calcula el camino más corto sin superar un riesgo medio ponderado de acoso r. El segundo algoritmo calcula el camino con el menor riesgo medio ponderado de acoso sin superar una distancia d.

4.2.1 Primer algoritmo

Explica el diseño del algoritmo para calcular el camino más corto sin superar una media ponderada de riesgo de acoso r y haz tu propia gráfica. No utilices gráfica de Internet, haz las tuyas propias. (En este semestre, el algoritmo podría ser DFS, BFS, una versión modificada de Dijkstra, una versión modificada de A^* , entre otros). El algoritmo se ejemplifica en la Figura 3.

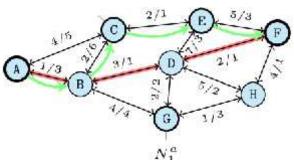


Figura 3: Resolución del problema del camino más corto restringido con la Búsqueda Primera Profunda (DFS). (Por favor, siéntase libre de cambiar esta figura si utiliza un algoritmo diferente).

4.2.2 Segundo algoritmo

Explica el diseño del algoritmo para calcular el camino con el menor riesgo medio ponderado de acoso sin superar una distancia d y haz tu propia gráfica. No utilices gráficas de Internet, haz las tuyas propias. (En este semestre, el algoritmo podría ser DFS, BFS, una versión modificada de Dijkstra, una versión modificada de A*, entre otros). El algoritmo se ejemplifica en la Figura 4.

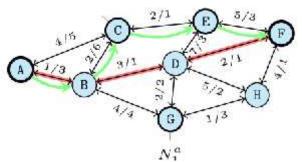


Figura 4: Resolución del problema del camino más corto restringido con la Búsqueda Primera Profunda (DFS). (Por

favor, siéntase libre de cambiar esta gráfica si utiliza un algoritmo diferente).

4.4 Análisis de la complejidad de los algoritmos

Explica, con tus propias palabras, el análisis, para el peor caso, utilizando la notación O. ¿Cómo ha calculado esas complejidades? Explique brevemente.

Algoritmo	Complejidad temporal
Nombre del algoritmo	O(V ² *E ²)
Nombre del segundo algoritmo (en caso de que haya probado dos)	

Tabla 1: Complejidad temporal del nombre de su algoritmo, donde V es... E es... (*Por favor, explique qué significan V y E en este problema*).

Estructura de datos	Complejidad de la memoria
Nombre de la estructura de datos	O(V*E*2 ^E)
Nombre de la segunda estructura de datos (en caso de que haya intentado dos)	O(2 E*2 V)

Tabla 2: Complejidad de memoria del nombre de la estructura de datos que utiliza su algoritmo, donde V es... E es... (*Por favor, explique qué significan V y E en este problema*).

4.5 Criterios de diseño del algoritmo

Explique por qué el algoritmo fue diseñado de esa manera. Utilice criterios objetivos. Los criterios objetivos se basan en la eficiencia, que se mide en términos de tiempo y memoria. Ejemplos de criterios NO objetivos son: "estaba enfermo", "fue la primera estructura de datos que encontré en Internet", "lo hice el último día antes del plazo", "es más fácil", etc. Recuerda: Este es el 40% de la calificación del proyecto.

5. RESULTADOS

En esta sección, presentamos algunos resultados cuantitativos sobre el camino más corto y el camino con menor riesgo.

5.1.1 Resultados del camino más corto

A continuación, presentamos los resultados obtenidos para el camino más corto, sin superar un riesgo medio ponderado de acoso r, en la Tabla 3.

Origen	Destino	Distancia más corta	Sin exceder r
Universidad EAFIT	Universidad de Medellín	??	0.84
Universidad de Antioquia	Universidad Nacional	???	0.83
Universidad Nacional	Universidad Luis Amigó	??	0.85

Tabla 3. Distancias más cortas sin superar un riesgo de acoso medio ponderado r.

5.1.2 Resultados de menor riesgo de acoso

A continuación, presentamos los resultados obtenidos para el trayecto con menor riesgo de acoso medio ponderado, sin superar una distancia *d*, en la Tabla 4.

Origen	Destino	Acoso más bajo	Sin exceder d
Universidad EAFIT	Universidad de Medellín	??	5,000
Universidad de Antioquia	Universidad Nacional	???	7,000
Universidad Nacional	Universidad Luis Amigó	??	6,500

Tabla 3. Menor riesgo de acoso ponderado sin superar una distancia d (en metros).

5.2 Tiempos de ejecución del algoritmo

En la Tabla 4, explicamos la relación de los tiempos medios de ejecución de las consultas presentadas en la Tabla 3.

Calcule el tiempo de ejecución de las consultas presentadas en la Tabla 3. Indique los tiempos de ejecución medios.

	Tiempos medios de ejecución (s)
Universidad EAFIT a Universidad de Medellín	100.2 s
De la Universidad de Antioquia a la	800.1 s

Universidad Nacional	
De la Universidad Nacional a la Universidad Luis Amigó	845 s

Tabla 4: Tiempos de ejecución del nombre del *algoritmo* (*Por favor, escriba el nombre del algoritmo, por ejemplo, DFS, BFS, un A* modificado*) para las consultas presentadas en la Tabla 3.

6. CONCLUSIONES

Explique los resultados obtenidos. ¿Son los caminos más cortos significativamente diferentes de los caminos con menor riesgo de acoso? ¿Qué utilidad tiene esto para la ciudad? ¿Son razonables los tiempos de ejecución para utilizar esta implementación en una situación real?

6.1 Trabajos futuros

Responda, ¿qué le gustaría mejorar en el futuro? ¿Cómo le gustaría mejorar su algoritmo y su aplicación? ¿Continuará este proyecto trabajando en la optimización? ¿En estadística? ¿Desarrollo web? ¿Aprendizaje automático? ¿Realidad virtual? ¿Cómo?

AGRADECIMIENTOS

Identifique el tipo de agradecimiento que desea escribir: para una persona o para una institución. Tenga en cuenta las siguientes pautas: 1. El nombre del profesor no se menciona porque es un autor. 2. No debe mencionar a los autores de los artículos con los que no se ha puesto en contacto. 3. Debe mencionar a los alumnos, profesores de otros cursos que le han ayudado.

A modo de ejemplo: Esta investigación ha sido apoyada/parcialmente apoyada por [Nombre de la Fundación, Donante].

Agradecemos la ayuda con [técnica particular, metodología] a [Nombre Apellido, cargo, nombre de la institución] por los comentarios que mejoraron en gran medida este manuscrito.

Los autores agradecen al profesor Juan Carlos Duque, de la Universidad EAFIT, por facilitar los datos de la Encuesta de Calidad de Vida de Medellín, de 2017, procesados en un *Shapefile*.

REFERENCIAS

<u>Las</u> referencias se obtienen utilizando el formato de referencia de la ACM. Lea las directrices de la ACM en http://bit.ly/2pZnE5g

 Alcaldía de Medellín, Se fortalecen las acciones de prevención de la violencia y el acoso contra las mujeres en el Centro de Medellín. Retrieved May 3, 2021, from Alejandro Noreña Jaramillo https://www.medellin.gov.co/irj/portal/medellin?N avigationTarget=contenido/9737-Se-fortalecen-las-acciones-de-prevencion-de-la-violencia-y-el-acoso-contra-las-mujeres-en-el-Centro-de-Medellin

2. International Journal of Computer Science and Information Technologies, Algorithm to Determine the Safest Route, Retrieved 2016

http://ijcsit.com/docs/Volume%207/vol7issue3/ijcsit20160703106.pdf

 Computers, Materials & Continua, Safest Route Detection via Danger Index Calculation and K-Means Clustering, Retrieved 2021

https://www.techscience.com/cmc/v69n2/43879/p df

4. International Journal of Engineering Research & Technology, User Specific Safe Route Recommendation System, Retrieved 2020

https://www.ijert.org/research/user-specific-safe-route-recommendation-system-IJERTV9IS100268.pdf

5. Actas de la conferencia AIP1867, Be-safe travel, una aplicación geográfica basada en la web para explorar una ruta segura en un área, Retrieved August 01, 2017

https://aip.scitation.org/doi/pdf/10.1063/1.4994426