

Algoritmo para la prevención del acoso sexual en las calles de Medellín mediante la búsqueda de rutas seguras.

Juan David Zapata Moncada
Universidad Eafit
Colombia
jdzapatam@eafit.edu.co

Samuel Rodriguez Hernández
Universidad Eafit
Colombia
Srodrigueh@eafit.edu.co

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

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

RESUMEN

El problema es la inseguridad en las calles, por el acoso, la delincuencia, los accidentes y todo aquello que perturba a los ciudadanos y que no se pueda transcurrir tranquilamente sin el temor de que suceda algo malo. Este problema es muy importante debido a que se ven perjudicados todas las personas y empresas que sitúan en los alrededores donde se frecuente dicho problema. Los problemas relacionados pueden ser aquellos robos a turistas, a tiendas, a personas que viven en esa misma zona y esto se produce por una gran tasa de desempleo.

¿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 obtenidos con los tres caminos).*

Palabras clave

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

1. INTRODUCCIÓN

Se plantea diseñar un algoritmo que determine una ruta segura y rápida para el beneficio del usuario, se realiza debido a que la tasa de criminalidad y acoso sexual han aumentado a unas grandes cifras y se requiere una pronta intervención para disminuir este gran problema. Este trabajo se realiza para construir una sociedad más segura.

1.1. Problema

Actualmente el acoso es uno de los problemas que encabezan la lista de los más graves y comunes en el mundo y miles de personas sufren secuelas psicológicas y físicas que en ocasiones les generan trastornos todo gracias al abuso de otro. Esto impacta la sociedad negativamente ya que esto perjudica el desempeño en el trabajo por lo que indirectamente las empresas se ven afectadas y el deterioro de las relaciones sociales a causa de la desconfianza.

1.2 Solución

La solución la obtenemos mediante el algoritmo dijkstra el cual va comparando valores de nodos vecinos para encontrar el camino que establezca la distancia y el valor del riesgo de acoso, esto lo hacemos sumando la distancia y el riesgo

multiplicado por la distancia($d+(d*h)$) con el valor resultante es con el que trabaja el algoritmo dijkstra. Lo elegimos porque nos pareció una solución práctica y poco compleja de implementar.

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

2.1 Prevención del acoso sexual a través de un algoritmo de búsqueda de rutas utilizando la búsqueda cercana

Este trabajo surge por la necesidad de encontrar puntos seguros y rutas con el menor riesgo de acoso posible en relación con una dirección y un desplazamiento de (x, y) punto inicial a (x, y) destino. Para llevar a cabo este objetivo buscaron la forma de crear un algoritmo de búsqueda de caminos para prevenir el acoso sexual y lo hicieron utilizando un mapa de calor, dividiendo el área en una cuadrícula de $\sim 1,74 \text{ km}^2$ cada cuadro a la cual se le asignaba una puntuación de riesgo que pueden ir de 0 hasta 4, siendo 0 el lugar más seguro y 4 siendo el más peligroso.

Para determinar la mejor ruta lo que se hizo fue que calcular el promedio de las puntuaciones de riesgo de los cuadros por donde pasan las rutas lo cual, para determinar la mejor implica encontrar la que tiene menor distancia euclidiana mediante su ecuación. Por otro lado, para que se pudiera saber si una ruta tocaba una cuadrícula realizaron al método del dibujo lineal de Bresenham, el cual, es un algoritmo que según la distancia de línea a punto determina si sombrear el punto de dicha intersección. Esto se hace con la intención de tomar los valores de riesgo de los cuadros donde están los puntos y que quede así:

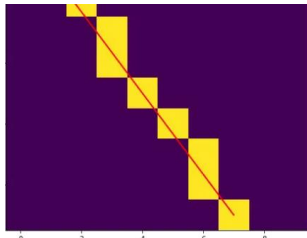


Figure 8: Illustration of Bresenham's line drawing method using skimage and matplotlib

Finalmente se puede ver que algunos cuadros no fueron incluidos en el cálculo por lo que optaron por incluir otro algoritmo que se encargue de eso y que su función es encontrar la “supercubierta” del segmento de línea. El resultado fue un algoritmo que recibía un inicio y un destino y te daba una ruta que prioriza la seguridad antes que la distancia más reducida.

2.2 Un sistema de integración y análisis de datos para la planificación de rutas seguras

Este trabajo surge por un gran índice de casos de acosos sexual contra las mujeres y además por el aumento de criminalidad causada por el desempleo y la pérdida de recursos por el periodo del COVID-19. Se da la idea de diseñar una aplicación que puede determinar las rutas más seguras y a su vez más rápidas. Unas de las maneras más eficaces para resolver el problema es desarrollar un modelo que nos brinde con exactitud qué tan seguro es un lugar. Este utilizaría varios insumos, que podrían influir en la tasa relativa de acoso en un área. A pesar de lo complicado que es determinar una ruta 100% segura se puede dar de que la ruta no sea la rápida posible y es por eso que debe haber un equilibrio.

Para darle solución se realizaron encuestas a varias mujeres preguntándoles que condiciones las haría sentir seguras y cuales las harían sentir inseguras.

El algoritmo se ensaya con suponiendo que hay 5 rutas:

y_1, y_2, y_3, y_4, y_5 , cada valor será de 0 a 100 siendo la máxima seguridad 500 y en cada ruta hay una diferente situación y se trata de determinar cuál es aquella que cumple con más de las especificaciones en las encuestas realizadas a varias mujeres.

2.3 Más allá de la ruta más corta: una encuesta sobre la navegación en rutas consciente de la calidad para

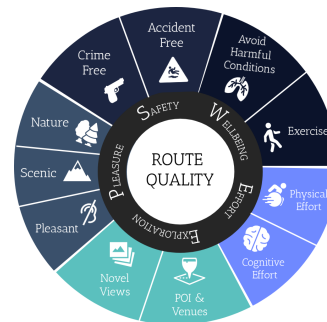
Este trabajo surge de la funcionalidad de grandes sistemas como por ejemplo Google Maps, queriendo profundizar su función dándole un mejor beneficio al usuario implementando algoritmos que den las rutas con menos probabilidad de riesgo. Se propone que la información que

se necesita para crear dicho algoritmo sea brindada por la propia sociedad en distintas redes sociales.

Se define una ruta de navegación consciente de la calidad sistema para peatones basado en la siguiente característica:

El sistema brinda al peatón una ruta que le permite moverse de un punto A hasta un punto B y cuando el usuario ingresa más puntos para visitar en la ruta, la tarea se convierte en encontrar una serie de rutas entre ubicaciones intermedias consecutivas X_i a X_{i+1} dentro de una ruta.

Con base en la literatura identificada, se proporciona una descripción general de los sistemas de navegación de ruta conscientes de la calidad que han sido propuesto en base a los atributos que se utilizan para representar las cualidades de una ruta, el tipo de datos que se han utilizado para implementar tales sistemas, los algoritmos utilizados para implementarlos y los enfoques de evaluación utilizados para implementar la ruta consciente de la calidad de los sistemas de navegación.



2.4 ruta segura: un modelo sólido para la predicción de rutas más seguras utilizando datos sobre delitos y accidentes

Este trabajo surge por la inseguridad en las ciudades y esto es causado por todos los casos de delincuencia. Incluso los mapas de Google pueden dar una ruta rápida pero no segura. Está claro que las personas locales a la zona conocen mejor las calles y sus peligros, pero los turistas no tienen ni idea y están confiando en una aplicación que no está diseñada para garantizar seguridad, por ende, el gobierno reúne datos sobre crímenes, accidentes, rutas, etc. para diseñar un sistema que pueda brindar una mayor función que la anterior mencionada.

Se toma como ejemplo para desarrollar el sistema la ciudad de New York. Se propone como solución en un primer caso predecir una ruta segura utilizando datos de delincuencia y accidentes, considerando un punto de origen y un punto destino, en segundo caso se divide la ciudad en varias regiones de riegos para dar mejores predicciones ya que se tienen en cuenta áreas de delincuencia más pequeñas y como tercer caso se calcula la puntuación de riesgo más baja para sugerir la ruta más segura de acuerdo a los datos obtenidos de los clústeres más cercanos.

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)¹ y se descargó utilizando la API² 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 GitHub³

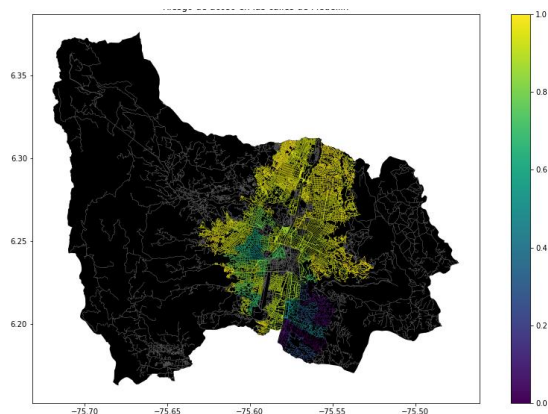


Figura 1. Riesgo de acoso sexual calculado como una combinación lineal de la fracción de hogares que se sienten inseguros y la fracción de hogares con ingresos inferiores a un salario mínimo, obtenidas de la Encuesta de Calidad de Vida de Medellín, de 2017.

3.2 Alternativas de caminos que reducen el riesgo de acoso sexual callejero y distancia

A continuación, presentamos diferentes algoritmos utilizados para un camino que reduce tanto el acoso sexual callejero como la distancia.

3.2.1 algoritmo lee

El algoritmo lo que busca es encontrar el camino más corto desde un punto a otro, el laberinto es construido como una matriz binaria y son las celdas que tengan el número 1 por donde se puede avanzar. Lo primero que el algoritmo hace es crear una cola vacía y poner en cola la celda de origen que tenga una distancia 0 y luego la marca como visitada y cada vez que se mueva de celda se aumentará la distancia en 1, además, si este es el nodo de destino la función retorna la distancia y si no se llega al destino se retorna false. También hay que tener en cuenta que las celdas donde halla un 1 se recorren primero

3.2.2 Encuentra el camino más corto en un laberinto

En este caso desde un principio se buscan todos los caminos posibles de un laberinto de posición inicial a destino. Esto se realiza explorando los 4 caminos posibles (arriba, abajo, derecha, Izquierda) y verifica recursivamente si conducen al destino o no, además la “longitud mínima” se compara y actualiza cada vez que se llega por un camino al destino y si se encuentra en un callejón sin salida hace un retroceso para explorar otros caminos, además no tiene en cuenta si una celda ya fue visitada por otra ruta.

3.2.3 Encuentra el camino más corto en un laberinto usando retroceso

Este es un algoritmo de retroceso en el cual tenemos una matriz binaria y nuevamente solo podemos avanzar si la celda cercana a la actual contiene un 1, el movimiento puede ser vertical u horizontal y va de 1 en 1. Este algoritmo lo que hace es que avanza al 1 más cercano y luego analiza si puede ir arriba y si no puede entonces agota todas las posibilidades y si ninguna sirve entonces se devuelve una celda para intentar otro camino. Al llegar al destino explora otros caminos para finalmente determinar cuál es el más corto

3.2.4 algoritmo dijkstra

¹ <https://www.openstreetmap.org/>

² <https://osmnx.readthedocs.io/>

dist, una matriz de distancias desde el nodo de origen ss a cada nodo en el gráfico, inicializado de la siguiente manera: $\text{dist}(\text{ss}) = 0$; y para todos los demás nodos vv, $\text{dist}(\text{vv}) = \infty$. Esto se hace al principio porque a medida que avanza el algoritmo, la dist desde la fuente hasta cada nodo vv en el gráfico se recalculará y finalizará cuando se encuentre la distancia más corta a vv. QQ, una cola de todos los nodos en el gráfico. Al final del progreso del algoritmo, QQ estará vacío. SS, un conjunto vacío, para indicar qué nodos ha visitado el algoritmo. Al final de la ejecución del algoritmo, SS contendrá todos los nodos del gráfico.

El algoritmo procede como sigue:

Si bien QQ no está vacío, extraiga el nodo vv, que aún no está en SS, desde QQ con la dist más pequeña (vv). En la primera ejecución, se elegirá el nodo de origen ss porque $\text{dist}(\text{ss})$ se inicializó en 0. En la siguiente ejecución, se elige el siguiente nodo con el valor de dist más pequeño. Agregue el nodo vv a SS, para indicar que se ha visitado vv. Actualice los valores dist de los nodos adyacentes del nodo actual vv de la siguiente manera: para cada nuevo nodo adyacente uu, si $\text{dist}(\text{vv}) + \text{peso}(\text{u,v}) < \text{dist}(\text{uu})$, se encontró una nueva distancia mínima para uu, así que actualice $\text{dist}(\text{uu})$ al nuevo valor de distancia mínima; de lo contrario, no se realizan actualizaciones en $\text{dist}(\text{uu})$.

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.

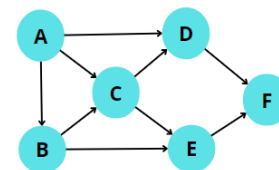
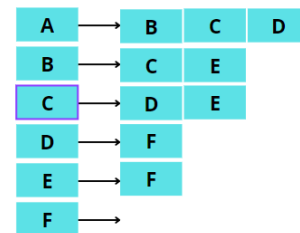
4.1 Estructuras de datos

La estructura de datos utilizada para almacenar los datos de la ciudad de Medellín fue una lista de adyacencia utilizando diccionarios a partir de la información del dataframe. Lo primero que se hizo fue llenar los espacios de riesgo del dataframe que no tenían valor con la media de los valores de “riesgos de acoso”, luego se determinaron cuáles eran los orígenes únicos del dataframe para posteriormente crear las llaves del diccionario con estos orígenes únicos. Finalmente, dentro de las llaves de origen se creó otra llave con los destinos la cual almacena un arreglo con dos valores: la distancia y el riesgo de acoso. Además, había que tener en cuenta si la calle era doble vía; si este valor en el dataframe

era “true” entonces ese destino debía ser tratado también como un origen así que se agregó a las llaves y dentro de este va otra llave que es la del origen la cual contiene los valores de distancia y riesgo de acoso. En las siguientes imágenes se puede ver:

1. La representación de una lista de adyacencia
2. La representación como un grafo de la lista de adyacencia
3. Como se implementa una lista de adyacencia con diccionarios en Python.

Lo único que no incluimos en las imágenes fue la información que pueden opcionalmente contener los enlaces y nodos.



```
grafo={
  "a"={"b","c","d"}
  "b"={"c","e"}
  "c"={"d","e"}
  d={"f"}
  e={"f"}
  f={}
}
```

4.2 Algoritmos

En este trabajo, proponemos un algoritmo para un camino que minimiza tanto la distancia como el riesgo de acoso sexual callejero.

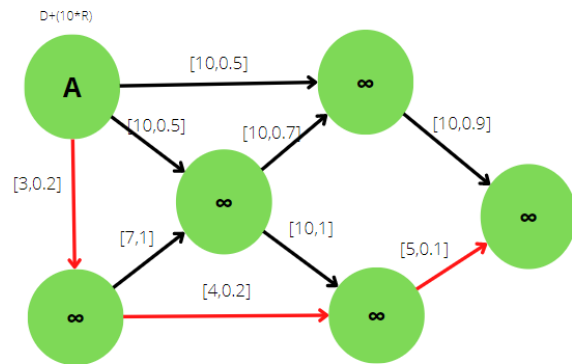
4.1 algoritmo dijkstra modificado para encontrar el mejor camino tomando en cuenta el riesgo de acoso sexual callejero.

Lo primero que ocurre es que el algoritmo recibe un origen, un destino y un grafo. En principio se definen 4 diccionarios uno para ir guardando las distancias otro para los riesgos de acoso uno que es con los valores del algoritmo y otro que es para hacer un rastreo del camino, también se define un

número muy grande y finalmente un arreglo vacío que al final va a contener el camino más óptimo.

Lo primero que se hizo fue darles un valor de infinito a todos los nodos exceptuando al origen, posterior a esto se realiza otro ciclo para encontrar el nodo que contenga menor valor. Luego de haberlo encontrado identificamos a los nodos “vecinos” para irles cambiando el valor que se inicializó en un número muy grande por el menor posible, además es aquí donde se van guardando los antecesores, las distancias y los riesgos. Finalmente, luego de que termine esta iteración se elimina el nodo de menor distancia que se había escogido al inicio del algoritmo y se vuelve a repetir este proceso hasta que se hayan sacado todos los nodos del grafo.

Por último, se ejecuta otro ciclo con el objetivo de extraer del diccionario del rastreo de camino la mejor opción y pasarla a un arreglo para luego ser mostrada por pantalla. Además, se hizo un try catch por si era imposible ir del origen al destino que se recibieron.



4.2.2 Cálculo de otros dos caminos para reducir tanto la distancia como el riesgo de acoso sexual callejero

Explica los otros dos caminos que reducen tanto la distancia como el riesgo de acoso sexual callejero y haga su propia gráfica. No utilice gráficas de Internet, haga las suyas. (En este semestre, el algoritmo podría ser DFS, BFS, Dijkstra, A*, entre otros). El algoritmo se ejemplifica en la Figura 4.



Figura 4: Mapa de la ciudad de Medellín donde se presentan tres caminos para peatones que reducen tanto el riesgo de

acoso sexual como la distancia en metros entre la Universidad EAFIT y la Universidad Nacional.

4.3 Análisis de la complejidad del algoritmo

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)	$O(E^3 * V * 2^V)$

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). No, no use ‘n’.

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). No, no sive ‘n’. Es decir, no usar ‘n’. No ‘n’.

4.4 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. Recuerde: Este es el 40% de la calificación del proyecto.

5. RESULTADOS

En esta sección, presentamos algunos resultados cuantitativos sobre los tres caminos que reducen tanto la distancia como el riesgo de acoso sexual callejero.

5.1 Resultados del camino que reduce tanto la distancia como el riesgo de acoso sexual callejero

A continuación, presentamos los resultados obtenidos de *tres caminos que reducen tanto la distancia como el acoso*, en la Tabla 3.

Origen	Destino	Distancia	Riesgo
Eafit	Unal	??	??
Eafit	Unal	???	??
Eafit	Unal	??	??

Tabla 3. Distancia en metros y riesgo de acoso sexual callejero (entre 0 y 1) para ir desde la Universidad EAFIT hasta la Universidad Nacional caminando.

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.

Cálculo de v	Tiempos medios de ejecución (s)
v = ??	100000.2 s
v = ??	800000.1 s
v = ??	8450000 s

Tabla 4: Tiempos de ejecución del nombre del *algoritmo* (Por favor, escriba el nombre del algoritmo, por ejemplo, DFS, BFS, A*) para cada uno de los tres caminos calculadores entre EAFIT y Universidad Nacional.

6. CONCLUSIONES

Explique los resultados obtenidos. ¿Son los caminos significativamente diferentes? ¿Qué utilidad tiene esto para la ciudad? ¿Son razonables los tiempos de ejecución para utilizar esta implementación en una situación real? ¿Qué camino recomendaría para una aplicación móvil o web?

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 archivo *Shapefile*.

Referencias

1. Sandeep, C., Gauri Shankar, V., Soni, S. 2019. Route-The Safe: A Robust Model for Safest Route Prediction Using Crime and Accidental Data. Recuperado agosto 13, 2022 de https://www.researchgate.net/publication/338096313_Route-The_Safe_A_Robust_Model_for_Safest_Route_Prediction_Using_Crime_and_Accidental_Data
2. Jigang, W., Jin, S., Ji, H. Srikanthan, T. 2011. Algorithm for Time-dependent Shortest Safe Path on Transportation Networks. Recuperado agosto 13, 2022 de <https://www.sciencedirect.com/science/article/pii/S1877050911001591>

3. Sarraf, R. and P., M., 2018. A Data Integration and Analysis System for Safe Route Planning. Proquest.com.
4. Ma, D. 2022. Preventing Sexual Harassment Through a Path Finding Algorithm Using Nearby Search. Omdena | Building AI Solutions for Real-World Problems.
5. Anon, Shortest Path in Maze using Backtracking – Pencil Programmer. Pencil Programmer
6. Dijkstra's Algorithm. 2022. Programiz.com. <https://www.programiz.com/dsa/dijkstra-algorithm>