



UNIVERSIDAD ANDRÉS BELLO

Facultad de Ingeniería

# **Mejoramiento de un algoritmo de ambigüedad espacial topológico**

Autor: Felipe Andrés Pérez Palavecino

Profesor Guía: Dra. Carola Blazquez

Profesor Co-guía: Dr. Roberto León Vásquez

Tesis para optar al Grado de  
Magíster en Ciencias de la Computación

Santiago - Chile

Abril - 2022



UNIVERSIDAD ANDRÉS BELLO

Facultad de Ingeniería

Título: **Mejoramiento de un algoritmo de ambigüedad espacial topológico**

Autor: **Felipe Andrés Pérez Palavecino**

Tesis para optar al grado de **Magíster en Ciencias de la Computación** de la Universidad Andrés Bello.

Comité de Evaluación

Dra. Carola Blazquez

Universidad Andrés Bello

\_\_\_\_\_

Guía

Dr. Roberto León

Universidad Andrés Bello

\_\_\_\_\_

Co-Guía

???

Universidad ???

\_\_\_\_\_

Examinador/a interno/a

???

Universidad ???

\_\_\_\_\_

Examinador/a externo/a



## Índice general

Índice de figuras	5
Índice de tablas	7
Resumen	9
Capítulo 1. Introducción	11
1.1. Objetivo general	12
1.2. Objetivos específicos	12
Capítulo 2. Revisión Bibliográfica	13
Capítulo 3. Marco Teórico	15
3.1. Map Matching	15
3.2. Algoritmo Map-Matching Topológico (TMMA)	17
Capítulo 4. Metodología	21
Capítulo 5. Resultados	29
Capítulo 6. Discusión	41
Capítulo 7. Conclusiones	43
Bibliografía	45



## Índice de figuras

2.1.Asignación perpendicular más cercana.	13
2.2.Radio entre los puntos n1 y n2.	14
3.1.Interferencia en Sistemas GPS [[6]].	15
3.2.Diagrama de flujo TMMA.	18
3.3.Diagramas de las mejoras realizadas.	19
4.1.Cantidad de registros dentro del periodo de la muestra.	21
4.2.Red Vial comuna de Renca.	22
4.4.Rutas seleccionadas.	24
4.5.Distribución de los set de datos seleccionados.	24
4.6.Diagrama y representación del cálculo del ángulo de la calle.	26
5.1.Algoritmo original: Porcentaje Acierto / Radios.	30
5.2.Comparativa Acierto/Tiempo: Alg. Original V/S Mejorado.	31
5.3.Resultado de proyección para el registro 66 (Serie 3) por el Alg. Original y Alg. Mejorado.	32
5.4.Registros con velocidad y heading.	35
5.5.Alg. Mejorado: TH 10°, TH 20°, TH 30°.	36
5.6.Alg. Mejorado: TV8km/h, TV16km/h, TV32km/h, TV64km/h.	38



## Índice de tablas

4.1.Set de datos utilizados.	22
4.2.Porcentaje de vías utilizadas.	25
5.1.Alg. original: Porcentaje de acierto.	29
5.2.Alg. original: Tiempo computacional (min).	29
5.3.Comparativa Acierto/Tiempo: Alg. Original V/S Mejorado.	31
5.4.Cantidad de ruteo entre pares de puntos.	33
5.5.Algoritmo Original V/S Algoritmo Mejorado, Serie 3 original y menor frecuencia.	34
5.6.Algoritmo Mejorado V/S Filtro $v > 5\text{km/h}$ .	35
5.7.Algoritmo Mejorado: TH $10^\circ$ V/S TH $20^\circ$ V/S TH $30^\circ$ .	37
5.8.Alg. Mejorado: TV8km/h, TV16km/h, TV32km/h, TV64km/h.	38





## Resumen

El problema de Map-Matching ocurre cuando se quiere integrar la posición GPS de un ente, principalmente un vehículo, con la red vial en un mapa digital. La principal motivación de esta investigación es implementar una mejora a un Algoritmo de Map-Matching Topológico (TMMA) ya desarrollado en un estudio anterior, aplicando principalmente dos mejoras que ayuden a obtener resultados en un menor tiempo computacional sin perjudicar la calidad de estos. El objetivo es poder reducir el trabajo realizado por el algoritmo asegurando un buen resultado, sin la necesidad de tener que estar variando los parámetros de entrada en la búsqueda de un mejor rendimiento. Para poder llevar a cabo esta investigación, se realizó un análisis previo de los datos de entrada para trabajar con información coherente que no interfiriera en los resultados finales. De esta manera, se seleccionaron ocho set de datos con mediciones GPS con frecuencia de 30 segundos, provenientes de vehículos de carga que transitaban por la comuna de Renca, Santiago en febrero de 2015. Dentro de las mejoras realizadas, se incorporó una obtención previa del radio de búsqueda óptimo que asegura que la gran mayoría de las mediciones GPS tengan una vía cercana. En segundo lugar, se incluyó un procedimiento encargado de filtrar las calles que no tienen la misma dirección en la que se iba moviendo el vehículo. Con estas mejoras desarrolladas, se ejecutaron diferentes experimentos con los ocho set de datos, que posteriormente se compararon con el algoritmo original respecto al tiempo de ejecución y calidad de la solución. Donde se logró apreciar que seis de las ocho rutas se resolvieron en un tiempo menor que con el algoritmo original. Asimismo, siete de las ocho rutas obtuvieron un porcentaje de acierto igual o mejor. Analizando los resultados con un mayor detalle, se logró apreciar que las mediciones que tenían una velocidad inferior a 5 km/h entregaban una dirección (heading) imprecisa. Por lo tanto, se decidió realizar un experimento que no considerara el filtro de las calles por su dirección, cuando un registro tuviera una velocidad inferior a 5 km/h. Sin embargo, los resultados obtenidos con la nueva mejora mostraron ser despreciable como para considerar esta modificación en el algoritmo. Por otro lado, se decidió evaluar cómo se comportaría con un set de datos al reducir la frecuencia de los

registros a la mitad, quedando con un set de datos registrado cada 1 minuto. En el cual se logró percibir que el tiempo computacional igualmente se redujo a la mitad, y el comportamiento del algoritmo mejorado siguió siendo superior al algoritmo original. Por último, se realizó una evaluación de sensibilidad para las variables de Tolerancia de Velocidad y de Heading, donde se logró comprobar la importancia que éstas tienen y como influyen directamente en el rendimiento del algoritmo. Se espera que para una siguiente investigación se puedan estandarizar estas variables, tal como se hizo con el radio de búsqueda, el cual se selecciona previamente de acuerdo al set de datos con el que se trabajará de forma independiente.

## Introducción

Durante las últimas décadas, los sistemas de posicionamiento global GPS (por sus siglas en inglés, Global Positioning System) han pasado a formar parte de nuestra vida cotidiana al encontrarse mayormente en nuestros dispositivos móviles. Estos sistemas de geolocalización en conjunto con sistemas de información geográfica (SIG) son altamente utilizados en el seguimiento y monitoreo de flotas de vehículos, los cuales se mueven por medio de diferentes vías en nuestro planeta. Ahora bien, los registros de puntos GPS al igual que la información en GIS pueden presentar diferentes errores. La red vial, por ejemplo, se encuentra en un constante dinamismo debido a la urbanización de sectores, cambios de sentido, y/o nuevas rutas. Por otro lado, la medición GPS no es del todo exacta, y se ve mayormente afectada en zonas urbanas donde el receptor se encuentra rodeado por muchos factores que generan interferencia. Como consecuencia, la asignación del punto GPS a una red digital de mapas puede quedar asociada a un tramo incorrecto de la vía, y es donde surge el problema de asignación de mapas (Map-Matching) o ambigüedad espacial [3]. Para resolver este problema, existen múltiples investigaciones con diferentes algoritmos, con el objetivo de integrar las mediciones GPS en un mapa digital, asignando los registros GPS a las rutas correctas por donde transitan los vehículos. En esta investigación, se opta por trabajar con los algoritmos de asignación de mapas topológicos (Topological Map-Matching Algorithms, TMMA), en virtud a la rápida implementación y modalidad de uso gracias a la utilización de información topológica (como la conectividad de enlaces viales, por ejemplo). Además, se utilizará un algoritmo ya implementado por los investigadores León y Blazquez en [8], el cual está desarrollado en Python con librerías de ArcGIS las cuales ayudan con la información topológica en el ruteo y asignación de puntos GPS. Este algoritmo tiene como entrada los siguientes parámetros: el radio de búsqueda y la tolerancia de velocidad. La principal motivación de esta investigación es analizar el TMMA y desarrollar una versión mejorada que asegure buenos resultados evitando el proceso de pruebas iterativas, y así obtener el resultado en un menor tiempo computacional.

### 1.1. Objetivo general

El objetivo general es analizar un algoritmo de ambigüedad espacial topológico ya desarrollado e implementar una versión mejorada que reduzca la cantidad de ejecuciones de evaluación asegurando buenos resultados en un menor tiempo computacional.

### 1.2. Objetivos específicos

Los objetivos específicos son los siguientes:

1. Analizar, depurar y definir los datos de entrada.
2. Evaluar el estado y rendimiento del algoritmo actual con los datos de entrada.
3. Proponer e implementar propuestas de mejora algorítmica.
4. Comparar y analizar los resultados con el algoritmo original y con las mejoras propuestas.

## Revisión Bibliográfica

A medida que avanza la utilización de dispositivos GPS en la cotidianidad, para el guiado de rutas, administración de flotas, o información de tráfico por ejemplo, el rubro académico se encuentra en constante desarrollo e investigación para poder resolver de mejor manera la problemática que se presenta en la asignación de los vehículos en las vías de un mapa. En el año 1973 French [5] presenta una investigación que utilizaba un sistema de guiado de ruta que con la ayuda de un algoritmo de asignación de mapas se ayudaba a obtener una mayor precisión en la navegación. Similar a lo que en el año 1996 ya se habla del problema de Map Matching para la asistencia en navegación [1], lo que hoy en día todos conocemos y

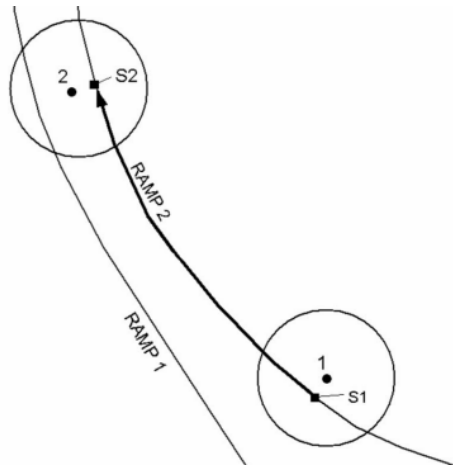


FIGURA 2.1. Asignación perpendicular más cercana.

utilizamos en aplicaciones como Google Maps o Waze. En esta investigación principalmente emplean técnicas geométricas y topológicas para poder asignar correctamente el registro GPS a la calle, proponiendo como sugerencia incorporar mayor información topológica para poder ayudar al algoritmo a decidir de mejor manera la proyección del punto a la calle más cercana. Por consiguiente, ya en 2005 [4], se trabaja con una técnica de resolución que incorpora información topológica para la asignación del punto GPS a la red vial con menor distancia

perpendicular dentro de un radio de búsqueda acotado [Figura 2.1]. Además, se verifica la factibilidad de conectividad por medio de un algoritmo de camino mínimo, entre los puntos asignados. Para el caso de que no sea factible la conexión, se itera con un punto subsiguiente dentro de un rango de limitación para no caer en una búsqueda interminable.

Posteriormente en la siguiente investigación [2], además de incluir ya la información mencionada, se agregan nuevas reglas de decisión para poder ayudar a seleccionar de mejor manera con el algoritmo usado, y se incorpora un área de búsqueda dinámica para encontrar una calle cercana al punto, la cual aumenta en caso de que no se encuentre ninguna calle dentro del radio de búsqueda. Otra propuesta de ejecución en modo post-proceso es la que presentan en [7], donde los autores proponen un algoritmo encargado de buscar las calles más cercanas de acuerdo con una alineación entre puntos, para que el radio de búsqueda sea proporcional en dirección a cómo se mueve el vehículo [Figura 2.2], más como una elipse que se va formando entre dos puntos, y no una circunferencia en cada registro, como se realiza habitualmente en otros estudios. El objetivo es poder reducir la cantidad de calles candidatas y enfocar la búsqueda

por el sector donde se movía el vehículo reduciendo la cantidad de cálculos. Recientemente, en [8] se incorpora una mejora destacable que consiste en que los puntos GPS que no lograron ser asociados a ningún camino, sean forzados a asignarse a la ruta que se en-

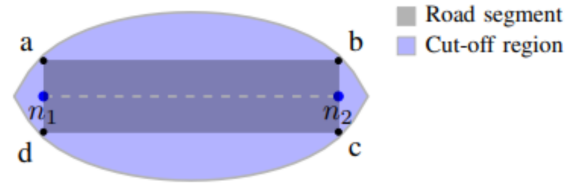


FIGURA 2.2. Radio entre los puntos  $n_1$  y  $n_2$ .

contraba entre puntos previos y posteriores, para no detenerse en la búsqueda iterativa de encontrar la proyección correcta dentro del radio de búsqueda.

## Capítulo 3

### Marco Teórico

Los sistemas GPS son los encargados de posicionar cualquier ente (vehículos, personas, mascotas, por ejemplo) en la Tierra que tenga un receptor encargado de recibir señales electromagnéticas que emiten diferentes satélites. Estas señales se pueden ver interferidas por diferentes factores que alteren la clara comunicación entre emisor y receptor [Figura 3.1], lo que genera un error en el posicionamiento y el registro de la persona en este caso queda incorrecto.

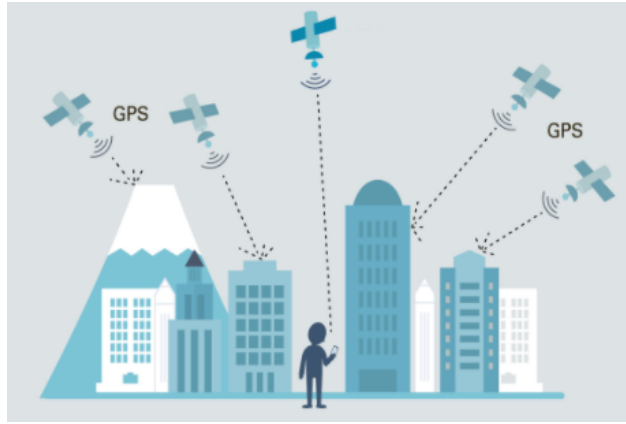


FIGURA 3.1. Interferencia en Sistemas GPS [[6]].

### 3.1. Map Matching

Un algoritmo de Map Matching, o de ambigüedad espacial, se encarga de asignar las mediciones GPS a la red vial, la cual se encuentra en un sistema de información geográfica (GIS). Para realizar esta asignación existen múltiples algoritmos y técnicas de Map-Matching (MMA) que se pueden encontrar en la literatura, y se dividen principalmente en las presentadas a continuación:

**3.1.1. Geométricas.** Esta asignación consiste principalmente en utilizar métodos geométricos para proyectar el punto a la calle más cercana en el mapa, lo que es simple y rápido. Sin embargo en la práctica y llevándolo a la vida real, conlleva una serie de inconvenientes,



ya que las redes viales tienen diseños muy delicados, como por ejemplo las restricciones de tránsito. Por lo que no siempre la calle más cercana es la proyección correcta.

**3.1.2. Topológicos.** Esta técnica se apoya principalmente en la información topológica de la red vial, como la conectividad entre segmentos, sentido de las vías, entre otros. Esto ayuda a determinar de mejor manera la proyección correcta del punto con la red vial y no se basa solo en la distancia entre el punto y la calle.

**3.1.3. Probabilísticos.** El algoritmo probabilístico en la mayoría de los casos utiliza el radio de error que presenta el dispositivo GPS, desde la posición obtenida. Con este atributo se evalúa la proximidad de los segmentos y de acuerdo con los criterios configurados, se asigna el punto en el mapa. Por ejemplo, en el estudio realizado en [9], el algoritmo probabilístico toma en cuenta la imprecisión del dispositivo cuando el vehículo circula a menor velocidad, lo que le ayuda a tener mejores proyecciones a baja velocidad, ideal para sectores urbanos donde las detenciones son más frecuentes.

**3.1.4. Avanzados.** La técnica de asignación de mapas avanzados, son algoritmos que utilizan conocimiento más delicado, los cuales pueden estimar por donde debería pasar un vehículo de acuerdo al patrón que tiene, por ejemplo. De hecho existen algoritmos que además de aplicarse solo en los puntos que se están asignando, al final repiten la asignación para toda la ruta para poder reducir el error general.

### 3.2. Algoritmo Map-Matching Topológico (TMMA)

Para abordar esta investigación se utilizó un algoritmo TMMA existente desarrollado por León y Blazquez en [8], el cual utiliza un radio de búsqueda en cada punto GPS para seleccionar las calles candidatas para proyectar el punto. Luego, se obtiene la distancia mínima entre pares de puntos, y se calcula la velocidad de viaje de acuerdo con la distancia entre ellos y el tiempo de viaje, a partir de la diferencia de tiempo registrada entre ambos puntos. Posteriormente, se calcula la velocidad promedio del vehículo en relación con la velocidad en ambos registros. Si esta velocidad promedio es coherente con la velocidad calculada dentro de una tolerancia, entonces la asignación es aceptada. En caso contrario, se selecciona otro punto candidato, y se repite el procedimiento hasta que sea válido. Si ninguna ruta es viable, entonces el algoritmo itera con puntos previos o posteriores hasta alcanzar un número máximo predefinido para que no quede iterando sin detenerse. Si aún así no encuentra la asignación correcta, entonces el punto GPS que está analizando es forzado a la ruta más cercana generada entre puntos precedentes y posteriores [Figura 3.2].

En el estudio actual, se realizaron principalmente dos modificaciones significativas al algoritmo para optimizar el proceso de map-matching. Inicialmente, se encarga de seleccionar un radio de búsqueda óptimo para la serie evaluada [Figura 3.3a], a partir de un valor pequeño que va aumentando hasta que el 99 % de los puntos GPS encuentre una calle cercana en el mapa, o hasta que el radio supere una cota superior predefinida en caso que algún punto esté fuera de la red vial. En segundo lugar, se incluyó un cruce entre el sentido que lleva el vehículo (heading) y la dirección que tienen las calles candidatas cercanas, para poder filtrar las cuales no tengan coherencia con el trayecto del vehículo y reducir la cantidad de candidatos para el ruteo [Figura 3.3b].

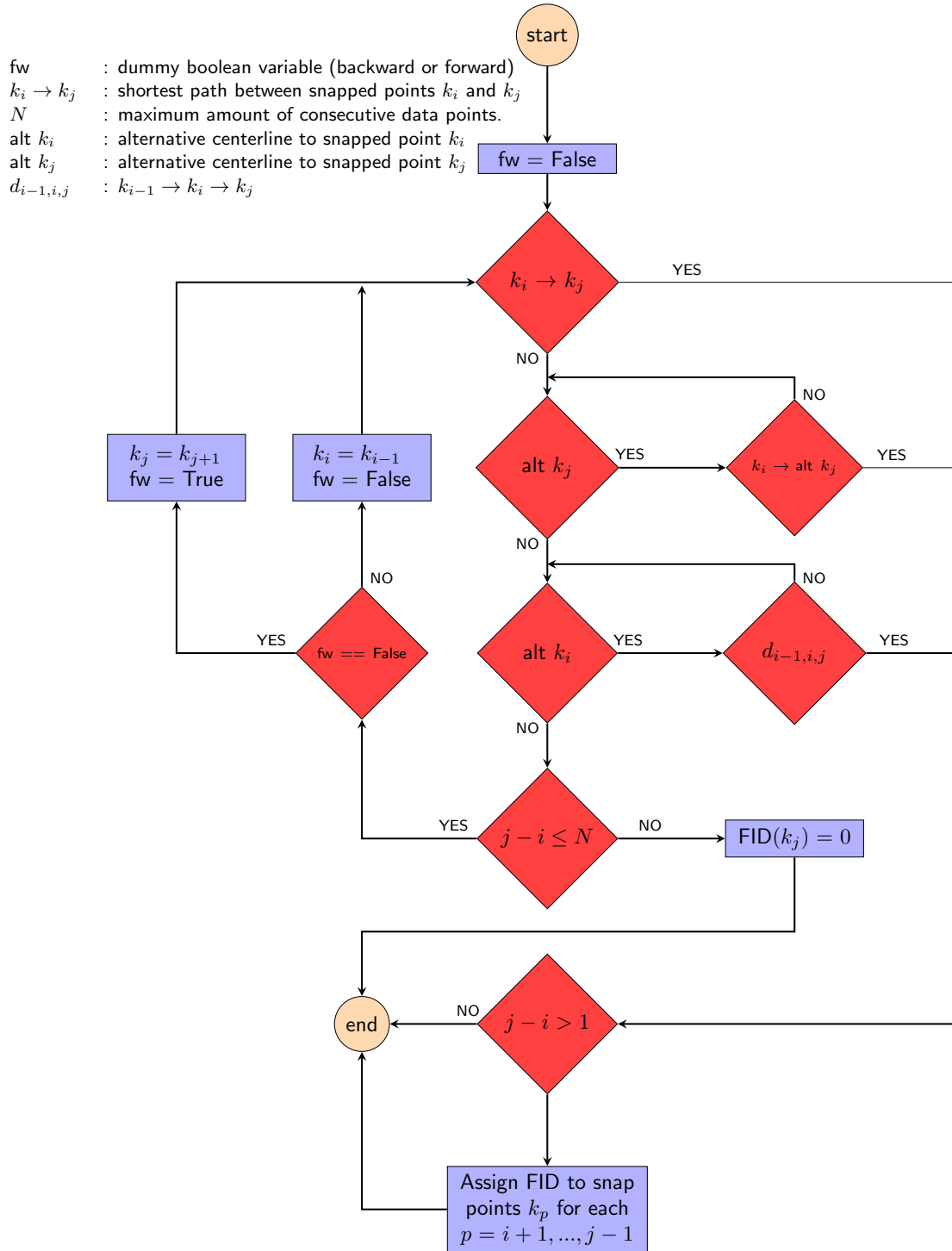
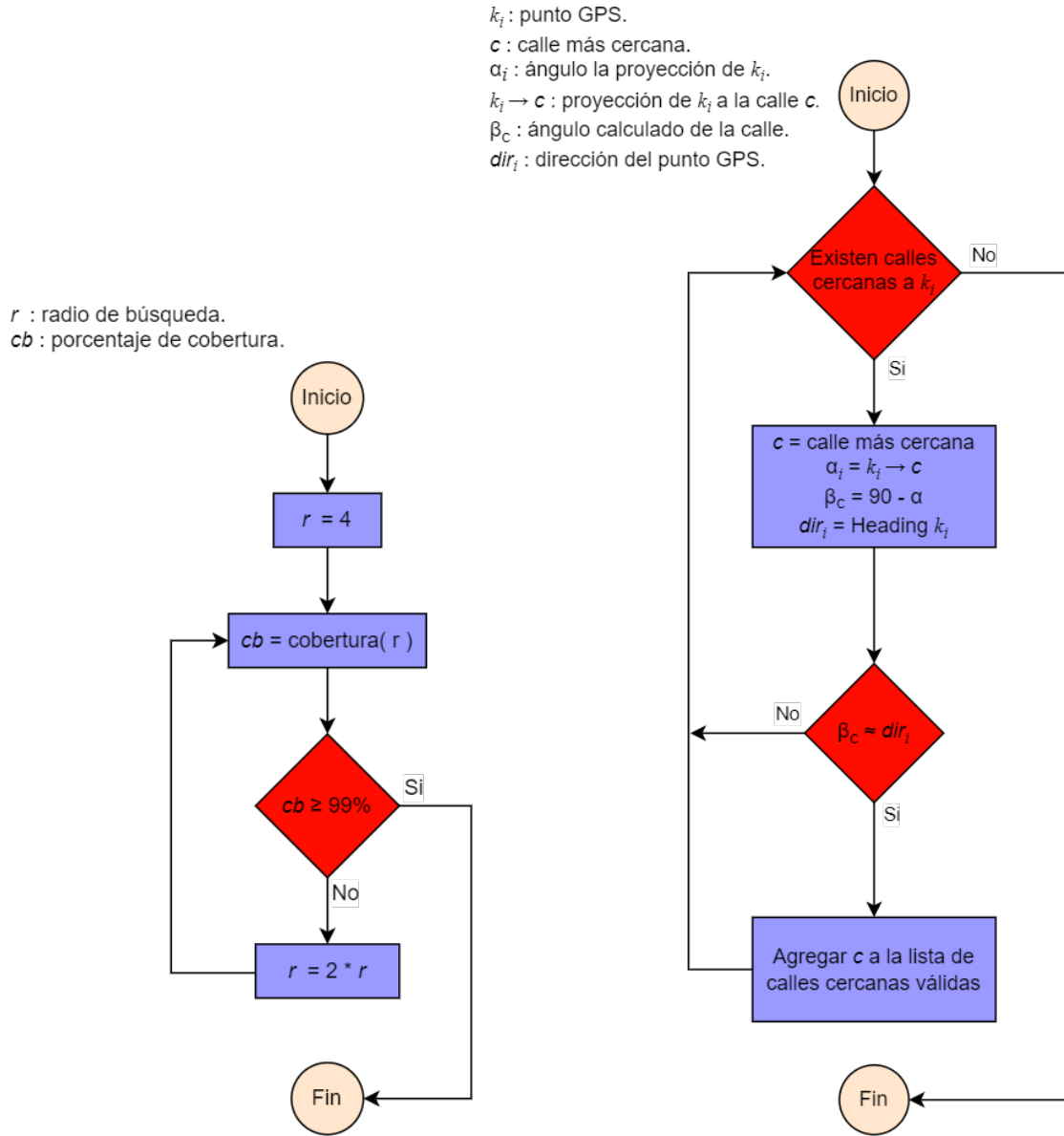


FIGURA 3.2. Diagrama de flujo TMMA.



(A) Diagrama obtención radio óptimo.

(B) Diagrama filtrado por Heading.

FIGURA 3.3. Diagramas de las mejoras realizadas.



## Capítulo 4

### Metodología

Para llevar a cabo este trabajo de investigación, se realizó un análisis y filtrado de datos de vehículos de transporte, teniendo como entrada 30 millones de registros mayormente entre fines de 2014 y comienzo del 2015 [Figura 4.1], con la siguiente información:

- Datetime: Fecha y hora en un mismo campo del registro GPS.
- Fecha: Día del registro.
- Hora: Hora del registro.
- Móvil: Identificación del vehículo.
- Latitud: localización del vehículo, en el eje Y (coordenadas geográficas).
- Longitud: localización del vehículo, en el eje X (coordenadas geográficas).
- Velocidad: Velocidad a la que se mueve el vehículo.
- Dirección (heading): Ángulo acimut en el que se dirige el vehículo, en relación al norte.



FIGURA 4.1. Cantidad de registros dentro del periodo de la muestra.

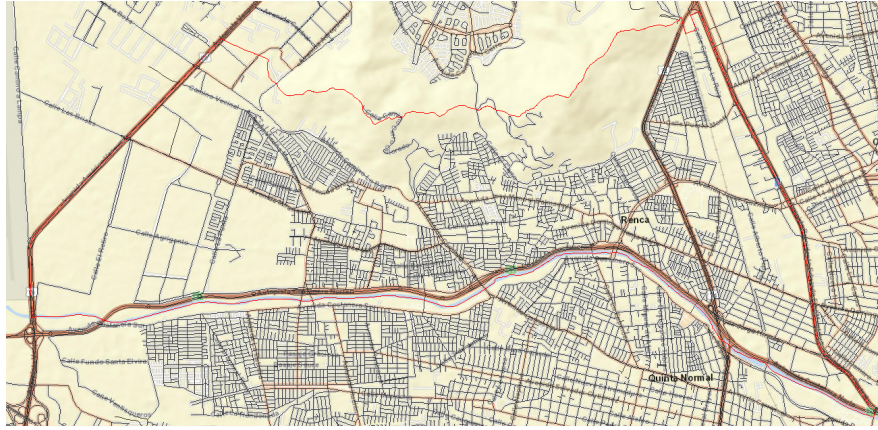


FIGURA 4.2. Red Vial comuna de Renca.

Por otro lado, se cuenta con la red vial de la comuna de Renca [Figura 4.2], con información relevante de sus diferentes vías, la cual fue utilizada para este estudio. Luego se realizó un cruce entre la data de entrada y la red vial de Renca, analizando las coordenadas del registro GPS, si se encontraba dentro del cuadrante de la red vial, quedando tan solo 1.49 millones de registros, de los cuales se separaron por vehículo y fecha, y se seleccionaron los que tenían mayor cantidad de registros consecutivos. Finalmente, para abordar el proyecto se seleccionó un set de datos acotado a ocho rutas que pasan por la comuna de Renca, con frecuencia de 30 segundos. En la Tabla 4.1 se puede ver el detalle de cada set, y en la Figura 4.4, se puede observar el mapa por donde pasan estas rutas.

Set	Nombre	Móvil	Fecha	Cant. puntos
(1)	T386_20150207_2302_2341	386	07-02-2015	40
(2)	T386_20150207_2793_2819	386	07-02-2015	27
(3)	T386_20150212_1389_1466	386	12-02-2015	78
(4)	T386_20150212_1712_1767	386	12-02-2015	55
(5)	T386_20150212_2640_2693	386	12-02-2015	54
(6)	T916_20150207_2161_2181	916	07-02-2015	21
(7)	T916_20150208_2161_2196	916	08-02-2015	32
(8)	T916_20150208_2241_2271	916	08-02-2015	31

TABLA 4.1. Set de datos utilizados.



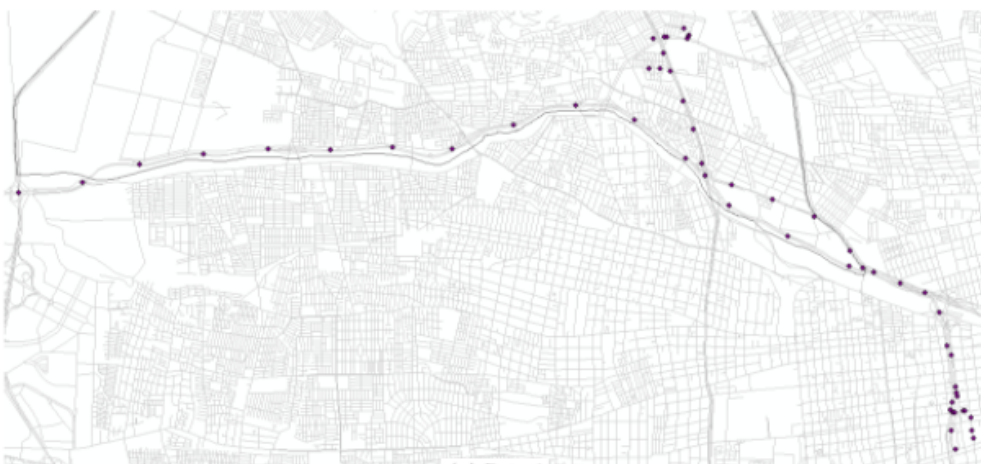
(A) Ruta 1



(B) Ruta 2



(C) Ruta 3



(D) Ruta 4



(E) Ruta 5



(F) Ruta 6



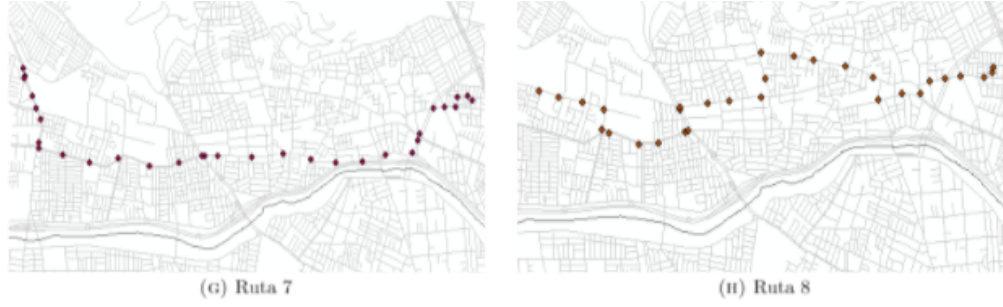


FIGURA 4.4. Rutas seleccionadas.

Para tener mayor contexto de los datos utilizados y entender el comportamiento de los vehículos se realizó una serie de análisis. Inicialmente, se realizó un muestreo de cómo se distribuye la velocidad y heading en los diferentes sets de datos, como se puede observar en [Figura 4.5]. En segundo lugar, se obtuvo la información de las vías por las cuales se movía el vehículo y en la Tabla 4.2 se puede apreciar el porcentaje de las vías utilizadas.

Con estas series, se ejecutaron múltiples pruebas en el algoritmo original con el objetivo de poder contrarrestar los resultados obtenidos posteriormente una vez aplicadas las mejoras propuestas. En esta prueba, se utiliza como parámetros de entrada una Tolerancia de Velocidad de 16 km/h, y un radio (buffer) de 4, 8, 16, 32 y 64 metros. Para obtener los diferentes tiempos del algoritmo original, a medida que se modificaba el radio, la Tolerancia de Velocidad quedó fija para no interferir en el análisis. Luego, se realizó una evaluación de sensibilidad que contemplaba la modificación de la Tolerancia de Velocidad.

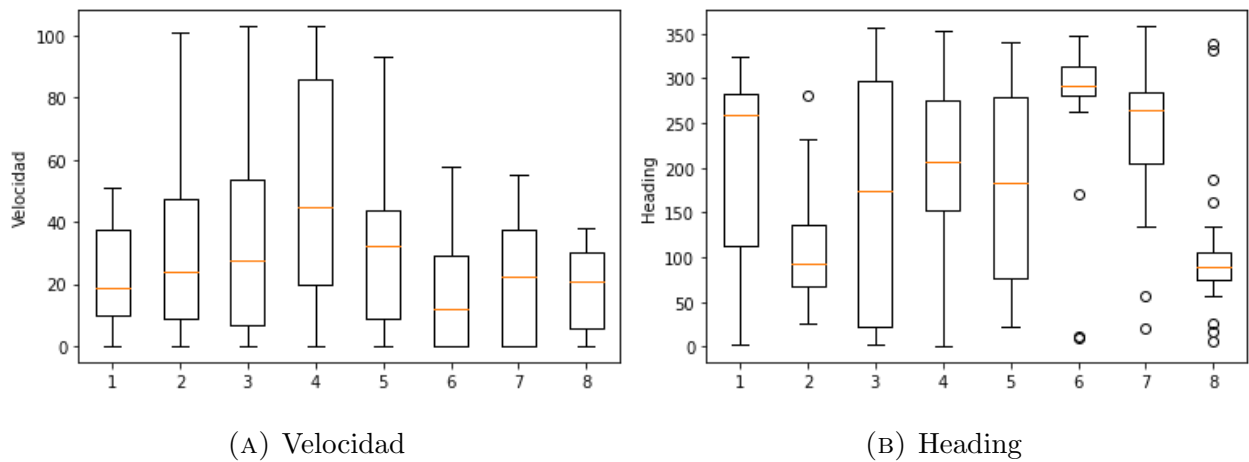


FIGURA 4.5. Distribución de los set de datos seleccionados.

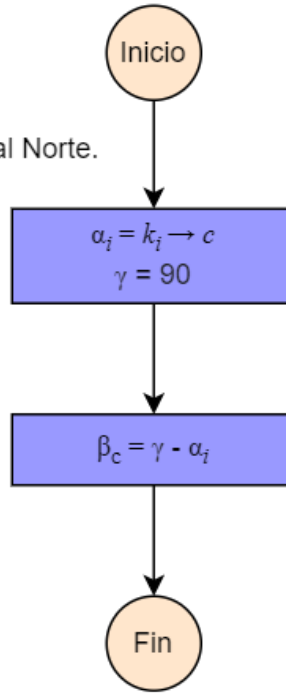
Set	Tipo Vía	Porcentaje	Set	Tipo Vía	Porcentaje
(1)	CALLE	50 %	(5)	AVDA	60 %
(1)	AVDA	39 %	(5)	CALLE	32 %
(1)	PSJE	11 %	(5)	PSJE	8 %
(2)	AVDA	56 %	(6)	AVDA	50 %
(2)	PSJE	33 %	(6)	CALLE	38 %
(2)	CALLE	11 %	(6)	PSJE	13 %
(3)	AVDA	61 %	(7)	AVDA	54 %
(3)	CALLE	35 %	(7)	CALLE	33 %
(3)	PSJE	4 %	(7)	PSJE	13 %
(4)	CALLE	52 %	(8)	CALLE	57 %
(4)	AVDA	39 %	(8)	AVDA	39 %
(4)	PTE	4 %	(8)	PSJE	4 %
(4)	PSJE	4 %			

TABLA 4.2. Porcentaje de vías utilizadas.

Después se desarrolló un pequeño algoritmo encargado de buscar un radio óptimo de selección de calles más cercanas para trabajar con un radio acotado que pueda asegurar que el 99 % de los puntos GPS tenga una calle cercana y con ello poder realizar ruteo entre todos los puntos. Por último, se aplica un filtro en el algoritmo, encargado de que al momento de seleccionar las calles cercanas, sólo incluya las que van en la misma dirección (con cierta tolerancia) en la que se dirige el vehículo, es decir, que si en una intersección tiene como calle cercana una calle perpendicular, esta no será considerada como una calle candidata por donde transita el vehículo. Para obtener el ángulo acimut de la calle, se realiza un cálculo por medio del ángulo de proyección que genera el punto GPS a la calle que se está evaluando, y como la proyección de punto a la calles es perpendicular, el acimut de la calle será la diferencia entre el ángulo de proyección y  $90^\circ$  que es el ángulo que se genera en la proyección perpendicular. Como se observa en la Figura 4.6, el vehículo (punto azul) lleva un heading de  $30^\circ$  y genera un ángulo de proyección a la calle ( $\alpha$ ) de  $119.3^\circ$ . Por consiguiente, al calcular la diferencia

entre este ángulo de proyección y la perpendicular ( $\gamma$ ) de  $90^\circ$ , se obtiene que la calle tiene una dirección ( $\beta$ ) de  $29.3^\circ$ .

$k_i \rightarrow c$  : proyección de  $k_i$  a la calle  $c$ .  
 $k_i$  : punto GPS.  
 $c$  : calle más cercana.  
 $\alpha_i$  : ángulo de proyección en relación al Norte.  
 $\beta_c$  : ángulo calculado de la calle.  
 $\gamma$  : proyección perpendicular de  $90^\circ$ .



(A) Diagrama obtención ángulo calle.



(B) Representación cálculo.

FIGURA 4.6. Diagrama y representación del cálculo del ángulo de la calle.

Una vez aplicadas estas modificaciones al algoritmo original, se realiza la ejecución de este algoritmo mejorado utilizando los sets de datos mencionados anteriormente con el objetivo de poder comparar los resultados obtenidos por el algoritmo original. Además, se incluyeron diferentes contadores unitarios y de tiempo, los cuales iban aumentando a medida que el algoritmo avanzaba su ejecución. El objetivo era contabilizar la cantidad de veces que se accedía a las diferentes funciones y medir el tiempo que tomaba el paso por dicha función. En último lugar, todos estos contadores fueron expuestos y almacenados para poder analizar con claridad el rendimiento y comportamiento del algoritmo.

Finalmente, se realizó un análisis de sensibilidad para las variables de Tolerancia de Velocidad y Tolerancia de Heading, para poder comprobar como influían en el rendimiento

del algoritmo mejorado. Asimismo tener una otra mirada del panorama que sirva para dar mayor claridad de como poder utilizar esta información en una investigación posterior que ayude a parametrizar dichas variables, tal como se realizó con el radio de búsqueda.



## Capítulo 5

### Resultados

Como se comenta anteriormente, la primera prueba realizada fue ejecutar el algoritmo original con el set de datos seleccionados, utilizando distintos radios con el objetivo de poder analizar posteriormente la mejora realizada. Se utilizó una tolerancia de velocidad fija de 16 km/h para no interferir en el análisis del radio. Luego, se ejecutó el nuevo algoritmo mejorado con todos los sets de datos para obtener los radios propuestos, que se puede ver en la primera columna de la Tabla 5.1, en la cual se encuentran los porcentajes de acierto que tuvo el algoritmo original con los diferentes radios que se ejecutaron las pruebas. Se puede observar, que en la mayoría de los casos, el radio propuesto por el nuevo algoritmo fue el que mejor se comportó para la prueba realizada con el algoritmo original.

r	Set	r4	r8	r16	r32	r64
8	(1)	95 %	<b>90 %</b>	90 %	85 %	75 %
8	(2)	56 %	<b>96 %</b>	96 %	96 %	74 %
16	(3)	56 %	83 %	<b>81 %</b>	78 %	69 %
16	(4)	29 %	65 %	<b>89 %</b>	80 %	73 %
32	(5)	76 %	87 %	83 %	<b>85 %</b>	78 %
16	(6)	10 %	24 %	<b>71 %</b>	71 %	67 %
16	(7)	25 %	66 %	<b>75 %</b>	75 %	66 %
16	(8)	52 %	87 %	<b>87 %</b>	87 %	81 %

TABLA 5.1. Alg. original:  
Porcentaje de acierto.

r4	r8	r16	r32	r64
1,54	<b>2,39</b>	3,57	2,33	3,41
1,31	<b>3,42</b>	5,51	3,08	3,35
3,45	38,20	<b>11,04</b>	9,38	18,32
2,14	5,14	<b>9,57</b>	6,38	9,57
3,43	5,60	7,37	<b>5,30</b>	7,30
1,36	4,23	<b>5,49</b>	4,52	8,05
1,25	3,56	<b>4,12</b>	3,15	4,03
1,23	2,18	<b>4,27</b>	2,24	2,52

TABLA 5.2. Alg. original:  
Tiempo computacional (min).

Esto se aprecia de mejor manera en la Figura 5.1, donde se distingue que a medida que el radio comienza a aumentar, el porcentaje de acierto del algoritmo igualmente aumenta.

Sin embargo, este aumento encuentra una cota superior debido a que el radio no entrega un mejores resultados, y luego en la mayoría de los casos comienza a disminuir su acierto.

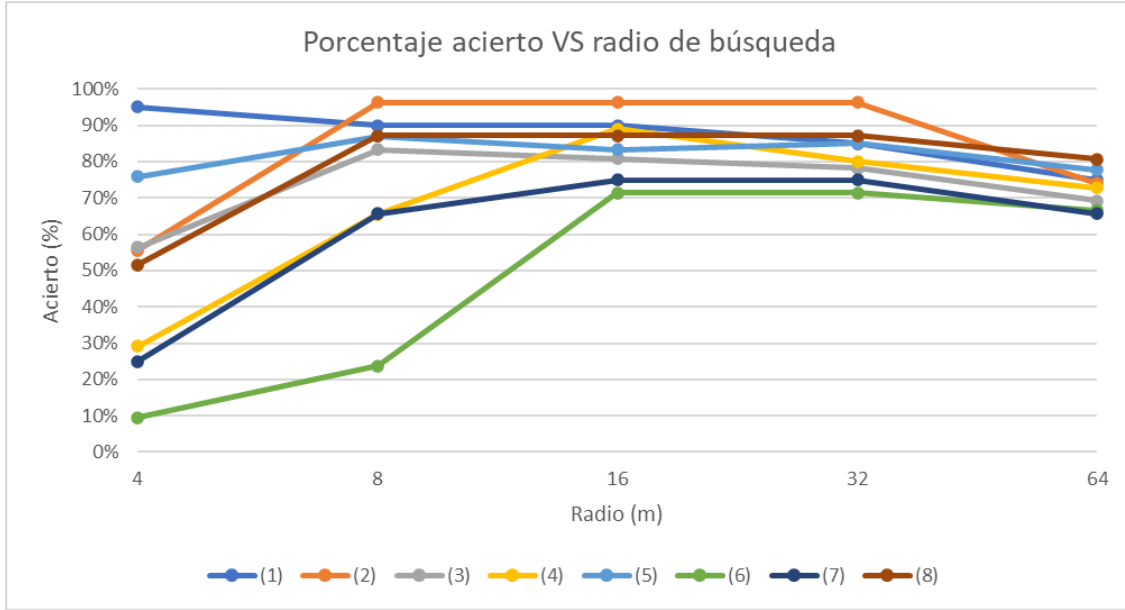


FIGURA 5.1. Algoritmo original: Porcentaje Acierto / Radios.

Lo que lleva a entender que al utilizar un radio muy grande, aumenta la cantidad de calles candidatas y el algoritmo deberá evaluar más opciones de caminos posibles por donde se movió el vehículo. A diferencia del radio adecuado con el que se debe trabajar, ya que con ello el algoritmo se encontrará acotado a trabajar con las calles que tengan mayor probabilidad de seleccionar la calle correcta por la cual se movía el vehículo.

Por último, analizando la ejecución del TMMA original, en la Tabla 5.2, se puede ver el tiempo computacional (en minutos) que tomó resolver el problema Map Matching. En ella, se percibe que el tiempo obtenido, para la mayoría de los casos, es el mejor dentro de la prueba. Que si bien existen casos donde hay tiempos menores, al mirar en la Tabla 5.1 se puede ver que el porcentaje de acierto es menor.

Continuando con el estudio, se realizó la ejecución del algoritmo mejorado con todos los set de datos. En esta ejecución, al igual que con el algoritmo original se utilizó una tolerancia de velocidad de 16 km/h para comparar los algoritmos sin interferencia. En la Figura 5.2, se puede observar el contraste entre ambos algoritmos, donde se logra apreciar que el nuevo

algoritmo mejorado, con los ocho sets de datos, siete obtuvieron igual o mejor porcentaje de acierto que el algoritmo original, y tan solo uno obtuvo menor porcentaje de acierto.

Set	Acierto (%)		Tiempo (min)	
	Original	Mejorado	Original	Mejorado
(1)	90 %	98 %	2,39	4,08
(2)	96 %	96 %	3,42	3,25
(3)	81 %	87 %	11,04	9,56
(4)	89 %	84 %	9,57	7,30
(5)	85 %	89 %	5,30	6,48
(6)	71 %	71 %	5,49	3,27
(7)	75 %	78 %	4,12	3,47
(8)	87 %	94 %	4,27	3,45

TABLA 5.3. Comparativa Acierto/Tiempo: Alg. Original V/S Mejorado.

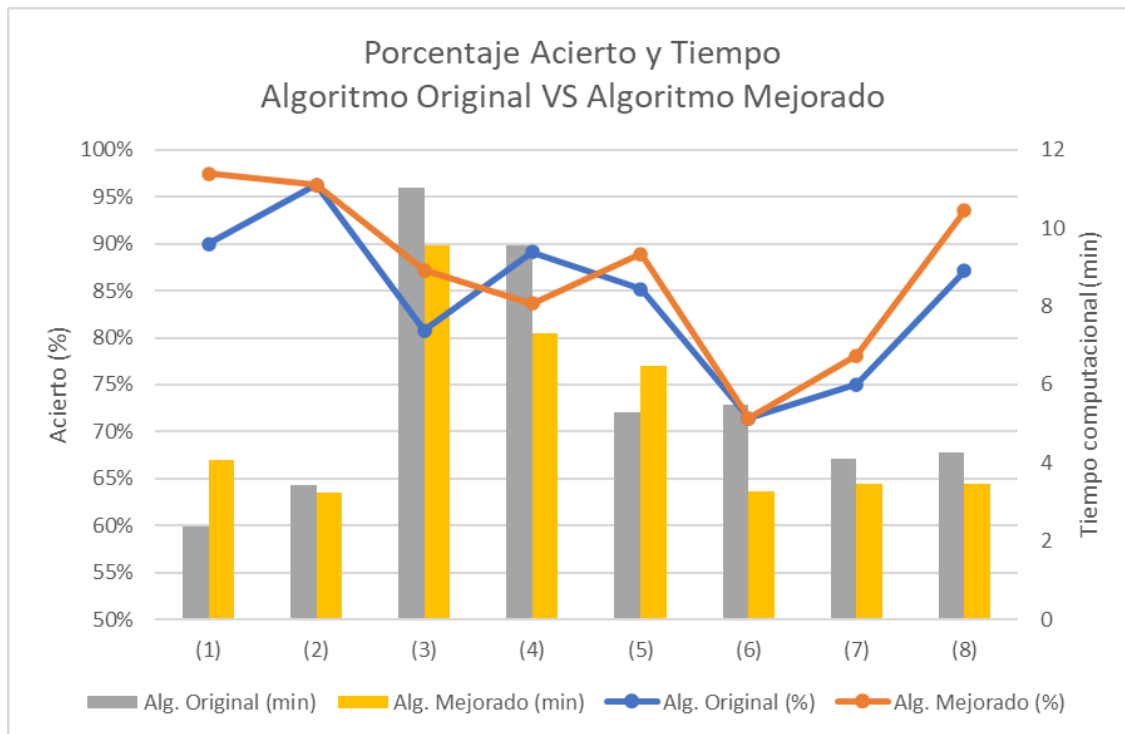


FIGURA 5.2. Comparativa Acierto/Tiempo: Alg. Original V/S Mejorado.



Igualmente en la Figura 5.2 y Tabla 5.3, se puede observar el tiempo de ejecución de ambos algoritmos, donde en la mayoría de los casos este disminuye para la versión mejorada. Tan solo en dos casos, el tiempo se ve que aumenta, pero se logra visualizar que su porcentaje de acierto de igual forma aumentó. Se puede entender que el filtrado por el heading ayuda a reducir la cantidad de calles candidatas y el algoritmo debe evaluar menos calles posibles por donde pasó el vehículo, esto se aprecia de mejor manera en el posterior análisis de las funciones realizadas por el algoritmo.



FIGURA 5.3. Resultado de proyección para el registro 66 (Serie 3) por el Alg. Original y Alg. Mejorado.

A continuación, podemos observar con un ejemplo gráfico [Figura 5.3] de como los algoritmos entregan una proyección de un registro GPS, punto 66 de la Serie 3 (punto rojo). El cual indica que el vehículo se movilizaba por la calle 51721, justo debajo de él. La proyección entregada por el Algoritmo Mejorado, representado por un punto verde, fue asignado correctamente a la calle 51721. A diferencia de la proyección que generó el Algoritmo Original, ya que fue incorrectamente asociado a la calle 51613. Esto nos deja un poco más claro, el por qué en la Figura 5.2 la serie 3 obtuvo mejor porcentaje de acierto en el Algoritmo Mejorado, que en el Algoritmo Original. Para este caso en particular, podemos observar que el filtrado de calle por Heading fue de gran ayuda para seleccionar la calle correcta por donde circulaba el vehículo.

Por otra parte, se realizó un conteo del trabajo realizado por las diferentes funciones que componen el algoritmo para poder entender su comportamiento y ver que funciones eran las que principalmente más trabajo realizaban. Esta evaluación fue aplicada en el algoritmo original y el algoritmo mejorado, para ver si las mejoras realizadas ayudaban a reducir el trabajo realizado, lo que influiría directamente en la disminución del tiempo de ejecución. Específicamente, la función que genera el ruteo entre dos puntos es directamente proporcional al radio de búsqueda utilizado. En otras palabras, si se utiliza un radio mayor (como por ejemplo el utilizado de 64 metros), tendrá muchas calles candidatas con las cuales debe realizar ruteo. Por esta razón, la selección de un radio óptimo y acotado, en conjunto con el filtrado por medio del heading, son de gran ayuda para poder reducir la cantidad de calles candidatas, y a su vez reducir la cantidad de veces que el algoritmo debe calcular la ruta entre dos puntos. En la Tabla 5.4, se puede apreciar cómo aumenta la cantidad de veces que se llama a la función de ruteo en el algoritmo original a medida que aumenta el radio, que comparado con el algoritmo mejorado se ve reducido significativamente contrarrestando con su símil del radio óptimo utilizado (destacado en negrita).

Set	Alg. Original					Alg. Mejorado
	r4	r8	r16	r32	r64	Radio Opt.
(1)	56	<b>74</b>	135	206	317	56 ( <i>r8</i> )
(2)	74	<b>142</b>	202	279	343	61 ( <i>r8</i> )
(3)	184	253	<b>425</b>	900	1914	251 ( <i>r16</i> )
(4)	102	198	<b>450</b>	608	992	254 ( <i>r16</i> )
(5)	190	234	331	<b>499</b>	724	159 ( <i>r32</i> )
(6)	112	217	<b>277</b>	488	842	105 ( <i>r16</i> )
(7)	58	152	<b>166</b>	261	386	71 ( <i>r16</i> )
(8)	51	69	<b>145</b>	190	257	62 ( <i>r16</i> )

TABLA 5.4. Cantidad de ruteo entre pares de puntos.

Como los resultados del algoritmo mejorado presentan un buen comportamiento respecto al algoritmo original, se decidió evaluar qué resultados entregaría si se utilizara un set de datos con menor frecuencia (MF). Para ello, se tomó el set con mayor cantidad de registros (Set 3)

y se quitó un registro por medio, pasando de tener la muestra tomada cada 30 segundos, a una con frecuencia de 1 minuto.

Set	Acierto ( %)		Tiempo (min)	
	Original	Mejorado	Original	Mejorado
(3)	81 %	87 %	11,04	9,56
(3 MF)	79 %	82 %	5,11	4,56

TABLA 5.5. Algoritmo Original V/S Algoritmo Mejorado,  
Serie 3 original y menor frecuencia.

A continuación, se ejecutaron ambos algoritmos con el nuevo set (Set 3 MF) y se compararon con el set original (Set 3). Los resultados se pueden ver en la Tabla 5.5, donde se observa que el tiempo computacional se reduce aproximadamente a la mitad. Esto se puede atribuir a que la cantidad de registros del nuevo set tiene la mitad de puntos, por ende el tiempo es proporcional a la cantidad de registros. Si bien el porcentaje de acierto disminuye en ambos algoritmos, el algoritmo mejorado sigue obteniendo un mejor rendimiento, alcanzando un mejor porcentaje de acierto que el algoritmo original

Además, revisando los log de la ejecución del algoritmo mejorado, y apoyándose con la aplicación de ArcMap para ver los datos gráficamente, se logró identificar que cuando el vehículo se encontraba detenido, o con una velocidad ( $v$ ) inferior a los 5 km/h, la medición del heading entregaba un valor incorrecto. Como por ejemplo en la Figura 5.4, se puede ver en el primer registro de la Serie 1 (punto celeste) con velocidad 0 km/h, que tiene un heading de  $40^\circ$ , pero su calle más cercana tiene una dirección de  $97^\circ$ . Por esta razón, se decidió hacer una prueba del algoritmo mejorado donde se omita el filtro de heading para los registros que tengan una velocidad inferior a 5 km/h con el objetivo de no dejar fuera las calles candidatas por culpa de un heading incorrecto, y así ver si la precisión del algoritmo aumentaba. Los resultados se pueden ver en la [Tabla 5.6], donde se compara este nuevo algoritmo con el algoritmo mejorado.

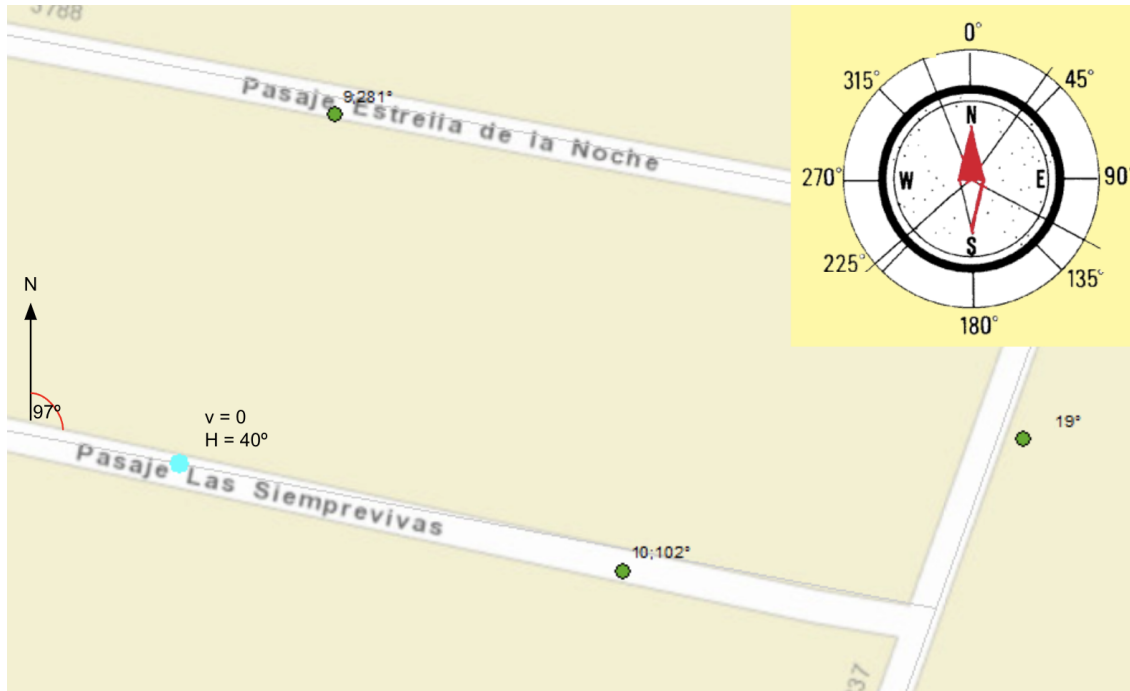


FIGURA 5.4. Registros con velocidad y heading.

Set	Acierto (%)		Tiempo (min)	
	Mejorado	Filtro $v > 5\text{km/h}$	Mejorado	Filtro $v > 5\text{km/h}$
(1)	98 %	98 %	4,08	3,42
(2)	96 %	96 %	3,25	3,32
(3)	87 %	85 %	9,56	9,44
(4)	84 %	85 %	7,30	8,50
(5)	89 %	89 %	6,48	8,04
(6)	71 %	71 %	3,27	4,27
(7)	78 %	78 %	3,47	4,02
(8)	94 %	97 %	3,45	4,18

TABLA 5.6. Algoritmo Mejorado V/S Filtro  $v > 5\text{km/h}$ .

Se puede observar que la modificación realizada no influye positivamente en los resultados, ya que en la mayoría de los set de datos el porcentaje de acierto es el mismo que obtenía el Algoritmo Mejorado. Incluso, para la Serie 3, la modificación impactó negativamente,

reduciendo en un 2% su acierto, ya que al momento de seleccionar una calle por donde se movilizaba el vehículo a una velocidad inferior a los 5 km/h, el algoritmo no pudo seleccionar la calle correcta con la ayuda del filtrado por heading, generando una ruta incorrecta que afectó a la selección posterior. Afirmando que esta variación realmente no aporta de manera efectiva al Algoritmo Mejorado, por lo que no fue considerada.

Finalmente, se realizó un análisis de sensibilidad con la tolerancia de heading y tolerancia de velocidad, para evaluar cómo influyen estas variables en el comportamiento final del algoritmo mejorado. La primera prueba fue con la Tolerancia de Heading (TH), la que influye directamente en el filtrado de calles candidatas por medio del heading que tiene la calle con la dirección que lleva el vehículo, ayudando al algoritmo a reducir la cantidad de calles que deba evaluar por donde realmente se movía el móvil. La prueba se realizó con las siguientes Tolerancias de Heading: TH 10°, TH 20° y TH 30°. Mientras que la Tolerancia de Velocidad se mantuvo constante en 16 km/h para no interferir en el experimento, y el radio de búsqueda utilizado fue el radio encontrado por el algoritmo mejorado para cada serie, como se explicó anteriormente. Esto quiere decir que para la TH 30°, si el vehículo iba con un heading de 45°, se aceptarían calles que tuvieran un ángulo entre 30° y 60°.

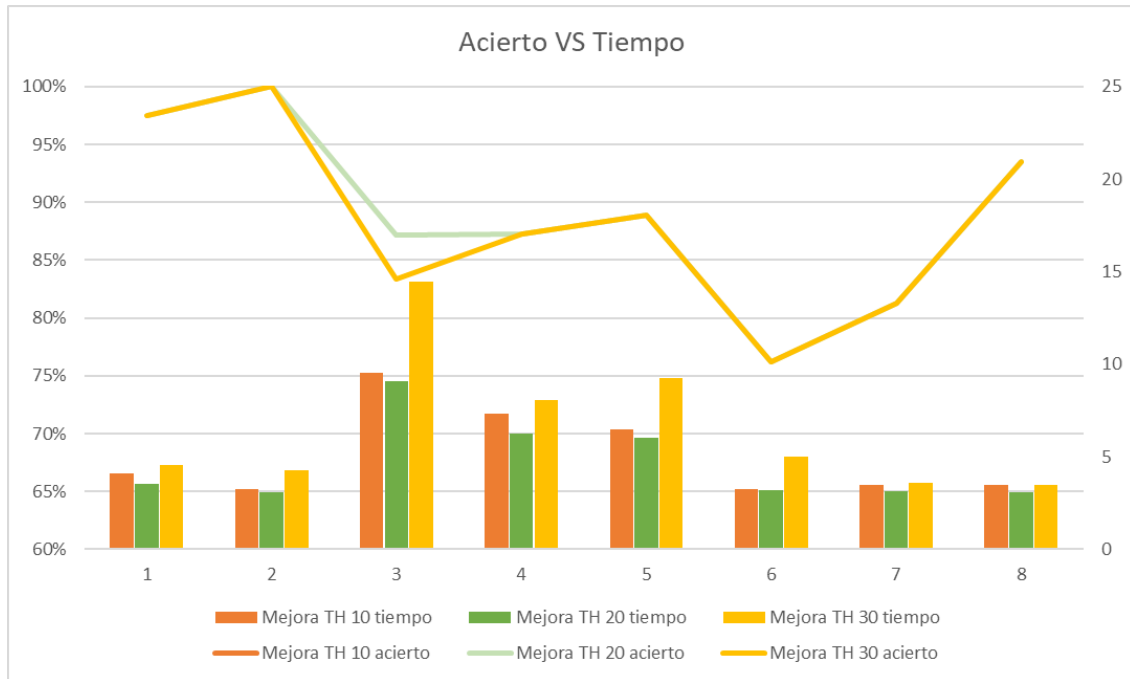


FIGURA 5.5. Alg. Mejorado: TH 10°, TH 20°, TH 30°.

En la Figura 5.5, se puede observar la comparativa de los resultados obtenidos para las tres variaciones de TH en el algoritmo con los ocho set de datos. Mirando el porcentaje de acierto conseguido para las tres TH en [Figura 5.5] y la [Tabla 5.7], tan solo en la serie 3, el algoritmo con TH 30° obtuvo un peor porcentaje de acierto, pero en el resto de sets se obtiene el mismo acierto para TH 10° y TH 20°. En cuanto al tiempo obtenido en cada prueba, podemos comentar que con TH 20° fue donde se obtuvo tiempos menores que el resto de las pruebas para todas las series.

Set	Acierto (%)			Tiempo (min)		
	TH 10°	TH 20°	TH 30°	TH 10°	TH 20°	TH 30°
(1)	98 %	98 %	98 %	4,08	3,55	4,53
(2)	96 %	100 %	100 %	3,25	3,06	4,29
(3)	87 %	87 %	83 %	9,56	9,10	14,48
(4)	84 %	87 %	87 %	7,30	6,27	8,05
(5)	89 %	89 %	89 %	6,48	6,05	9,26
(6)	71 %	76 %	76 %	3,27	3,18	5,03
(7)	78 %	81 %	81 %	3,47	3,14	3,58
(8)	94 %	94 %	94 %	3,45	3,07	3,49

TABLA 5.7. Algoritmo Mejorado: TH 10° V/S TH 20° V/S TH 30°.

Para la segunda evaluación, de la Tolerancia de Velocidad (TV) se efectuó con las siguientes Tolerancias de Velocidad: TV 8 km/h, TV 16 km/h, TV 32 km/h y TV 64 km/h. Esto quiere decir, para ser más claros, que para la primera tolerancia si el vehículo circula a 50 km/h la velocidad permitida calculada por el algoritmo será entre 46 km/h y 54 km/h. Los resultados que se obtuvieron se pueden ver en [Figura 5.6] y [Tabla 5.8], donde a simple vista podemos asegurar que la modificación de estos parámetros influyen directamente en el comportamiento del algoritmo mejorado, como por ejemplo en el caso de la serie 4 se tiene una variación del acierto de un 60 % con TV 8 km/h hasta un 89 % con TV 32 km/h. Continuando con el análisis del porcentaje de acierto, se puede inferir que la Tolerancia de Velocidad que mejor comportamiento tuvo fue TV 32 km/h, ya que en la mayoría de las series obtuvo un mejor porcentaje. Por otro lado, podemos advertir que TV 32 km/h y TV

64 km/h son las que obtienen mejor resultado, con una variación cercana a los 48 segundos en el mayor de los casos entre ellas y mejorando entre 1 y 4 minutos respecto a las Tolerancias de Velocidad TV 8 km/h y TV 16 km/h.

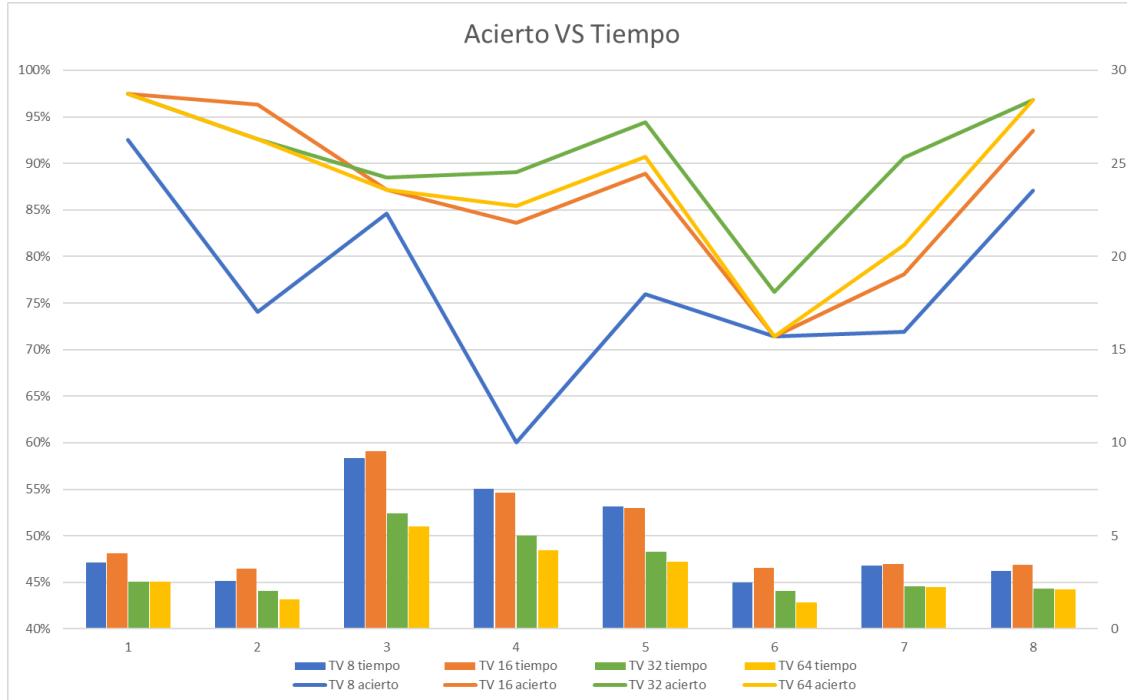


FIGURA 5.6. Alg. Mejorado: TV8km/h, TV16km/h, TV32km/h, TV64km/h.

Set	Acierto (%)				Tiempo (min)			
	TV8	TV16	TV32	TV64	TV8	TV16	TV32	TV64
(1)	93 %	98 %	98 %	98 %	3,58	4,08	2,54	2,54
(2)	74 %	96 %	93 %	93 %	2,56	3,25	2,06	1,59
(3)	85 %	87 %	88 %	87 %	9,19	9,56	6,20	5,51
(4)	60 %	84 %	89 %	85 %	7,54	7,30	5,01	4,22
(5)	76 %	89 %	94 %	91 %	6,58	6,48	4,16	3,58
(6)	71 %	71 %	76 %	71 %	2,47	3,27	2,03	1,41
(7)	72 %	78 %	91 %	81 %	3,42	3,47	2,27	2,23
(8)	87 %	94 %	97 %	97 %	3,11	3,45	2,16	2,13

TABLA 5.8. Alg. Mejorado: TV8km/h, TV16km/h, TV32km/h, TV64km/h.

En definitiva, si miramos un poco más atrás la distribución de los datos y lo comparamos con los resultados obtenidos en este análisis de sensibilidad para ambas variables, se logra apreciar que cuando una serie tiene mayor dispersión, independiente de la variable que se analice, la tolerancia seleccionada toma mayor importancia. En otras palabras, si la distribución de los datos, velocidad o heading, es muy dispersa, la tolerancia seleccionada puede afectar en un 29 % en el acierto del algoritmo (Tolerancia de Velocidad, serie 4), o llegar a demorar un 37 % más (Tolerancia de Heading, serie 3). A diferencia de cuando en la distribución de los datos podemos apreciar que se encuentran más agrupados, la variación de la tolerancia, no refleja grandes cambios en el comportamiento del algoritmo.





## Capítulo 6

### Discusión

Observando de manera general los resultados obtenidos en esta investigación, se pueden sacar ciertas afirmaciones al respecto. En un comienzo se destaca la selección previa del radio óptimo para cada serie evaluada por el Algoritmo Mejorado, asegurando porcentajes de acierto superiores a los obtenidos por el Algoritmo Original, sin la necesidad de tener que disponer de tiempo extra para realizar una evaluación de sensibilidad del radio. Otra característica importante que se distingue en el Algoritmo Mejorado, es el filtrado de calles candidatas por medio del Heading que lleva el vehículo, ya que ayudó notablemente en a reducir la cantidad de ruteo entre un punto y otro, por ende el algoritmo debe realizar un menor trabajo en encontrar la ruta correcta por donde circulaba el transporte.

Por otro lado, los resultados obtenidos de la evaluación de una serie con menor frecuencia se pudo confirmar que el Algoritmo Mejorado seguía obteniendo mejor resultado que el Algoritmo Original. Por lo que se podría evaluar si se quiere perder cierto porcentaje, pero mejorando los tiempos cercanos al 50 %, ya que se logró apreciar que el tiempo que tomaba el algoritmo era proporcional a la cantidad de puntos. Esto quiere decir que si la cantidad de puntos se reduce a la mitad, el tiempo final se verá disminuído a la mitad aproximadamente.

Finalizando con los experimentos realizados, por medio de un análisis de sensibilidad para las variables, tolerancia de velocidad y tolerancia de heading, se puede ver como los resultados obtenidos pueden mejorar aún más si se seleccionan los parámetros adecuados para el algoritmo. Y por otra parte, si miramos la distribución de los datos de entrada podemos encontrar que cuando no existe mucha dispersión, estas variables no influyen demasiado en los resultados obtenidos. Lo que incita a que en una posterior investigación se realice una revisión con mayor profundidad, con el objetivo de lograr que el algoritmo seleccione previamente estas variables asegurando el mejor resultado acierto-tiempo, tal como se hizo con el radio en este Algoritmo Mejorado.



## Conclusiones

A medida que la asignación de mapas va tomando cada vez más importancia para el seguimiento vehicular, es importante contar con algoritmos que tengan un buen comportamiento al momento de resolver el problema.

En esta investigación, se ha presentado una propuesta de mejora a un algoritmo que ya obtenía buenos resultados, con el objetivo de poder reducir el trabajo realizado y por consiguiente obtener resultados en un menor tiempo computacional. Aplicando una mejora de selección previa del parámetro de radio de búsqueda, que permite apoyar de forma correcta la selección del radio a utilizar de acuerdo a la ruta con la que se trabajará. Haciendo una mirada general en los resultados, podemos comprobar que tienen coherencia con lo que propone el nuevo algoritmo versus los resultados obtenidos en las pruebas con el algoritmo original, ya que en su mayoría se obtiene un buen resultado con lo propuesto. Por otro lado, con el filtrado de calles candidatas de acuerdo a su ángulo y el heading del vehículo, se logró reducir significativamente la cantidad de ruteo que realiza el algoritmo, lo que reduce directamente el tiempo que le lleva resolver el problema.

Finalizando con las pruebas realizadas, se logra apreciar que por medio de un muestreo previo de los datos, se puede decidir de mejor manera las variables a utilizar y si la selección de éstas afectará finalmente en el comportamiento del algoritmo. Idealmente, lo que se espera es poder llegar al punto de que la selección de estas variables sea independiente para cada ruta, similar a como se seleccionó el radio de búsqueda, para que el algoritmo pueda entregar un mejor rendimiento, filtrando de mejor manera y por ende obtener un mejor resultado. Por último, sería ideal poder analizar el algoritmo propuesto con otros set de datos de diferente comportamiento, como también con un mayor volumen, lo que sería de gran ayuda para identificar nuevas propuestas de mejoras.



## Bibliografía

- [1] David Bernstein, Alain Kornhauser, et al. An introduction to map matching for personal navigation assistants. Technical report, New Jersey TIDE Center, 1996.
- [2] C Blazquez, P Miranda, and A Ponce. Performance of a new enhanced topological decision-rule map-matching algorithm for transportation applications. *Journal of applied research and technology*, 10(6):929–940, 2012.
- [3] C. Blazquez and A. Vonderohe. Effects of controlling parameters on performance of a decision-rule map-matching algorithm. *Journal of transportation engineering*, 135(12):966–973, 2009.
- [4] Carola A Blazquez and Alan P Vonderohe. Simple map-matching algorithm applied to intelligent winter maintenance vehicle data. *Transportation Research Record*, 1935(1):68–76, 2005.
- [5] Robert L French and Guy M Lang. Automatic route control system. *IEEE Transactions on Vehicular Technology*, 22(2):36–41, 1973.
- [6] E İçen. Küresel ve bölgesel konumlama sistemleri, teknolojileri ve uygulamaları. *Havacılık Ve Uzay Teknolojileri Uzmanlığı Tezi*, 2018.
- [7] Prashant Kumar, Radha Krishna Ganti, and Gaurav Raina. A new technique to find candidate links for map matching for transportation applications. In *2016 8th International Conference on Communication Systems and Networks (COMSNETS)*, pages 1–6. IEEE, 2016.
- [8] R. León, C. Blazquez, and V. Depassier. Improvements of a topological map-matching algorithm in post-processing mode. In *2020 39th International Conference of the Chilean Computer Science Society (SCCC)*, pages 1–5. IEEE, 2020.
- [9] Washington Y Ochieng, Mohammed Quddus, and Robert B Noland. Map-matching in complex urban road networks. *Revista Brasileira de Cartografia*, 55(2), 2003.