

Una instancia específica Enfoque de ajuste de parámetros Uso de lógica difusa para un posprocesamiento topológico Algoritmo de correspondencia de m

Carola Alejandra Blazquez

Universidad Andrés Bello, Viña del Mar, Chile. +56 32 284 5579. Correo electrónico: cblazquez@unab.cl

Jana Ries

Universidad de Portsmouth, Portsmouth, Reino Unido. Correo electrónico: Jana.Ries@port.ac.uk

Roberto Jesús León

Universidad Andrés Bello, Viña del Mar, Chile. Correo electrónico: roberto.leon@unab.cl

Pablo Andrés Miranda

*Escuela de Ingeniería Industrial de la Pontificia Universidad Católica de Valparaíso,
Valparaíso, Chile. Correo electrónico: pablo.miranda@pucv.cl*

XXXXXX

Abstracto — Los algoritmos de coincidencia de mapas (MMA) se desarrollan para resolver las ambigüedades espaciales que surgen en el proceso de asignación de mediciones de GPS a una red de carreteras digitales. Existen escasos enfoques sistemáticos de ajuste de parámetros en la literatura para optimizar el rendimiento de MMA. Por lo tanto, se propone un marco novedoso para una calibración sistemática de los parámetros de un MMA de posprocesamiento. El enfoque de calibración consiste en una estrategia de ajuste de parámetros (IPTS) específica de la instancia que emplea principios de lógica difusa. La herramienta IPTS difusa propuesta determina valores de parámetros específicos del algoritmo basados en información específica de la instancia a priori para la ejecución del MMA. Finalmente, la herramienta IPTS propuesta puede ajustarse a dos preferencias particulares de los tomadores de decisiones sobre el rendimiento del algoritmo, a saber, la calidad de la solución y el tiempo computacional.

I. Introducción

GRAMO Se emplea principalmente para adquirir datos de posicionamiento. La integración de GPS con Sistemas de Información Geográfica (SIG) es parte de la innovadora tecnología avanzada de sistemas de posicionamiento global (GPS) aplicada por Intelligent Transportation Systems (ITS) para mejorar la movilidad y la seguridad, reducir los impactos ambientales y el consumo de energía, y mejorar la calidad de vida general de las personas. Cuando se integran las mediciones de GPS con una red de carreteras en GIS, estas mediciones (representadas como puntos de datos) generalmente se proyectan al segmento de la línea central de la carretera más cercana para determinar la carretera en la que ocurren los eventos e incidentes, se ubican las entidades puntuales o se desplaza un vehículo.

Un problema de coincidencia de mapas (MMP), también conocido como ambigüedad espacial o desajuste espacial, surge cuando las mediciones de GPS se asignan incorrectamente al segmento de la carretera. El MMP se produce debido a complejidades de la red de carreteras, recopilación de datos GPS inadecuada, mapas de carreteras digitales y mediciones GPS inexactos, frecuencia de muestreo inadecuada o combinaciones de estas situaciones [1] - [3]. Por lo tanto, puede parecer que los objetos se encuentran ubicados o se mueven a lo largo de líneas centrales de carreteras incorrectas, particularmente en estructuras de carreteras complejas como rampas, carreteras divididas, pasos superiores e inferiores, carreteras paralelas e intersecciones.

Como resultado, este problema puede influir negativamente en cualquier evaluación, análisis, planificación o toma de decisiones posteriores. La clara necesidad de reducir los incidentes (accidentes), el tiempo de viaje, la congestión y los costos de transporte, al tiempo que se aumenta el servicio al cliente, la productividad, la eficiencia y la seguridad en numerosas aplicaciones de ITS, como la navegación de vehículos y la guía de ruta, la gestión de flotas y la prevención de accidentes, son solo algunas de las principales motivaciones para resolver el MMP determinando las ubicaciones correctas de los vehículos.

Se han desarrollado una variedad de algoritmos de correspondencia de mapas (MMA) para resolver el MMP. Estos MMA emplean parámetros de control o ponderación que determinan el rendimiento algorítmico. Los valores de parámetros inadecuados de estos algoritmos, junto con otras fuentes de errores que pueden surgir del mapa de carreteras digital o de las mediciones del GPS, pueden tener un gran impacto en el rendimiento de MMA. Los métodos de calibración básicos (por ejemplo, prueba y error) se emplean comúnmente para identificar valores de parámetros de buen desempeño antes de ejecutar cualquier tipo de MMA. Los algoritmos deben ejecutarse generalmente para una amplia gama de valores de parámetros, lo que resulta en una amplia experimentación y consumo de recursos. Tradicionalmente, los valores de los parámetros se calibran para los datos recopilados en un entorno ambiental específico, como áreas urbanas o rurales densas [4]. Por eso, el rendimiento informado y los valores de los parámetros de las MMA pueden estar condicionados y sesgados de acuerdo con el conjunto de datos de prueba empleado. Además, existe escasa literatura sobre metodologías de calibración para obtener los mejores valores de parámetros que optimicen

el rendimiento de las MMA. Por ejemplo, Velaga et al [5] determinaron coeficientes de ponderación óptimos para los datos recopilados en áreas rurales que minimizan el error de coincidencia de mapas. En otro estudio, Oran y Jaillet [6] desarrollaron una formulación de peso de proximidad precisa para cualquier tipo de MMA.

El objetivo y principal contribución de esta investigación es desarrollar un enfoque de ajuste de parámetros sistemático basado en principios de lógica difusa para un MMA topológico de posprocesamiento desarrollado por Blazquez et al. [7]. Este enfoque aborda la falta de técnicas de ajuste sistemáticas en la literatura sobre MMA y comprende el comportamiento de rendimiento de las MMA en respuesta a información específica de la instancia.

II. Revisión de antecedentes y literatura

A. Aplicaciones de transporte y MMA

Las MMA se han estudiado ampliamente durante décadas mediante el empleo de diferentes niveles de complejidad que van desde técnicas geométricas simples hasta metodologías complejas y avanzadas de inferencia, filtrado y modelado matemático, como filtros de Kalman, lógica difusa, filtrado de partículas, modelos ocultos de Markov y estadísticas bayesianas [2], [7] - [12]. Los MMA basados en redes topológicas se han utilizado particularmente en aplicaciones ITS debido a su simplicidad, eficiencia, rápida y fácil implementación [7] - [9], [12]. Estos MMA se basan en el cálculo de las rutas más cortas y / o el tiempo de viaje más corto entre pares de puntos GPS coincidentes utilizando la topología de la red de carreteras (es decir, conectividad de enlace y contigüidad) y restricciones de giro. Una ventaja de las MMA topológicas es que proporcionan posiciones lo más precisas posible sin la necesidad de grandes memorias [7], [8], [13] - [16].

Los MMA de posprocesamiento determinan la carretera correcta por la que viaja un usuario o vehículo una vez que la totalidad o parte del conjunto de datos GPS está disponible. En este estudio, un MMA de posprocesamiento se aplica a una instancia que se compone de un conjunto de puntos de datos GPS que representa una ruta completa del vehículo [4], [9]. Por ejemplo, se necesitan MMA de posprocesamiento para obtener las rutas correctas al recopilar datos de viaje multimodal con receptores GPS para el modelado de elección de ruta en un entorno GIS [1], [14], [17] - [19]. Blazquez y Vonderohe [2] emplearon un MMA topológico en modo de posprocesamiento para determinar la línea central de la carretera correcta en la que ha viajado una flota de vehículos de invierno para calcular las medidas de rendimiento para los responsables de la toma de decisiones y los planificadores. Los sistemas automáticos de peaje en carretera basados en el kilometraje total conducido por vehículo equipado con tecnología GPS son otro ejemplo que requiere MMA de posprocesamiento para evitar errores en la tarificación por uso de la carretera [20]. Por lo tanto, la ejecución de MMA en modo de posprocesamiento ha sido ampliamente

estudiado e implementado en aplicaciones de transporte y logística [21].

B. Algoritmo de correspondencia de mapas topológicos

El MMA de posprocesamiento empleado en este estudio resuelve el MMP al determinar la línea central de la carretera correcta en la que viaja un vehículo. Primero, este MMA selecciona todas las líneas centrales de la carretera dentro de un búfer alrededor de cada punto de datos GPS analizado y lo proyecta ortogonalmente a la línea central de la carretera más cercana (también conocido como "ajuste"). Posteriormente, se emplea el algoritmo Dijkstra's para resolver el problema de la ruta más corta entre los puntos previamente ajustados y los actuales utilizando la topología de la red y las restricciones de giro. Una ruta se considera viable y se aceptan ubicaciones de puntos de datos ajustados si la diferencia entre la velocidad de viaje calculada y la velocidad promedio registrada del vehículo para los puntos de datos está dentro de una tolerancia de rango de velocidad. Si se rechaza una ruta, los puntos de datos se ajustan a líneas centrales de carreteras alternativas contenidas en sus zonas de influencia. Los caminos más cortos se vuelven a calcular y las velocidades se comparan una vez más. Si no existen otros candidatos a la línea central de la carretera dentro de las zonas de amortiguamiento o no se obtienen rutas viables, el algoritmo prueba las rutas factibles entre los puntos de datos anteriores y posteriores [13]. Este estudio utiliza los principales parámetros específicos del algoritmo del MMA de posprocesamiento: tamaño de búfer (BZ) y tolerancia de rango de velocidad (ST). BZ se emplea para seleccionar segmentos de carretera como candidatos para ubicaciones de ajuste, y ST proporciona un rango para aceptar o rechazar trayectorias entre un par de puntos de ajuste al comparar velocidades de vehículos registradas y calculadas. Luego, el algoritmo prueba las rutas factibles entre los puntos de datos anteriores y posteriores [13]. Este estudio utiliza los principales parámetros específicos del algoritmo del MMA de posprocesamiento: tamaño de búfer (BZ) y tolerancia de rango de velocidad (ST). BZ se emplea para seleccionar segmentos de carreteras como candidatos para ubicaciones de ajuste, y ST proporciona un rango para aceptar o rechazar trayectorias entre un par de puntos de ajuste al comparar velocidades de vehículos registradas y características de instancia relevantes del problema. Este estudio

La figura 1 muestra el diagrama de flujo del MMA con el diseño de secuencia de pasos. Primero, el MMA calcula la ruta más corta entre pares de puntos de datos GPS ajustados k_i y k_j . Si esta ruta no es factible, entonces el algoritmo obtiene candidatos alternativos a la línea central de la calzada dentro del búfer alrededor del punto de datos k_j . En consecuencia, las rutas factibles se prueban entre puntos de datos k_i y el punto recién roto $alt k_j$. Si no se encuentran caminos viables, entonces candidatos alternativos a la línea central del camino dentro de la zona de influencia alrededor del punto de datos k_i se obtienen, los caminos más cortos se recalculan y las velocidades se comparan entre los puntos vecinos inmediatos k_{i-1} y k_i . La variable booleana ficticia fw ayuda a avanzar hacia atrás o hacia adelante para probar rutas con puntos de datos consecutivos hasta que se resuelva el MMP o *norte*. Se han probado puntos de datos consecutivos.

C. Estrategias de configuración de parámetros

Se sabe que los algoritmos heurísticos funcionan de acuerdo con la calibración de un conjunto de parámetros específicos del algoritmo. Este llamado problema de configuración de parámetros para un algoritmo A busca para determinar los valores de los parámetros $\{^*$ que optimiza el rendimiento algorítmico $f(i, \{^*\})$, donde I es la instancia problemática que se está considerando [22]. Las estrategias para abordar el problema de configuración de parámetros pueden ser

clasificado en *control de parámetros* y *ajuste de parámetros*, en el que la configuración de los valores de los parámetros se integra en el algoritmo o se determina a priori a la ejecución del algoritmo, respectivamente. Las estrategias de control de parámetros (PCS) suelen adoptar la forma de una estrategia adaptativa que permite reevaluar los parámetros {en función del rendimiento actual del algoritmo. Considerando que, las estrategias de ajuste de parámetros (PTS) se basan principalmente en un Diseño de Experimentos (DOE), como en Coy et al. [23], Adenso-Díaz y Laguna [24] y Dobslaw [25] para encontrar la mejor configuración de parámetros antes de la ejecución de un algoritmo.

Los PCS requieren la modificación del propio algoritmo para integrar una estrategia reactiva que generalmente depende del comportamiento de rendimiento del algoritmo. Estos enfoques están predominantemente justificados si parte de la información estructural de la instancia del problema no se conoce antes de ejecutar el algoritmo. Este concepto se puede aplicar dinámicamente para ajustar los parámetros de MMA en diferentes puntos de los pasos algorítmicos.

Por el contrario, PTS tiene como objetivo proporcionar valores de parámetro robustos que se puedan aplicar a una amplia variedad de casos [25]. En el contexto de la optimización combinatoria, estudios recientes han propuesto diseños de Estrategias de ajuste de parámetros específicos de instancia (IPTS) utilizando información específica de instancia dentro de estrategias de ajuste de parámetros [26] - [28]. Al hacerlo, IPTS evalúa la información específica de la instancia a priori para la ejecución de un algoritmo y, posteriormente, determina los mejores valores de parámetro $\{^*$ para el algoritmo. Una ventaja de estas estrategias es la comprensión explícita de las características de instancia relevantes del problema. Este estudio propone un marco para calibrar los parámetros del algoritmo de un MMA topológico utilizando IPTS.

III. Marco IPTS para un MMA de posprocesamiento

A. Estructura del marco

Este estudio presenta un marco IPTS para determinar un conjunto de valores de parámetros específicos del algoritmo basados en información específica de la instancia. La herramienta IPTS proporciona valores de parámetros $\{^*$ para un caso particular I [27]. Posteriormente, este estudio tiene como objetivo determinar los mejores valores de los parámetros $\{^*$ para un conjunto dado de mediciones GPS. La figura 2 muestra el marco de calibración de parámetros propuesto. Su estructura se basa en encontrar una configuración de parámetro $\{^*$ para una instancia determinada I antes de ejecutar un MMA. Ejemplo I se define como un conjunto de puntos de datos GPS que utiliza el MMA de posprocesamiento para determinar la ruta subyacente del vehículo. La herramienta IPTS se realiza para cualquier instancia I , permitiendo ajustar los parámetros de MMA en función de las características de la instancia.

La instancia se caracteriza por dos conjuntos de información, a saber, información específica del escenario y específica de los datos. El primero puede incluir información relevante sobre el entorno y las características del vehículo.

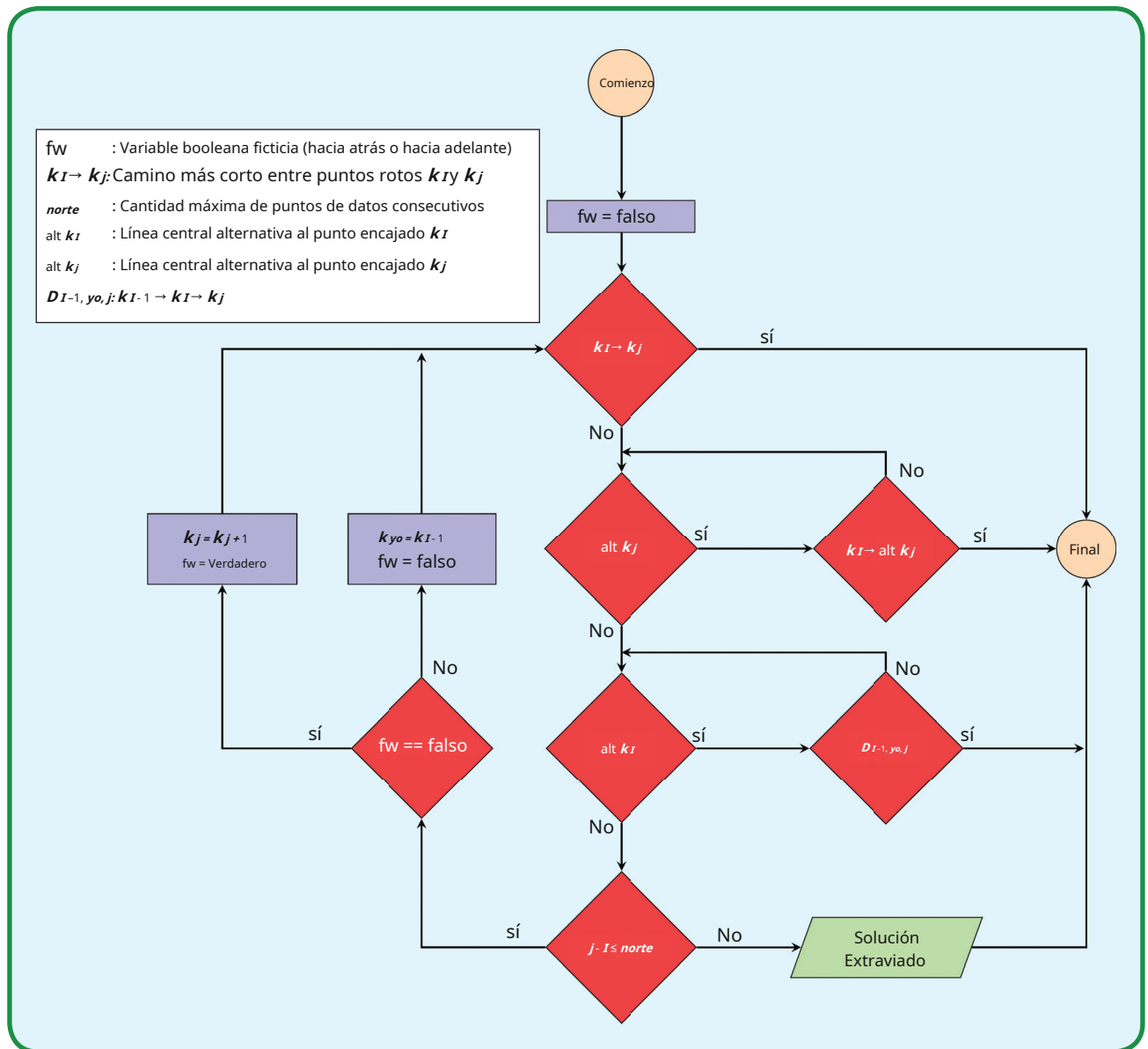


Figura 1 Diagrama de flujo del MMA de posprocesamiento.

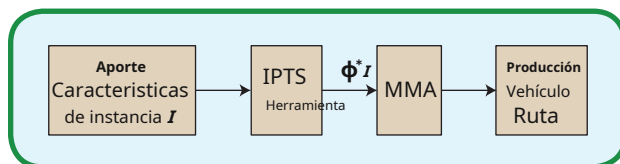


Figura 2 Diagrama del marco IPTS.

Parte de esta información permanece constante a lo largo de la ruta del vehículo, como el tipo de vehículo o las condiciones climáticas, mientras que otra información puede variar a medida que el vehículo se desplaza, como la velocidad máxima de la carretera, el tipo de entorno construido circundante (por ejemplo, el uso del suelo), el tipo de carretera, etc. Por el contrario, la información específica de los datos está relacionada con la medición del GPS.

mentos (por ejemplo, número de satélites detectados, velocidad del vehículo) y frecuencia de muestreo (por ejemplo, 2 segundos o 5 minutos).

B. Diseño del IPTS de lógica difusa

Este estudio propone un diseño IPTS que utiliza un sistema de inferencia difusa para determinar un conjunto de valores de parámetros específicos del algoritmo basados en información específica de la instancia. Zadeh [29] introdujo la lógica difusa como un concepto alternativo a la teoría de conjuntos clásica para permitir tratar con información no nítida. La lógica difusa permite una membresía parcial de información en múltiples conjuntos que se describen mediante términos lingüísticos y se ha aplicado en estrategias de ajuste de parámetros en la literatura relacionada (por ejemplo, en la forma de un sistema de inferencia difusa [27]). Este sistema de inferencia difusa proporciona

una alta flexibilidad de manera que los elementos específicos relevantes que definen un algoritmo (por ejemplo, parámetros de control, conceptos de diseño) se pueden ajustar fácilmente, lo que afecta el rendimiento del algoritmo. Particularmente, en el contexto de MMP, se puede emplear información específica de instancia para ajustar los parámetros de MMA.

La base de la lógica difusa es la capacidad de proporcionar una etiqueta lingüística para cualquier valor de parámetro nítido, asignando cualquier valor a uno o más conjuntos difusos con un cierto grado de pertenencia. Este grado de membresía se evalúa mediante una función de membresía que se crea para cada conjunto difuso, como 'bajo', 'medio' y 'alto'. El sistema de inferencia difusa propuesto, que incluye un conjunto de funciones de pertenencia para cada variable y una base de reglas, se basa en un diseño manual. Este diseño se obtiene de un análisis estadístico. Por tanto, los conjuntos difusos son una interpretación subjetiva de los niveles de parámetros individuales.

El marco propuesto consta de las siguientes dos etapas de diseño:

1) Investigación del impacto de la información específica de instancia y algoritmo en el rendimiento de MMA
Una etapa inicial analiza el impacto de los cambios en la información específica de la instancia y los parámetros específicos del algoritmo, y los posibles efectos de interacción en el rendimiento de MMA. En otras palabras, un análisis estadístico indica si la variación de un parámetro, con todos los demás iguales, conduce a un cambio estadísticamente significativo en el desempeño (efecto principal), y si el impacto de una variable independiente varía según el nivel de una variable independiente. segunda variable independiente (interacción bidireccional).

Se propone un diseño factorial en combinación con un análisis de regresión basado en un conjunto de instancias de problemas de entrenamiento. Este análisis estadístico proporciona una comprensión de la importancia de los parámetros específicos de la instancia y del algoritmo con respecto al rendimiento de MMA.

2) Diseño de la herramienta Fuzzy IPTS

Para diseñar la herramienta IPTS difusa, cada característica de instancia se considera una variable de entrada y los mejores valores de parámetro (*) se denominan variable de salida. En el *fuzzificación* proceso de la herramienta IPTS difusa, la información específica de la instancia se asigna a un conjunto de funciones de membresía que están asociadas con términos lingüísticos mediante la identificación de un grado de membresía *norte* que se caracteriza por (1) con *norte* siendo el número de funciones de pertenencia.

$$norte = [0, 1], \quad j \in \mathbb{N}, \text{ norte} \quad (1)$$

Posteriormente, las reglas difusas se diseñan en función de los conocimientos extraídos del análisis estadístico anterior. Cada regla comprende información sobre las variables de entrada y salida. Cada regla considera una combinación única de conjuntos difusos a los que se asigna cada variable. Un conjunto de SI-ENTONCES

Las reglas se activan en función de los grados de membresía identificados de la siguiente manera:

$$\text{Si } X_1 \text{ es } A_1 \text{ y } X_2 \text{ es } A_2, \text{ Luego } Y_1 \text{ es } B_1 \text{ y } Y_2 \text{ es } B_2$$

Dónde, X_1 y Y_1 son las variables de entrada y salida, respectivamente, y A_1 y B_1 son los términos lingüísticos asociados con los conjuntos difusos. Por ejemplo, "SI la frecuencia de muestreo (SF) es alta, ENTONCES el tamaño del búfer (BZ) es medio".

Finalmente, la herramienta IPTS difusa *desenmascara* esta información para derivar valores precisos para cada parámetro de salida (es decir, cada parámetro del algoritmo) mediante el uso de la construcción de un sistema de inferencia difusa de tipo Mamdani. los *defuzzificación* El proceso puede aplicar el método de centroide descrito por (2), en el que los centros de gravedad de los conjuntos difusos se emplean para determinar valores nítidos [30].

$$z = \frac{\sum_{no(z)} z dz}{no(z)} \quad (2)$$

El resultado de la herramienta IPTS descrita anteriormente es un conjunto de mejores valores de parámetros (*). El MMA utiliza este conjunto como información de entrada para generar rutas de vehículos completas, como se muestra en la Fig. 2. Finalmente, el rendimiento de la herramienta IPTS difusa se evalúa con un conjunto de instancias de prueba utilizando los mejores valores de parámetros para cada instancia en la ejecución. del MMA de posprocesamiento.

IV. Resultados

A. Datos

Los datos empleados en este estudio fueron recopilados por vehículos de mantenimiento invernal (también conocidos como quitanieves) en el condado de Portage, Wisconsin, Estados Unidos. Se recopilaron dieciséis casos diferentes (es decir, rutas de vehículos) durante una temporada de invierno. La Fig.3 presenta una muestra de datos de los puntos de datos GPS recopilados en un intercambio de carreteras principal, y la

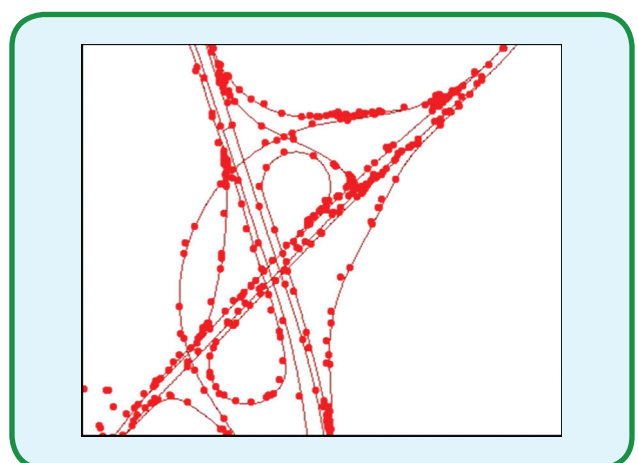


Fig. 3 Muestra de datos para el condado de Portage.

mapa de la línea central de la carretera con una escala nominal de 1: 12.000. Los datos medidos capturaron coordenadas, marcas de tiempo y velocidades del vehículo. Se utilizan veinticuatro instancias como conjunto de datos de entrenamiento para investigar los efectos principales y de interacción de cualquier parámetro involucrado. Se utilizan dieciocho instancias como conjuntos de datos de prueba al ejecutar el MMA con los valores de los parámetros de salida proporcionados por la herramienta difusa IPTS. Dos parámetros específicos de la instancia, a saber, la frecuencia de muestreo (SF) y la proporción de carreteras de alta velocidad (HS), se identifican inicialmente como características de instancia relevantes en este estudio. El MMA se ejecutó en Python 2.7.8 con una biblioteca arcpy del software ArcGIS.

B. Diseño del marco IPTS

Etapas 1: Diseño de experimentos (DOE)

El conjunto de instancias de entrenamiento se resolvió para diferentes combinaciones de parámetros, como se muestra en la Tabla I. SF y HS son variables de entrada (es decir, información específica de la instancia) del marco IPTS propuesto, mientras que BZ y ST son variables de salida (es decir, algoritmo- parámetros específicos) que son determinados por la herramienta IPTS. Los valores de SF probados son 10, 20 y 30 segundos, y el porcentaje de HS a lo largo de las carreteras varía, agrupando alrededor del 30%, 55% y 70% para las instancias de entrenamiento. Este último se calcula como una relación entre el número de puntos GPS recopilados a lo largo de las carreteras y todos los tipos de carreteras (carreteras, carreteras locales y rampas). La selección de los niveles de BZ y ST fue guiada por un estudio experimental de parámetros del algoritmo realizado en [31].

Debido a la naturaleza de posprocesamiento del MMA, el resultado de ajuste correcto o incorrecto de cada punto de datos se conoce de antemano. El rendimiento de la MMA se evaluó comparando el resultado de la rotura de cada punto de datos con su ruta real, con "1" que representa una rotura correcta y "0" en caso contrario. Calidad de la solución $SQ(i)$ del algoritmo por ejemplo I está determinada por la proporción de instantáneas correctas usando (3) con $metro_i$ siendo el número de puntos ajustados correctamente, y $norte_i$ siendo el número total de puntos de datos de instancia I .

$$SQ(i) = \frac{metro_i}{norte_i} \quad (3)$$

Cuadro I. Niveles de diseño factorial.

Parámetro	Nivel 1	Nivel 2	Nivel 3
Instancia: parámetros específicos			
Frecuencia de muestreo (SF) [ft]	10	20	30
Cuota de carreteras de alta velocidad (HS)	0,3	0,55	0,7
Parámetros específicos del algoritmo			
Tamaño de búfer (BZ) [pies]	20	40	60
Tolerancia de velocidad (St) [mi / h]	15	25	35

Se utilizó una regresión lineal para investigar la significación estadística de los efectos de interacción principal y bidireccional con respecto al rendimiento de MMA. El análisis se ha realizado para dos preferencias diferentes de los tomadores de decisiones con respecto al rendimiento algorítmico, es decir, la mejor calidad de la solución (SQ) y tiempo computacional ($TIEMPO$).

La Tabla II muestra los resultados del análisis de regresión con los coeficientes de los factores significativos y el nivel de significancia correspondiente. los R^2 valores para SQ y $TIEMPO$ son 0,66 y 0,53, respectivamente. Esta tabla destaca que ST no es estadísticamente significativo para ninguna de las dos preferencias de desempeño (es decir, SQ y $TIEMPO$) con los efectos principales y de interacción. Por lo tanto, ST se descartó de un análisis adicional. Por el contrario, BZ se muestra como un parámetro clave en los resultados. Se obtiene una calidad de solución mejorada y tiempos de cálculo más largos cuando BZ aumenta.

Con respecto a la información específica de la instancia, las Figuras 4 y 5 muestran que las instancias con mayores valores de HS dan como resultado mejores SQ rendimiento y tiempos de cálculo más cortos en comparación con instancias con valores de HS más pequeños. Más importante aún, el algoritmo da como resultado un menor SQ para los casos que tienen una proporción igual de carreteras de alta y baja velocidad, es decir, $HS = 0,5$. En aplicaciones particulares de mantenimiento invernal, la dificultad de resolver instancias con $HS = 0,5$ puede estar relacionada con la entrada y salida de carreteras de alta velocidad en intercambios o nodos críticos en la infraestructura. Este hallazgo apunala la necesidad de ajustar el algoritmo de acuerdo con este tipo de características de la carretera.

Cuadro II. Resultados del análisis de regresión lineal.

	Preferencia de rendimiento	
	Calidad de la solución (<i>SQ</i>)	Computacional Tiempo (<i>TIEMPO</i>)
Variables independientes	Coeficientes estandarizados Beta	
Efectos principales		
Frecuencia de muestreo (SF)	0,119	- 0,761 ***
Compartir carreteras de alta velocidad (HS) Tamaño de búfer (BZ)	0,454 ***	- 0.379 **
	1,407 ***	0,487 **
Tolerancia de velocidad (St)	0,020	- 0,169
Efectos de interacción		
HS # SF	- 0,111	0,288 *
HS # St	- 0.013	0.105
HS # BZ	- 0,623 ***	0,011
SF # St	- 0,020	0,135
SF # BZ	- 0,179	- 0.370 **
St # BZ	- 0,004	- 0,004

*** pag 1 0.01, ** pag 1 0.05, * pag 1 0.1.

*** pag 1 0,01, ** pag 1 0,05, * pag 1 0,1.

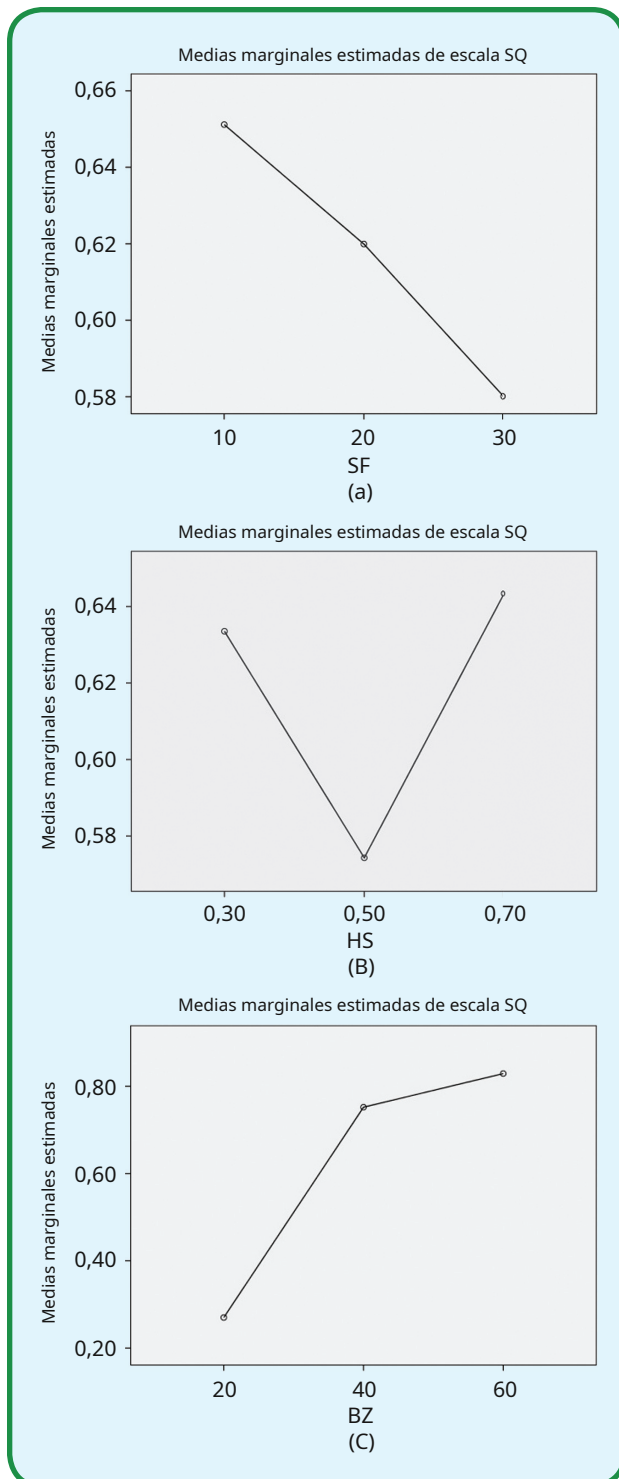


FIGURA 4 Efectos principales significativos de la frecuencia de muestreo (SF), la proporción de carreteras de alta velocidad (HS) y el tamaño del búfer (BZ) en la calidad de la solución (*SQ*).

El efecto de interacción de HS y BZ en *SQ* se observa en la Fig. 6a) debido a la similitud en el comportamiento algorítmico cuando BZ toma valores de 40 y 60. Más precisamente, el rendimiento converge para valores altos de HS. En contraste, Fig.6a)

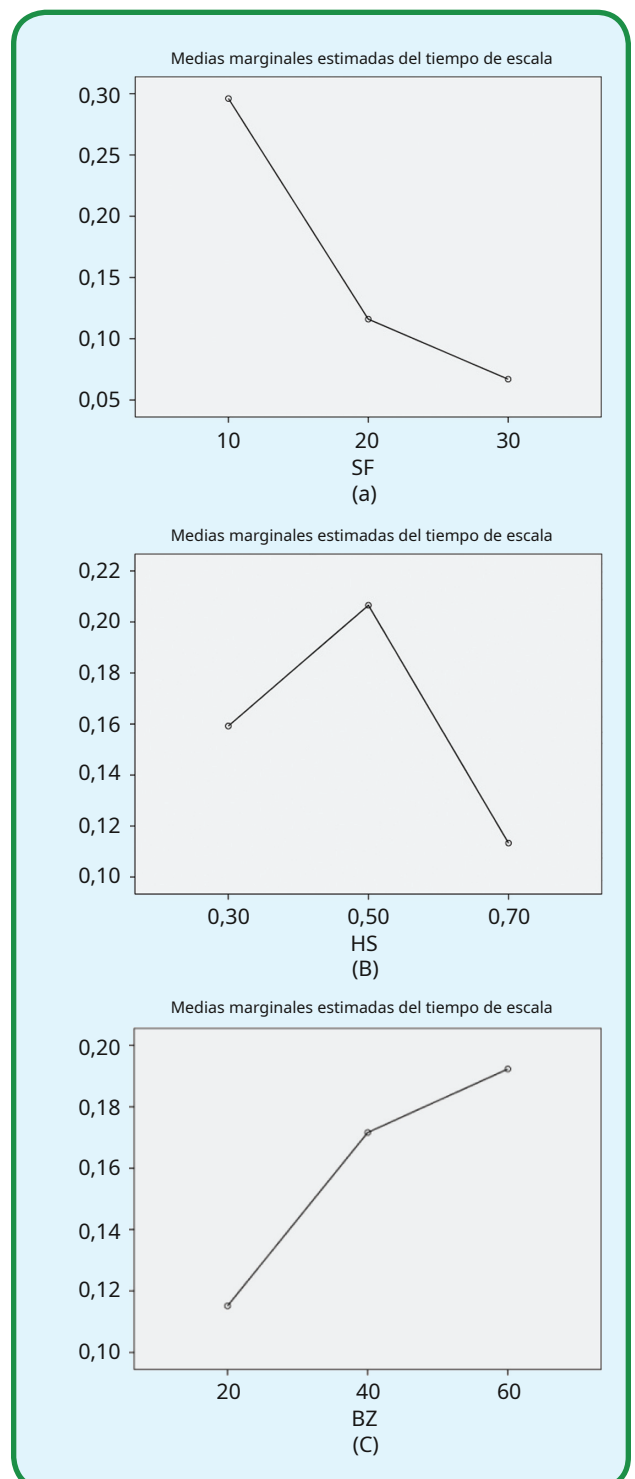


FIGURA 5 Efectos principales significativos de la frecuencia de muestreo (SF), la proporción de carreteras de alta velocidad (HS) y el tamaño del búfer (BZ) en el tiempo computacional (*TIEMPO*).

yb) indica que valores más pequeños de BZ conducen a un desempeño sustancialmente peor en *SQ* y tiempos de cálculo más cortos, respectivamente. Este resultado puede estar relacionado con no

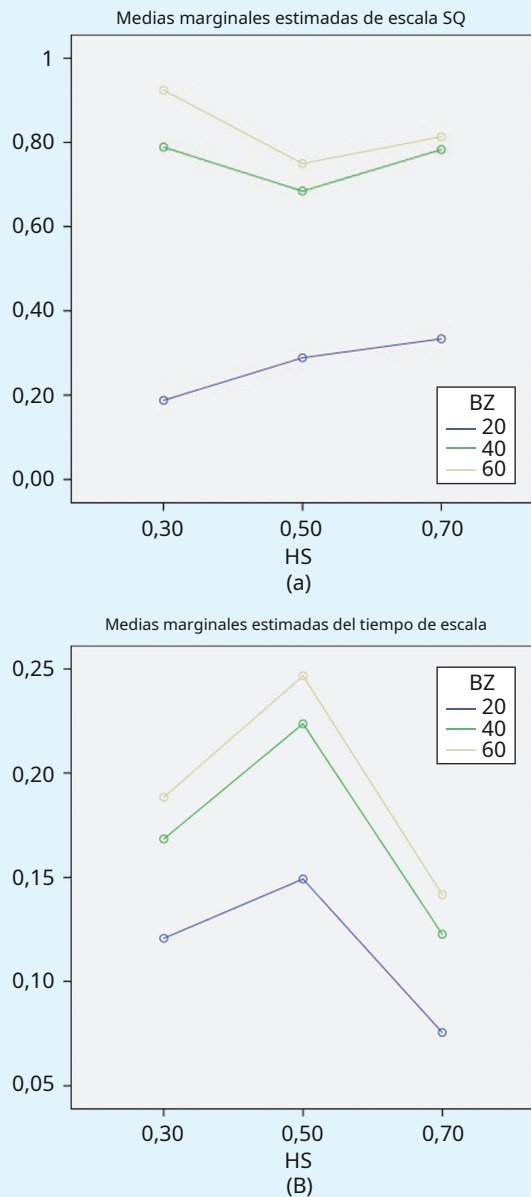


FIGURA 6 Efectos de la interacción de las carreteras de alta velocidad (HS) y el tamaño de la zona de amortiguamiento (BZ) sobre (a) la calidad de la solución (SQ) y (b) tiempo computacional (TEMPO).

encontrar cualquier camino candidato dentro de los tamaños de búfer más pequeños.

Etap 2: Diseño de la herramienta Fuzzy IPTS

La Tabla III muestra el diseño de las funciones de pertenencia de forma triangular para la entrada (parámetros específicos de la instancia) y la salida (parámetros específicos del algoritmo) de la herramienta IPTS. El análisis de regresión se utilizó para informar el diseño manual de una base de reglas para la herramienta IPTS difusa. La Tabla IV muestra la base de reglas correspondiente del sistema IPTS difuso propuesto. Cada regla consta de la información

Cuadro III. Funciones de membresía para entrada y salida.

Aporte			
	Min	Medio	Max
Frecuencia de muestreo (SF) [s]			
Pequeño	0	10	20
Medio	10	20	30
Grande	20	30	40
Cuota de carreteras de alta velocidad (HS) [%]			
Bajo	10	30	55
Medio	35	55	75
Elevado	50	70	90
Producción			
	Min	Medio	Max
Tamaño de búfer (BZ) [pies]			
Muy pequeña	0	0	20
Pequeño	0	20	40
Medio	20	40	60
Grande	40	60	80
Muy grande	60	80	100

Cuadro IV. Base de reglas para el IPTS difuso.

Aporte		Producción	
Parámetros específicos de la instancia		Parámetros específicos del algoritmo	
		Tamaño de búfer (BZ)	
Muestreo Frecuencia (SF)	Cuota de alto Carreteras rápidas (HS)	Calidad de la solución (SQ)	Computacional Tiempo (TEMPO)
Pequeño	Bajo	Grande	Muy pequeña
Medio	Bajo	Grande	Pequeño
Grande	Bajo	Grande	Pequeño
Pequeño	Medio	Muy grande	Muy pequeña
Medio	Medio	Muy grande	Muy pequeña
Grande	Medio	Muy grande	Muy pequeña
Pequeño	Elevado	Grande	Muy pequeña
Medio	Elevado	Grande	Pequeño
Grande	Elevado	Grande	Pequeño

sobre dos variables de entrada (es decir, SF y HS) y la información sobre una variable de salida (es decir, BZ). Por ejemplo, si SF es *pequeño* y HS es *alto*, entonces BZ se establece en *grande* al buscar la mejor calidad de solución (SQ) y está configurado para *muy pequeña* al buscar un tiempo computacional bajo (TEMPO).

Tras la evaluación de HS y SF para cualquier instancia I , la herramienta IPTS asigna un cierto valor BZ para obtener la preferencia de SQ y $HORA$, como se muestra en la vista de la superficie en la Fig. 7 (a) y (b), respectivamente. La Fig. 7 (a) sugiere que los tamaños de las zonas de amortiguamiento se establecen particularmente altos para los casos difíciles que tienen un cierto equilibrio entre las carreteras de alta velocidad y otras (HS 0.5). Por preferencia SQ , el impacto de HS domina y SF no interactúa con este efecto. Por el contrario, los resultados de la Fig. 7 (b) indican que BZ se establece sustancialmente más bajo para estos casos difíciles debido al impacto en el tiempo de cálculo. Debido al efecto de interacción con SF, los valores de BZ se establecen particularmente bajos para frecuencias de muestreo pequeñas, ya que el impacto en el tiempo de cálculo es más fuerte que para valores de SF más grandes.

C. Ejecución de la herramienta Fuzzy IPTS antes de MMA

El IPTS difuso diseñado se probó con 18 instancias. La Tabla V muestra el número total de puntos de datos GPS y los valores SF y HS para cada uno de estos casos. Esta tabla demuestra la diversidad en el conjunto de datos de prueba. La herramienta difusa IPTS se utilizó para identificar el parámetro BZ específico del algoritmo para cada instancia de prueba a priori para la ejecución del MMA dada la preferencia de SQ y $HORA$, como se muestra en la Tabla VI. A efectos comparativos, la Tabla VII

Tabla V. Características de la instancia de prueba.

Ejemplo	Número de Puntos de datos	Frecuencia de muestreo (SF) [s]	Cuota de alto Carreteras rápidas (HS)
1	358	10	0,37
2	179	20	0,37
3	179	20	0,37
4	120	30	0,38
5	119	30	0,36
6	119	30	0,39
7	232	10	0,89
8	117	20	0,89
9	117	20	0,89
10	78	30	0,87
11	78	30	0,90
12	77	30	0,90
13	176	10	0,56
14	88	20	0,57
15	88	20	0,56
dieciséis	58	30	0,59
17	59	30	0,54
18	59	30	0,56

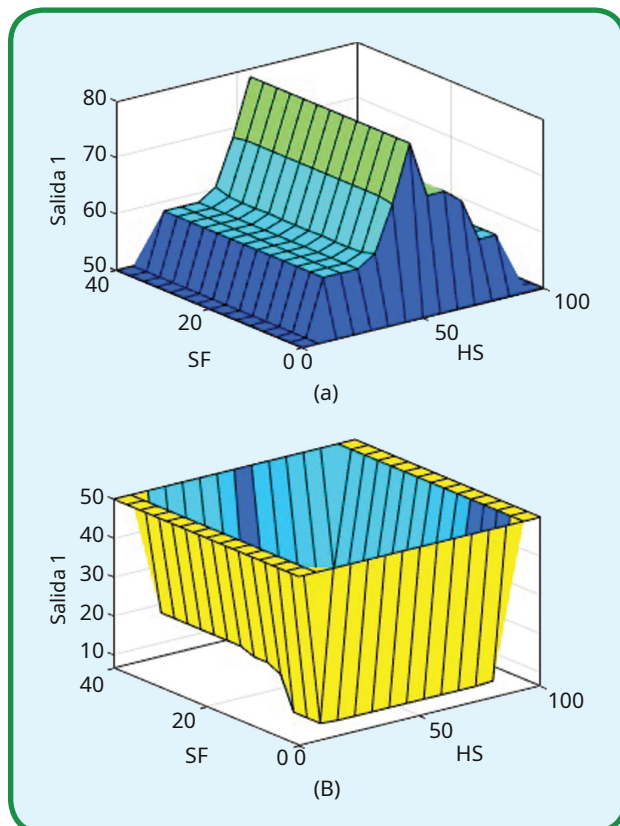


FIGURA 7 Vista de superficie para frecuencia de muestreo (SF) y participación de carreteras de alta velocidad (HS) en tamaño de búfer (BZ) para (a) calidad de la solución (SQ) y (b) tiempo computacional ($TIEMPO$).

Cuadro VI. Valores de parámetros específicos del algoritmo BZ para las instancias de prueba de los resultados de la herramienta IPTS.

Ejemplo	Tamaño de búfer difuso (BZ) [pies]	
	Calidad de la solución (SQ)	Tiempo computacional ($TIEMPO$)
1	63	7
2	63	20
3	63	20
4	64	20
5	62	20
6	66	19
7	60	10
8	60	20
9	60	20
10	60	20
11	50	50
12	50	50
13	74	6
14	73	15
15	74	14
dieciséis	72	dieciséis
17	75	12
18	74	14

Cuadro VII. Resultados con un tamaño de búfer fijo de 60 [pies] para las instancias de prueba.

Ejemplo	Calidad de la solución (<i>SQ</i>)	Tiempo computacional (<i>TIEMPO</i>) [s]
1	91,90%	15.454
2	92,18%	6.813
3	92,18%	7.268
4	88,33%	4.548
5	89,92%	5.039
6	91,60%	5.595
7	91,81%	9.223
8	91,45%	5.302
9	90,60%	4.809
10	87,18%	2.892
11	91,03%	3.308
12	85,71%	3.161
13	94,89%	7.634
14	94,32%	3.529
15	89,77%	4.165
dieciséis	91,38%	2.446
17	93,22%	2.600
18	86,44%	2.629

muestra los resultados para cada instancia para el valor fijo de BZ de 60 [pies] sugerido en [32].

La Tabla VIII presenta los resultados de comparación para el desempeño del MMA usando los valores BZ propuestos por el Fuzzy IPTS y el valor BZ fijo de 60 [pies]. Esta tabla muestra mejoras sutiles con respecto a *SQ*. Los resultados confirman que la ejecución de la herramienta IPTS antes del MMA ayuda a seleccionar estratégicamente los mejores valores de los parámetros del algoritmo para cada instancia analizada en lugar de

Cuadro VIII. Comparación de resultados de desempeño.

	Corregir tamaño de búfer (BZ) de 60 [pies]	Búfer difuso Tamaño (BZ) (Preferencia = <i>SQ</i>)	Búfer difuso Tamaño (BZ) (Preferencia = Hora)
	Calidad de la solución (<i>SQ</i>) [%]		
Mediana	91,42%	91,53%	55,34%
Significar	90,77%	90,51%	52,95%
Estándar Desviación	0,025%	0,035%	0,174%
	Tiempo computacional (<i>TIEMPO</i>) [s]		
Mediana	4,68	4,67	3,29
Significar	5,36	5,34	3,49
Estándar Desviación	3,18	3,17	1,51

empleando un enfoque de prueba y error para calibrar los parámetros del algoritmo.

V. Conclusiones

Existe una investigación limitada en el área de métodos de calibración de parámetros para MMA, como un enfoque de prueba y error o algoritmos descendentes. Además, hay poca información disponible sobre el impacto de las características específicas de la instancia y cualquier impacto potencial en el rendimiento de MMA. Para llenar este vacío, este estudio propone un enfoque sistemático que se basa en una estrategia de ajuste de parámetros específicos de la instancia (IPTS) antes de ejecutar un MMA topológico de posprocesamiento. La lógica difusa se utiliza en el diseño de IPTS, lo que permite la consideración de información imprecisa para respaldar su idoneidad en el contexto de la correspondencia de mapas.

En particular, el IPTS proporciona los mejores valores de parámetros de algoritmo adaptados a información específica de la instancia, en los que se evitan esfuerzos experimentales adicionales al tiempo que se obtienen valores de parámetros que optimizan el rendimiento algorítmico. Esto manifiesta una de las principales contribuciones de esta investigación. Además, este estudio permite que el ajuste de parámetros se ajuste a una preferencia de rendimiento particular,

es decir, maximizar la calidad de la solución o minimizar el tiempo de cálculo. La investigación futura prevé proporcionar una combinación de ambas preferencias de rendimiento, como en [22], que es relevante en la implementación de MMA en tiempo real.

El marco propuesto integra información específica de la instancia en la determinación de parámetros MMA específicos del algoritmo utilizando principios de lógica difusa. La principal ventaja de la herramienta IPTS difusa propuesta es el nivel de flexibilidad con el que se puede rediseñar o ajustar a otros tipos de conjuntos de datos recopilados de diferentes aplicaciones de transporte (por ejemplo, gestión de flotas). Esta herramienta también puede utilizar algoritmos adicionales y parámetros específicos de la instancia (por ejemplo, configuración del entorno), lo que revela otra contribución de esta investigación. Las investigaciones futuras deberían estudiar la interacción entre el tamaño, la diversidad y las características de los datos de entrenamiento y el rendimiento de la herramienta IPTS.

Para diseñar la herramienta IPTS, se presenta un enfoque manual basado en expertos que utiliza información de un análisis estadístico, que alimenta la configuración de la herramienta IPTS. Este enfoque puede modificarse utilizando un enfoque semiautomático o automático que permita una menor dependencia del conocimiento experto con respecto al algoritmo [33] y, por lo tanto, puede ser aplicable para ajustar los parámetros de cualquier tipo de MMA.

La investigación adicional incluye la implementación de la herramienta IPTS para calibrar los parámetros MMA para cada punto de datos GPS individual. Posteriormente, el MMA se ejecuta para una instancia completa considerando diferentes valores de parámetros para cada punto de datos. Además, la herramienta IPTS puede integrarse en el MMA, para ser ejecutada en aplicaciones en tiempo real. Además, se pueden incorporar más avances algorítmicos en el MMA, incluido el uso del

criterio de tiempo de viaje más corto para resolver el MMP, particularmente en entornos de tráfico congestionado, como en [34] - [36].

Reconocimiento

Los autores agradecen las contribuciones del Dr. Darren Scott de la Universidad McMaster y el Departamento de Carreteras del Condado de Portage por proporcionar los datos.

Sobre los autores



Dra. Carola Alejandra Blazquez recibió su Ph.D. Licenciado en Ingeniería de Información Geoespacial de la Universidad de Wisconsin-Madison, EE.UU. en 2005. En 2006, el Dr. Blazquez se incorporó a la facultad del Departamento de Ciencias de la Ingeniería de la Universidad Andrés Bello.

en Chile, donde actualmente ocupa el cargo de profesora. Sus intereses de investigación incluyen sistemas de transporte inteligentes (ITS), GIS para transporte (GIS-T) y seguridad en el transporte.



Jana Ries recibió su doctorado en Investigación Operativa del Departamento de Matemáticas de la Universidad de Portsmouth, Reino Unido. Es profesora titular del Grupo de Gestión de Sistemas y Operaciones de la Universidad de Portsmouth. Su principal investigación está en

diseño algorítmico y calibración de parámetros para algoritmos heurísticos en optimización combinatoria con especial interés en el desarrollo de sistemas de apoyo a la decisión en las áreas de transporte y fabricación.



Dr. Roberto Jesús León recibió su Ph.D. en Ingeniería Informática de la Universidad Técnica Federico Santa María en Valparaíso, Chile en 2016. Actualmente se desempeña como profesor asistente en la Universidad Andrés Bello en Viña del Mar, Chile. Su respuesta

Los intereses de búsqueda incluyen computación científica, simulación numérica y geometría computacional.



Dr. Pablo Andrés Miranda recibió su Ph.D. en Investigación en Logística y Operación de la Pontificia Universidad Católica de Chile en 2004. Actualmente, se desempeña como profesor asociado en la Pontificia Universidad Católica de Valparaíso, Chile. Sus intereses de investigación son

modelos y algoritmos para abordar problemas del mundo real en la industria, centrándose en la operación, el transporte, la logística y la gestión de la cadena de suministro.

Referencias

- [1] M. Bierlaire, J. Chen y J. Newman, "A probabilistic map matching method for smartphone GPS data", *Transporte Res. Parte C*, vol. 26, págs. 78-98, 2013.
- [2] C. Blazquez y A. Vonderohe, "Algoritmo simple de correspondencia de mapas aplicado a datos inteligentes de vehículos de mantenimiento invernal", *Transporte Res. Rec. J. Transporte Res. Rec.*, vol. 1935, págs. 68-76, 2005.
- [3] W. Chen, Z. Li, M. Yu e Y. Chen, "Efectos de los errores del sensor en el rendimiento de la correspondencia de mapas", *J. Navegación*, vol. 58, págs. 273-282, 2005.
- [4] R. Dalumpines y D. Scott, "Coincidencia de mapas basada en SIG: desarrollo y demostración de un algoritmo de coincidencia de mapas de posprocesamiento para la investigación del transporte", en *Avanzando la ciencia de la geoinformación para un mundo cambiante. Notas de conferencias en geoinformación y cartografía*, S. Geertman, W. Reinhardt y F. Toppen, Eds. Berlin: Springer-Verlag, 2011.
- [5] N. Velaga, M. Quddus y A. Bristow, "Desarrollo de un algoritmo de correspondencia de mapas topológicos basado en el peso mejorado para sistemas de transporte inteligentes", *Transporte Res. Parte C*, vol. 17, no. 6, págs. 672-683, 2009.
- [6] A. Oran y P. Jaillet, "Una formulación precisa de ponderación de proximidad para algoritmos de coincidencia de mapas", en *Proc. X Workshop Posicionamiento Navegación y Comunicación*, 2013, págs. 1-6.
- [7] C. Blazquez, P. Miranda y A. Ponce, "Rendimiento de un nuevo algoritmo de correspondencia de reglas de decisión topológica mejorado para aplicaciones de transporte", *J. Appl. Res. Technol.*, vol. 10, págs. 929-940, 2012.
- [8] C. Blazquez, "Un algoritmo de coincidencia de mapas topológicos de reglas de decisión con múltiples datos espaciales", en *Sistemas globales de navegación por satélite: señal, teoría y aplicaciones*, J. Shuanggen, Ed. Croacia: Intechweb, 2012, págs. 215-240.
- [9] M. Hashemi y H. Karimi, "Una revisión crítica de los algoritmos de emparejamiento de mapas en tiempo real: problemas actuales y direcciones futuras", *Computación. Medio ambiente. Urbano*, vol. 48, págs. 153-165, 2014. [10] G. Taylor y G. Blewitt, *Unificación GIS-GPS de Posicionamiento Inteligente*. West Sussex: Wiley, 2006.
- [11] M. Ren y H. Karimi, "Un mapa oculto basado en el modelo de Markov algoritmo para la navegación en silla de ruedas", *J. Navegación*, vol. 62, no. 3, págs. 383-395, 2009.
- [12] Y. Zhao, *Sistemas de navegación y localización de vehículos*. Norwood, MA: Casa Artech, 1997.
- [13] C. Blazquez y A. Vonderohe, "Efectos de los parámetros de control en rendimiento de un algoritmo de correspondencia de mapas de reglas de decisión", *Amer. Soc. Ing. Civil J. Transporte Ing.*, vol. 135, no. 12, págs. 966-973, 2009. [14] F. Marchal, J. Hackney y K. Axhausen, "Efficient map matching de grandes conjuntos de datos del sistema de posicionamiento global", *J. Transporte Res. Junta*, vol. 1935, no. 1, págs. 93-100, 2005.
- [15] M. Quddus, W. Ochieng y R. Noland, "Current map-matching algoritmos para aplicaciones de transporte: vanguardia y direcciones de investigación futuras", *Transporte Res. Parte C*, vol. 15, págs. 312-328, 2007.
- [16] Z. Wang y Z. Yang, "Investigación sobre la correspondencia de mapas de regiones típicas basado en el análisis topológico", en *Proc. 2do Int. Conf. Tecnología y Automatización de Computación Inteligente*, Zhangjiajie, China, 2009. [17] D. Papinski, D. Scott y S. Doherty, "Exploring the route choice proceso de toma de decisiones: una comparación de las rutas planificadas y observadas obtenidas mediante GPS basado en personas", *Transporte Res. Parte F*, vol. 12, págs. 347-358, 2009.
- [18] T. Rasmussen, J. Ingvarsson, K. Halldorsdottir y O. Nielsen, "Im-métodos probados para deducir los tramos y el modo del viaje de las encuestas de viaje utilizando dispositivos GPS portátiles: un estudio de caso del área metropolitana de Copenhague", *Computación. Reinar. Urbano*, vol. 54, págs. 301-313, 2015.
- [19] L. Shen y R. Stopher, "Review of GPS travel survey and GPS data-métodos de procesamiento", *Transporte Rev.*, vol. 34, no. 3, págs. 316-334, 2014.
- [20] K. Sheridan, T. Dyjas, C. Botteron, J. Leclere, F. Dominic y G. Marucco, "Demandas de carreteras. Una solución GNSS asistida utiliza el servicio de acceso a datos EG-NOS", *Mundo GPS*, vol. 22, no. 3, págs. 28 de 2011.
- [21] C. Blazquez y P. Miranda, "A real-time map topological matching metodología para estudios de comportamiento de viaje basados en GPS / GIS", en *Tecnologías móviles para la recopilación y análisis de datos de viajes y actividades*, S. Ra-souli y H. Timmermans, Eds. Nueva York, NY, EE. UU.: IGI, 2014, págs. 155-173.
- [22] J. Ries, P. Beullens y D. Salt, "Pa-ajuste de parámetro basado en lógica difusa", *EUR. J. Oper. Res.*, vol. 218, págs. 305-315, 2012.
- [23] SP Coy, BL Golden, GC Runger y EC Wasil, "Using experi-diseño mental para encontrar ajustes de parámetros efectivos para la heurística", *J. Heurística*, vol. 7, no. 1, págs. 77-97, 2001.
- [24] B. Adenso-Díaz y M. Laguna, "Ajuste de algoritmos mediante fracc-diseños experimentales nacionales y búsqueda local", *Oper. Res.*, vol. 54, no. 1, págs. 99-114, 2006.

- [25] F. Dobslaw, "Un marco de ajuste de parámetros para metaheurísticas basadas sobre el diseño de experimentos y redes neuronales artificiales", en *Proc. En t. Conf. Computación Matemática y Computación Natural*, 2010, págs. 1-4.
- [26] M. Kern, "Adaptación de parámetros en la búsqueda heurística: una población enfoque", Ph.D. tesis, Univ. Essex, 2006.
- [27] J. Ries y P. Beullens, "Un diseño semiautomático de ajuste difuso de parámetros para metaheurísticas basadas en la inducción del árbol de decisiones", *J. Oper. Res. Soc.*, vol. 66, no. 5, págs. 782-793, 2015.
- [28] R. Pavon, F. Díaz, R. Laza y V. Luzon, "Ajuste automático de parámetros con un sistema de razonamiento bayesiano basado en casos", *Expert Syst. Appl.*, vol. 36, no. 2, págs. 3407-3420, 2009.
- [29] LA Zadeh, "Fuzzy sets", *Inf. Control.*, vol. 8, págs. 338-353, 1965.
- [30] O. Cordón, "Una revisión histórica de los métodos de aprendizaje evolutivo para Sistemas basados en reglas difusas tipo Mamdani: diseño de sistemas difusos genéticos interpretables", *En t. J. Appl. Razonamiento*, vol. 52, págs. 894-913, 2011. [31] CA Blazquez, "Algoritmo de reglas de decisión para la correspondencia de mapas en trans-portación", Ph.D. tesis, Univ. Wisconsin-Madison, 2005.
- [32] CA Blazquez, J. Ries y PA Miranda, "Towards a parameter tuning enfoque para un algoritmo de coincidencia de mapas", en *Proc. IEEE Int. Conf. Electrónica y seguridad vehicular*, Viena, 2017, págs. 85-90.
- [33] J. Ries, P. Beullens e Y. Wang, "Ajuste de parámetros específicos de instancia para metaheurística", en *Algoritmos de optimización metaheurística en ingeniería, negocios, economía y finanzas*, T. Vasant, Ed. Her- shey: IGI Global, 2012, págs. 136-170.
- [34] J. Tang, Y. Song, H. Miller y X. Zhou, "Estimación de las trayectorias espacio-temporales, tiempos de permanencia e incertidumbres de la trayectoria a partir de los datos de trayectoria del vehículo: un método geográfico del tiempo", *Transp. Res. Parte C*, vol. 66, págs. 176-194, 2016.
- [35] T. Hunter, "Estimación del estado de latencia baja a gran escala de cyberphysical sistemas con aplicación a la estimación de tráfico", Ph.D. tesis, Univ. California, 2014.
- [36] W. Li, D. Nie, D. Wilkie y MC Lin, "Estimación del tráfico en toda la ciudad dinámica a través de rastreos GPS dispersos", *IEEE Intell. Transp. Syst. Revista.*, vol. 9, no. 3, págs. 100-113, 2017.

Nombre: Carola Alejandra Blazquez

Dirección postal: Universidad Andrés Bello, Facultad de Ingeniería, Quillota 980, Viña del Mar, Chile

Teléfono: +56 32284 5579

Correo electrónico: cblazquez@unab.cl

* Investigación financiada con fondos del proyecto de la Universidad Andrés Bello Proyecto N ° DI-10-18 / REG, y Proyecto Fondecyt N ° 1140811.