

El análisis de los patrones de movimiento en las caracterizaciones de trayectorias de objetos espaciales

The analysis of the movement patterns in the characteristics of the trajectories of spatial objects.

Diego Fernando Rodríguez Lamus¹ & Álvaro Enrique Ortiz Dávila²

Para citar este artículo: Rodríguez-Lamus, D.F. & Ortiz-Dávila, A.E. (2019) El análisis de los patrones de movimiento en las caracterizaciones de trayectorias de objetos espaciales. *UD y la Geomática*, (14), 34-40

DOI: <https://doi.org/10.14483/23448407.15235>

Fecha de recepción: 14 de mayo de 2019

Fecha de aceptación: 01 de diciembre de 2019

RESUMEN

La representación de objetos en movimiento en bases de datos espacio-temporales ha dado lugar a considerar un nuevo elemento denominado trayectoria, con el cual se puede analizar, consultar y utilizar el movimiento mismo del objeto. El modelamiento de las trayectorias en bases de datos de objetos en movimiento permite simular escenarios del mundo real y realizar tomas de decisiones más acertadas y efectivas, al tener en cuenta aspectos como patrones y anomalías en el comportamiento de agrupaciones de las trayectorias analizadas. El presente artículo muestra los aspectos fundamentales de las trayectorias, diferentes tipos de análisis de patrones, detección de anomalías y los usos potenciales que pueden tener cada uno de ellos, con referencias a ejemplos de aplicación.

Palabras clave: movimiento, bases de datos de objetos en movimiento, trayectorias, patrones, anomalías.

ABSTRACT

The representation of objects in motion in spatio-temporal databases has led to consider a new element, called trajectory, with which you can analyze, consult and use the movement itself of the object. The modeling of the trajectories in databases of objects in movement, allows to simulate real world scenarios and make more accurate and effective decision making, taking into account aspects such as patterns and anomalies in the behavior of groupings of the trajectories analyzed. The present article shows the fundamental aspects of trajectories, different types of pattern analysis, anomalies detection, and the potential uses that each of them can have, with references to application examples.

Keywords: movement, moving object databases, trajectories, patterns, anomalies.

1 Universidad Distrital Francisco José de Caldas, Bogotá D.C., Colombia, dfrodriguez@correo.udistrital.edu.co

2 Universidad Distrital Francisco José de Caldas, Bogotá D.C., Colombia, aortizd@udistrital.edu.co

Introducción

Con el auge de los dispositivos móviles y por ende de los sensores incorporados a ellos, como los que reciben la señal de los satélites de posicionamiento global (GNSS), que indican la posición en centésimas de segundo, se ha abierto un abanico de posibilidades para extraer comportamientos de objetos con base en su movimiento. Una de las opciones es la construcción y análisis de trayectorias a partir de los datos de posición arrojados cada determinado intervalo de tiempo por algún dispositivo receptor, lo cual permite extraer conocimiento de dichos datos como el descubrimiento de patrones, detección de anomalías, agrupación (*clustering*), etc., donde la detección de anomalías cumple un papel muy importante, al facilitar el descubrimiento de comportamientos anormales en un conjunto de datos con respecto al resto del conjunto o a otro que se tenga como modelo.

En el presente artículo se exponen de forma general las caracterizaciones de las distintas definiciones fundamentales de las trayectorias, desde su conceptualización hasta su implementación en bases de datos espacio-temporales, además de los distintos tipos de análisis que se pueden realizar sobre estas, como su agrupación (*clustering*), detección de patrones y de anomalías, exponiendo los posibles usos prácticos en situaciones cotidianas para los tomadores de decisiones en sistemas de movilidad que involucren cualquier tipo de objetos en movimiento (personas, vehículos, aeronaves, etc.).

Caracterizaciones

A continuación, se explican brevemente las caracterizaciones de las trayectorias de objetos espaciales.

Bases de datos espaciales

Son aquellas que permiten extender los sistemas manejadores de bases de datos relacionales para que se pueda realizar toda clase de consultas geométricas de una forma sencilla, teniendo en cuenta que la base de datos maneja lenguaje SQL. Para extender los modelos de bases de datos es necesario crear los correspondientes tipos de datos para las formas geométricas, procedimientos y funciones para realizar cálculos, índices para espacio multidimensional para elaborar el mapeo desde el lenguaje de consulta a los componentes geométricos (Güting y Schneider, 2005).

Bases de datos de objetos en movimiento

Se refieren a aquellas bases de datos espacio-temporales que permiten almacenar y manipular entidades en movimiento, para responder cualquier pregunta acerca del

movimiento. En este grupo se puede describir el estado actual del movimiento, así como también su historia total, donde se puede devolver en cualquier instante para conocer el estado en ese momento.

En estas bases de datos, se trata de representar los cambios continuos de los objetos, en vez de los discretos, que ha sido la forma tradicional de abordar el problema. Existen dos clases de datos susceptibles a estos cambios continuos: el punto y la región (polígono) (en el presente trabajo se habla solo del primero). La abstracción para pasarlo al modelo físico de la base de datos se denomina *moving point*, que representa cualquier objeto que se mueva en un plano en un espacio multidimensional, del cual interesa el cambio en la posición, y todas sus cantidades derivadas como velocidad y aceleración (Güting y Schneider, 2005).

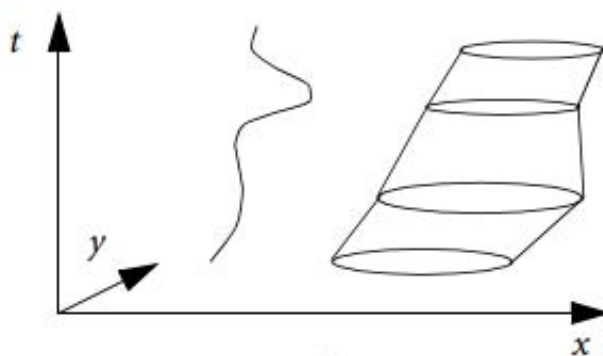


Figura 1. Representación del movimiento continuo de los objetos en movimiento en el espacio tiempo

Fuente: Güting y Schneider (2005).

Modelamiento en base de datos de los objetos en movimiento

Para el modelamiento de las bases de datos se utilizan datos espacio-temporales como *moving point* (*mpoint*) o *moving region* (*mregion*), los cuales son unos valores que capturan el desarrollo temporal de un objeto en el tiempo. En cuanto a la representación geométrica, se representan como objetos en 3D, como se observa en la figura 1 (Güting y Schneider, 2005).

La propuesta descrita permite representar los datos de forma continua y también evaluar de forma discreta el movimiento, ya que se puede realizar cualquier consulta sobre un instante específico (discreto) o sobre un intervalo de tiempo determinado (continuo). La representación física en una base de datos se trata de una tabla con un tipo de dato como lo es *mpoint*, que es el de interés en el presente trabajo:

Ruta (id: integer, desde: string, hasta: string, vehiculo: integer, recorrido: mpoint)

La trayectoria en la base de datos es la proyección de un punto en movimiento sobre el plano, llegando a ser una línea.

Trayectorias

Para entender qué es una *trayectoria*, es necesario definir *movimiento*, el cual consiste en el “cambio en el tiempo de la posición u orientación de un cuerpo” (Encyclopædia Britannica, 2017). En ese sentido, el movimiento de un objeto depende esencialmente de dos componentes: posición y tiempo, lo cual implica que para modelarlo, es necesario tener en cuenta los componentes mencionados.

Para analizar cada uno de los componentes, se tiene que la *posición* es definida como la localización en el espacio de un cuerpo u objeto, el cual se representa, por convención, a través de un sistema de coordenadas, lo que requiere por lo menos un par de coordenadas para establecer la ubicación; por otro lado, el tiempo no es más que una magnitud de orden físico que mide la duración o separación entre eventos o acontecimientos. Entonces, el movimiento es el cambio de localización de un cuerpo en el espacio entre eventos, siendo despreciable la duración entre estos, con lo cual se tiene una magnitud de tipo continuo.

Teóricamente, el movimiento de un objeto puede ser definido como, el mapeo desde un espacio temporal- I (unidimensional) a uno geográfico- S (bidimensional):

$$I \subseteq \mathbb{R} \rightarrow S \subseteq \mathbb{R}^2: t \rightarrow l(t)$$

Donde, $l(t)$ representa la ubicación donde un objeto se encuentra en un tiempo t . La trayectoria de un objeto en movimiento se define así:

$$T_{real} = \{(t, l(t)) | t \in I\} \subset \mathbb{R}^2 \times \mathbb{R}$$

Donde, principalmente, se representa una trayectoria como un conjunto ordenado de parejas de tiempo y posición, tal que t pertenezca al espacio temporal definido, donde como resultado se representa una curva continua en tres dimensiones (Pelekis y Theodoridis, 2014).

Para representar las trayectorias en las computadoras, es necesario realizar una discretización de estas, a partir de

la definición de un conjunto de parejas posición y tiempo finito, que sea representativo para la definición de la trayectoria, desembocando en algo así:

$$T = \{ \langle p_1, t_1 \rangle, \langle p_2, t_2 \rangle, \dots, \langle p_n, t_n \rangle \}$$

Donde, $p_i \in \mathbb{R}^2, t_i \in \mathbb{R}, 1 \leq i \leq n$ y $t_1 < t_2 < \dots < t_n$.

Al realizar la correspondencia entre el y T , se puede considerar que existe pérdida de información ya que se lleva a cabo una compresión de la información que por definición descarta datos. Existe una pregunta acerca del movimiento que hace concluir que nuestro conocimiento acerca de este es de naturaleza continua: ¿Dónde estuvo localizado un objeto en un tiempo entre dos puntos tomados?, donde la respuesta incluye cierta incertidumbre que depende de la tasa de muestreo y los parámetros del movimiento. Se utilizan técnicas de interpolación, siendo la más popular la interpolación lineal, debido a su fácil implementación.

Las trayectorias y su modelación en ambientes computacionales son otra forma de ver el movimiento de los objetos, ya que aquí las técnicas computacionales se unen para entender el movimiento y hacer preguntas sobre estas, que fácilmente un ordenador puede resolver, además de ser más eficiente para la recuperación y almacenamiento de estos datos.

Patrones sobre trayectorias

Un patrón se define como un comportamiento regular hallado en los datos, es decir, cuando estos se repiten de una manera predecible o sistemática. En el caso de las trayectorias, existen principalmente cuatro clases de patrones (Yu, 2015):

- a. *Patrones de movimiento simultáneo*. Estos consisten en el descubrimiento de un grupo de objetos que se mueven juntos por un determinado periodo de tiempo, como *rebaños, convoy, enjambre, compañía viajera* y *reunión*. Estos patrones son útiles en el estudio de migración de especies de animales, vigilancia militar y detección de eventos en el tráfico. Las principales características son la forma o densidad de un grupo, el número de objetos en el grupo y la duración del patrón.

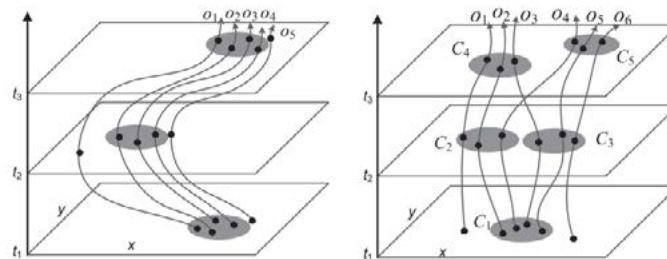


Figura 2. Ejemplos de patrones de movimiento simultáneo. Izq.: rebaño, convoy y enjambre; der.: reunión

Fuente: Yu (2015).

- b. *Cluster de trayectorias*. Este tipo de patrones funciona para encontrar caminos o tendencias representativas que son compartidas por diferentes objetos en movimiento, a través del agrupamiento de trayectorias similares en *clusters* o grupos. La propuesta general radica en la representación de una trayectoria a través de un vector de características, indicando la similitud entre dos trayectorias por la distancia entre sus vectores de características. Una de las restricciones es la dificultad que se tiene para generar el vector de características debido a que cada trayectoria es un mundo totalmente diferente, en cuanto propiedades como la longitud, forma, tasa de muestreo, número de puntos y su orden, por tanto, es difícil comprimir todas las propiedades espacio-temporales de una trayectoria en un vector de características.
- c. *Patrones secuenciales*. El objetivo de esta rama es encontrar los patrones secuenciales en una sola o múltiples trayectorias; donde un patrón secuencial se interpreta como un determinado número de objetos en movimiento que viajan siguiendo una secuencia común de ubicaciones en intervalos de tiempo similares. Dadas dos trayectorias A y B:

$$A: l_1 \xrightarrow{1.5h} l_2 \xrightarrow{1h} l_7 \xrightarrow{1.2h} l_4, \quad B: l_1 \xrightarrow{1.2h} l_2 \xrightarrow{2h} l_i$$

Donde, l_i representa los lugares visitados, que A y B comparten (l_i) (Yu, 2015), y las flechas entre ellos representan el intervalo de tiempo que ocupan en el desplazamiento entre la secuencia de lugares. Se puede observar que las trayectorias presentadas comparten lugares de recorrido,

así como tiempos de recorrido similares. Cuando la ocurrencia de tal secuencia en un espacio definido excede un límite, un patrón secuencial es detectado.

Para detectar este tipo de patrones, se requiere definir cuáles son los lugares por los que deben pasar los objetos dentro de su secuencia de movimiento, por lo cual se debe establecer el tipo, si es un punto específico o una región que puede facilitar el análisis, así se obtienen resultados consistentes.

- d. *Patrones periódicos*. El último tipo de patrón susceptible de ser descubierto en trayectorias es el periódico, que se fundamenta en que los objetos en movimiento suelen tener patrones periódicos de actividad. Por ejemplo, la rutina diaria de una persona, que sale de su casa hacia el trabajo en la mañana para luego en horas de la tarde realizar la misma trayectoria, pero en dirección contraria; la migración anual de los animales de un lugar a otro; etc. Estos patrones son utilizados para descubrir comportamientos intrínsecos de los objetos en movimiento, y para tratar de predecir movimientos futuros debido a la cantidad de datos históricos para hallarlos (Li *et al.*, 2010). Un objeto en movimiento no repite exactamente los mismos puntos en sus recorridos ni el mismo intervalo de tiempo para llegar de un punto a otro, por lo que se debe realizar una aproximación a las dimensiones espacial y temporal para hallar la información oculta, esto radicado en que los datos de entrada deben suministrarse con una cierta tolerancia o rango de acción para detectar este tipo de patrones.

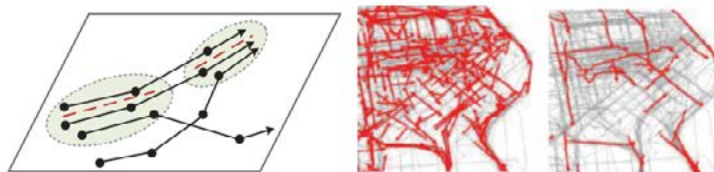


Figura 3. Agrupamiento de trayectorias basado en segmentos parciales. Izq.: clusters de segmentos; centro: microclusters; der.: macroclusters

Fuente: Yu (2015).

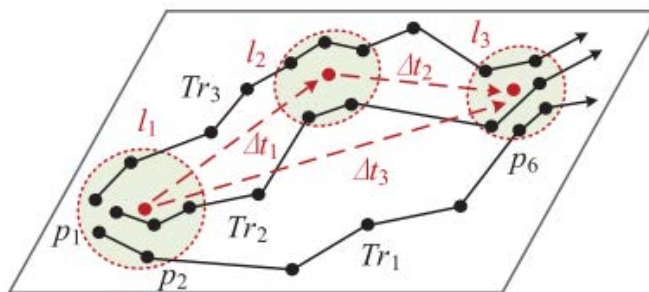


Figura 4. Patrones secuenciales

Fuente: Yu (2015).

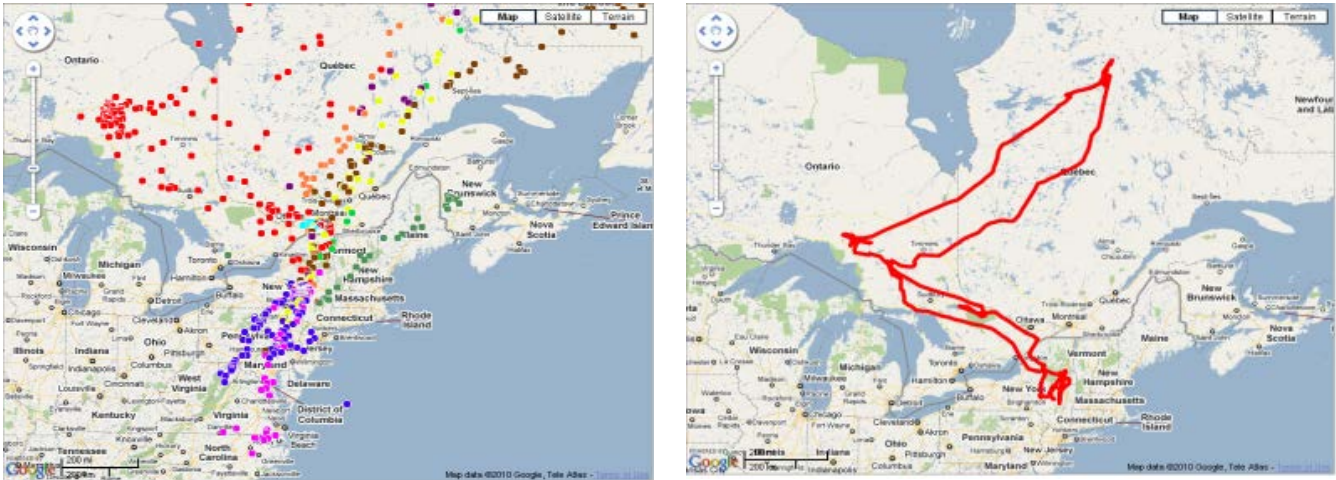


Figura 5. Detección de patrones periódicos. Izq.: datos muestreados; der.: patrón periódico hallados

Fuente: Li et al. (2010).

Metodología para hallar patrones en trayectorias

En Rodríguez y Ortiz (2019), se plantea la unión de dos propuestas teóricas encontradas y su implementación sobre de bases de datos de objetos en movimiento utilizando el motor de base de datos Oracle 11g y la extensión para objetos en movimiento HERMES.

La metodología propuesta consiste en una serie de pasos:

- *Encontrar puntos característicos.* Permite reducir una trayectoria a su mínima representación posible sin perder la información relevante que describe el recorrido y las características más importantes, lo cual hace que se tenga un dato más sencillo que representa un menor tiempo de ejecución y mayor eficiencia para los algoritmos posteriores. Esta reducción recurre a la distancia perpendicular y angular entre los segmentos que componen una trayectoria.
- *Construcción y cálculo de dimensiones para los mínimos cubos envolventes.* Permite descomponer una trayectoria en cubos que son los de menor tamaño que puede encerrar cada uno de los segmentos que la conforman. Estos cubos son descritos a través de dos magnitudes denominadas *volumen* y *traslape*; de esta manera es posible detallar las trayectorias mediante cifras para hallar las medidas utilizadas en el algoritmo de detección de clúster, reemplazando los datos crudos que representan un alto costo computacional.
- *Cálculo de similitud y distancia.* A partir de las dos magnitudes que describen los mínimos cubos envolventes (volumen y traslape), se calcula la medida de similitud entre dos segmentos, donde esta última se utiliza para calcular a su vez la distancia de similitud con la cual finalmente se determina a qué grupo o *cluster* pertenece determinado segmento.

- *Detección de patrones a partir del algoritmo DB-SCAN adaptado a trayectorias.* El último paso de la propuesta, es hallar los *clusters* o patrones que se presentan en los datos, tomando el algoritmo DB-SCAN adaptado a trayectorias (Jae-Gil, Han y Whang, 2007), que presenta la particularidad de no trabajar con puntos si no con subtrayectorias, y el cual manejó dos parámetros: a) *eps*, indica el tamaño del vecindario alrededor de una subtrayectoria núcleo, y b) *MinTrs*, que se refiere a la mínima cantidad de trayectorias que puede tener un *cluster*, descartando los que no cumplan con la condición.

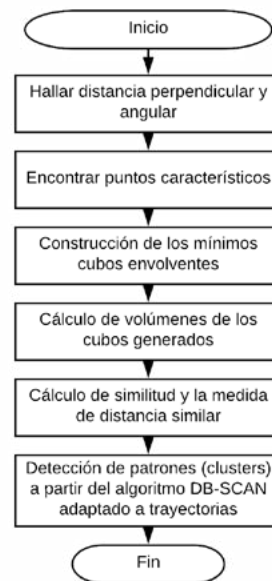


Figura 6. Diagrama de flujo que indica el procedimiento para hallar patrones en trayectorias utilizando el algoritmo DB-SCAN

Fuente: elaboración propia.

Usos potenciales de trayectorias, agrupaciones, patrones y anomalías

Trayectorias

En el contexto de los objetos en movimiento, las trayectorias son las que representan el desplazamiento de estos, a partir de puntos de muestreo de coordenadas en un instante de tiempo determinado.

Entonces, dentro de los usos potenciales de las trayectorias se encuentra la representación del desplazamiento de cualquier objeto que se mueva en un espacio y una ventana de tiempo determinada. Algunos ejemplos pueden ser los siguientes:

- *Sistemas de transporte*, donde los objetos en movimiento son cada uno de los vehículos que lo componen, donde cada uno puede describir una o varias trayectorias en una ventana de tiempo determinado. En este grupo se puede incluir Transmilenio (de Bogotá) donde cada uno de los articulados representa un objeto en movimiento y las rutas que hacen se refieren a las trayectorias. Esto se puede extrapolar a sistemas de transporte por aire y mar.
- *Movimientos de migración de animales*. Las especies migratorias tienen rutinas de movimiento de un sitio al otro del planeta, dependiendo de la dinámica de las estaciones. En este tema, los objetos en movimiento son cada uno de los individuos que pertenecen a la especie, mientras que las trayectorias son los recorridos que realizan para llegar a su punto de destino (Li, Wu y Crofoot, 2013).

Agrupaciones

Las agrupaciones simbolizan aquellos grupos de trayectorias que tienen un comportamiento similar en el

espacio-tiempo, es decir, comparten un mismo espacio geográfico en un intervalo similar de tiempo, definiendo el intervalo de tiempo como relativo (diferencia de tiempos entre un punto de inicio y fin). Ejemplos de agrupaciones se pueden encontrar fácilmente en el reino animal con las especies migratorias, ya que por lo general siempre lo hacen en grandes grupos que comparten un espacio y un tiempo. O como ocurre en Rodríguez y Ortiz (2019), quienes hallan los *clusters* o grupos en rutas de Transmilenio que comparten características similares en las dimensiones espaciales y de tiempo.

Patrones

A partir de las agrupaciones se pueden generalizar patrones de movimiento (Jae-Gil, Han y Whang, 2007) que permitan a los expertos en un dominio determinado entender de una forma más abstracta el movimiento de cientos o miles de objetos en movimientos. Como ejemplo y siguiendo con el trabajo llevado a cabo por Rodríguez y Ortiz (2019), a partir de los *clusters* es posible extraer uno o varios patrones que se pueden graficar de una manera más amena sobre un mapa, en este caso, el patrón hallado para la ruta G43-K43 del sistema Transmilenio entre la estación de San Mateo (Soacha) y el portal El Dorado (Bogotá) (figura 7).

Anomalías

Las anomalías se refieren al problema de encontrar patrones en los datos que no forman parte de un comportamiento esperado (Chandola, Banerjee y Kumar, 2009). Específicamente para el ámbito de las trayectorias existen dos tipos generales de anomalía:

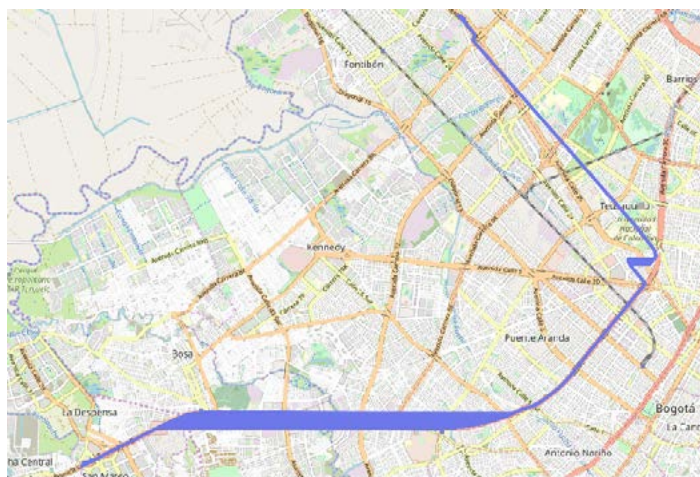


Figura 7. Patrón hallado para la ruta G43-K43

Fuente: elaboración propia.

- Las trayectorias anómalas, que son movimientos de objetos que representan algún desplazamiento anómalo ya sea en tiempo o espacio, en su totalidad.
- Las subtrayectorias anómalas, que se refieren a que la trayectoria puede tener un comportamiento que obedezca al esperado, pero en algunos segmentos de esta, por alguna razón, no se obedece lo planeado, por lo que se llega a un nivel de detalle más profundo que en el primer tipo.

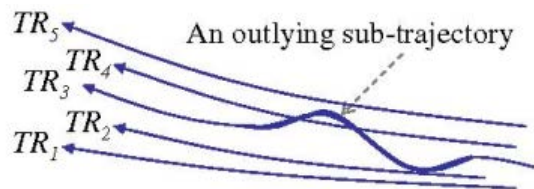


Figura 8. Ejemplo de una subtrayectoria anómala

Fuente: Lee, Han y Li (2008).

Para detectar anomalías se deben tener dos entradas, el patrón *normal* de los datos, y el patrón a evaluar; estas dos entradas deben compararse y si existe algún punto que varié mucho con respecto al patrón *normal* se puede afirmar que existe una anomalía.

Retomando el trabajo de Rodríguez y Ortiz (2019), se esperan encontrar anomalías de tipo espacio-temporal en las subtrayectorias que componen los patrones, por ejemplo, alguna anomalía que permita detectar que tramos de las rutas se están subutilizando o sobre utilizando en cuanto a medidas como velocidad que se puede calcular a partir de las diferencias entre tiempo y velocidad, así como las aceleraciones derivadas de la anterior magnitud. Podría responder preguntas del tipo:

- ¿Qué tramos del recorrido pueden ser susceptibles de mejoras para aumentar la velocidad? (Subutilización).
- ¿Qué tramos deberían tener especial cuidado con la velocidad por las condiciones de los alrededores o de la misma calzada? (Sobreutilización).

Conclusiones

En la actualidad, el tema de las bases de datos espacio-temporales no es muy conocido ni muy utilizado, por lo que estudiando sus posibles aplicaciones en situaciones de la vida real se puede incentivar su uso con todas las ventajas que esto conlleva para el estudio del movimiento, en especial, que se intenta respetar su naturaleza continua, con lo cual es posible tener una interpretación continua o discreta, dependiendo del tipo de análisis que se requiere.

Las bases de datos espacio-temporales tradicionales estudian el movimiento de forma discreta, a diferencia de las de objetos en movimiento, lo cual permite analizar el desplazamiento en cualquier instante de tiempo respetando la naturaleza continua del desplazamiento de un objeto. Lo anterior posibilita estudios mucho más profundos en las trayectorias que describen los objetos en movimiento pudiendo mejorar la toma de decisiones en sistemas de movilidad, al tener información más cercana a la real, que se da de forma natural al movimiento.

Como se expone en las posibles aplicaciones de las trayectorias, se pueden hallar anomalías del tipo de subutilización o sobreutilización de alguna vía por algún objeto en movimiento, en cuanto a su velocidad, siendo una señal o un estudio que otorga a las autoridades de tránsito herramientas para ser más eficientes en la delimitación de los límites de velocidad en las distintas vías de una ciudad.

Referencias bibliográficas

- Chandola, V., Banerjee, A. y Kumar, V. (2009). Anomaly Detection: A Survey. *ACM Computing Surveys*, 41(3), 1-58.
- Encyclopædia Britannica (2017). *Motion*. Recuperado de <http://academic.eb.com.bdigital.udistrital.edu.co:8080/levels/collegiate/article/motion/53959>
- Güting, R.H. y Schneider, M. (2005). *Moving Objects Databases*. San Francisco, EE. UU.: Morgan Kaufmann Publishers.
- Jae-Gil, L., Han, J. y Whang, K.-Y. (2007). Trajectory Clustering: A Partition-and-Group Framework. *SIGMOD'07*. Beijing, China.
- Lee, J.-G., Han, J. y Li, X. (2008). Trajectory Outlier Detection: A Partition-and-Detect Framework. En *IEEE 24th International Conference on Data Engineering, 2008. ICDE 2008* (pp. 140-149). Cancún.
- Li, Z., Wu, F. y Crofoot, M. (2013). Mining Following Relationships in Movement Data. *Proc. 2013 IEEE Int. Conf. on Data Mining*. Dallas, TX.
- Li, Z., Ji, M., Lee, J.-G., Tang, L.-A., Yu, Y., Han, J. y Kays, R. (2010). MoveMine: Mining Moving Object Databases. En *Proceedings of the 2010 International Conference on Management of Data, SIGMOD '10* (pp. 1203-1206). Indianapolis, Indiana.
- Pelekis, N. y Theodoridis, Y. (2014). *Mobility Data Management and Exploration*. Nueva York: Springer.
- Rodríguez L., D.F. y Ortiz D., A.E. (2019). Metodología para hallar patrones de movimiento en trayectorias sobre DBMS. *Revista Ibérica de Sistemas e Tecnologías e Informação*, 636-647.
- Yu, Z. (2015). Trajectory Data Mining: An Overview. *ACM Transactions on Intelligent Systems and Technology*, 6(3), 29:1-29:41.

