

Movilidad urbana y accidentalidad en Cali.

Arango Moreno, Isabella Naranjo Jaramillo, Julián Andrés

24 de noviembre de 2025

Resumen

En este artículo se analizará la estructura topológica del conjunto de datos de siniestralidad en la ciudad de Cali mediante técnicas de Análisis Topológico de Datos (ATD). La secuencia hecha en este proyecto fue, después de limpiar la base de datos, tomar sus direcciones y con ayuda del software QGIS obtener sus coordenadas para luego generar la red de accidentes dentro de la ciudad y sus alrededores, luego, se generaron complejos de Vietoris-Rips. Ya con estos se probaron distintos valores ε que nos muestran las distintas conexiones que se forman entre los lugares donde ocurrieron los accidentes. Finalmente se generaron diagramas de persistencia H_0 y H_1 , los cuales permiten identificar la robustez de los componentes y patrones de conectividad de los datos. Este enfoque brinda una visión vanguardista que puede complementar los métodos de análisis de siniestralidad clásicos, brindándonos una caracterización topológica de los distintos factores que influyen en los accidentes de tránsito.

Palabras clave: Análisis Topológico de Datos (ATD), accidentalidad, Vietoris-Rips, conectividad.

1. Introducción

El análisis de datos de siniestralidad en Cali es fundamental para entender los patrones, las variables y dinámica de accidentalidad en el casco urbano. Siendo Cali la tercera ciudad más poblada del país y la segunda con mayor número de muertes por accidentes de tránsito registradas [2], tener distintos caminos al momento de estudiar este fenómeno es algo crucial. Tradicionalmente estos estudios se hacían con regresión espacial OLS y GWR [1] y con análisis descriptivo de contenido [3].

El Análisis Topológico de Datos (ATD) nos ofrece una vía alterna muy poderosa para caracterizar patrones en nuestra información. Particularmente el análisis de persistencia nos permite estudiar el nacimiento y muerte de

conexiones presentes en los datos, lo que nos brinda información sobre su robustez y conectividad. La propuesta de analizar la accidentalidad en una ciudad colombiana bajo el ATD, es una idea vanguardista debido a que este tipo de exploración se ha hecho principalmente en biología, medicina y ciencias de la computación.

En este trabajo utilizamos técnicas topológicas para analizar los accidentes de tránsito de la ciudad de Cali desde 2016 hasta 2024. Para hacer esto, con ayuda del software QGIS, generamos sobre el mapa de Cali los lugares donde se presentaban los accidentes, a partir de esta nube de puntos generamos complejos de Vietoris-Rips (más adelante se argumentará esta elección) donde su parámetro ε va variando desde 300 hasta 800 metros, mostrando las relaciones de cercanía que se generan entre los distintos accidentes. Luego, sobre este complejo se calcula la homología persistente (H_0) que nos permite identificar la escala en la que los puntos comienzan a fusionarse y los accidentes generan conglomerados o clústeres urbanos, finalmente se calcula la homología persistente de dimensión 1 (H_1) que genera los ciclos, de los cuales cada uno representa un agujero topológico, es decir, un área rodeada por accidentes pero con poca o nula dentro de ellos (por ejemplo un parque o una manzana de casas).

2. Marco Teórico

El presente Marco Teórico tiene como objetivo establecer la base conceptual y contextual necesaria para la comprensión y desarrollo de este análisis. Se estructura en cuatro secciones fundamentales: la Revisión de Literatura Previa sobre el fenómeno de estudio; la exposición de las Teorías Base que sustentan la aproximación metodológica y analítica; la Definición de Conceptos Clave que serán empleados a lo largo del texto; y finalmente, la Contextualización Teórica del Problema de investigación dentro del estado del arte actual.

2.1. Revisión de Literatura Previa

Este punto es crucial para situar el estado del arte del ATD y análisis de siniestralidad en transporte. Una revisión de los trabajos más relevantes en estas áreas revela la falta de estudios que intersecten estos dos campos de trabajo en una ciudad colombiana. Aún así, hay estudios previos de accidentalidad en Cali que demuestran una relación no lineal entre los accidentes y su ubicación [1], además de identificar zonas críticas (puntos calientes o *hotspots*), además de identificar factores como impericia e imprudencia, daños

en la vía y falta de señalización [3]; además como se mencionó anteriormente desde el ATD principalmente se encuentran estudios enfocados a la medicina, biología y ciencias de la computación, a pesar de esta tendencia, de las pocas investigaciones que se pueden encontrar en esta área con una inclinación hacia el tráfico [4] es una comparación entre filtración de VRIN y Vietoris-Rips para identificar cambios significativos en la movilidad poblacional en ventanas de tiempo de una semana. Debido a la desconexión entre estas dos áreas, este artículo sobre accidentalidad desde el enfoque topológico pretende aportar una idea en la intersección entre estas dos áreas.

2.2. Contextualización Teórica del Problema

A pesar de que en [1] se hace un análisis geográfico de los accidentes de tránsito en Cali, no llega a abordar con suficiente amplitud el problema, mientras que en [4] nos da una breve indicación sobre cómo podemos abordar los datos geográficos desde el ATD, tampoco llega a encajar completamente con la intención de este trabajo, así que esto demuestra la pertinencia y novedad de esta investigación, accidentalidad vial en Cali desde el enfoque del Análisis Topológico de Datos.

3. Metodología

La metodología usada usa técnicas de análisis geográfico, teoría de grafos y herramientas de Análisis Topológico de Datos. El proceso consta de _ etapas principales: Limpieza, preprocesamiento de los datos y transformación de palabras, conversión de direcciones a datos geográficos y generación del mapa, generación de complejos Vietoris-Rips y cálculo de homología persistente en dimensión 0 y 1 (H_0 y H_1) mediante distancia euclídea con su diagrama de barras.

3.1. Preprocesamiento, limpieza y transformación

Siendo la base usada, datos rellenados a mano por agentes de tránsito, es normal encontrar errores de reacción, de ortografía o maneras distintas de llamar las vías donde se encontraron los accidentes, por esto primero se omitieron las entradas donde el dato de dirección estaba vacío o no correspondía con un lugar, después se descartaron los datos que no se iban a utilizar para este análisis (tipo de vehículo, medio de reporte, etc.), finalmente se cambiaron algunas palabras por razones como manera de decirles (Autopista Sur - Calle 10, Autopista Simón Bolívar - Calle 25, etc.), errores ortográficos y

abreviaciones (Cra - Carrera, Trans - Transversal, etc.) para que el software de identificación geográfico arrojara resultados sin confusión.

3.2. Conversión a datos geográficos y generación del mapa

Ya teniendo la base de datos con las direcciones y lugares donde están registrados los accidentes, con ayuda del software QGIS estos datos se pasan a coordenadas (latitud y longitud), para posteriormente en un mapa marcarlos y generar una visualización de los accidentes sobre la ciudad.

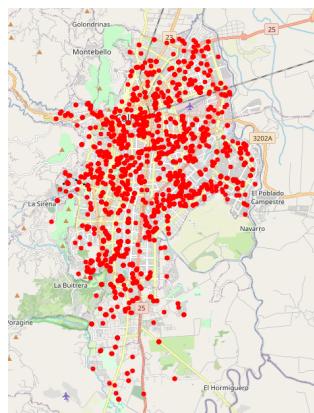


Figura 1: Mapa de los accidentes en Cali

Adicionalmente también generamos un mapa de calor para ver los lugares de mayor concentración (zona crítica, punto caliente o *hotspot*) de los accidentes.

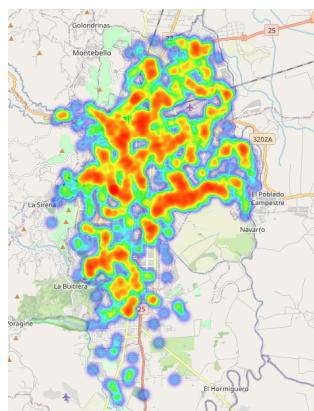


Figura 2: Mapa de calor de accidentes en Cali

3.3. Generación de los complejos

Con los puntos ya sobre el mapa de la ciudad, el paso a seguir es la generación de complejos, para este caso se decidieron usar complejos de Vietoris-Rips porque estos son ideales para realizar un análisis exploratorio en una cantidad de datos media, además de solo ser necesaria la matriz de distancias que es fácilmente calculable teniendo las coordenadas de los puntos y al tener una sensibilidad media al ruido, los accidentes en la periferia de la ciudad no afectarían demasiado el análisis.

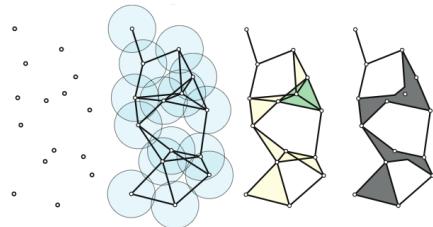


Figura 3: Ejemplo de Complejos Vietoris-Rips

4. Resultados

Inicialmente realizamos análisis con $\varepsilon = 300$ metros, lo que nos dio la siguiente gráfica:

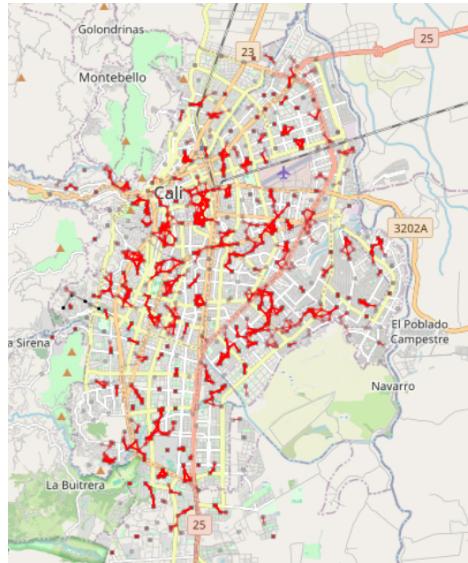


Figura 4: Complejo Vietoris-Rips $\varepsilon = 300$ metros

Luego se hizo con 500 y 800 metros:

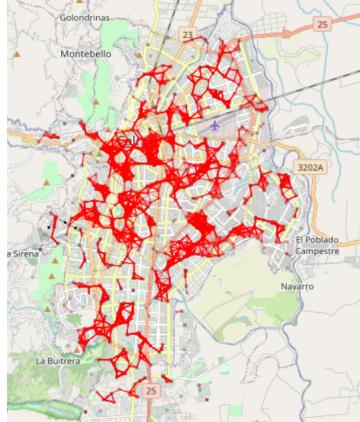


Figura 5: Vietoris-Rips $\varepsilon = 500$ m

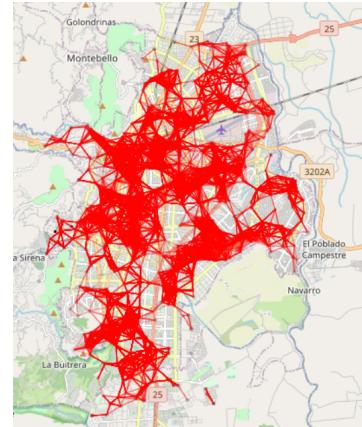


Figura 6: Vietoris-Rips $\varepsilon = 800$ m

De estas gráficas obtenemos los siguientes datos:

ε en metros	Aristas (conexiones a ε)	$ \Delta $ (Número de triángulos)
300	88.289	3'370.120
500	125.325	4'526.744
800	219.873	9'380.779

Cuadro 1: Datos de los Complejos

4.1. Cálculo de homología persistente

Finalmente, cuando ya se tienen las distintas conexiones que se generan frente a la variación del parámetro ε se genera el diagrama de persistencia:

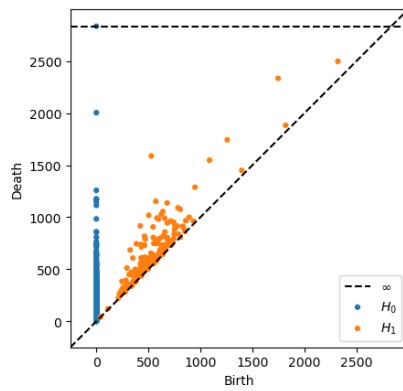


Figura 7: Diagrama de persistencia

5. Análisis y Discusión

El análisis de los complejos Vietoris–Rips generados a partir de los puntos georreferenciados de accidentalidad en Cali revela patrones espaciales consistentes con la estructura vial de la ciudad, pero que además ofrecen una lectura topológica que complementa y amplía lo reportado por estudios tradicionales. La variación del parámetro ε permite observar la forma en que las relaciones de cercanía evolucionan desde lo local hasta configuraciones de mayor escala, lo que facilita la interpretación de cómo los accidentes se organizan en el espacio urbano.

Primero, al trabajar con valores pequeños ($\varepsilon = 300$) el complejo resultante muestra un conjunto altamente fragmentado, con escasas conexiones que tienden a agruparse en torno a sectores específicos. Esto sugiere que los accidentes se presentan inicialmente como eventos aislados, reflejando la naturaleza puntual de la siniestralidad. A pesar de esto, la aparición de pequeñas agrupaciones locales indica zonas donde los accidentes se presentan con mayor densidad, lo cual coincide con la identificación de *hotspots* reportados en la literatura, como frente al colegio Santa Librada y barrios como Antonio Nariño.

Al aumentar el parámetro un poco ($\varepsilon = 500$) la estructura del complejo cambia significativamente, ya que los clústeres comienzan a fusionarse, generando conglomerados de mayor tamaño. Esto sugiere que a una escala de algunos cientos de metros, existen patrones de movilidad y densidad vial que conectan los puntos de accidental en zonas continuas, posiblemente asociadas a tramos viales con alto flujo o intersecciones críticas. La transición entre estructuras dispersas a conglomerados más sólidos refuerza la idea de que la accidentalidad no es un hecho aleatorio, sino que se organiza alrededor de componentes urbanos persistentes.

Finalmente, al llevar el parámetro a una mayor distancia ($\varepsilon = 800$), los complejos muestran una conectividad mucho más densa, aumentando notablemente el número de aristas y triángulos. En este punto, los clústeres locales convergen hacia macrocomponentes que abarcan extensas áreas del casco urbano. Este comportamiento indica que los patrones de accidentalidad están íntimamente relacionados entre sí en escalas superiores, algo que estudios previos sugieren desde enfoques netamente estadísticos, pero que ahora se hace evidente desde la estructura topológica de los datos. Estos resultados sugieren que la morfología de la ciudad (sus vías principales, troncales y vías articuladas) genera zonas conectadas donde la siniestralidad se acumula y persiste en distintas escalas.

En cuanto los diagramas de persistencia:

1. H_0 - Conectividad local

- Los puntos azules se concentran en $Birth \approx 0$, lo que significa que cada accidente empieza como un punto aislado.
- A medida que avanza ε (unos cientos de metros), los puntos comienzan a fusionarse: esto revela la escala a la que los accidentes empiezan a formar conglomerados o clústeres urbanos.
- Las barras o puntos con $Death$ muy alta indican los grupos más grandes y persistentes, típicamente las zonas de mayor concentración de accidentes.
- Algunos puntos azules llegan hasta $Death \approx 2700$ metros, lo cual sugiere que los principales conglomerados de accidentes se unen en escalas de unos 2.5km

2. H_1 - Ciclos o vacíos

- Los puntos naranjas aparecen más dispersos y se alejan de la diagonal hasta $Birth \approx 500 - 1500$ metros, $Death 1000 - 2500$ metros.
- Cada punto representa un “agujero topológico” es decir, un área rodeada pero con poca o nula ocurrencia en el centro, podría indicar la presencia de barrios o parques.

Los conglomerados persistentes en H_0 muestran que la red de accidentes en Cali se estructura en torno a corredores de alta amovilidad, lo cual coincide con los *hotspots* identificados en estudios previos [1] [3]. Sin embargo la persistencia de componentes hasta aproximadamente 2500 metros sugiere una conectividad más amplia que la reportada por modelos geoestadísticos, indicando que la distribución espacial de accidentes puede estar influenciada por patrones urbanos de escala mayor, como zonas céntricas de intercambio vial o ejes troncales.

Finalmente, los resultados muestran que la accidentalidad en Cali no se distribuye de manera uniforme ni aleatoria, sino que exhibe una estructura conectada donde los patrones locales se integran en estructuras mayores a medida que la escala aumenta. Esto respalda la idea de que los fenómenos urbanos asociados a la movilidad y siniestralidad, deben analizarse desde múltiples escalas y con herramientas que captén la variabilidad local y conectividad global. El uso de ATD demuestra ser una herramienta valiosa para este propósito, ofreciendo una caracterización topológica que puede complementar significativamente los estudios tradicionales sobre seguridad vial en contextos urbanos.

6. Conclusiones

El presente trabajo aplicó técnicas de Análisis Topológico de Datos (ATD) para estudiar la estructura espacial de la accidentalidad vial en la ciudad de Cali entre 2016 y 2024. A partir de la construcción de complejos Vietoris–Rips y del análisis de homología persistente en dimensiones 0 y 1, fue posible identificar patrones robustos de conectividad, aglomeración y vacíos urbanos que caracterizan de manera más profunda la distribución espacial de los siniestros viales.

En primer lugar, los resultados muestran que la accidentalidad en Cali presenta una estructura altamente no aleatoria. Los conglomerados persistentes identificados en H_0 revelan que, desde escalas pequeñas ($\approx 300m$) hasta escalas más amplias ($\approx 2700m$), existen zonas donde los accidentes tienden a agruparse de manera consistente. Estos conglomerados coinciden con los corredores viales que la literatura previa reconoce como críticos, lo que valida la pertinencia del enfoque topológico para el estudio urbano de la siniestralidad.

En segundo lugar, la homología en dimensión 1 permitió detectar ciclos que representan áreas rodeadas por accidentes pero con baja ocurrencia en su interior. Estos vacíos topológicos sugieren la presencia de parques, zonas institucionales, conjuntos residenciales o áreas con baja exposición vehicular, ofreciendo una lectura espacial que complementa los estudios tradicionales basados en densidad o regresiones espaciales. La detección de estos patrones aporta evidencia sobre cómo la morfología urbana y la distribución de usos del suelo influyen en la dinámica de los siniestros.

Adicionalmente, el uso de complejos Vietoris–Rips se mostró adecuado para este tipo de análisis debido a la escala intermedia del conjunto de datos y a su balance entre sensibilidad al ruido y capacidad de capturar estructuras significativas. Los valores crecientes de ε permitieron observar transiciones claras entre patrones locales y patrones globales, lo que destaca la utilidad del ATD para estudiar fenómenos urbanos que se manifiestan de manera multiescalar.

Sin embargo, es importante reconocer algunas limitaciones del estudio. La calidad de los datos georreferenciados depende de la coherencia en el registro de las direcciones y del proceso de geocodificación, lo que puede introducir incertidumbre en la ubicación exacta de algunos eventos. Asimismo, el análisis se basó exclusivamente en distancias euclidianas, lo cual no captura completamente la compleja geometría de la red vial de la ciudad.

Aun así, los resultados demuestran que el ATD es una herramienta prometedora para el estudio de la seguridad vial en contextos urbanos. Su capacidad para identificar patrones estructurales sin depender de modelos paramétricos

cos lo convierte en un complemento valioso para métodos georreferenciados (como GWR, OLS o KDE). Este enfoque abre la puerta a nuevas líneas de investigación, entre ellas el análisis temporal de la persistencia de patrones de accidentalidad, la comparación entre distintos tipos de accidentes o la integración de complejos alternativos como los complejos VRIM, filtraciones basadas en grafos viales o implementación con una métrica no-euclídea, además, consideramos que la adición de características temporales de los accidentes y estructurales sobre la vía, podrían ser relevantes al momento de estudiar este tipo de casos, más aún con herramientas como el ATD que permite encontrar patrones en varias dimensiones según la morfología de los datos.

En conclusión, el análisis topológico permitió caracterizar de forma innovadora la estructura espacial de la accidentalidad en Cali, revelando patrones de conectividad, aglomeración y vacíos que capturan la complejidad del fenómeno de manera multiescalar. Estos hallazgos no solo aportan una nueva perspectiva metodológica, sino que también pueden contribuir al diseño de políticas de movilidad más informadas y estratégicas, enfocadas en reducir los riesgos viales en las zonas donde los patrones topológicos muestran mayor persistencia.

Referencias

- [1] Alexis, O.R.: Modelo espacial de muertes por accidentes de tránsito en. Tech. rep. (2004)
- [2] ANSV: Ansv alerta a los municipios con mayor incremento en muertes por siniestros viales (Agosto 2023)
- [3] Arboleda, J.L.A., Zuluaga, J.A.M., Giraldo, D.S.R.: Factores desencadenantes e impacto de accidentes de tránsito por parte de uniformados de la policía metropolitana de la ciudad de Santiago de Cali. Tech. rep. (2024), <https://co.creativecommons.net/tipos-de-licencias/>
- [4] Gonzalez, M.A.F.: Análisis topológico de datos: Estudio de la movilidad poblacional en la ciudad de Hermosillo. Ph.D. thesis, Centro de Investigación en Matemáticas, A.C. (11 2021)