

Modelo de manejo de datos espacio – temporales para la analítica visual en el área de siniestralidad de Bogotá

Camilo Andrés Nemocón Farfán
Secretaría de movilidad
Bogotá, Colombia

Resumen— Este artículo da a conocer el desarrollo una plataforma web que permite realizar consultas espacio – temporales con datos de siniestralidad, a partir de un modelo que administra la información en tres aspectos de espacio, tiempo y datos, así como de sus niveles de granularidad, lo cual permite la visualización, navegación y exploración de los datos de movilidad de la ciudad de Bogotá mediante la analítica visual.

Palabras Claves— Analítica visual, datos espacio – temporales, sistemas urbanos, movilidad, interacción, visualización, modelo de consultas.

I. INTRODUCCION

Bogotá es una de las ciudades con mayor siniestralidad en Colombia, debido al crecimiento acelerado de vehículos y de la ciudad de Bogotá, así mismo desde el 2003 existe una tasa anual de 600 muertos por siniestros viales.

Por tanto, para la generación de propuestas y medidas que permitan reducir los accidentes de tránsito en Bogotá, es necesario ver y evaluar el comportamiento actual de la ciudad, para así responder de una manera más efectiva las solicitudes de información en torno al área de siniestralidad, ya que en la actualidad la metodología y procesos utilizados no permiten realizar un análisis rápido sobre todos los datos que se tienen.

Los datos que usan los expertos en siniestros en la Secretaría de Movilidad de Bogotá (SDM), son datos espacio – temporales, es decir datos que varían o existen dependiendo de un determinado momento en el tiempo y de una específica ubicación [1].

A partir del uso de estos datos, se desarrolla de una herramienta que permita la integración de la información espacio - temporales, siguiendo un protocolo que permite determinar la calidad de los datos, para su posterior uso dentro de un sistema de modelado y visualización de datos [2], ya que a partir de la validación de los datos se pueden tomar decisiones y realizar análisis por medio de esta aplicación web, la cual permite la visualización y exploración de gran cantidad de datos de siniestralidad para identificar las zonas y posibles factores que generan o influyen en los siniestros, filtrando la información de acuerdo a sus atributos como clase, gravedad y localidad, para así analizar y evaluar el comportamiento de los

siniestros en diferentes periodos de tiempo en la ciudad de Bogotá.

II. TRABAJO RELACIONADO

En el área de movilidad y sistemas de información geográficos encontramos trabajos de investigación para la integración de datos espacio – temporales, como es el trabajo Intra-Urban Spacial Connectivity [3], en donde a partir de los datos de ubicación de los vehículos en diferentes horas del día de la semana, se determinan patrones de comportamiento que permiten el análisis de los trayectos y así generar puntos de interés (hogares, trabajos, shopping, centros deportivos, etc.) dependiendo del día y la hora.

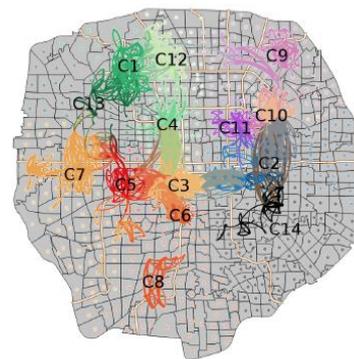


Figura 1. Delineating Intra-Urban Spacial Connectivity Patterns.

Otro proyecto entorno a los trayectos de vehículos se encuentra Urban Traffic Simulation [4], en donde se visualiza la movilidad vehicular permitiendo generar relaciones de origen y destino, los cuales se muestran por medio de una línea con el color del área del destino y los vehículos tienen este mismo color para determinar hacia donde se dirigen y así ver la congestión a partir de la dirección y la cantidad de los vehículos dirigiéndose a estas zonas.

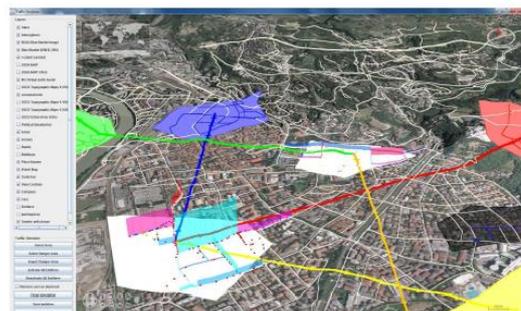


Figura 2. Urban Traffic Simulation.

En cuanto a la visualización del recorrido de los vehículos tenemos el trabajo de Andrienko [5], en donde muestra el trayecto realizado por los vehículos y el tiempo promedio de permanencia en diferentes lugares comparando en estos lugares el tiempo de parada entre camiones y carros mediante un diagrama de barras en cada punto. También muestra los trayectos desde un origen hasta un fin mediante flechas, donde el grosor de la flecha corresponde al valor de algún atributo (velocidad promedio, volumen de vehículos, modo de viaje) seleccionado por el usuario.

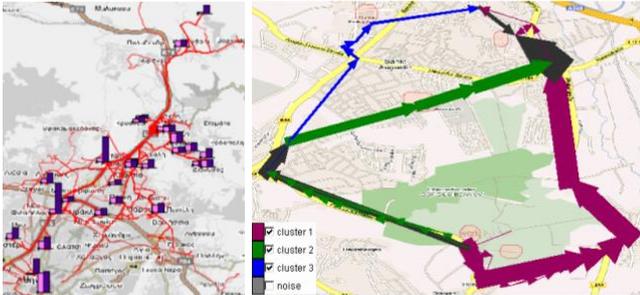


Figura 3. Analysis of Movement Data.

Estos mismos autores continúan su investigación con la agregación espacio – temporal [6], donde el mapa se divide en una cuadrícula en la cual se visualizan los datos a través de barras direccionales, donde el sentido y color de cada barra muestra la dirección de los carros (analogía rosa de los vientos) y donde la longitud de la barra representa la cantidad de carros en esa dirección. En cada punto se puede ver también un histograma con cuadrados de colores que representan la velocidad media que se presentó en cada punto en un periodo de tiempo determinado por el usuario.

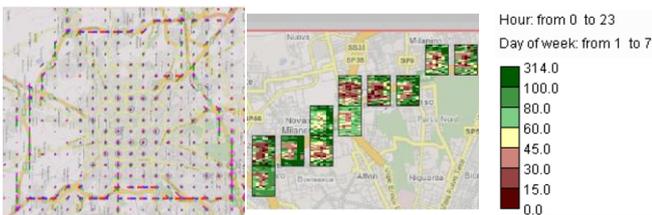


Figura 4. Spatio-Temporal Aggregation.

METODOLOGÍA

Esta investigación se centra en la producción de una herramienta interactiva de analítica visual, la cual se desarrolla en torno al diseño centrado en el usuario generando un proceso iterativo donde a partir de las pruebas que se realicen con los expertos, se va implementando funcionalidades dentro de la herramienta.

Los datos a manejar y visualizar son datos espacio – temporales, los cuales le permiten a los usuarios identificar y ubicar factores importantes, para que el experto pueda ver el comportamiento y analizar la siniestralidad de la ciudad a partir de la segmentación por periodos de tiempo (año, mes, día y hora), permitiendo el análisis, comparación y visualización de la información, mediante la plataforma web con la cual se puede realizar de forma fácil las consultas y visualizar los datos, apoyando las tareas de análisis de exploración de datos para los expertos.

MODELO E IMPLEMENTACIÓN

MODELO DE MANEJO DE DATOS:

A partir de un proyecto previo para la exploración de datos espacio – temporales [7] se implementa este sistema de visualización interactivo que consta de un modelo que administra la información en una estructura de desacoplamiento en los tres aspectos de espacio, tiempo, y datos, así como de sus niveles de granularidad. De esta manera, los niveles pueden ser mostrados y explorados de lado a lado para facilitar su análisis combinado. Este modelo permite la interacción por medio del semantic zoom y desagregación de los datos viéndolos desde diferentes niveles.

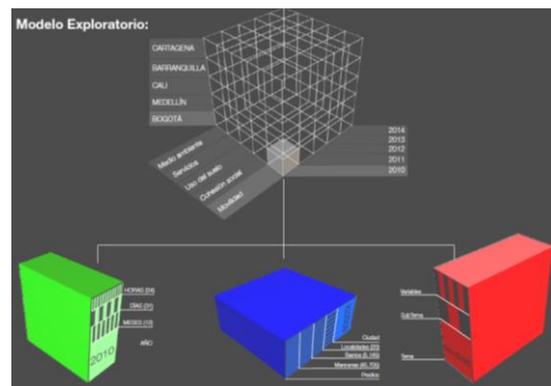


Figura 5. Modelo Exploratorio.

Con base en este cubo puedo generar una estructura de selección y exploración, ya que permite establecer una navegación para la lectura, búsqueda y filtrado en datos espacio temporales y sus diferentes granularidades, donde cada celda en el cubo representa un objeto espacial asociado a un atributo de siniestralidad (actores, gravedad, clase) y a un periodo en el tiempo, dentro del cual hay niveles de granularidad.

La navegación por el cubo a nivel general se puede realizar sobre el eje “X”, para la exploración de atributos de siniestros, sobre el eje “Y” para los lugares y sobre el “Z” podemos encontrar el tiempo y permite la selección y filtro de datos.

Este modelo se puede representar en forma de grafo, en donde el núcleo del grafo está compuesto por los componentes de Espacio, Tiempo y Atributos, los cuales poseen sus propios niveles de granularidad, generando una estructura basada en un árbol por cada componente, representando de esta forma la jerarquía de las granularidades de los objetos del modelo con un sistema no direccional entre los componentes, lo que permite la relación libre del núcleo del grafo y la producción de consultas a partir de las respuestas de otras consultas, permitiendo una exploración de los datos de acuerdo al interés de los expertos.

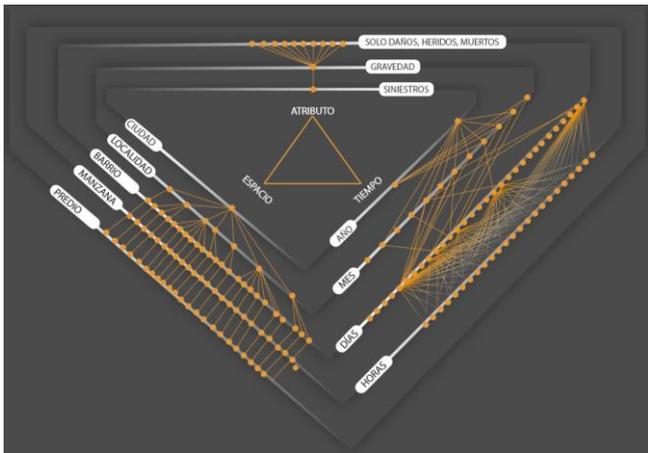


Figura 6. Grafo del sistema.

El modelo permite generar respuestas sobre los filtros de los datos realizados por el usuario, y sobre estos resultados se pueden realizar nuevas consultas, permitiéndole al usuario analizar los resultados con otros parámetros, planteando nuevas hipótesis y generar un análisis sobre cada una. Este sistema de grafos y consultas permite devolverse a anteriores consultas para retomarlas y plantear otras hipótesis con otros parámetros, generando finalmente un historial de las consultas realizadas por cada usuario.

Para la implementación del modelo en el área de siniestralidad se utilizaron datos espacio temporales enfocados específicamente a los accidentes de tránsito en la ciudad de Bogotá. Los datos corresponden a los años del 2007 al 2017, donde de cada año encontramos un componente espacial representado por la longitud y latitud donde ocurrió el siniestro, un componente temporal dado por un intervalo de tiempo de 24 horas, con los días de cada mes, desde enero hasta diciembre, de cada año.

Finalmente se encuentran los atributos compuestos por cuatro categorías, por un lado, encontramos la clase, la cual posee las variables de atropellos, autolesión, caída de ocupante, choque, incendio, otro y volcamiento. Otro atributo es la gravedad, constituida por las variables de solo daños, con heridos, con muertos. Por otro lado, encontramos el atributo de localidades, en donde cada dato del siniestro dice la localidad en donde éste ocurrió. El último atributo son los actores implicados en el siniestro, donde encontramos, automóviles, bicicletas, buses, camionetas, motocicletas entre otras.

Por tanto, se generó una base de datos compuesta por 3 tablas, obedeciendo al modelo propuesto, donde una tabla es la del tiempo, compuesta por el “Tiempo” con sus granularidades (id, año, meses, días, horas) con cerca de 77.500 registros, otra tabla denominada “Espacio” (id, código de la localidad, dirección del siniestro, latitud, longitud) constituida por 586.042 registros, y una última tabla que es “Atributos” (actores, gravedad y clase) la cual posee también 586.042 registros.

IMPLEMENTACIÓN:

A partir del modelo exploratorio definido anteriormente, se construye una aplicación web de visualización e interacción para la formulación de consultas en datos espacio – temporales, el cual posee componentes que se encargan de construir y manejar los núcleos del grafo del sistema y la relación entre ellos permitiendo generar consultas y responder a estas de forma visual e interactiva.

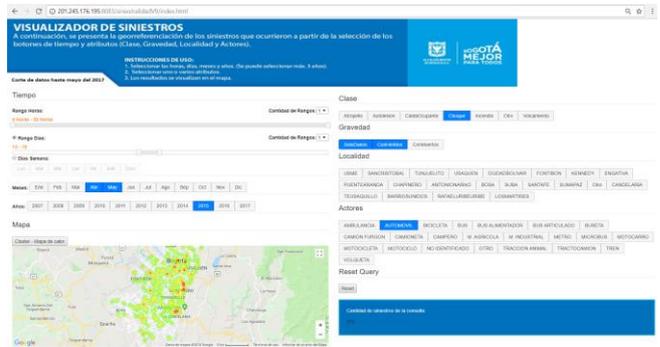


Figura 7. Aplicativo web. (<http://201.245.176.195:8083/siniestralidadV9/index.html>)

La visualización web está compuesta en la parte superior por las instrucciones, y le continúan dos áreas de consulta, una al costado derecho, en donde encontramos la selección de atributos de clase, gravedad, localidad y actores con sus correspondientes variables.

Clase						
Atropello	Autolesion	CaídaOcupante	Choque	Incendio	Otro	Volcamiento
Gravedad						
SoloDaños	ConHeridos	ConMuertos				
Localidad						
USME	SAIVCRISTOBAL	TUNJUELITO	USAQUEN	CIUDADBOLIVAR	FONTIBON	KENNEDY
ENGATNA	PUENTEARANDA	CHAPINERO	ANTONIOVARINO	BOSA	SUBA	SANTAFE
Otro	CANDELARIA	TEUSAQUILLO	BARRIOSUNIDOS	RAFAELURIBEURBE	LOSMARTIRES	
Actores						
AMBULANCIA	AUTOMOVIL	BICICLETA	BUS	BUS ALIMENTADOR	BUS ARTICULADO	BUSETA
CAMION FURGON	CAMIONETA	CAMPERO	M. AGRICOLA	M. INDUSTRIAL	METRO	MICROBUS
MOTOCARRRO	MOTOCICLETA	MOTOCICLOTA	NO IDENTIFICADO	OTRO	TRACCION ANIMAL	TRACCIONCAMION
TREN	VOLQUETA					

Figura 8. Área de consulta de atributos.

Mediante estos botones el usuario puede seleccionar uno o más categorías del tipo de siniestro para así observar la cantidad de siniestros ocurridos por cada categoría.

Al costado izquierdo encontramos el filtrado de información temporal en donde se puede realizar consultas temporales en sus distintas granularidades, por tanto, en la parte inferior de esta área, se encuentran los botones correspondientes a los años, donde el sistema coloca automáticamente la cantidad de botones de acuerdo a la cantidad de años que hay en la base de datos, luego encontramos la granularidad de los días, que permite seleccionar un día a la semana, un día del mes o un rango de días del mes, así mismo permite generar varios slider para poder seleccionar varios rangos de días en el mes, así mismos como rangos de horas, permitiendo la mayor flexibilidad en la selección temporal. Sin embargo, para el uso de éste sistema se debe realizar la consulta con un orden específico el cual es la selección, de horas, días,

mese y finaliza con el año, generando la consulta cuando se termina la selección del año.

Figura 9. Área de consulta temporal.

La última parte de la visualización esta compuesta por el mapa de resultados, en donde se evidencia puntualmente los lugares de interés para el usuario y muestra la información individual de la ubicación de cada uno de los siniestros a partir de la selección de tiempo y atributos realizada por el experto.

Este componente permite ver un mapa de calor para así proporcionar información de la concentración de siniestros o también posee la opción de mostrar clusters, en donde agrupa los siniestros a partir del semantic zoom que realiza el usuario evidenciando y mostrando el número de siniestros dentro del cluster.

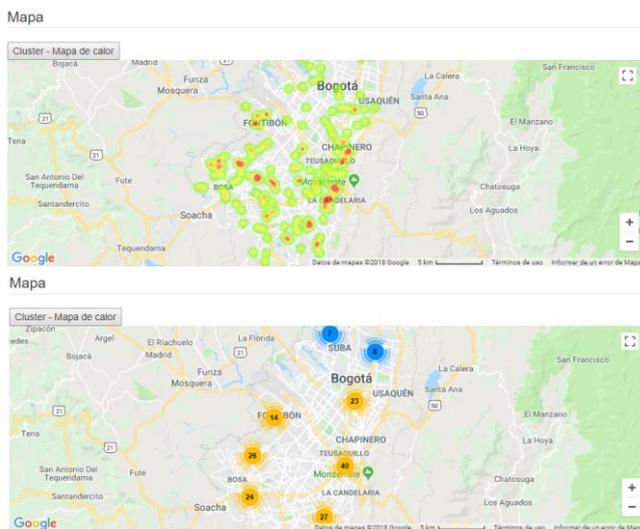


Figura 8. Área de consulta espacial.

Finalmente, a costado derecho inferior encontramos la cantidad de siniestros, donde es éste el resultado de la consulta, y realmente es el dato más usado por los expertos en siniestralidad, donde les piden constantemente la cantidad de siniestros ocurridos por localidades, actores, clase o gravead y en distintos periodos de tiempo y este valor corresponde a los mismos siniestros geo-referenciados en el mapa.



Figura 9. Área de valor total de siniestros.

III. CONCLUSIONES Y TRABAJO FUTURO

En conclusión, el modelo de manejo de datos permite realizar consultas rápidas y definidas de una gran cantidad de datos, y permitió su implementación en una herramienta web que tiene un área de consultas y un área de respuestas que va acorde con las funciones planteadas por el modelo.

Así mismo tanto la aplicación web como el modelo de manejo de datos permitieron ver una vista general de los datos que se encuentran en la base de datos, permitiéndole al usuario evaluar cuales son los datos que posee en cada una de sus granularidades, permitiéndole al experto una introducción de los datos que posee para realizar su análisis e informes de seguridad vial,

La aplicación muestra los resultados de las consultas en un tiempo promedio de 60 milisegundos, sin embargo, el usuario debe realizar las consultas en orden.

Esta herramienta permitió cargar y visualizar 586.041 accidentes viales georreferenciados ocurridos entre el año 2007 y 2017 en Bogotá.

Como trabajo a futuro se propone aumentar las funcionalidades gráficas para la representación de los resultados a parte del diagrama de barras, scatterplot, entre otros elementos, utilizando la librería de d3js. Así mismo se plantea realizar consultas sobre los resultados visualizados en el mapa.

Generar un historial y un sistema de anotaciones en donde se puede evidenciar todas las decisiones de consulta y sus respectivas respuestas para que el usuario pueda realizar nuevas consultas a partir de las consultas realizadas por otro experto.

REFERENCIAS

- [1] Space, time and visual analytics Space, Gennady Andrienko*, Natalia Andrienko, 2010.
- [2] Visualization of Time-Oriented Data. Wolfgang Aigner, S. M, 2011.
- [3] Delineating Intra-Urban Spatial Connectivity Patterns by Travel-Activities: A Case Study of Beijing, China. Chaogui Kang, 2015.
- [4] Analytics Tool For Urban Traffic Simulation, Alberto Debiasi, 2013.
- [5] Visual Analytics Tools for Analysis of Movement Data, Gennady & Natalia Andrienko, 2007.
- [6] Spatio-temporal Aggregation for Visual Analysis of Movement, Gennady & Natalia Andrienko, 2008.
- [7] Exploración de datos espacio – temporales para la analítica visual en sistemas urbanos, Camilo Nemocón, 2017.