

Exploración de datos espacio - temporales para la analítica visual en sistemas urbanos.

Camilo Andrés Nemocón Farfán
 José Tiberio Hernández
 Grupo de Investigación Imagine
 Dpto. Ingeniería de Sistemas y Computación
 Universidad de Los Andes
 Bogotá, Colombia

Resumen— Este artículo da a conocer el desarrollo una plataforma web que permite realizar consultas espacio - temporales, a partir de un modelo que administra la información en tres aspectos: 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, sistemas urbanos, movilidad, interacción, visualización, modelo de consultas.

I. INTRODUCCION

El crecimiento acelerado de la población y de la ciudad de Bogotá, requiere de una planeación urbana que satisfaga las nuevas necesidades de servicios e infraestructura de la ciudad, teniendo en cuenta los planes de gobierno, los cuales cambian constantemente y dependen de las políticas elaboradas por cada mandato.

Por tanto, para un buen desarrollo y planeación urbana en torno al crecimiento de la ciudad es necesario ver y evaluar el comportamiento actual de la ciudad para así tomar las mejores decisiones, planeaciones y políticas en torno a los sistemas urbanos. Para ello los expertos necesitan visualizar simultáneamente gran cantidad de datos y explorarlos para descubrir patrones, relaciones, tendencias u otras características que apoyen su posterior proceso de análisis, planteamiento de hipótesis y toma de decisiones.

Los datos que usan los expertos en planeación urbana, 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], ya que a partir del plan de desarrollo realizado en el tiempo de mandato de un gobierno se sacan datos que evidencian las actividades realizadas, las cuales pueden ser visualizados por medio de sistemas de información geográficos interactivos en plataforma web para explorar estos datos.

Para el proceso de exploración de datos espacio – temporales, se necesita una herramienta de visualización interactiva web que le permita al experto la selección de los factores que son importantes para él y evidenciar su

comportamiento en el espacio y en el tiempo, lo cual, le permite tener una visión general de la situación, con el propósito de poder ver el estado actual y la evolución de la ciudad en términos de sistemas urbanos.

Por tanto es importante generar una herramienta que permita la integración de los datos espacio - temporales dentro de un sistema de información geográfica, siguiendo un protocolo que permita 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 aplicación web que vinculen la analítica visual para generar hipótesis y conclusiones que apoyen la tarea de planeación urbana.

II. TRABAJO RELACIONADO

En el área de sistemas urbanos y analítica visual encontramos trabajos de investigación para la integración de sistemas de información con fines de simulación y planeación, como es el trabajo de Wang cuya aplicación llamada CAL3QHC [3], integra los sistemas de información geográfica como ArcGis, con simuladores de Trafico y simuladores de calidad de aire, generando visualizaciones pertinentes para los expertos, sin embargo el usuario debe generar los datos en cada simulador para ser integrados en la aplicación, lo cual genera un aumento en el tiempo para la preparación de los datos para la visualización de cada tema a analizar y no permite la visualización web centrado y limitando la información y la toma de decisiones.

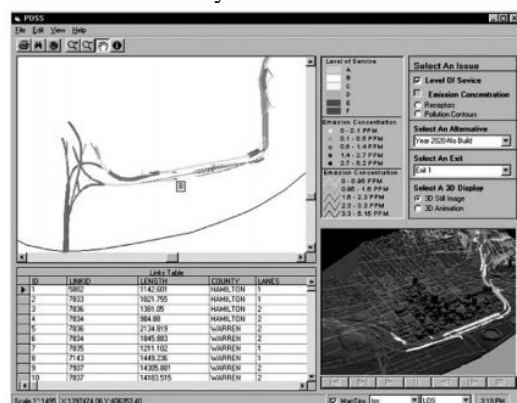


Figura 1. CAL3QHC Sistema de análisis del tráfico.

Otro proyecto de vinculación de analítica visual y Big data, es VisGets [4], en donde se centran en el desarrollo de una interfaz web para formular consultas sobre los datos, donde muestra un listado con las respuestas de las consultas utilizando widgets de consulta interactivos, sin embargo la visualización de las respuestas no es clara y es limitado para la visualización de datos heterogéneos.



Figura 2. Interfaz de VisGets.

También encontramos otra aplicación llamada Europeana Connect [5], en donde se puede obtener información relevante frente a tópicos históricos dependiendo de la ubicación y el momento en el tiempo consultado, permitiendo visualizar espacial y temporalmente la información de un tópico específico, sin embargo, no permite la relación de temas que podrían ser importantes para el usuario.

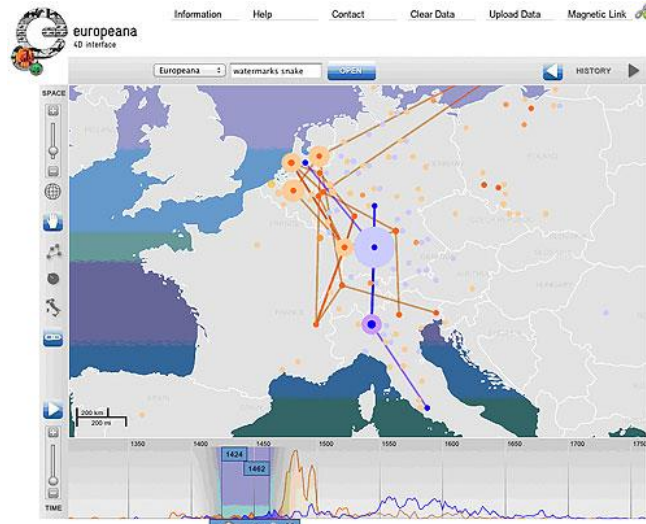


Figura 3. Interfaz de Europeana Connect.

Por otro lado un proyecto que plantea no solo una visualización interactiva sino un modelo de manejo de datos enfocado hacia la planeación urbana es el desarrollado por J.Triana [6] mediante un modelo que integra información Visum, Arc Gis y Excel para la visualización y combinación de los escenarios del flujo del tráfico con el uso del suelo.

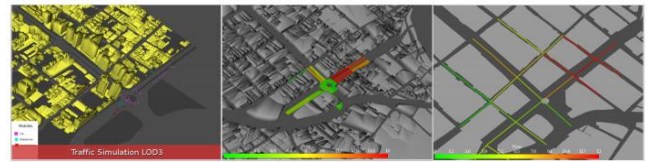


Figura 4. Simulación del tráfico.

III. DISEÑO E IMPLEMENTACIÓN

DISEÑO DEL MODELO DE MANEJO DE DATOS:

Se utilizó como base y referencia el modelo de consultas de D. Peuquet [7], en donde permite considerar todos los componentes de los datos y ayuda a expresar restricciones de consulta espacio - temporales por medio de las preguntas que facilitan la expresión de las consultas. A partir de éste, se desarrolló un modelo que maneja los datos permitiendo la navegación, filtrado y exploración de datos manejando cada uno de sus componentes de forma independiente, con lo cual el experto en el área de sistemas urbanos puede seleccionar tres tipos de componentes de interés (espacio, tiempo y atributos) para observar los datos de un periodo de tiempo a otro, a partir de una selección de una o varias ubicaciones en donde se visualizan únicamente las variables de contexto de interés.

El modelo Exploratorio es en una estructura que genera el desacoplamiento del espacio, tiempo y atributos y sus niveles de granularidad, permitiendo la visualización y exploración para la lectura, búsqueda y filtrado de los datos por medio del semantic zoom y la desagregación de los datos permitiendo acceder a éstos desde sus diferentes niveles.

El modelo apoya la tarea de análisis de exploración de datos mostrando la información solicitada y eliminando de la vista del usuario los datos que no se ajusten a las restricciones de consulta, permitiendo además la combinación entre los niveles de granularidad que poseen los datos temporales y espaciales ya que cada celda del cubo representa un objeto espacial asociado a un atributo en un paso en el tiempo y dentro de cada uno se puede mostrar sus niveles de granularidad.

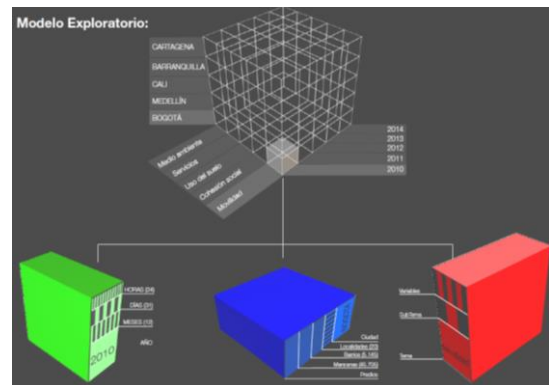


Figura 5. Modelo Exploratorio.

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.

IMPLEMENTACIÓN Y APLICATIVO WEB:

La herramienta interactiva web permite la implementación del modelo para resolver la tarea de análisis de la exploración de los datos dando resultados visuales e interactivos de los datos. La exploración de los datos se realiza a través de consultas mediante el uso de preguntas, las cuales se generan a partir del filtrado de los ejes temporales, espaciales o de atributos y de la selección de uno o varios datos en sus distintas granularidades.

El desarrollo de la aplicación web se realizó con base de datos SQLite utilizando procesos que permiten una descarga y búsqueda rápida de las consultas que pueda realizar el usuario sobre una base de datos que inicialmente tenía un peso de 98MB y se optimizó a 18MB, esta optimización se realiza mediante el navegador, el cual descarga la versión comprimida, permitiendo una carga de la página rápida y luego se descomprime con Javascript, mostrándole al usuario un “overlay” o capa de inicio en donde se muestra un temporizador, que evidencia el progreso de la descarga y descompresión de los datos, y la configuración de la base de datos para posteriormente empezar a interactuar con la página.

Las tecnologías utilizadas fueron Javascript y SQL.js, que se implementaron junto a SQLite, donde éste sistema es finalmente un manejador de base de datos en memoria donde todo el proceso corre en memoria por el lado del cliente, de ésta forma todas las consultas y procesos se realizan en el navegador del cliente y el servidor solo funciona como un repositorio de archivos e información, ya que toda la lógica se implementa en el index.html de la página web.

El sistema utiliza los índices de las tablas de la base de datos para realizar las consultas permitiendo una velocidad de respuesta entre consulta y consulta de 200 milisegundos o menos, mediante los id de espacio, tiempo y atributos, y utilizando como componente un time out que se resetea cuando se interactúa con cualquier elemento de la interfaz. La interfaz se realizó con Bootstrap, el cual permite darle la estructura a la página para que sea responsive y me permita ubicar los elementos de consulta y respuesta. Así mismo se utilizó la librería de JQuery para el manejo de los componentes gráficos e interactivos de la interfaz.

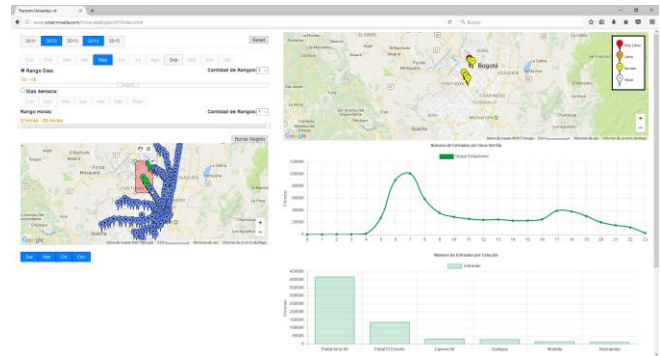


Figura 6. Aplicativo web.
(www.wiseinmedia.com/UniandesExploraST/index.html)

Basados en el proyecto “Visualization of origin-destination matrices” [8] donde se analiza el comportamiento del tráfico y el transporte a partir del uso de dos mapas coordinado y su conexión a un bar chart, se plantea el área espacial y el mapa focalizado, generando una interacción directa entre los datos generales y los filtrados por el usuario que se ven reflejados en los gráficos de utilidad.

Por tanto la visualización web está compuesta por dos áreas, al costado izquierdo encontramos el área de consultas y al costado derecho está el área respuestas, donde en el área de consultas el usuario se centra en la selección de ítems de espacio, tiempo y atributos para filtrar de esa forma la información que es mostrada en el área de resultados, en donde se evidencia puntualmente los lugares de interés para el usuario y muestra la información individual de estos en la parte superior de esa área y en la parte inferior se muestran los resultados en conjunto, para que el usuario pueda comparar los resultados de su selección y exploración de forma conjunta vs. el resultado individual de cada componente espacial seleccionado.

Área de consultas:

En el área de consultas se ubican los tres tipos de componentes de interés (espacio, tiempo y atributos) con sus respectivas granularidades, permitiendo al usuario tener una vista general de los datos que posee y sobre estos poder realizar las consultas en los tres componentes, donde cada componente tiene una forma de visualización e interacción diferente para realizar consultas y filtrados sobre los datos.

-Temporal:

 This screenshot shows the temporal query interface. At the top, there are buttons for the years 2011, 2012, 2013, 2014, and 2015, along with a 'Reset' button. Below this is a month selection bar with buttons for 'Ene', 'Feb', 'Mar', 'Abr', 'May', 'Jun', 'Jul', 'Ago', 'Sep', 'Oct', 'Nov', and 'Dic'. The 'May' button is currently selected. Underneath, there are two main filter sections. The first is 'Rango Dias' (Day Range), which includes a radio button for 'Dias Semana' (Week Days) and a range selector currently set to '13 - 15'. To the right of this is a dropdown menu for 'Cantidad de Rangos' (Number of Ranges) with options 1, 2, 3, 4, and 5. The second filter section is 'Rango Horas' (Hour Range), which includes a radio button for 'Dias Semana' and a range selector currently set to '7 horas - 14 horas'. To the right of this is another dropdown menu for 'Cantidad de Rangos' with options 1 and 2.

Figura 7. Área de consulta temporal.

Este componente se basó en el proyecto STRAD (Spacio-Temporal-Radar) [9] en donde se puede realizar consultas

temporales en sus distintas granularidades, por tanto, en la parte superior 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 y utilizando los modos exploratorios del modelo.

-Espacial:

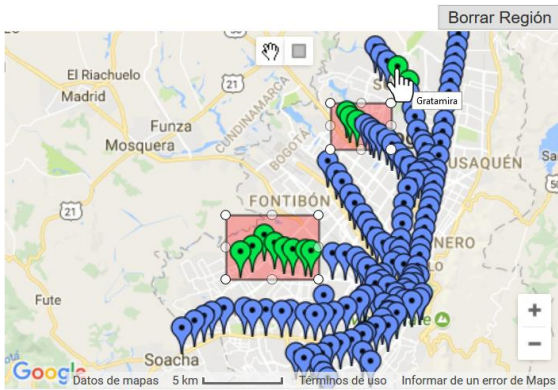


Figura 8. Área de consulta espacial.

Aquí nos permite ver los puntos espaciales de las cuales poseemos datos, ubicándolos sobre un mapa, a partir de la latitud y la longitud. Los puntos rojos corresponden a los puntos seleccionados por el usuario y los azules los que no han sido seleccionados.

El usuario tiene dos opciones de selección espacial, una permite la selección puntual de ubicaciones específicas y la otra selección permite la selección de ubicaciones dentro de un área rectangular.

-Atributos:



Figura 9. Área de consulta de atributos.

Aquí encontramos los botones correspondientes a los atributos de los datos, donde, en nuestro caso de estudio se encuentran los accesos a las estaciones. Mediante estos botones el usuario puede seleccionar uno o más accesos para observar la cantidad de personas que ingresaron por cada acceso.

Área de Resultados:

En el área de resultados esta compuesto por dos elementos un mapa focalizado y un sector de gráfico de utilidad que se encuentran abajo del mapa. Ambos sectores se activan simultáneamente cuando el usuario realiza las consultas de espacio, tiempo y atributos, evidenciando los resultados

generados por cada selección que se realiza en el área de consulta.

-Mapa focalizado:

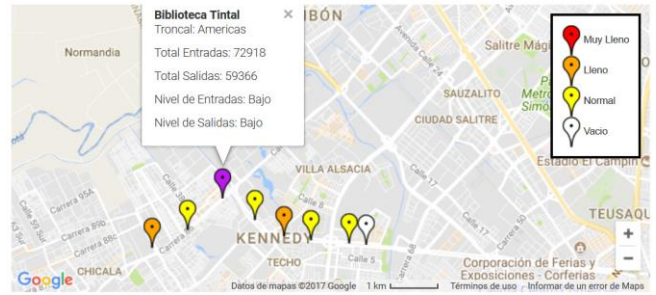


Figura 10. Mapa focalizado.

En la parte superior del área de resultados se encuentra un mapa en donde se visualizan las ubicaciones espaciales de interés del usuario, es decir, las estaciones seleccionadas por el usuario en el área de consulta espacial. De esa forma aparecen las estaciones con la información de interés para el usuario de forma individual, por tanto el usuario puede seleccionar cada uno de estas estaciones y le mostrará las características y la información puntual de cada ubicación.

Así mismo cada punto espacial visualizada tiene un color que determina la cantidad de un atributo ordenando de mayor a menor, el valor del atributo en cada punto espacial, de esta forma el atributo con mayor valor generará un color rojo, el siguiente atributo con un valor menor tendrá el color naranja, amarillo y blanco el atributo con menor valor entre los puntos espaciales seleccionados y comparados.

-Gráficos de utilidad:

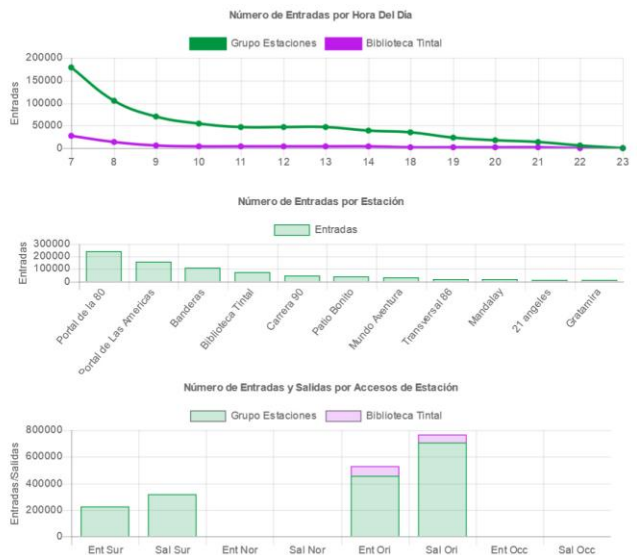


Figura 11. Gráficos de utilidad.

Las gráficas de utilidad son los diagramas que se encuentran abajo del mapa de resultados, donde estas gráficas permiten distintos análisis sobre los mismos datos, lo cual permite ver diferentes enfoques de la información y su retroalimentación automática de acuerdo a la interacción que se tenga sobre el área de consulta y el mapa focalizado.

Las gráficas de utilidad representan el resultado colectivo o total de las ubicaciones seleccionadas por el usuario versus los atributos seleccionados por el usuario, sin embargo estas cambian dinámicamente, cuando el experto toca una ubicación en el mapa focalizado, ya que en el área de utilidad se va a ver reflejado los resultados de esa ubicación en particular vs. los datos acumulados de todas las ubicaciones seleccionadas, para así evaluar el comportamiento de la ubicación en comparación con las demás.

IV. CONCLUSIONES Y TRABAJO FUTURO

La herramienta web cumplió con el objetivo de implementar el modelo y permitir al usuario realizar consultas complejas y detalladas mediante una expresividad visual, a través del diseño de la visualización y el uso de controles intuitivos y fáciles de utilizar.

El sistema web es versátil y liviano, por tanto, necesita de poca infraestructura, bajo presupuesto y bajo tiempo de montaje.

El modelo de consultas existentes permitió ver una vista general de los datos en todas sus granularidades, sin tener que entrar a explorarlos, permitiéndole al experto una introducción de los datos que posee para realizar su análisis.

Como trabajo se propone aumentar las funcionalidades gráficas para la representación de los resultados a parte del diagrama de barras y el scatterplot.

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] Gennady Andrienko*, Natalia Andrienko: Space, time and visual analytics Space, 2010.
- [2] Wolfgang Aigner, S. M. *Visualization of Time-Oriented Data*. Springer, 2011.
- [3] Xinhao Wang, “Integrating GIS, simulation models, and visualization in traffic impact analysis”, *Computers, Environment and Urban Systems*, pp. 496, 2005.
- [4] M. Dork, S. Carpendale, “VisGets: Coordinated Visualizations for Web-based Information Exploration and Discovery,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 14, no.6, Oct 2008.
- [5] Max Kaiser, “EuropeanaConnect,” *eContentplus* programme, a multiannual Community program to make digital content in Europe more accessible, pp. 65, 1994.
- [6] Triana, J. A., Zeckzer, D., & Hernandez, J. T. (2013, August). A novel data model to empower a visual analytics platform for urban systems. In *Computing Colombian Conference (8CCC), 2013 8th* (pp. 1-6). IEEE.
- [7] D. Peuquet, “It’s about time: Framework for the representation of temporal dynamics in geographic information systems”, *Annals of the Association of American Geographers* Vol. 84, No. 3, Sep. 1994.
- [8] Ibarra, J. C., Triana, J. A., & Hernández, J. T. (2016, April). Visualization of origin-destination matrices using a connection barchart and coordinated maps. In *Pacific Visualization Symposium (PacificVis), 2016 IEEE* (pp. 189-193). IEEE.
- [9] Fernandez-Prieto, D., Naranjo-Valero, C., Hernández, J. T., & Hagen, H. (2017). STRAD Wheel: Web-Based Library for Visualizing Temporal Data. *IEEE Computer Graphics and Applications*, 37(2), 99-105.