

INSTITUTO TECNOLÓGICO DE COSTA RICA

BASES DE DATOS GEOESPACIALES

Alejandro Rojas
Saúl Zamora

profesor
Kevin Moraga

1 Bases de datos espaciales

Las bases de datos espaciales son bases de datos optimizadas para almacenar datos que representan objetos definidos sobre un espacio geométrico y realizar consultas sobre dichos datos. La mayoría de las bases de datos espaciales permiten representar objetos geométricos simples como puntos, líneas y polígonos. Algunas permitan el manejo de estructuras más complejas como objetos 3D, coberturas topológicas, redes lineares y redes irregulares trianguladas. Mientras que las bases de datos tradicionales han sido desarrolladas para manejar múltiples tipos de datos numéricos y de caracteres, dichas bases de datos requieren funcionalidad adicional para procesar datos espaciales eficientemente y generalmente se agregan tipos de datos como *geometry* o *feature* para manejar datos espaciales.

2 Consultas espaciales

Consultas espaciales son el tipo de consulta manejada por los sistemas de bases de datos espaciales. Dichas consultas difieren de las no-espaciales en dos aspectos importantes:

- Las consultas espaciales usan datos geométricos como puntos, líneas y polígonos.
- Estas consultas consideran la relación espacial entre las geometrías mencionadas.

3 Tipos de consultas espaciales

Los nombres de funciones para consultas difieren de un sistema a otro. La siguiente lista contiene funciones comúnmente añadidas a *PostGIS*, el cual es una extensión de PostgreSQL para el manejo de datos espaciales (el término *geometry* se refiere a un punto, línea, caja u otra figura de 2 o 3 dimensiones):

- Distance(geometry, geometry) : number
- Equals(geometry, geometry) : boolean
- Disjoint(geometry, geometry) : boolean
- Intersects(geometry, geometry) : boolean
- Touches(geometry, geometry) : boolean
- Crosses(geometry, geometry) : boolean
- Overlaps(geometry, geometry) : boolean
- Contains(geometry, geometry) : boolean

- `Length(geometry)` : number
- `Area(geometry)` : number
- `Centroid(geometry)` : geometry

4 Bases de datos geoespaciales

Una base de datos geoespacial es una base de datos que almacena datos geográficos, tales como países, divisiones administrativas, ciudades e información relacionada. Dichas bases de datos pueden ser útiles para sitios web que deseen identificar la ubicación de sus visitantes con el fin de personalizar más la visita.

5 Características de las bases de datos espaciales

Los sistemas de bases de datos usan índices para buscar rápidamente valores y la forma en la que la mayoría de sistemas indexan la información no es óptima para consultas espaciales. En lugar de eso, las bases de datos espaciales utilizan índices espaciales para acelerar las operaciones de base de datos.

Adicionalmente a las consultas típicas de SQL como *SELECT*, las bases de datos espaciales pueden realizar una amplia variedad de operaciones espaciales:

- Medidas espaciales: procesa largo de líneas, áreas poligonales, distancias entre geometrías, etc.
- Funciones espaciales: modificar las características existentes de un objeto para crear nuevas, por ejemplo, agregando un buffer alrededor de un objeto, etc.
- Predicados espaciales: permite consultas de falsoverdadero sobre relaciones entre objetos, como la superposición de un objeto sobre otro, la ubicación de un punto dentro de un área.
- Constructores geométricos: creación de nuevas geometrías, usualmente mediante la especificación de los vértices que definen la figura.
- Funciones de observación: consultas que retornan información específica sobre una figura, como la ubicación del centro de un círculo.

Algunas bases de datos soportan solo versiones simplificadas o modificadas de estas operaciones, especialmente en casos como los de sistemas NoSQL como MongoDB y CouchDB.

6 Índices espaciales

Los índices espaciales son los utilizados por las bases de datos espaciales para optimizar las consultas. Los índices convencionales no manejan las consultas espaciales de forma eficiente, tales como que tan lejos está un punto de otro o si todos los puntos caen dentro de un área determinada. Los métodos de indexación espacial más comunes incluyen:

- R-tree: método preferido para la indexación de datos espaciales. Los objetos (formas, líneas y puntos) son agrupados usando un *rectángulo de enlace mínimo* o MBR por sus siglas en inglés (*Minimum Bounding Rectangle*). Los objetos se agregan al MBR dentro del índice, lo que lleva a que su tamaño aumente de forma mínima.
- HHCode
- Grid
- Z-order
- Quadtree
- Octree
- UB-tree
- R+ tree
- R* tree
- Hilbert R-tree
- X-tree
- kd-tree
- m-tree
- Método de punto de acceso (*Point access method*)
- Particionamiento espacial binario o BSP-tree por sus siglas en inglés *Binary space partitioning*: subdividir el espacio en hiperplanos.

Figure 1: Ejemplo de R-tree para rectángulos bidimensionales

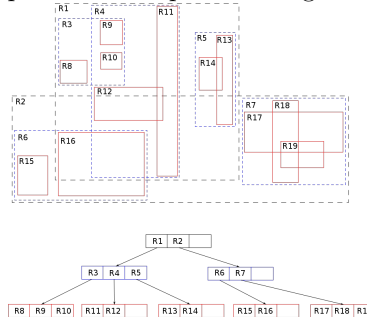
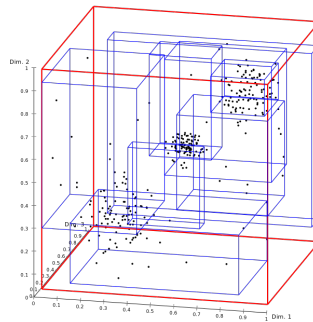


Figure 2: Visualización de un R*-tree para puntos 3D



7 Usos de las bases de datos espaciales

Uno de los usos más populares para las bases de datos espaciales son los sistemas de información geográfica o GIS (por sus siglas en inglés *Geographic Information Systems*) son sistemas diseñados para capturar, almacenar, manipular, analizar, manejar y presentar datos espaciales o geográficos.

Estas aplicaciones permiten a los usuarios crear consultas interactivas, analizar información espacial, editar datos en mapas y presentar los resultados de todas estas operaciones.

En los últimos años han proliferado sistemas de libre uso con fácil acceso a software de mapeo, tales como OpenStreetMap. Dichos servicios dan al público acceso a enormes cantidades de datos geográficos. Otros servicios, como Google Maps, ofrecen al usuario APIs para crear aplicaciones más personalizadas, con diferentes opciones de licencias para uso comercial.

8 Prototipo de Sistema de Información Geográfica

La principal característica de un SIG es que está diseñado para trabajar con datos referenciados con respecto a coordenadas espaciales o geográficas así como trabajar con distintas bases de datos de manera integrada, permitiendo generar información gráfica (mapas) útil para la toma de decisiones. Estos mapas ayudan a condensar varios aspectos de la realidad de una zona, con el objetivo de reconocer la existencia de patrones espaciales sobre algún fenómeno de interés. [9] Para generar estos mapas es necesario considerar algunos aspectos básicos:

- En primer lugar, comprender la interrogante a la cual se desea dar respuesta con la creación de un mapa, ya que al entender mejor el problema será más sencillo determinar que análisis serán necesarios.
- Para resolver la interrogante planteada es necesario obtener datos fiables y sobre todo que sean adecuados para nuestros objetivos. Hay que considerar que los datos correspondan con toda el área de interés y se tenga toda la información que sea de utilidad para la toma de decisiones.
- Es importante generar un diagnóstico de los datos para conocer el tipo de información que se ha obtenido, la distribución en la que se encuentra y de ser necesario ordenarla de acuerdo a nuestras necesidades para así diseñar o establecer una metodología del análisis.

Para el desarrollo de este prototipo se utilizó Open Street Map como fuente de datos, además de tres herramientas de gestión de SIGs, explicadas a continuación, todas de código abierto y disponibles para su descargar en los links referenciados en la bibliografía [6] [7] [8]. Debido a que cada herramienta está debidamente documentada en sus sitios respectivos, se omitirá lo relacionado a instalación y configuración de cada una, explicando únicamente la utilidad que presenta dentro del contexto de un SIG y de este prototipo.

8.1 Open Street Maps

OpenStreetMap (también conocido como OSM) es un proyecto colaborativo para crear mapas libres y editables. Los mapas se crean utilizando información geográfica capturada con dispositivos GPS móviles, ortofotografías y otras fuentes libres. Esta cartografía, tanto las imágenes creadas como los datos vectoriales almacenados en su base de datos, se distribuye bajo licencia abierta Licencia Abierta de Bases de Datos (en inglés ODbL). [4] Para efectos de este prototipo se utilizara la información relevante a la zona de Alajuela, Costa Rica.

8.2 QGIS

QGIS (anteriormente llamado Quantum GIS) es un Sistema de Información Geográfica de código libre para plataformas GNU/Linux, Unix, Mac OS, Microsoft Windows y Android. Este permite manejar formatos raster y vectoriales a través

de las bibliotecas GDAL y OGR, as como bases de datos. [10] Algunas de sus caractersticas son:

- Soporte para la extensión espacial de PostgreSQL, PostGIS.
- Manejo de archivos vectoriales Shapefile, ArcInfo coverages, Mapinfo, GRASS GIS, etc.
- Soporte para un importante número de tipos de archivos raster (GRASS GIS, GeoTIFF, TIFF, JPG, etc.)

QGIS posee un sistema de componentes o plugins que permiten realizar una extensa gama de operaciones sobre datos geográficos, sin embargo, nos interesa una en particular llamada QuickOSM la cual nos permite extraer información directamente de OpenStreetMap aplicando una consulta con la categoría y la zona donde queremos trabajar. OpenStreetMap maneja un sistema de etiquetas mediante pares Llave-valor, por ejemplo, Key: Amenity y Value: Hotels, buscaría la información correspondiente a comodidades, específicamente hoteles, que se encuentren en la zona especificada.

8.3 GeoServer

GeoServer es un servidor de código abierto escrito en Java. Permite a los usuarios compartir y editar datos geospaciales. Diseñado para la interoperabilidad, publica datos de las principales fuentes de datos espaciales usando estándares abiertos.[7] Este pretende operar como un nodo a través de una Infraestructura de Datos Espaciales libre y abierta para ofrecer datos geoespaciales, tal y como ha hecho Apache HTTP Server ofreciendo un servidor web abierto y libre para publicar HTML.

Este tendrá la función de conectarse a la base de datos de PostGIS y enviar las consultas realizadas por el usuario, así como mostrar los datos correspondientes que devuelva la base de datos.

8.4 Leaflet

Leaflet es una librería JavaScript opensource para crear mapas interactivos en un entorno móvil.[8]. Algunas de las ventajas de la API de Leaflet son:

- Facilidad de uso
- Soporte tanto a navegadores web como a dispositivos móviles.
- API bien documentada

Finalmente utilizaremos la librería Leaflet para mostrar los datos de manera interactiva y atractiva para el usuario, ademas de brindar soporte de capas (layers) las cuales pueden ser removidas y agregadas a necesidad.

9 Posibles Aplicaciones

Los Sistemas de Información Geográfica no solo permiten mostrar información, si no tomar decisiones y acciones en base a esta, un ejemplo de una posible aplicación consistiría en la creación de un SIG que permita a los usuarios de Costa Rica buscar disponibilidad acerca de servicios de salud, médicos, fisioterapeutas, entre otros. Esto es, supongamos que se desea obtener una cita con un dentista en una zona en particular del país, para un día y un rango de horas específicas, el SIG podría filtrar la información de clínicas dentales para verificar cuáles de estas tienen disponibilidad en ese horario para atender a un cliente. Además, podría permitir al usuario realizar la solicitud de la cita inmediatamente, y el mismo sistema se encargaría de comunicar los datos del cliente a la clínica, acelerando procesos usualmente tediosos y beneficiando tanto a los usuarios que necesitan de atención médica, como a los centros de servicios de salud que deseen aumentar su clientela.

References

- [1] En.wikipedia.org. (2017). Spatial database. [online] Available at: https://en.wikipedia.org/wiki/Spatial_database
- [2] En.wikipedia.org. (2017). R-tree. [online] Available at: <https://en.wikipedia.org/wiki/R-tree>
- [3] En.wikipedia.org. (2017). Geographic information system. [online] Available at: https://en.wikipedia.org/wiki/Geographic_information_system
- [4] En.wikipedia.org (2017). OpenStreetMap. [online] Available at: <https://en.wikipedia.org/wiki/OpenStreetMap>
- [5] OpenStreetMap. (2017). OpenStreetMap. [online] Available at: <https://www.openstreetmap.org/#map=15/10.0131/-84.2072>
- [6] Qgis.org. (2017). Bienvenido al proyecto QGIS!. [online] Available at: <https://www.qgis.org/es/site/index.html>
- [7] Geoserver.org. (2017). GeoServer. [online] Available at: <http://geoserver.org/>
- [8] Leafletjs.com. (2017). Leaflet an open-source JavaScript library for interactive maps. [online] Available at: <http://leafletjs.com/index.html>
- [9] Gulfprogram.ucsd.edu.. (2017). Importancia de los Sistemas de Información Geográfica (SIG) en la Conservación. [online] Available at: <http://gulfprogram.ucsd.edu/blog/coastal-and-marine/importancia-de-los-sistemas-de-informacion-geografica-sig-en-la-conservacion/>
- [10] Es.wikipedia.org (2017). QGIS. [online] Available at: <https://es.wikipedia.org/wiki/QGIS>