

Sistemas de Información Geográfica: Revisión de su Estado Actual

Nieves R. Brisaboa, José A. Cotelo Lema, Antonio Fariña,
Miguel R. Luaces, José R. R. Viqueira

Laboratorio de Bases de Datos.
Facultade de Informática. Universidade da Coruña
Campus de Elviña s/n. 15071. A Coruña. España.
[fbrisaboa.fari.luaces.joserios@udc.es](mailto:brisaboa.fari.luaces.joserios@udc.es),
jaclema@mail2.udc.es

Abstract. Los *sistemas de información geográfica* (SIG) ofrecen un entorno adecuado para la captura, almacenamiento y gestión tanto de información alfanumérica como de la información necesaria para representar sobre el plano los objetos manejados por la aplicación. Éstos sistemas han experimentado un gran desarrollo en los últimos años, gracias entre otras cosas a las mejoras en rendimiento de los ordenadores convencionales y a la aparición en el mercado de diversas herramientas de desarrollo SIG, las cuales tratan de suministrar al desarrollador un conjunto de utilidades que le asistan en el desarrollo de aplicaciones SIG lo más adaptadas posible al dominio de aplicación para el que han sido diseñadas. En este artículo describimos los conceptos básicos de esta tecnología y analizamos su evolución y las líneas de investigación actuales, tratando de describir el estado del arte de un campo poco conocido y en el que en general reciben poca formación los profesionales de la informática, a pesar de que en el mercado existe una demanda cada vez más grande de este tipo de aplicaciones.

1 Introducción

El incremento exponencial de la potencia de los equipos informáticos ha posibilitado que en la actualidad no sólo sea factible trabajar con grandes volúmenes de datos alfanuméricos, sino que además también se les pueda asociar información sobre aspectos geográficos (espaciales) de los objetos a los que se refieren, de modo que, por ejemplo, se pueda representar gráficamente sobre mapas o esquemas gráficos. Igualmente, han mejorado drásticamente las capacidades de consulta gráfica de toda esta información, gracias no sólo a ese incremento de potencia, sino también a la mejora en las capacidades gráficas de los sistemas informáticos.

Los *sistemas de información geográfica* (SIG) tratan de dar un paso más sobre los sistemas de información tradicionales para pasar a ofrecer un entorno adecuado para la captura, almacenamiento y gestión tanto de información alfanumérica (como hacían los sistemas tradicionales) como de información geográfica. Por información geográfica entendemos en este contexto información referente a la localización en el espacio de los objetos sobre los que queremos almacenar información. Esta información geo-

gráfica puede ser tan simple como la localización sobre un mapa de cada uno de los hospitales de un país o el área de un municipio ocupada por una parcela, o tan compleja como la distribución del territorio de un país en función del tipo de cultivo a que se dedica o de la salinidad de sus suelos. Además, las funcionalidades que ofrezca un SIG han de permitir una eficiente explotación de toda la información en él almacenada, lo cual no sólo incluye la capacidad para realizar *operaciones espaciales* sobre los datos geográficos sino también la de consultar y analizar gráficamente esta información. El aspecto gráfico adquiere un papel especialmente relevante en estos sistemas, ya que las relaciones entre datos geográficos o entre éstos y datos alfanuméricos se pueden hacer mucho más fácilmente identificables para el usuario mediante una adecuada representación gráfica.

Aunque pueda parecer una puntualización superflua, es conveniente distinguir claramente lo que es un SIG de lo que es una herramienta de desarrollo de SIG. Una herramienta de desarrollo SIG ofrece las funcionalidades necesarias para el almacenamiento y gestión de la información alfanumérica y geográfica, así como un conjunto de herramientas para la captura de datos y realización de consultas de esos mismos datos. Sin embargo, no ofrecen un entorno de captura y consulta adaptado a la información específica que se pretende almacenar en el SIG. Un símil de esta sutil y a la vez importante diferencia la tenemos entre un sistema gestor de bases de datos (SGBD) y un sistema de información tradicional. Si bien el SGBD ofrece el soporte para el almacenamiento y gestión eficiente de la información y las herramientas básicas (léase lenguajes de definición de datos, actualización y consulta) para la captura y consulta de datos, no se puede considerar a éste un sistema de información, sino que el sistema de información será una aplicación (o conjunto de aplicaciones) que haga uso de las funcionalidades ofrecidas por el SGBD para adaptarlo y presentar al usuario un entorno mucho más específico y sencillo.

Ejemplos significativos de áreas de aplicación de los SIG son la gestión catastral, de redes de saneamiento, de infraestructuras viarias o de comunicaciones, la navegación asistida por ordenador, o la toma de decisiones con respecto a la distribución de recursos, entre otros. Para el desarrollo de éstos sistemas, existen en la actualidad herramientas de desarrollo SIG que ofrecen al desarrollador la funcionalidad básica y las capacidades de desarrollo necesarias para poder construir un SIG adaptado al entorno de aplicación deseado.

Ante este panorama, parece necesario realizar una descripción precisa del *estado del arte* en el campo de los sistemas de información geográfica, que permita al no iniciado entender los conceptos en los que se basan y las funcionalidades que se deben esperar de ellos, así como tener una visión más clara de las implicaciones de los modelos de datos ofrecidos por las diferentes herramientas de desarrollo, para poder así realizar una elección más consciente de la misma. Para ello, un punto clave es distinguir claramente las dos filosofías reinantes en las herramientas de desarrollo SIG actuales, que se corresponden a lo que llamaremos herramientas orientadas a cartografía y herramientas orientadas a bases de datos espaciales.

El resto del artículo está estructurado del siguiente modo. La sección 2 muestra los conceptos básicos manejados en el dominio de los SIG. La sección 3 analiza las dos filosofías reinantes en las herramientas de desarrollo SIG actuales, a la vez que muestra la evolución experimentada por este tipo de herramientas. La sección 4 muestra las técnicas de representación de objetos geográficos típicas de cada una de estas filosofías.

as. Las secciones 5 y 6 analizan las necesidades de análisis espacial y de interfaces de consulta de los sistemas SIG, respectivamente. La sección 7 muestra las principales líneas de investigación en SIG. Por último, la sección 8 concluye el artículo.

2 Conceptos básicos

A la hora de comprender las características y funcionalidades de los SIG es necesario que previamente se tengan claros ciertos conceptos básicos sobre las características de la información que éstos van a manejar. La información de tipo geográfico puede clasificarse claramente en dos categorías, dependiendo de si ésta representa características del espacio geográfico (atributos referentes al espacio geográfico) o por el contrario representa propiedades de los objetos gestionados (atributos referentes a los objetos).

2.1 Conceptos básicos referentes al espacio geográfico

Se denomina espacio geográfico al espacio de coordenadas (normalmente R^2) en el que los datos geográficos son representados, el cual suele ser o bien cartesiano (en caso de que el SIG utilice un modelo plano del mundo) o bien geodésico. Sobre este espacio geográfico es a menudo interesante almacenar cierta información, consistente en atributos alfanuméricos de los cuales se asocia un valor de su dominio a cada punto del espacio. Esta información que queremos almacenar sobre el espacio geográfico es lo que denominaremos *atributos del espacio geográfico* (en adelante AEG). Un AEG puede ser continuo (si el valor asociado a los puntos del espacio puede variar de modo gradual a lo largo del espacio geográfico) o discreto (si el conjunto de valores que puede tomar el atributo asociado al espacio geográfico es discreto). Un ejemplo de AEG discreto es el *tipo de cultivo*, mientras que un ejemplo de atributo continuo sería la *salinidad del suelo*, la *temperatura* o la *presión atmosférica*.

2.2 Conceptos básicos referentes a los objetos

Sobre el espacio geográfico se representan los objetos geográficos, que son objetos que el SIG debe manejar y representar gráficamente. Así tenemos:

- *Objeto o entidad geográfica*. Es un objeto sobre el que la aplicación SIG guarda no sólo información alfanumérica sino también información geográfica que permita representarlo gráficamente sobre un mapa. Por ejemplo, si consideramos el objeto geográfico "Ciudad de A Coruña", además de atributos alfanuméricos como el "nombre", "población", etc. puede tener un atributo geográfico "Localización del Ayuntamiento" que represente la posición del edificio de su ayuntamiento y otro atributo geográfico "área" que represente el área del espacio geográfico ocupada por la misma. En el momento de visualización el GIS puede, por ejemplo, utilizar el atributo "Localización del Ayuntamiento" para representar la ciudad en presentaciones a escala 1:25.000 ó menor, y utiliza el atributo "área" para representar

a la ciudad en presentaciones a escalas mayores, donde esta información puede ser más útil. Es conveniente no confundir aquí la *escala* (de visualización) con la *precisión* de la representación de los atributos geográficos, que es la resolución con la que la información geográfica es almacenada en el SIG.

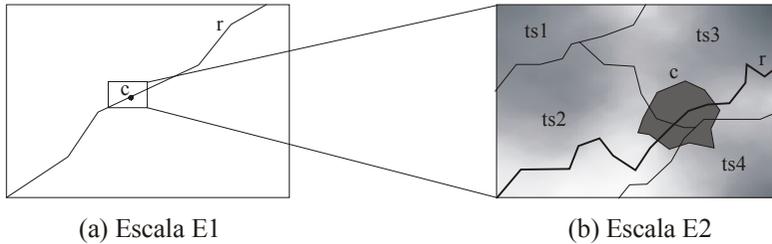


Fig. 1. Ejemplos de objetos geográficos y atributos del espacio geográfico a dos escalas distintas

- *Atributo geográfico.* Es un atributo que representa información referente a una característica geográfica del objeto al que pertenece (posición, extensión, etc.). Es un subconjunto no vacío y posiblemente infinito del espacio geográfico. Los *atributos geográficos* se representan mediante figuras geográficas, y su tipo se corresponde con el tipo de figura geográfica que se utiliza para representarlo.
- *Figura geográfica.* Se usan para representar sobre el plano de forma gráfica atributos geográficos de un objeto. En el dominio de los SIG se utilizan diversos tipos de figuras geográficas, siendo las más comunes las siguientes:
 - *Punto:* El valor de una figura geográfica de este tipo se corresponde con un elemento (punto) del espacio geográfico. Un ejemplo de atributo geográfico representable mediante una figura geográfica de este tipo es la posición de un objeto, como por ejemplo la localización geográfica de una oficina de correos.
 - *Línea:* Una figura geográfica de este tipo se corresponde con un conjunto de curvas en el espacio geográfico, donde una curva es una secuencia de puntos del espacio contiguos. Un ejemplo de atributos geográficos que pueden ser representados mediante una *línea* es el curso de un río o el trazado de una carretera.
 - *Región:* El valor de una figura geográfica de este tipo se corresponde con un conjunto de áreas del espacio geográfico. Ejemplos de atributos geográficos representables mediante una *región* son el área del espacio ocupada por una parcela de terreno o la zona de un país afectada por las nevadas en un día dado.
 - *Geográfico:* El valor de una figura geográfica de este tipo se corresponde con un conjunto de puntos y/o líneas y/o regiones del espacio geográfico. Aunque menos usado que los anteriores, este tipo de figura geográfica también es de utilidad en ciertos dominios de aplicación para representar atributos geográficos.
 - *Partición:* Una figura geográfica de este tipo representa un elemento de una partición del espacio geográfico en áreas disjuntas. Un ejemplo de

atributo geográfico de tipo partición es el área de un ayuntamiento de una provincia, en la que los ayuntamientos no pueden superponerse y el área ocupada por el conjunto de todos sus municipios ha de coincidir exactamente con todo el área de la provincia.

Como ejemplo, en la Fig 1 pueden verse los atributos geográficos correspondientes a varios objetos geográficos a dos escalas distintas, así como la representación de dos atributos del espacio geográfico. Más en concreto, se observa cómo a la escala E1 el objeto geográfico "c" (que representa una ciudad) se representa mediante su atributo geográfico "Localización del Ayuntamiento", de tipo *punto*, mientras que el objeto geográfico "r" se representa mediante un atributo geográfico de tipo *línea*. A la escala E2, el objeto "c" pasa a ser representado mediante su atributo geográfico "área", de tipo *región*, mientras que "r" continúa representándose mediante el mismo atributo de tipo *línea* (aunque al usar una escala mayor en la presentación podemos visualizar la misma con más detalle). A esta escala E2 puede verse además el atributo del espacio geográfico (AEG) discreto "tipo de suelo", que divide el espacio en cuatro regiones con valores ts1, ts2, ts3 y ts4 respectivamente, y el AEG continuo "salinidad" (colores más oscuros representan valores de salinidad del suelo más baja).

3 Herramientas de desarrollo SIG

En los últimos años han aparecido en el mercado numerosas herramientas de desarrollo de SIG, las cuales ofrecen al desarrollador un entorno orientado a facilitar el desarrollo de aplicaciones SIG. Estas herramientas se caracterizan por ofrecer al desarrollador las funcionalidades básicas para la gestión y almacenamiento de información tanto alfanumérica como geográfica, junto con ciertas facilidades para el desarrollo de interfaces de captura y consulta de los datos, de modo que éste pueda enfocar su esfuerzo en adaptar el interfaz de la aplicación a las necesidades específicas del SIG desarrollado.

3.1 Filosofías seguidas por las herramientas de desarrollo SIG

Las herramientas de desarrollo de SIG se ajustan a una de las dos filosofías o enfoques fundamentales que existen a la hora de ver la relación entre información alfanumérica y geográfica. La filosofía seguida por la herramienta condiciona algunas decisiones de diseño en el modelo de datos usado, así como la gestión y almacenamiento de los mismos, introduciendo con ello ciertas restricciones y limitaciones que han de ser tomadas en cuenta a la hora de elegir la herramienta a utilizar. Estas dos filosofías son:

- *Herramientas orientadas a cartografía*. Es el enfoque seguido por las herramientas de desarrollo SIG evolucionadas a partir de sistemas de gestión cartográfica o sistemas CAD (como por ejemplo Microstation), en los cuales la información geográfica se almacena en ficheros estructurados en capas, cada una de los cuales contiene todos los objetos del plano de una determinada clase (por ejemplo, el

fichero contiene una capa de carreteras, una de ríos y una de parcelas). En este enfoque, la componente geográfica de los objetos es el elemento principal en torno al cual se agrupa la información a manejar, de modo que a cada objeto de una capa se le asociará además, probablemente, alguna información alfanumérica (por ejemplo, esa objeto es un núcleo urbano, al que se asociará su nombre o el número de habitantes). Como resultado, se asume que un objeto geográfico contendrá un único atributo geográfico y un conjunto quizás vacío de atributos alfanuméricos, dando un trato diferencial a cada tipo de atributos.

- *Herramientas orientadas a bases de datos espaciales.* Es el enfoque seguido por aquellas herramientas de desarrollo SIG que han evolucionado desde el dominio de los SGBD. En este enfoque un atributo geográfico es un atributo más del objeto que se pretende modelar, sin hacer ninguna distinción especial entre éstos y los atributos alfanuméricos. En un SGBD esto implica que será posible utilizar predicados espaciales (relacionando atributos espaciales) en una consulta SQL, o que será posible realizar JOINS espaciales (cuando los atributos por los que se realiza el JOIN de las dos tablas son espaciales).

3.2 Evolución de las herramientas SIG

Las técnicas utilizadas para afrontar los aspectos de gestión y almacenamiento de la información manejada por los SIGs han ido evolucionando a lo largo de los años, conforme la tecnología de desarrollo de estos sistemas ha ido madurando. Además, la filosofía seguida por las herramientas de desarrollo SIG utilizadas también ha tenido un peso importante en la arquitectura adoptada. Fundamentalmente, podemos distinguir tres etapas o generaciones en la evolución de los SIG [4]:

- **Primera generación.** En las herramientas SIG de primera generación la información geográfica (e índices asociados) se almacena en ficheros cuyo formato es particular del SIG, por lo que éste es la única herramienta que puede ser utilizada para interpretar y manipular los datos. Las desventajas de estos sistemas son varias:
 - Sólo pueden tratar de forma nativa información espacial, por lo que en caso de necesitar combinar datos geográficos con datos tradicionales, se necesita que una aplicación funcionando sobre el SIG realice esta tarea.
 - El modelo de datos y la interfaz de la herramienta SIG son propietarios.
 - La aplicación SIG misma debe gestionar las cuestiones relacionadas con seguridad, recuperación, e integridad de los datos, ya que la herramienta SIG no da soporte.
- **Segunda generación.** La segunda generación de herramientas SIG nació del deseo de integrar los datos geográficos con los de una base de datos tradicional. Las dos técnicas que se utilizaron en mayor medida son la *arquitectura en capas* y la *arquitectura dual*.
En la arquitectura en capas, la funcionalidad del SIG se implementa sobre un SGBD comercial sin ninguna variación, tal y como se muestra en la Figura 2(a). La principal ventaja de estos sistemas es que los datos geográficos son gestionados por el SGBD, lo que permite utilizar su funcionalidad de control de transacciones,

seguridad, recuperación de fallos, etc. Hay dos posibles estrategias para representar los valores geográficos:

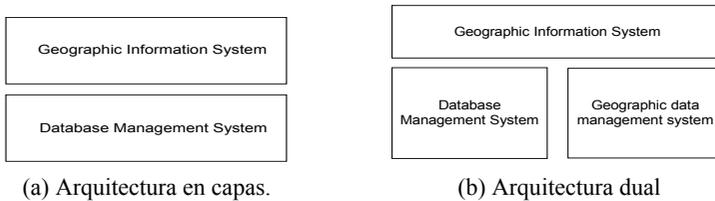


Fig. 2. Arquitecturas de segunda generación.

- Dividir la representación de un atributo geográfico en tuplas de forma que cada tupla almacene una componente de ésta. La desventaja de esta aproximación es que cada operación espacial debe reconstruir el objeto antes de realizar la operación, lo que resulta muy costoso.
- Utilizar un SGBD que proporcione *objetos largos* y almacenar en ellos los objetos geográficos. Esta aproximación es más eficiente, pero todavía presenta problemas debido a que los tipos de datos geográficos son opacos para el SGBD, lo que provoca que sus características no pueden ser utilizadas para optimizar las consultas y que no se pueda utilizar con ellos índices espaciales avanzados (únicamente un B-Tree con z-elementos).

Por otra parte, en la arquitectura dual (Figura 2(b)), se utiliza un SGBD para almacenar los datos alfanuméricos y un segundo subsistema que se encarga de gestionar los valores geográficos. Ambos subsistemas se integran a continuación mediante una capa que proporciona, además, el interfaz de usuario. Esta arquitectura permite utilizar la representación más apropiada para los valores geográficos, además de permitir el tipo de índice más apropiado para ellos. Un problema que presenta esta arquitectura es que la realización de consultas se complica, y la optimización global de las mismas es imposible. Además, el subsistema de gestión de datos geográficos debe implementar de nuevo funcionalidad estándar de los SGBD, tal como control de transacciones, recuperación de fallos, etc.

La arquitectura dual es característica de las herramientas orientadas a cartografía, donde lo que se hace es asociar a la información geográfica que se almacena en un sistema la información alfanumérica almacenada en el otro. Un ejemplo de este tipo de herramientas es Microstation Geographics, que permite el desarrollo de SIGs que almacenan la información geográfica en ficheros de Microstation convencionales y la información alfanumérica en un SGBD convencional. Por su parte, la arquitectura de capas es característica de las primeras generaciones de herramientas orientadas a bases de datos espaciales. Un ejemplo de esta arquitectura lo tenemos en las primeras extensiones espaciales de Oracle.

De forma general, los problemas que presentan las herramientas SIG de segunda generación son los siguientes:

- El SIG obliga al usuario a seguir el esquema de base de datos que se implementa, por lo que la integración con datos preexistentes es compleja y el sistema pierde en flexibilidad.

- La estructura y la semántica de los datos geográficos son totalmente desconocidos para el SGBD, con lo cual ciertos beneficios de los SGBD no pueden ser disfrutados por el SIG (indexación, optimización de consultas, etc..).
 - El administrador del SGBD debe conocer el funcionamiento del SIG y cómo interactúa éste con el SGBD para poder gestionar las bases de datos.
- **Tercera generación.** La tercera generación de herramientas SIG, de la cual comienzan a aparecer en los últimos tiempos los primeros productos comerciales, consiste en sistemas que utilizan los nuevos SGBD extensibles, los cuales pueden ser extendidos mediante módulos. Estos módulos incorporan nuevos tipos de datos, nuevos operadores y nuevos índices que permiten que el SGBD entienda de forma nativa los datos espaciales y sus operadores, gestionándolos de forma eficiente. Esto permite integrar totalmente el SIG sobre el SGBD, trasladando toda la responsabilidad de gestión de datos a éste último, permitiendo al usuario realizar consultas mediante un lenguaje de consulta estándar extendido (p.e. SQL), y permitiendo que las herramientas de optimización del SGBD sean utilizadas de forma efectiva.

Afortunadamente, las herramientas GIS orientadas a bases de datos espaciales parecen estar migrando claramente hacia esta última generación, y así las últimas extensiones espaciales de Oracle, entre otras, ya lo han hecho. Por su parte, las herramientas orientadas a cartografía, aunque más reacias a migrar a esta tercera generación por su especial visión de la relación entre información geográfica y alfanumérica, sí lo están haciendo al menos parcialmente, principalmente en cuanto a usar los SGBD para almacenar también la información espacial y así hacer uso de las funcionalidades ofrecidas por éstos en cuanto a seguridad, recuperación, e integridad de los datos.

4 Representación de los datos

Si bien hasta este punto nos hemos centrado en la visión de las herramientas de desarrollo SIG desde un punto conceptual, hablando de objetos con atributos geográficos y representaciones de éstos en el espacio R^2 , lo cierto es que a la hora de representar esos atributos geográficos en un sistema informático no queda más remedio que restringirse a representaciones finitas, por lo que se vuelve necesario aproximar el conjunto R^2 mediante algún tipo de malla finita de puntos. Dicha malla puede ser modelada, por ejemplo, con el conjunto finito R_m^2 , donde $R_m = \{-m, \dots, -1, 0, 1, \dots, m\}$.

Los tipos de datos geográficos, definidos en el nivel conceptual mediante subconjuntos de R^2 , se suelen representar sobre esta malla mediante aproximaciones lineales [10]. Sin embargo, el modo en que estas representaciones finitas son definidas, la semántica de sus valores e incluso el número de tipos definidos varía de unos autores a otros. Así, en [14] se utilizan aproximaciones lineales basadas en *quantums*, en las cuales un valor espacial es representado mediante un conjunto finito de puntos y segmentos paralelos a los ejes de coordenadas. Por su parte, en [12] los tipos de datos espaciales se representan mediante elementos de un *realm*, el cual es un conjunto finito (no fijo) de puntos y segmentos tal que los segmentos pertenecientes a éste no se intersecan entre sí. En la Figura 3 se muestra un ejemplo de ambas representaciones.

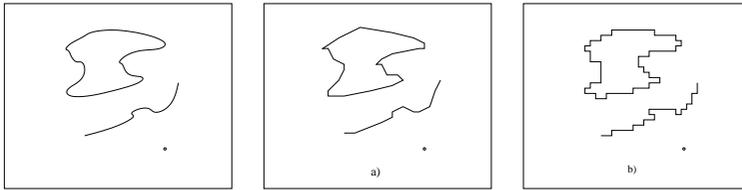


Fig. 3. Representación de valores espaciales mediante *realms* y mediante *quantums*.

Por otra parte, a la hora de almacenar los objetos geográficos, la solución usada depende de la filosofía seguida por la herramienta de desarrollo GIS. Para ilustrar la diferencia entre ambos enfoques se muestra a continuación cómo el modelo Relacional, como ejemplo de modelo conceptual usado en sistemas de información convencionales, puede ser utilizado para facilitar el modelado de información geográfica siguiendo cada una de las dos filosofías.

CAPAS:

CIUDAD_E1 (cod_ci, cod_mun, nombre, punto)	CARRETERAS (cod_car, línea)
CIUDAD_E2 (cod_ci, cod_mun, nombre, region)	SALINIDAD (salinidad, particion)
TIPO_SUELO (tipo, particion)	

RELACIONES:

MUNICIPIO (cod_municip, poblacion)

Fig. 4. Representación en capas en herramientas orientada a cartografía

En las herramientas de desarrollo SIG orientadas a cartografía, la información se almacena por *capas*. Una *capa* se define como una relación (en el sentido dado a esta palabra en el modelo relacional) que contiene en su esquema un atributo cuyo tipo de dato es uno de los tipos de dato geográficos soportados. Cada capa contendrá por tanto toda la información referente a un tipo de objeto geográfico (un objeto por tupla). Los atributos del espacio geográfico (AEGs), tanto continuos como discretos, también generan sus correspondientes capas, una por cada AEG. Una capa, además de soportar las operaciones usuales del álgebra relacional sobre sus atributos no geográficos, ofrece también un conjunto de operaciones específicas entre capas, referentes a sus atributos geográficos (como se explicará más detalladamente en la sección 5).

En la Figura 4, se muestra cómo podría modelarse en este tipo de herramientas el ejemplo de la Figura 1 mediante un conjunto de capas y relaciones.

Nótese que aquellas relaciones entre capas que pueden resolverse mediante relaciones entre atributos geográficos (por ejemplo, qué ciudades comunica la carretera “1”) no generan ninguna tabla nueva, sino que son implícitas (gracias a que están definidos sobre el mismo espacio geográfico), y pueden obtenerse aplicando operadores espaciales. Extensiones del modelo relacional siguiendo esta filosofía han sido definidas, entre otros, en [5] y [19].

En las herramientas de desarrollo SIG orientadas a bases de datos espaciales no se hace ningún tratamiento especial a las entidades o relaciones que contienen atributos geográficos. En el caso de los atributos del espacio geográfico, por su parte, éstos se

representan de modo similar a como se hace en las herramientas orientadas a cartografía. Al contrario que en la filosofía de herramientas orientadas a cartografía, aquí las operaciones aplicables a relaciones que contienen atributos geográficos en su esquema son exactamente las mismas que las aplicables a aquellas que no los contienen, consiguiendo de esta manera la integración total en la manipulación de todos los tipos de datos. En la Figura 5 se muestra como el modelo entidad-relación del ejemplo de la Figura 1 produce un conjunto de relaciones con tipos de dato espaciales.

Extensiones de otros modelos lógicos, como por ejemplo el relacional anidado [1], objeto relacional [20] (utilizado en las extensiones espaciales de la mayoría de sistemas de bases de datos comerciales [17, 13, 4]), orientado a objetos [21], etc. han sido propuestas en la literatura de bases de datos espaciales. La inclusión de tipos de dato espaciales en un modelo orientado a objetos se incluye en el estándar OpenGIS [16].

RELACIONES:

CIUDAD (cod_ci, cod_mun, nombre, punto_e1, region_e2)	
TIPO_SUELO (tipo, paticion)	SALINIDAD (salinidad, particion)
CARRETERAS (cod_car, linea)	MUNICIPIO (cod_municip, poblacion)

Fig. 5. Representación en herramientas orientada a bases de datos espaciales

5 Manipulación de los datos: álgebras y lenguajes

En una herramienta de desarrollo SIG, además de la capacidad de almacenamiento de información alfanumérica y geográfica hay un aspecto clave a cuidar, que es la capacidad de explotación de la información almacenada, tanto alfanumérica como geográfica. Para ello es necesario incluir en el sistema de información geográfica un conjunto de herramientas de análisis espacial.

5.1 Funcionalidad requerida

La gran diversidad de aplicaciones de los sistemas de información geográfica hace imposible ofrecer un conjunto de herramientas que soporten toda la funcionalidad requerida. La solución a este problema se centra en suministrar un conjunto de herramientas básicas que sean utilizadas en un gran número de aplicaciones, y proporcionar además un entorno extensible en el que el usuario pueda definir nuevas herramientas adaptadas a su aplicación concreta. Las herramientas básicas que deberán ser suministradas en la herramienta de desarrollo SIG son, entre otras, las siguientes:

- Herramientas para comprobar las distintas relaciones topológicas entre atributos geográficos [2, 8] (por ejemplo comprobar si las figuras A y B son disjuntas, si A está contenida en B, etc.).
- Cálculo de características de tipo numérico a partir de datos de tipo espacial, como por ejemplo el cálculo de áreas de regiones, longitudes de líneas, distancias entre figuras geográficas, etc.

- Construcción de figuras geográficas a partir de otras. Ejemplos de este tipo de herramientas son la intersección, unión y diferencia de figuras, la creación de áreas de influencia, cálculo del borde de regiones, puntos de inicio y final de líneas, etc.
- Manipulación conjunta de datos alfanuméricos y geográficos, como son las operaciones de *reclasificación* (permite fusionar los valores geográficos de una capa o relación que contengan la misma información en los atributos alfanuméricos y cuyos atributos geográficos sean adyacentes) *overlay* (permite agrupar en una sola capa o relación de salida la información que contienen dos capas o relaciones de entrada) y los *joins espaciales* (*joins* usando condiciones espaciales).
- Análisis espaciales de tipo más complejo y específico, como por ejemplo el cálculo del camino más corto entre dos puntos en una red de líneas.
- Estudio estadístico de atributos del espacio geográfico de tipo continuo, como puede ser la *interpolación* (construcción de una capa o relación de temperaturas a partir de un conjunto finito de mediciones) y la *discretización* (a partir de una capa o relación de temperaturas dividir el espacio en zonas donde la temperatura sea mayor o menor que cero grados centígrados).

5.2 Consultas en herramientas orientadas a cartografía

Como ya se ha visto anteriormente, en las herramientas orientadas a cartografía los objetos geográficos se almacenan en *capas*. Una capa no es más que una relación en la que se incluye un atributo geográfico. Una capa no admite el mismo conjunto de operaciones sobre atributos geográficos que el álgebra relacional proporciona sobre atributos alfanuméricos, por lo que para permitir la correcta explotación de los datos almacenados en capas debe definirse un álgebra especializada para ellas.

De una manera general, las operaciones que debe contener un álgebra de capas son las siguientes:

- *Selección*: Permite seleccionar el conjunto de tuplas de una capa que cumple una determinada condición. Esta condición puede ser una condición sobre atributos alfanuméricos, sobre el atributo geográfico o incluso mixta.
- *Proyección*: Permite eliminar de una capa uno o varios atributos no geográficos.
- *Unión, diferencia e intersección*: Dadas dos capas con el mismo esquema (mismos tipos de datos en sus atributos), permiten calcular la unión, diferencia e intersección de sus tuplas.
- *Cálculo convencional*: Esta operación permite calcular una o varias columnas de información alfanumérica mediante la aplicación de una determinada función a los atributos de una capa. Las nuevas columnas de información pueden ser añadidas a la propia capa.
- *Cálculo geográfico*: Permite sustituir el atributo geográfico de una capa por el resultado de aplicar una función al mismo.
- *Overlay*: Dadas dos capas $C1(a, G)$ y $C2(b, G)$, donde a y b representan un conjunto de atributos alfanuméricos. Esta operación produce una capa $C0(a, b, G)$, conteniendo los siguientes conjuntos de tuplas:
 - Para las intersecciones de figuras geográficas de $C1$ y $C2$ se almacenan los atributos alfanuméricos (a, b) .

- Para las partes de figuras geográficas de C1 que no intersecan con figuras geográficas de C2, se almacenan los atributos alfanuméricos (a, null).
- Para las partes de figuras geográficas de C2 que no intersecan con figuras geográficas de C1, se almacenan los atributos alfanuméricos (null, b).

Nótese que tanto la operación *producto cartesiano* como la operación *join*, aplicadas a dos capas, producirían como resultado una relación con dos atributos de tipo geográfico, que no cumple con la definición de capa. Por ello, estas operaciones no pueden aplicarse sobre dos capas, aunque sí sobre una capa y una relación alfanumérica. Un ejemplo de un álgebra de capas se presenta en [5].

5.3 Consultas en herramientas orientadas a bases de datos espaciales

En las herramientas orientadas a bases de datos espaciales se apuesta por un manejo integrado de todos los tipos de información almacenados en el sistema. Esto, en el ámbito del álgebra relacional se traduce en que las operaciones aplicables a relaciones que sólo contengan atributos alfanuméricos han de ser también aplicables a aquellas que también contengan atributos geográficos.

La extensión de la funcionalidad del álgebra relacional se realiza mediante la:

- Definición de un conjunto de *predicados espaciales*, los cuales pueden ser utilizados como condiciones en las operaciones de *selección* y *join*.
- Definición de un conjunto de *funciones espaciales*. Estas funciones, permiten calcular información alfanumérica a partir de atributos geográficos y calcular valores espaciales (geográficos) a partir de otros valores espaciales (union, diferencia, creación de zonas de influencia, etc.). Al igual que los predicados, pueden ser incluidas en las expresiones de las condiciones utilizadas en las operaciones de selección y join.
- Definición de nuevas operaciones relacionales. Estas nuevas operaciones permitirán el manejo integrado de la información alfanumérica y geográfica contenida en una o dos relaciones. Ejemplos de estas operaciones son de nuevo la *reclasificación* y el *overlay*, cuya funcionalidad ya ha sido especificada en la sección anterior. La diferencia entre estas operaciones y las operaciones definidas para capas es que este caso las operaciones deben poder ser aplicadas a relaciones que no contengan atributos geográficos.

Extensiones del álgebra relacional para el manejo de información espacial se presentan en [9, 14]. Otros modelos, como el orientado a objetos o el objeto-relacional, han sido también elegidos para su extensión para el tratamiento de datos geográficos. Tanto en el modelo objeto-relacional como en el orientado a objetos, la extensión viene dada por la definición de nuevos tipos de datos espaciales, en forma de clases abstractas de objetos, en las que no sólo se incluye la estructura de datos, sino que también se definen métodos para su manipulación. Entre los métodos definidos en los tipos de datos espaciales se incluyen tanto predicados como funciones espaciales. Por último, en [12] se define un álgebra para la manipulación directa de valores de tipos de dato espaciales. Esta álgebra puede ser incluida en muchos de los modelos lógicos existentes.

5.4 Lenguajes de consulta

En lo que se refiere a lenguajes de consulta, la mayoría de los sistemas de información geográfica comerciales proporcionan amplios conjuntos de comandos, incluidos o no en una interfaz gráfica de usuario, que implementan operaciones entre capas (en el caso de una relación con más de un atributo geográfico, ésta puede verse como varias capas, una por cada atributo geográfico). Éste es el caso de paquetes comerciales como Arc/Info 7 y Geomedia 3.0. La expresividad y facilidad de uso de estos conjuntos de comandos, a pesar de estar incluidos en entornos gráficos de usuario, está lejos de la proporcionada por los lenguajes clásicos de consulta a bases de datos, como son el SQL.

En el mundo de las bases de datos espaciales son muchas las extensiones de SQL que se presentan en la literatura [7, 18]. Centrándonos de nuevo en el modelo relacional, SQL92 se extiende con nuevos tipos de datos, nuevos predicados y funciones espaciales para incluir en las expresiones usadas en consultas (sobre todo en la cláusula WHERE de las sentencias SELECT), y nuevas operaciones, como OVERLAY, que se añaden a las ya existentes UNION, JOIN, etc.

Similarmente, las extensiones del modelo objeto-relacional deberían considerar el estándar SQL99, mientras que las extensiones del modelo orientado a objetos deberían basarse en el lenguaje OQL y el estándar *OpenGIS*. Este estándar define un conjunto básico de tipos de datos geográficos que los SIG y bases de datos espaciales deberían de ser capaces de manejar y las operaciones espaciales sobre ellos que deberían de suministrar.

6 Interfaces

Dos son los aspectos que debemos considerar en cuanto a las interfaces de los SIG: las interfaces de usuario y las interfaces con otros sistemas (interconectividad).

6.1 Interfaces de usuario

La finalidad principal de un SIG es la de servir como herramienta de análisis y ayuda a la toma de decisiones, para lo cual debe permitir el análisis cualitativo de los datos geográficos. A diferencia de los sistemas gestores de bases de datos tradicionales, en los que se almacenan datos alfanuméricos que pueden ser introducidos fácilmente mediante el teclado y cuyos resultados pueden ser mostrados textualmente, los datos que manejan los sistemas de información geográfica deben ser mostrados en una interfaz gráfica. Además, tal y como se remarca en [6], el análisis cualitativo, que se basa a menudo en la comparación visual de situaciones diferentes, es difícil de conseguir a no ser que las situaciones a comparar sean visibles simultáneamente, por lo que las interfaces de SIG deben ser dispuestas de tal manera que varios mapas e información alfanumérica sean visibles simultáneamente y puedan ser comparados directa e indirectamente. Esto se consigue en los SIGs mediante el concepto de *capas*.

Una capa muestra un conjunto de objetos geográficos o un atributo del espacio geográfico (AEG). Una capa con objetos geográficos será transparente allí donde no

hay ningún elemento de este tipo, de modo que se puedan ver las capas inferiores. En las zonas donde hay algún elemento éste se puede representar como un objeto opaco, translúcido o bien mediante su contorno. En el caso de una capa que represente un AEG, el funcionamiento será similar. Un AEG discreto puede representarse como si existiese un objeto geográfico por cada posible valor tomado por el atributo alfanumérico asociado, representando los puntos del espacio que tienen asociado el mismo valor. Para los AEGs continuos, existen dos opciones, que son representar éstos por niveles (es decir, discretizándolos) o bien representarlos mediante gamas de colores, de modo que cada punto tenga un color proporcional al del valor alfanumérico asociado a él. El SIG incorpora un gestor de capas que permite al usuario configurar el orden en el que se muestran las capas, añadir y eliminar capas, de forma que permite construir de modo dinámico un *mapa* con la información. Además, se permite al usuario almacenar de forma permanente configuraciones interesantes que pueden ser recuperadas más adelante, con la particularidad de que no se almacenan los datos que componen el mapa, sino la estructura del mismo, con lo que un cambio en los datos provoca un cambio en el mapa. Es importante remarcar que las propiedades de los elementos cartográficos dibujados pueden ser modificadas fácilmente y asignadas en función de los datos alfanuméricos. Por tanto, los objetos geográficos se encuentran completamente desacoplados de los objetos cartográficos.

La naturaleza dual de los datos almacenados, geográficos y alfanuméricos, también es tenida en cuenta en estos sistemas, que permiten mostrar resultados alfanuméricos, en particular los de un objeto geográfico determinado. Esto hace que los SIG deban proporcionar un método para seleccionar un objeto geográfico y, dado que los objetos pueden solaparse, esto debe ser considerado por el sistema y proponer una solución que permite determinar precisamente qué elemento se desea seleccionar. Además, es conveniente que la interfaz permita realizar una selección “burda”, o “grosera”, es decir, que no obligue a apuntar a los objetos con precisión milimétrica, sino que permitan seleccionar un objeto pulsando en un punto “relativamente” cercano a él. Los métodos de consulta se apoyan también en las interfaces gráficas y las herramientas de selección, pues los parámetros geográficos de las consultas suelen ser seleccionados o introducidos mediante la interfaz.

La escala de los mapas es un parámetro fundamental en el mundo cartográfico, y debe estar presente en los SIGs. Una diferencia fundamental es que los mapas de un SIG no se encuentran a una escala determinada, sino que ésta puede ser cambiada de forma dinámica. Es importante que el SIG muestre en todo momento la escala a la cual se visualizan los datos, y que permita conocer las coordenadas de los objetos geográficos. Un aspecto relacionado, pero a la vez completamente diferente, es la resolución con la que la información geográfica es almacenada en el SIG, la cual ha de ser la adecuada para el uso que se vaya a dar a dicha información, ya que limitará el nivel de detalle con el que ésta podrá ser mostrada.

Una característica importante en los SIG es la capacidad de mostrar representaciones diferentes de los objetos geográficos en escalas diferentes. Esto se conoce como *problema de la generalización* [22, 23, 24]. Las posibles opciones a la hora de construir nuevos objetos geográficos a partir de los existentes son las siguientes:

- Simplificar el contorno de un objeto geográfico. Si la escala utilizada es tal que es imposible mostrar todo el detalle del contorno de un objeto geográfico, puede reducirse el detalle del mismo para acelerar el proceso de dibujado.

- Asignar una nueva representación a un objeto en función de la escala, por ejemplo, las ciudades que se representan con superficies en un mapa de gran detalle, se representan con puntos en un mapa de pequeña escala. Es posible, incluso, que se permita que objetos desaparezcan a no ser que tengan importancia semántica.
- Agregar objetos que representan una partición del plano, por ejemplo agregando las provincias de un país para mostrar únicamente el contorno del país en mapas de pequeña escala.
- Seleccionar algunos de los elementos de un objeto geográfico complejo, como por ejemplo seleccionar sólo algunos edificios de una ciudad.

6.2 Interconectividad

Conforme el acceso a la información a través de redes de datos (como Internet) ha dejado de ser algo excepcional para pasar a estar a la orden del día, la capacidad de los sistemas SIG para interrelacionar información almacenada en distintos sistemas y procedente de diferentes fuentes se ha vuelto un requerimiento imprescindible. Una herramienta de desarrollo SIG debería permitir la fácil interconexión de éste con nuevas fuentes de información, situadas en ordenadores completamente distintos y que utilizan distintos sistemas de almacenamiento, y permitir utilizar esta información como si estuviese almacenada localmente. Esto implica por un lado que el método de acceso a la fuente de los datos ha de ser lo más transparente posible al usuario, pero también que la estructura concreta de las capas a analizar (en cuanto a qué atributos han de contener) no debería de tener que ser prefijada de antemano. Asimismo, el sistema SIG debería ser capaz de trabajar con datos representados en diferentes sistemas de coordenadas e interrelacionarlos igual que si todos los datos estuviesen utilizando el mismo.

Dado que un importante grupo de las aplicaciones que trabajan con información geográfica son los sistemas CAD, y la interoperabilidad con sistemas existentes es siempre algo deseable, es también importante que los SIG permitan utilizar, de forma cómoda, datos provenientes de éstos sistemas.

Por último, no se debe descuidar la capacidad para hacer pública la información gestionada por el SIG, lo cual implica, por un lado la capacidad de exportar la información gestionada siguiendo los principales estándares (por ejemplo, OpenGIS [16]), y por otro la posibilidad de ofrecer al público la capacidad de realizar consultas al sistema SIG (por ejemplo, a través de la WEB).

7 Nuevas líneas de investigación

Aunque a lo largo de los últimos años se han dado grandes pasos en el desarrollo de herramientas SIG, aún quedan funcionalidades deseables que incorporar a éstas. Entre los aspectos a investigar y/o incorporar están los siguientes:

- Indexación. Aunque se ha realizado una gran cantidad de investigación en el área de indexación de información espacial, la expectativa de que los SIG tengan que manejar cantidades de información geográfica cada vez más grandes implica la ne-

cesidad de seguir mejorando los métodos de indexación existentes, para así lograr una mayor eficacia de los mismos.

- Modelos de datos. El uso de representaciones finitas para la información geográfica genera no pocos problemas de robustez y consistencia de datos, así como en la integración de información representada a diferentes resoluciones. Aunque ya se han propuesto mejoras para reducir estos problemas [3, 12, 14, 15], todavía queda bastante camino por recorrer en esta dirección.
- Tiempo en SIG y BB.DD. espacio-temporales. Un aspecto interesante y un reto para la comunidad SIG es la incorporación de la dimensión temporal, tanto en la representación y gestión de información espacio-temporal que cambia a intervalos discretos como aquella que evoluciona de forma continua en el tiempo [11].
- Interoperabilidad. Aspectos como la integración de información geográfica, posiblemente a distinta resolución, y proveniente de distintas fuentes requiere todavía más investigación para poder mejorar las capacidades de interconexión de sistemas SIG.
- Indexación y *joins* en atributos heterogéneos (por ejemplo, figuras geográficas del tipo *geográfico*).
- Manipulación de atributos espaciales y redes. La mayoría del trabajo en el dominio SIG se ha centrado en la visión del espacio como un almacén de objetos, pero todavía queda mucho trabajo que hacer en la gestión de atributos del espacio geográfico (especialmente los continuos) y en el desarrollo de herramientas SIG orientadas a la gestión de redes.
- Manejo de información difusa.

8 Conclusiones

Se espera que los sistemas SIG sufran gran impulso en los próximos años debido a su capacidad de representar gráficamente objetos geográficos, demostrando así las relaciones espaciales entre ellos fácilmente y mejorando las funcionalidades y facilidad de uso de muchas aplicaciones. Por otro lado, la evolución de las herramientas de desarrollo SIG facilita el desarrollo de aplicaciones SIG eficaces y adaptadas a cada dominio de aplicación. Las nuevas líneas de investigación descritas anteriormente dotarán además a los SIG de mayores funcionalidades que las herramientas actuales, principalmente en ciertos entornos de aplicación (catastro, gestión de redes, navegación, etc.)

Por ello consideramos que éste será un campo con una gran demanda de profesionales en los próximos años y lamentamos que los recién licenciados carezcan, en general, de formación en el manejo y programación de este tipo de aplicaciones.

Referencias

- [1] E P. F. Chan, R. Zhu. *QL/G – A Query Language for Geometric Data Bases*. Proc. First International Conference on GIS, Urban Regional and Environmental Planning, Samos, Greece, 271-286, 1996.

- [2] E. Clementini, P. Di Felice, P. van Oosterom. *A Small Set of Formal Topological Relationships Suitable for End-User Interaction*. Advances in Spatial Databases, D. Abel, B. C. Ooi (eds.), Third International Symposium, SSD'93, Lecture Notes in Computer Science 692, Springer-Verlag, Singapore, 277-295, 1993.
- [3] J. A. Coteló Lema, R. H. Güting. *Dual Grid: A New Approach for Robust Spatial Algebra Implementation*. FernUniversität Hagen, Informatik-Report 268, May 2000.
- [4] J. R. Davis. *IBM's DB2 Spatial Extender: Managing Geo-Spatial Information within the DBMS*. <http://www.ibm.com/software/data/pubs/papers/>. Technical Report IBM Corporation, May 1998.
- [5] V. Delis, Th. Hadzilacos, N. Tryfona. *An Introduction to Layer Algebra*. Proc. 6th International Symposium on Spatial Data Handling (SDH'94), 1020-1041, 1994.
- [6] M. J. Egenhofer. *Interaction with Geographic Information Systems via Spatial Queries*. Journal of Visual Languages and Computing, Vol. 1, pp. 389-413, Academic Press Limited, 1990.
- [7] M. J. Egenhofer. *Spatial SQL: A Query and Presentation Language*. Transactions on Knowledge and Data Engineering, Vol. 6(1), pp. 86-95, 1994.
- [8] M. J. Egenhofer, J. R. Herring. *Categorizing Binary Topological Relations Between Regions, Lines, and Points in Geographic Databases*. Technical Report, Department of Surveying Engineering, University of Maine, 1990.
- [9] M. Gargano, E. Nardelli, M. Talamo. *Abstract data types for the logical modeling of complex data*. Information Systems 16(6), 565-583 (1991).
- [10] R. H. Güting. *Special Issue on Spatial Database Systems: An Introduction to Spatial Database Systems*. VLDB Journal: Very Large Data Bases, 3(4), pp. 357-399, October 1994.
- [11] R.H. Güting, M.H. Böhlen, M. Erwig, C.S. Jensen, N.A. Lorentzos, M. Schneider, and M. Vazirgiannis. *A Foundation for Representing and Querying Moving Objects*. FernUniversität Hagen, Informatik-Report 238, September 1998. Revised version to appear in ACM Transactions on Database Systems.
- [12] R. H. Güting, M. Schneider. **Realm-Based Spatial Data Types: The ROSE Algebra**. VLDB Journal 4, 100-143, 1995.
- [13] *Illustra 2D Spatial Database (Release 1.3) Guide*. October 1994.
- [14] N. A. Lorentzos, N. Tryfona, J. R. Rios Viqueira. *Relational Algebra for Spatial Data Management*. Proc. of the International Workshop Integrated Spatial Databases: Digital Images and GIS (ISD'99), pp 192-208, Lecture Notes in Computer Science 1737, June 1999.
- [15] V. Muller, N. W. Paton, A. A. A. Fernandes, A. Dinn, M. H. Williams. *Virtual Realms: An Efficient Implementation Strategy for Finite Resolution Spatial Data Types*. Proc. 7th International Symposium on Spatial Data Handling, M. Kraak and M. Molenaar (eds.), Taylor & Francis, 697-710, 1997.
- [16] *OpenGIS[®] Abstract Specification (Version 4)*. June 1999.
<http://www.opengis.org/techno/specs.htm>
- [17] *Oracle 8i Spatial User's Guide and Reference. Release 8.1.5*. February 1999.
- [18] N. Roussopoulos, S. Kelley. *PSQL: An Inexpensive GIS Solution for Oracle*. Advanced Communications Technology Inc. Technical Report, 1995.
- [19] M. Scholl, A. Voisard. *Thematic Map Modeling*. A. Buchmann et al. (Eds.), Design and Implementation of Large Spatial Databases, Lecture Notes in Computer Science 409, Springer-Verlag, Berlin, 1989.
- [20] T. Vijlbrief, P. van Oosterom. *The GEO++ system: An Extensible GIS*. Proc. 5th Int. Symp. on Spatial Data Handling (SDH'92), Charleston, SC, 1992, 40-50.
- [21] A. Voigtmann, L. Becker, K. Hinrichs. *An Object-Oriented Data Model and Query Language for Geographic Information Systems*. Technical Report 15/95-I, University of Munich, 1995.

- [22] A. Voisard. *Towards a Toolbox for Geographic User Interfaces*. Second International Symposium in Advances in Spatial Databases (SSD), pp. 75-97, 1991.
- [23] A. Voisard. *Designing and Integrating the User Interface of Geographic Database Applications*. ACM International Conference on Advanced Visual Interfaces (AVI), pp. 133-143, 1994.
- [24] A. Voisard. *Mapgets: A Tool for Visualizing and Querying Geographic Information*. Journal of Visual Languages and Computing, Vol. 6, Academic Press, December 1995.