

ARTÍCULO

ANÁLISIS ESPACIAL POR MEDIO DE UN SISTEMA DE INFORMACIÓN GEOGRÁFICA DISTRIBUIDO

Miguel Torres, Marco Moreno-Ibarra, Rolando Menchaca-Méndez, Rolando Quintero y Giovanni Guzmán
Centro de Estudios de Cómputo
Instituto Politécnico Nacional

Análisis espacial por medio de un sistema de información geográfica distribuido

Resumen

Con la idea de mejorar la recuperación espacial, por medio de un Sistema de Información Geográfica (SIG), ha sido diseñada una aplicación SIG distribuida. Esta aplicación se enfoca principalmente en la realización de operaciones de análisis espacial, a través de una herramienta denominada Módulo de Análisis Espacial (SAM por sus siglas en inglés). Este módulo contiene varios modelos espaciales y matemáticos, que fueron diseñados e implementados para detectar áreas con riesgo de deslave e inundación. Los usuarios pueden llevar a cabo un análisis para recuperar los datos espaciales de diferentes sitios, utilizando una especificación basada en el lenguaje XML. Los datos geospaciales están almacenados en una base de datos geográfica, dentro de un SIG empresarial. Asimismo, todos los procesos de análisis espacial se ejecutan en el servidor de aplicación, para mejorar el rendimiento del análisis y la recuperación de información. La aplicación SIG ha sido implementada en el módulo ArcMap del sistema ArcGIS.

Palabras clave: SIG, XML, deslaves, inundaciones, ambiente distribuido, base de datos espacial

Spatialanalysis by meansof a distributed geographical information system

Abstract

To improvethes patial retrieving by means of Geographical Information Systems (GIS), a distributed GIS-Application has beendesigned. This a pplication is mainly focused on makings patialanalysis, throughout Spatial Analyzer Module (SAM). SAM contains severals patial and mathematical models that have been designed and implemented for detectin glands lides and flooding areas. Users can makeananalysistoretrievethe spatial data from differents ite using a specification based on Extensible MarkupLanguage (XML). Geospatial data are stored in a geographic database in to the Enterprise GIS. Thus, allspatialanalysisprocesses are executed in theserverapplicationtoimprovethe performance of the análisis and information retrieving. The GIS-Application has been implementedi nto ArcMap module of ArcGIS system.

Keywords: GIS, XML, landslides, floodingareas, distributedenvironment, spatialdatabase.

Introducción

Hoy en día los Sistemas de Información Geográfica (SIG) son herramientas poderosas y útiles como medios para la representación de información, la visualización y la investigación; o bien en aplicaciones para la toma de decisiones. [1] Actualmente, el análisis espacial inteligente es una de las principales necesidades que se presentan como tendencia en el área de la Geocomputación. El análisis espacial juega un papel muy importante en diversas situaciones. Muchas veces la información se extiende en diferentes lugares, lo cual hace que el problema crezca, debido a que los datos espaciales presentan diferentes formatos y especificaciones, tales como escala, proyección, sistema de referencia, tipos de representación, tipo de SMD, temática y fecha. Por estas razones, la heterogeneidad de los datos espaciales complica el análisis espacial como tal.

Por lo tanto, se propone un esquema distribuido, basado en la tecnología XML, para desarrollar una especificación espacial estándar, que sea utilizado para recuperar los datos espaciales. Este esquema

permitirá el intercambio, la transferencia y el almacenamiento de la información geográfica dentro de un repositorio geográfico.

Diversas propuestas y proyectos han sido diseñados para recuperar los datos espaciales. Sin embargo, en la mayoría de los casos los datos recuperados representan la información en un formato raster (jpg, gif). [2] Esta información dificulta la realización del análisis espacial, porque las características intrínsecas no son exploradas. Algunos proyectos importantes se han diseñado para representar información geográfica en la Web (web-mapping), pero otras aplicaciones están cerradas únicamente para manejar datos espaciales. [3]

En este artículo se propone un mecanismo para realizar análisis espacial por medio del Módulo de Análisis Espacial (SAM por sus siglas en inglés), el cual es utilizado para realizar algunas operaciones de análisis espacial. Este módulo puede recuperar objetos geográficos, utilizando un ambiente distribuido. SAM es una herramienta diseñada para fenómenos naturales similares, con métodos especiales enfocados en detectar áreas con riesgo de deslaves e inundación. La tecnología XML se utiliza para desarrollar una especificación codificada, basada en las propiedades cualitativas y cuantitativas de los datos espaciales. El objetivo principal de esta aplicación es encontrar las soluciones necesarias para llevar a cabo operaciones de análisis espacial y recuperar datos geográficos en un ambiente distribuido. Este es un gran reto de las nuevas tendencias en el campo de la Geocomputación.

Este artículo de investigación está organizado de la siguiente manera. En la sección 2 se presenta la arquitectura de la aplicación SIG y se describe la funcionalidad de los módulos del sistema. Algunos de los resultados obtenidos se muestran en la sección 3. Finalmente en la sección 4 se describen nuestras conclusiones relacionadas con el trabajo.

Arquitectura de la Aplicación SIG

La aplicación SIG desarrollada, presenta una arquitectura cliente servidor. Esta herramienta está compuesta por los siguiente módulos: un SIG Empresarial, un Módulo de Comunicación, un Módulo de Análisis Espacial, una Base de Datos y un Módulo de Administración de XML. En la figura 1 se muestra la arquitectura de la herramienta.

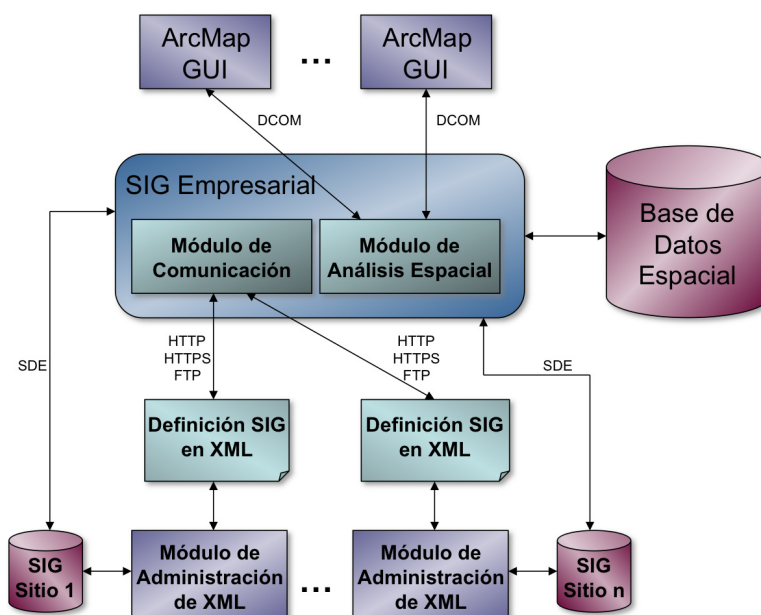


Figura 1. Arquitectura de la Aplicación SIG.

El mecanismo funcional de la aplicación SIG es el siguiente: los usuarios de la herramienta ArcMap necesitan realizar una petición. Esta petición se envía por medio de la tecnología DCOM al módulo de comunicación que se encuentra embebido en el SIG Empresarial. Este módulo procesa la petición y envía los parámetros vía los protocolos http o HTTPS.

En el sitio remoto de algún SIG, el módulo de administración de XML verifica la definición XML inicial y consulta la definición XML local para ubicar y comparar las características cualitativas y cuantitativas de los datos espaciales, tales como escala, proyección, referencia espacial, tipo de representación, temática, tipo de SMBD y datos atributivos. Si la definición XML coincide con la definición XML local, los datos espaciales serán recuperados y enviados directamente por medio

del Motor de Base de Datos Espacial (SDE). Los objetos geográficos serán almacenados en la base de datos espacial, en donde podrán ser analizados por el SAM.

Base de datos espacial en la aplicación SIG

La aplicación SIG está compuesta por una base de datos geográfica o espacial, la cual ha sido diseñada e implementada a través de Geodatabase. Este contenedor es un mecanismo de almacenamiento provisto por ArcGIS, el cual está enfocado en generar repositorios geográficos independientes. [4]

La base de datos espacial es un ambiente que sirve para modelar las características de los fenómenos geográficos. El Geodatabase proporciona un modelo topológico que está integrado por las clases de los elementos. Este modelo es similar a las coberturas espaciales. El modelo del contenedor es soportado por una base de datos relacional orientada a objetos. Por tanto, se considera un mecanismo híbrido entre las técnicas de modelación relacional y orientada a objetos. Con el uso de esta tecnología, los usuarios pueden acceder a los atributos espaciales y descriptivos a partir de diferentes fuentes, por medio del SDE. [5]

El mecanismo híbrido define una interfaz abierta al sistema de base de datos y permite manejar información geográfica en una forma intrínseca. En este caso, el comportamiento de los objetos espaciales es definido por el sistema. Las entidades son representadas como objetos espaciales con propiedades, comportamientos y relaciones entre sí. La base de datos espacial ha sido diseñada con el sistema ArcInfo. Todos estos componentes están involucrados en el análisis para determinar áreas de deslave e inundación. Los datos espaciales que coinciden con la especificación XML, son recuperados y almacenados en la base de datos espacial.

El modelo de este repositorio ha sido desarrollado utilizando UML (UnifiedModelingLanguage), el cual genera objetos del tipo COM. Estos objetos implementan el comportamiento y el esquema, en donde los objetos son creados, almacenados y manejados dentro de la base de datos. Este esquema establece la relación entre objetos y la base de datos espacial. Por otra parte, éste define las relaciones espaciales entre los objetos en diferentes clases, dependiendo de las primitivas de representación (líneas, puntos y polígonos). En la figura 2 se muestran los elementos que componen la base de datos espacial.

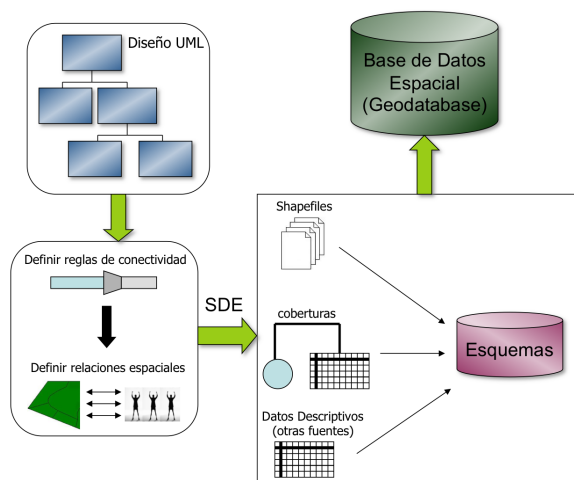


Figura 2. Elementos que componen la base de datos espacial.

Módulo de comunicación en la aplicación SIG

Es muy común que el análisis espacial involucre una gran cantidad de fuentes de datos. Estas fuentes de datos son localizadas en diferentes sitios, que pertenecen tal vez a diversas organizaciones. Por ejemplo, en el caso de las áreas de inundación, el análisis involucra datos como volumen de agua superficial, geología e infraestructura, entre otros. Todos estos recursos son localizados en una base de datos de una organización única.

Para solucionar este problema, se han invertido muchos esfuerzos para llevar a cabo interoperabilidad entre diferentes proveedores y SIG distribuidos. Una de las alternativas es GML (Geographic Markup Language). [6] [7] Por lo tanto, una especificación desarrollada por la OPC (organismo dedicado a la automatización e interoperabilidad) está enfocada en la transferencia y el almacenamiento de información geográfica basada en XML. Sin embargo, esta recomendación no brinda un soporte completo en la mayoría de los SIG comerciales. Es difícil encontrar una solución para el problema en un tiempo corto, ya que es necesario que los proveedores de SIG acuerden la construcción de traductores a partir de las representaciones de sus datos propietarios a un estándar. El problema para interoperar multi-proveedores y SIG distribuidos puede dividirse en tres partes: a) Localización de los datos requeridos para realizar un análisis espacial; b) Acceso remoto a los datos relevantes, y c) Transformación de diferentes representaciones de datos.

En este trabajo se propone una solución parcial pero útil, que ataca los dos primeros problemas. Por lo tanto, proponemos desprender el primer componente del problema, porque en muchos casos los SIG involucran un análisis que sea distribuido. Para localizar los datos requeridos y realizar un análisis espacial, hemos incluido un módulo denominado Administración de XML. Este

módulo se encarga de obtener la geo-información a conocer, si una fuente de datos en particular es relevante en un análisis espacial específico. [8] La arquitectura se muestra en la figura 1. La geo-información obtenida es entonces codificada en un documento XML, por lo cual en la tabla 1 se muestra un ejemplo de los contenidos del documento XML. Parte de la información es consultada desde el SIG y otra parte es cargada manualmente por el administrador del SIG.

El módulo de comunicación se encuentra localizado en el SIG empresarial, así como las consultas de los documentos XML de los diferentes sitios para desplegar su información al usuario. De esta forma los usuarios pueden seleccionar las fuentes de información que sean relevantes a su análisis. La comunicación entre el SIG empresarial y los Módulos de Administración de XML, puede realizarse utilizando diferentes protocolos estándar, tales como http, FTP o HTTPS, en caso de necesitarse una transacción segura. Cuando las fuentes de datos son seleccionadas, el Módulo de Comunicación utiliza el SDE para obtener los datos espaciales.

```
<?xmlversion="1.0"?>
<Spatial_Data>
<Description>
<Layer_Name>topo</Layer_Name>
<Theme>Topography</Theme>
<Elab_By>GeoLab-CIC-PN</Elab_By>
<Elab_Date>02/08/1999</Elab_Date>
<Last_Update>12/08/2001</Last_Update>
<Type>Line</Type>
<Topology>Y</Topology>
</Description>
<Geographical_Properties>
<Projection>UTM 14</Projection>
<Datum>NAR_D</Datum>
<Units>METERS</Units>
<Spheroid>GRS1980</Spheroid>
<Boundary>
<Xmin>397041.431</Xmin>
<Xmax>685954.665</Xmax>
<Ymin>2432826.985</Ymin>
<Ymax>3097482.722</Ymax>
</Boundary>
<Scale>1:50000</Scale>
</Geographical_Properties>
<DBMS_Properties>
<Provider>ArcInfo</Provider>
<Table_Name>topo.aat</Table_Name>
<Data_Type>Coverage</Data_Type>
<Attributes>Topo#, TopoD, FNODE#,
            TNODE#, Length, LPOLY#,
            RPOLY#, Height
</Attributes>
<Num_Records>420</Num_Records>
</DBMS_Properties>
</Spatial_Data>
```

Tabla 1. Documento XML que contiene la descripción de un SIG.

Módulo de Análisis Espacial en la aplicación SIG

SAM es un módulo especial que ha sido diseñado para realizar procedimientos de análisis espacial. El objetivo principal de SAM es identificar zonas con altas probabilidades de deslave e inundación. SAM utiliza datos vectoriales para llevar a cabo el análisis. Un procedimiento específico se ha diseñado para cada uno de estos fenómenos. Este análisis se basa en la utilización de diferentes datos espaciales relacionados al fenómeno estudiado. En la figura 3 se muestra la arquitectura del SAM.

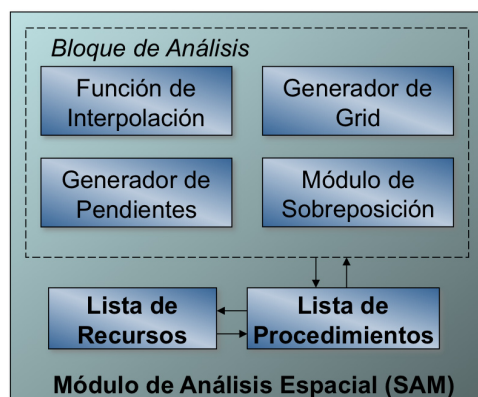


Figura 3. Arquitectura del Módulo de Análisis Espacial

Este módulo contiene tres componentes: Bloque de Análisis, Lista de Procedimientos y Lista de Recursos: a) El Bloque de Análisis está compuesto de un conjunto de procesos para realizar el análisis de los datos; b) Lista de Procedimientos, se encarga de almacenar la secuencia de pasos para encontrar áreas de riesgo, así como la descripción de los datos requeridos, y c) Lista de Recursos, la cual contiene la descripción de los datos atributivos y espaciales que pueden consultarse utilizando la red.

Bloque de Análisis

Contiene las funciones para detectar las áreas de riesgo. Estas funciones son las siguientes:

- **Función de Interpolación.** Es utilizada para diversas tareas. El método propuesto es un spline con curvatura mínima en dos dimensiones, a partir de un conjunto de puntos. Para propósitos computacionales, el espacio entero de la salida del grid se divide en bloques o regiones de igual tamaño. Estos se representan a través de una forma regular. La ecuación 1 muestra la función de spline que se ha utilizado: [10]

$$S(x, y) = T(x, y) + \sum_{j=1}^N \lambda_j R(r_j), \quad (1)$$

donde $j = 1, 2, \dots, N$

N es el número de puntos.

λ_j son los coeficientes obtenidos del sistema de ecuaciones.

R_j es la distancia del punto (x, y) al j -ésimo punto.

Para esta función es necesario proporcionar el conjunto de puntos y tolerancias, que dependen directamente de un caso de estudio particular.

- **Generador de Pendientes.** La superficie de la Tierra presenta muchas irregularidades y diferentes características topográficas. Por lo tanto, la función se utiliza para evaluar el comportamiento del fluido en algún área. En este proceso las pendientes son utilizadas para determinar la velocidad de los fluidos e identificar las zonas de acumulación. El proceso lo describe en detalle M. Moreno. [11] La figura 4 muestra el proceso de generación de la capa de pendientes.

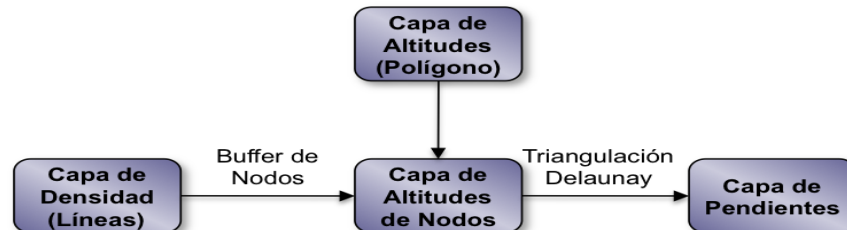


Figura 4. Proceso de generación de una capa de pendientes.

- **Generador de Grid.** Se utiliza para procesar algunos datos analizados, especialmente en la generación de mapas de densidad. Los grids vectoriales son regulares, de una magnitud $m \times m$, en donde m es el tamaño de la celda. La magnitud de la celda en el grid se determina por el fenómeno estudiado, escala y área de cobertura. Existen dos alternativas que pueden utilizarse para generar estos grids. Primeramente se especifica el inicio y el fin de las coordenadas del grid $((x_0, y_0), (x_1, y_1))$. Establece el número de divisiones requeridas para el grid. La segunda alternativa es especificar las coordenadas de inicio (x_0, y_0) , el tamaño de la celda, el número de columnas y renglones en el grid. [12] En la figura 5 se muestran las especificaciones del grid.

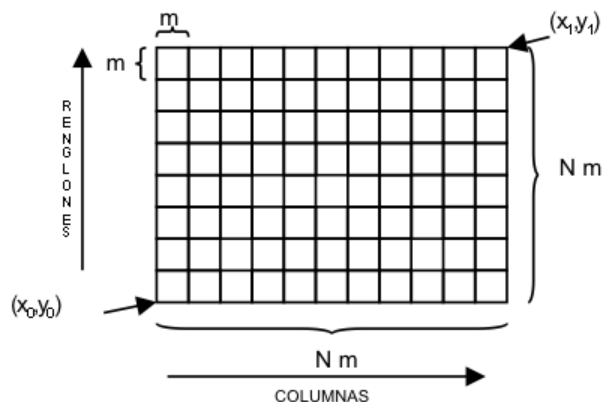


Figura 5. Proceso de generación del grid.

- **Funciones de sobreposición.** Este módulo se ha diseñado para realizar sobreposiciones topológicas, las cuales pueden ser utilizadas para identificar áreas de riesgo. Un conjunto de operaciones se han definido y aplicado al análisis espacial. Esto se realiza para establecer las condiciones y combinar capas de información diferente, utilizando operadores lógicos. Estas funciones combinan datos espaciales y atributivos. Además, las operaciones implementadas para la sobreposición topológica en esta aplicación, son: intersección, unión e identidad, las cuales se representan por \cap , \cup e I , respectivamente.

Lista de Procedimientos

Almacena el conjunto de procedimientos para cada uno de los procesos de análisis. Éste tiene una descripción del tipo de dato requerido y sus restricciones. Sin embargo, los usuarios pueden cambiar los criterios de selección. Este mecanismo proporciona una lista de funciones, como una alternativa para el análisis, en donde los parámetros pueden ser modificados. Las tareas de detección son las siguientes:

- **Mecanismo de Detección de Áreas de Inundación.** Para identificar las áreas con riesgo de inundación, es necesario construir una capa espacial de infiltración. [13] Con esta capa se puede definir la permeabilidad en el área de estudio. Por lo tanto, la capa de infiltración se diseña con la utilización de la geología. Posteriormente, estos datos espaciales son intersectados con las capas de densidad de drenaje y fracturación. Finalmente la infiltración se obtiene con base en el procedimiento que describe M Torres, et al. [12] Este proceso se muestra en la figura 6.

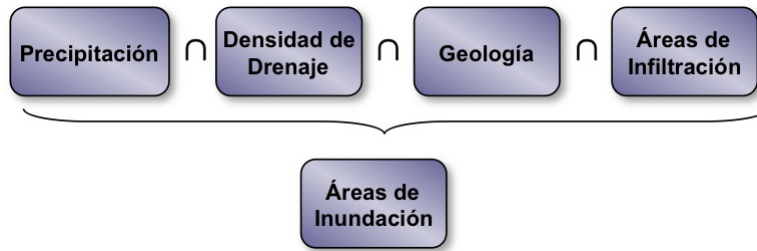


Figura 6. Mecanismo de detección de áreas de inundación.

- Mecanismo de Detección de Áreas de Deslave. Para identificar las áreas con riesgo de deslave, es necesario construir una capa de pendientes. [12] Las capas de geología, vegetación y uso de suelo juegan un papel importante en este análisis. Esto se debe a que la composición del suelo determina la afectación, cuando las condiciones climáticas cambian. Adicionalmente, las capas de cuencas, precipitación y densidad de drenaje, pueden utilizarse para encontrar posibles deslaves originados por las precipitaciones. En la figura 7 se muestra el proceso para detectar las áreas con riesgo de deslave.

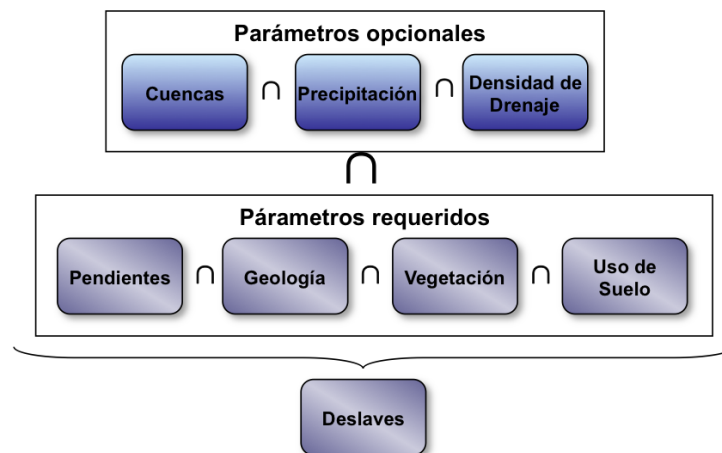


Figura 7. Mecanismo para detectar áreas de deslave.

Lista de Recursos

Este mecanismo almacena la descripción de los sitios remotos y SIG distribuidos. Esta descripción se utiliza para recuperar los datos espaciales desde la base de datos geográfica o espacial. Por lo tanto, los usuarios cuentan con la disponibilidad de seleccionar la lista de productos cartográficos digitales. Esta lista proporciona una vista general de las características cualitativas y cuantitativas de los datos espaciales.

Resultados

12 -xx

Con el uso de la aplicación SIG podemos contar con un mecanismo para detectar potencialmente áreas de riesgo, las cuales podrían ser afectadas por deslaves e inundaciones. La metodología diseñada ha sido aplicada al estado de Tamaulipas, México. Algunos resultados del uso de esta herramienta se presentan en esta sección.

Por ejemplo, en la figura 8 se muestran las siguientes capas espaciales recuperadas desde sitios SIG distribuidos: capas de densidad de drenaje, infraestructura y población.

Asimismo, la figura 9 muestra las áreas con riesgo de deslave e inundación. Estas áreas han sido recuperadas desde sitios remotos. Los datos espaciales están almacenados en una base de datos espacial local. Por lo tanto, el análisis espacial ha sido hecho en el SAM, proporcionando zonas con características particulares. Se puede observar que las áreas de riesgo están concentradas en el centro y sur del estado. También se pueden observar en la figura 9 las áreas clasificadas de acuerdo con el nivel de riesgo: bajo (azul), medio (amarillo) y alto (verde). De igual forma se utiliza una imagen Landsat TM para complementar y mejorar el análisis espacial.

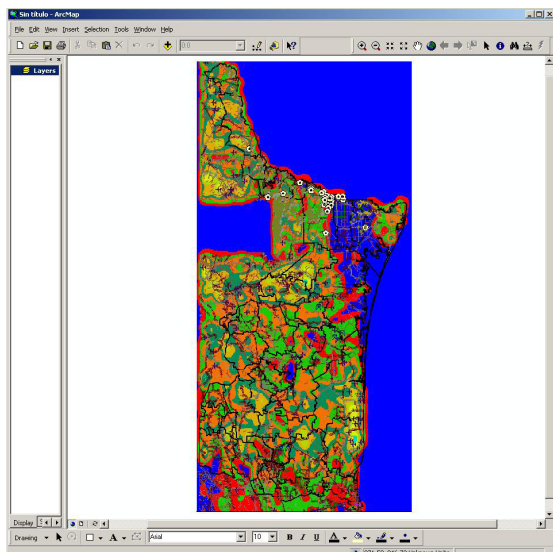


Figura 8. Capas de densidad de drenaje, infraestructura y poblaciones recuperadas de sitios SIG distribuidos

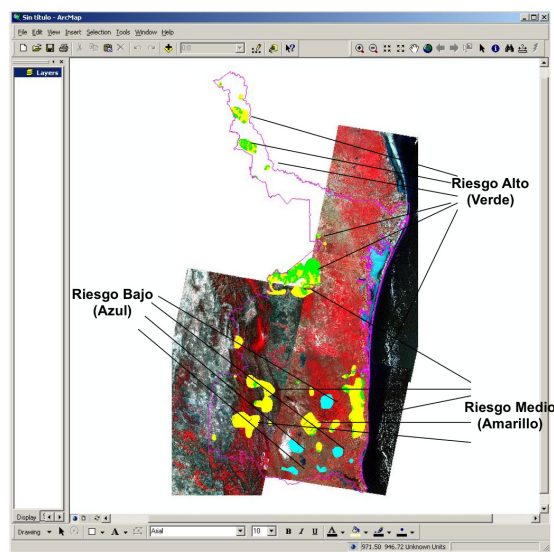


Figura 9. Áreas de riesgo en el estado de Tamaulipas, México

Finalmente, en la figura 10 se presenta una operación de zoom (acercamiento) sobre los datos espaciales recuperados de diferentes sitios SIG remotos. Se puede observar la infraestructura y la densidad de drenaje alrededor de la ciudad de Tampico.

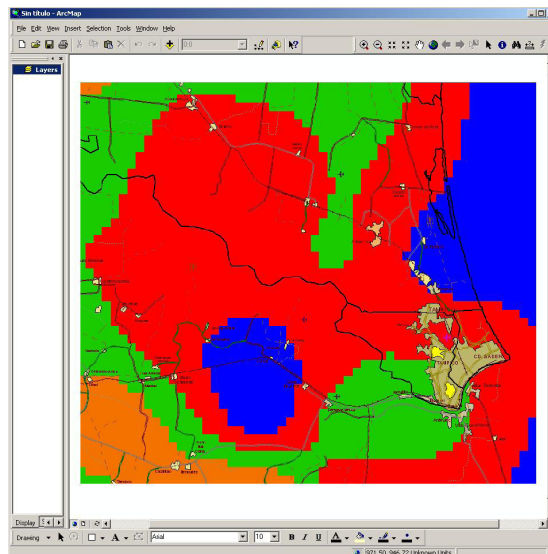


Figura 10. Capas temáticas recuperadas de un ambiente distribuido

Conclusiones

En este trabajo ha sido descrita una aplicación SIG que trabaja en un ambiente distribuido. Se propone una especificación basada en XML para solucionar parcialmente el problema de la heterogeneidad de los datos espaciales. Con esta aplicación los datos espaciales y descriptivos pueden ser recuperados desde diferentes sitios SIG, comparando sus propiedades cualitativas y cuantitativas, por medio de la descripción en XML.

La base de datos espacial ha sido diseñada e implementada en un SIG empresarial (ArcGIS). Este repositorio almacena todos los objetos geográficos recuperados desde diversos sitios SIG. El mecanismo proporciona un modelo topológico que se integra directamente por clases de elementos.

SAM es un módulo diseñado para realizar análisis espacial, relacionado con fenómenos naturales, tales como áreas con riesgo de deslave e inundación. Los usuarios pueden modificar los criterios de selección para contar con diferentes escenarios. Utilizando el SAM es posible definir la importancia de las características de los datos espaciales, que los usuarios pueden considerar o no relevantes para el análisis. Cuando la comunicación se establece, SAM requiere los datos espaciales por medio de la Lista de Recursos. Por lo tanto, la especificación basada en XML se compara con la definición XML local, con la finalidad de verificar los datos espaciales y recuperar la información de diferentes sitios SIG. Esta información es enviada a la base de datos espacial a través del mecanismo SDE. Las técnicas implementadas en la Lista de Procedimientos pueden modificarse y extenderse para aplicarse en el análisis de fenómenos diferentes.

Asimismo, se propone una descripción en XML, la cual puede ser utilizada como un estándar para solucionar parcialmente el problema de integración de los datos espaciales. Esta técnica es una alternativa para recuperar datos espaciales en un ambiente distribuido, ya que la definición en XML que ha sido diseñada, representa la semántica espacial de los objetos geográficos y puede utilizarse para encontrar soluciones relacionadas con la interoperabilidad de los datos espaciales.

5. Agradecimientos

Este trabajo ha sido auspiciado y apoyado por el Instituto Politécnico Nacional (IPN), el Consejo Nacional de Ciencia y Tecnología (CONACYT) y la Secretaría de Investigación y Posgrado del IPN, a través de los proyectos: 20110875, 20110988, 20110851, 20110881 y 20110302.

Referencias

1. M. F. Goodchild (2000) Perspective: Browsing metadata, where do we go from here? *GeoInfo Systems* 10(5): 30–31.
2. M. Torres, *Herramienta SIG de Escritorio para la Recuperación, Manejo y Análisis de datos espaciales* (CIC-IPN, México, D.F., 2001), (M.S. Thesis in Spanish).
3. M. Worboys, *Innovations in GIS Selected paper from the First National Conference on GIS Research* (Editorial GME, United Kingdom, 1994).
4. A. MacDonald, *Building a Geodatabase* (ESRI Press, United States, 2001).
5. ESRI Inc., *Understanding ArcSDE, Integrating GIS and RDBMS* (ESRI Press, USA, 1999).
6. OGC, Geography Markup Language (GML), Implementation Specification, version 2.1.1, 25 April 2002, <http://www.opengis.net/gml/02-009/GML2-11.html>
7. R. Sekhar & P.K. Srimani, A Specification of a Spatial Query Language over GML, *Proceedings of the 9th ACM International Symposium on Advances in Geographical Information Systems*, Atlanta, Georgia, USA, 2001, 112-117.
8. T. Arnold-Moore, M. Fuller, A. Kent, R. Sacks-Davis & N. Sharman, Architecture of a Content

- Management Server for XML Document Applications, *Proceedings of the 1st International Conference on Web Information Systems Engineering (WISE 2000)*, Hong Kong, 2000, 77-88.
9. R. Sacks-Davis, T. Arnold-Moore & J. Zobel, Database Systems for Structured Documents, *IEICE Transactions on Information and Systems*, 34-D(11), 1995, 37-48.
10. L. Mitás & H. Mitasova, General Variational Approach to the Interpolation Problem, *Journal of Computers and Mathematics with Applications*, 16(12), 1988, 983 - 992.
11. M. Moreno, *La Generalización Automática de la Información Geográfica Multiescala* (CIC-IPN, México, D.F., 2001), (M.S. Thesis in Spanish).
12. M. Torres, M. Moreno & S. Levachkine, SIGETAM: Herramienta GIS de Escritorio enfocada a la Detección de Zonas de Riesgo de Deslave e Inundación, *1st International Congress of Informatics and Computing*, Guadalajara, México, 2002, 156-168.
13. A. Linsley, K. Kohler & S. Paulos, *Hidrology for Engineers* (McGraw Hill, United States, 1998).