

Clasificación de Componentes SIG

Gabriela Gaetán, Viviana Saldaño
Grupo UNPA-UACO
Universidad Nacional de la Patagonia Austral
Unidad Académica Caleta Olivia
Santa Cruz, Argentina

Alejandra Cechich y Agustina Buccella
Grupo de Investigación GIISCo
Departamento de Ciencias de la Computación
Universidad Nacional del Comahue
Neuquén, Argentina

Resumen

Hoy en día existen muchos enfoques que proponen metodologías para el desarrollo de software. Entre ellos, el Desarrollo de Software basado en Componentes (DSBC) propone una manera más rápida y segura de desarrollar aplicaciones. Por otro lado, en cuanto a la información que se modela en las mismas, la información geográfica esta ganando más espacio. Es sabido que cada vez es mayor la cantidad de entidades u organizaciones en el mundo que están utilizando la información geográfica como base para sus sistemas de información. Así, tanto la creación de nuevos componentes geográficos como la necesidad de los mismos para incluirlos en Sistemas de Información Geográficos (SIG) han generado una explosión en el campo del DSBC para SIG. En este trabajo, nos centramos en ambas áreas creando nuevas herramientas que faciliten la búsqueda, publicación y selección de componentes tanto para los proveedores como para los consumidores de los mismos.

1. Motivación

En nuestro trabajo se combinan dos áreas de la Ciencia de la Computación para dar soluciones a la tarea de desarrollo de software. Por un lado, nos interesa el *Desarrollo de Software basado en Componentes (DS-BC)* ya que intenta minimizar tiempos de desarrollo y asegurar productos más confiables utilizando el *reuso* como herramienta fundamental de modelado [13, 14]. Principalmente está basado en la utilización de piezas prefabricadas reusando así tareas de análisis, diseño, e implementación de los servicios que proveen. A su vez, además de la reducción del tiempo de de-

sarrollo y costos, la calidad del producto final mejorará debido a que estas piezas ya han sido probadas y validadas.

Por otro lado, el área de los *Sistemas de Información Geográficos (SIG)* incluye a todos aquellos productos que poseen tanto información geográfica como herramientas para manipularla. Una definición completa de los SIG ha sido propuesta en [9] en donde “*un SIG es un sistema basado en computadora diseñado para modelar, capturar, almacenar, manipular, consultar, recuperar, analizar y visualizar información eficientemente, donde parte de la misma es de naturaleza geográfica*”.

Con esta definición se ve claramente que los SIG son más que herramientas para producir mapas. Mientras que en la cartografía tradicional el mapa es la base de datos, en los SIG el mapa solo es una proyección de una vista particular de una base de datos geográfica en un momento dado. De esta manera el usuario de un SIG posee un número ilimitado de alternativas de análisis y de alternativas para realizar mapas desde puntos de vista diferentes de acuerdo a diferentes aspectos de la información [3]. Las características exclusivas de la información geográfica generan requerimientos funcionales especiales en términos de modelos conceptuales y lógicos, estructura de datos, acceso a métodos, técnicas de análisis y procedimientos de visualización. Los mismos deben considerarse a la hora de iniciar un desarrollo de software para sistemas geográficos.

La combinación de estas dos áreas, DSBC y SIG, trae consigo varios aspectos a tener en cuenta tanto para la clasificación y selección de componentes como para los servicios requeridos de los mismos. Así hemos definido y trabajado en dos modelos, *de demanda* y *de oferta*, definiendo requerimientos especiales y soluciones de acuerdo a las restricciones y problemas que involucran cada uno de ellos.

El *modelo de demanda* se refiere a los requerimientos o servicios que necesitan los clientes de SIG (empresas que desarrollan SIG), los cuales, en un desarrollo basado en componentes, buscan en la Web componentes que brinden servicios que se adecúen a las necesidades de sus aplicaciones. Para esto, nuestro trabajo se centra en definir las herramientas necesarias para consultar en un sistema integrado los componentes que brindan esos servicios. Aquí, es crucial el nivel de especificidad y la forma en que los clientes solicitan dicha información para encontrar los componentes deseados.

El *modelo de oferta* se refiere a los requerimientos o servicios que brindan los desarrolladores de componentes para SIG. En general los componentes ofrecidos se publican en la Web con información asociada indicando aspectos técnicos y funcionales. Esto es llamado la documentación del componente la cual debe ser lo suficientemente clara, compacta y específica para ser bien comprendida por clientes.

De esta manera, para que los que demandan componentes se entiendan con aquellos que los ofertan, es fundamental la forma en que los componentes son publicados, llamado *proceso de publicación*, y la forma en que

los mismos son seleccionados de acuerdo a las necesidades de los clientes, llamado *proceso de selección* [10]. A su vez, para que ambos procesos se comuniquen surge la necesidad de utilizar un *modelo de mediación* que facilite tanto la recuperación de la información como la selección automatizada de componentes candidatos [6].

2. Contribuciones a la Fecha

Durante el año 2008 se ha trabajado en ambos modelos, de oferta y demanda, proponiendo metodologías que mejoren la forma de llevar a cabo las tareas incluídas en ellos.

Dentro del modelo de demanda hemos definido una taxonomía de servicios para SIG que brinde mejoras al proceso de selección y búsqueda de componentes en ese dominio. En este trabajo, publicado en [12], se estudiaron varias propuestas en la literatura que definen diferentes taxonomías y hasta incluso ontologías para clasificar la gama de servicios geográficos que pueden solicitarse [8, 11, 15]. Además se analizó la norma ISO 15115/OGC (Services) [2] y la Arquitectura de Servicios de OpenGIS (Service Architecture) [1] las cuales proveen un framework para crear software que permite a los usuarios acceder y procesar datos geográficos a partir de diferentes fuentes de datos a través de una interfaz genérica. Esta arquitectura se basa en el Modelo de Referencia para Procesamiento Distribuído Abierto (RM-ODP) y está conformada por un conjunto de componentes, conexiones y topologías. Utilizando toda esta información más la información obtenida de relevamientos efectuados a clientes SIG pertenecientes a organizaciones de la zona, se definió una taxonomía de servicios SIG, la cual cuenta con cuatro capas que incluyen un conjunto de servicios clasificados en cada una de ellas.

Dentro del modelo de oferta hemos definido otra taxonomía conteniendo la información necesaria para describir un componente y así ayudar al proceso de publicación. Este trabajo fue publicado en [5] y se basó principalmente en analizar cincuenta componentes SIG disponibles en la Web¹²³ y clasificarlos según la taxonomía propuesta. La misma se compone de 21 categorías distribuidas en tres grupos. El objetivo principal de este trabajo fue analizar qué información está realmente presente en la documentación de los catálogos de componentes para saber con que datos contar a la hora de implementar el modelo de mediación.

¹<http://www.componentsource.com/>

²<http://freegis.org/>

³<http://freshmeat.net/>

3. Trabajo Futuro

Para ambos modelos, oferta y demanda, se está actualmente trabajando en la creación de metodologías que extraigan la información necesaria para completar las dos taxonomías definidas en los trabajos previos. La idea principal es conocer qué información está disponible y qué información se necesita para poder implementar luego el modelo de mediación.

Para esto, se están analizando técnicas de lenguaje natural [7] que extraigan información de los catálogos de componentes en la Web y de los casos de uso (en UML [4]) para aplicarlas a ambos modelos.

4. Impacto y Tranferencia

El mayor impacto del proyecto se centra en la formación de recursos humanos. A la fecha algunos integrantes del grupo están iniciando sus maestrías en la Universidad Nacional de La Plata bajo la supervisión de investigadores de UNC. Al mismo tiempo, se transfieren técnicas y métodos de investigación que permitan el crecimiento sostenido de la actividad, tanto individualmente como grupalmente. La futura consolidación de estas actividades y relaciones permitiría pensar en futuras transferencias al medio y en mayor impacto desde el punto de vista de contribución al conocimiento científico.

5. Investigadores

Este proyecto reúne aproximadamente 6 (seis) investigadores, entre los que se cuentan docentes de la Universidad Nacional del Comahue (UNC) y de la Universidad de la Patagonia Austral (UNPA) y alumnos de la UNPA. Algunos de los docentes-investigadores de UNC-UNPA han comenzado sus carreras de postgrado y otros están terminado. Durante el año 2009 se avanzará en las 2 tesis de maestrías iniciadas con maestrandos en UNPA supervisados por investigadores de UNC.

Contacto: Ing. Gabriela Gaetán (ggaetan@uaco.unpa.edu.ar, gabrielagaetan@yahoo.com.ar)

Referencias

- [1] Open gis consortium. The OpenGIS Abstract Specification: Service Architecture, 2002.
- [2] Geographic information. Services, International Standard 19119, ISO/IEC, 2005.
- [3] P. Burrough and R. McDonnell. *Principles of Geographical Information Systems*. Oxford University Press, 1998.

- [4] Martin Fowler and Kendall Scott. *UML distilled (2nd ed.): a brief guide to the standard object modeling language*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2000.
- [5] G. Gaetán, A. Buccella, and A. Cechich. Un esquema de clasificación facetado para publicación de catálogos de componentes sig. In *Proceedings of the CACIC'08: XIV Congreso Argentino en Ciencias de la Computación*, Chilecito, La Rioja, Argentina, 2008.
- [6] L. Iribarne. *Un Modelo de Mediación para el Desarrollo de Software Basado en Componentes COTS*. Phd thesis, Univerdad de Almería, 2003.
- [7] Z. Kedad, N. Lammari, E. Métais, F. Meziane, and Y. Rezgui, editors. *Natural Language Processing and Information Systems*. Lecture Notes in Computer Science 4592. Springer-Verlag, 2007.
- [8] W. Li, S. Zhao, H. Sun, and X. Zhang. Ontology-based qos driven gis grid service discovery. In *Proceedings of the SKG'06: Second International Conference on Semantics, Knowledge, and Grid*, 2006.
- [9] M.A. Rodríguez Luaces. *A Generic Architecture for Geographic Information Systems*. PhD thesis, Univerdade da Coruña, 2004.
- [10] V. F. Lucena. *Flexible Web-based Management of Components for Industrial Automation*. Phd thesis, Stuttgart University, 2002.
- [11] M. Lutz. Ontology-based service discovery in spatial data infrastructures. In *Proceedings of the GIR'05: Proceedings of the 2005 workshop on Geographic information retrieval*, pages 45–54, New York, NY, USA, 2005. ACM.
- [12] V. Salda no, A. Buccella, and A. Cechich. Una taxonomía de servicios geográficos para facilitar la identificación de componentes. In *Proceedings of the CACIC'08: XIV Congreso Argentino en Ciencias de la Computación*, Chilecito, La Rioja, Argentina, 2008.
- [13] C. Szyperski. *Component Software-Beyond Object-Oriented Programming*. Addison-Wesley, 1998.
- [14] K. Wallnau, S. Hissam, and R. Seacord. *Building Systems from Commercial Components*. Addison-Wesley, 2002.
- [15] P. Yue, L. Di, P. Zhao, W. Yang, G. Yu, and Y. Wei. Semantic augmentations for geospatial catalogue service. In *Proceedings of the IGARSS'06: Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium*, Denver, USA, 2006.