

Geometría Computacional y Bases de Datos

Maria Gisela Dorzán, Susana Esquivel, Edilma Olinda Gagliardi,
Pablo Rafael Palmero y Maria Teresa Taranilla

Departamento de Informática, Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis, Argentina
{mgdorzan, esquivel, oli, prpalmero, tarani }@unsl.edu.ar

Gregorio Hernández Peñalver

Departamento de Matemática Aplicada, Facultad de Informática
Universidad Politécnica de Madrid, España
gregorio@fi.upm.es

Maria Mercedes Guasch y Maria Rosana Piergallini

Escuela de Informática y Tecnología
Universidad Nacional del Noroeste de la Provincia de Buenos Aires, Argentina
mariarosanapier@hotmail.com, mercedes_guasch@yahoo.com

Resumen

La línea de investigación *Geometría Computacional y Bases de Datos* se enmarca en el proyecto *Tecnologías Avanzadas de Bases de Datos* de la Universidad Nacional de San Luis. En la línea de investigación, vinculamos temáticas de las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas debido a que en diversos campos de aplicación relacionados con bases de datos no estructuradas, surge la necesidad de construir y utilizar objetos y estructuras geométricas que tengan propiedades deseables. Algunos de los problemas estudiados son de naturaleza NP-dura y por tal motivo utilizamos técnicas metaheurísticas para encontrar soluciones aproximadas para ellos.

Palabras clave: Bases de datos, Geometría Computacional Metaheurísticas, Bases de datos espacio-temporales.

Contexto

El proyecto *Tecnologías Avanzadas de Bases de Datos* de la Universidad Nacional de San Luis, tiene como objetivo principal el estudio

de bases de datos avanzadas, en este marco se investiga el diseño y desarrollo de herramientas para administrar eficientemente sistemas de bases de datos no estructuradas.

En el proyecto existen tres líneas de investigación, orientadas al desarrollo de nuevos modelos para administrar y recuperar información almacenada en repositorios de datos no estructurados, donde los escenarios de exploración requieren modelos más generales tales como las bases de datos espacio-temporales, bases de datos de texto, espacios métricos, entre otros.

En la línea de investigación *Geometría Computacional y Bases de Datos*, nos dedicamos al estudio de objetos y estructuras geométricas de utilidad en diversas áreas de aplicación, por ejemplo, robótica, visión artificial, computación gráfica, sistemas de información geográfica, computación móvil, entre otras, y que se relacionan con las bases de datos mencionadas.

En diversas aplicaciones del campo de las Ciencias de la Computación se requiere de repositorios no tradicionales, que conllevan a nuevos modelos de bases de datos para administrar y buscar información en ellos. Así,

surge la necesidad de estudiar modelos como las bases de datos espacio-temporales. También, se requiere la construcción de diferentes objetos y estructuras geométricas que cumplan con propiedades deseables. En particular, algunos de los problemas relacionados con la optimización de estructuras geométricas son problemas NP-duros, por tanto en la búsqueda de soluciones aproximadas hemos utilizado técnicas metaheurísticas. Así, hemos vinculado temáticas que surgen de las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas.

Desde el año 2002, desarrollamos nuestro trabajo de investigación en forma conjunta con investigadores afines de proyectos locales y de la Universidad Politécnica de Madrid, en el marco de los convenios de cooperación institucional. En este trabajo, presentamos los temas en estudio, junto con las propuestas más recientes de interés.

Introducción

Diversas aplicaciones requieren guardar y consultar información histórica y actual, relacionada a los cambios de forma y/o posición que tuvieron los objetos estudiados en diferentes escenarios a lo largo del tiempo. Por este motivo, se requieren modelos de bases de datos espacio-temporales para tales requerimientos y es necesario contar con herramientas teóricas, de base, que permitan modelar estos tipos de datos, realizar operaciones sobre ellos, definir lenguajes de consulta, analizar su expresividad, entre otras propiedades.

La Geometría Computacional estudia el diseño de algoritmos eficientes para resolver problemas de tipo geométrico. Una tarea fundamental en esta disciplina consiste en identificar conceptos, propiedades y técnicas que apuntan al desarrollo de algoritmos eficientes. Esto conduce al análisis y estudio de estructuras de datos geométricas y problemas relacionados con dichas estructuras [BKOS97]. Estos problemas pueden ser estudiados desde un punto de vista combinatorio o algorítmico. Los aspectos

algorítmicos de los problemas aparecen en la búsqueda de soluciones, exactas o aproximadas, para una estructura geométrica particular.

En este contexto aparecen problemas que son NP duros, o bien, son problemas para los cuales no se conocen algoritmos eficientes que permitan alcanzar soluciones exactas. De todos modos, es de interés obtener soluciones a tales problemas, utilizando métodos de carácter heurístico, aunque sean aproximadas a las óptimas. Debido a esto, utilizamos técnicas metaheurísticas, ya que las mismas permiten encontrar soluciones cercanas a las óptimas. Una metaheurística se puede considerar como una metodología de alto nivel que puede ser utilizada como una estrategia que guía las heurísticas subyacentes para resolver problemas de optimización específicos, combinando inteligentemente diferentes conceptos de diversos campos. Estos métodos son simples de implementar y han demostrado ser exitosos en encontrar, de forma eficiente, buenas soluciones para problemas de optimización NP-duros [MF04] [BFM97].

Algunos de los problemas que estudiamos corresponden a ciertas estructuras geométricas, obtenidas a partir de un conjunto de puntos u objetos en el plano, y para las cuales se busca optimizar alguna propiedad que permita valorar la calidad de dichas estructuras. Los criterios de optimización consisten en minimizar o maximizar alguna propiedad de las estructuras geométricas. Los problemas de optimización relacionados a triangulaciones, pseudo-triangulaciones y poligonizaciones, son de interés debido a que son utilizadas en diversos campos por ejemplo, en visibilidad y vigilancia, detección de colisiones, problemas de rigidez de estructuras, entre otros.

Algunos de los problemas de optimización estudiados son: la Triangulación de Peso Mínimo (*Minimum Weight Triangulation, MWT*) y la Pseudo-Triangulación de Peso Mínimo (*Minimum Weight Pseudo-Triangulation, MWPT*). Dichos problemas minimizan la suma de las longitudes de las aristas proveyendo así una medida de calidad para determinar cuán “buena” es la estructura. La complejidad del cálculo de MWT fue uno

los problemas abiertos más estudiados en Geometría Computacional hasta que Mulzer y Rote demostraron que la construcción de MWT es un problema NP-duro [MR06]. Por otra parte, la complejidad de la pseudo-triangulación de peso mínimo aún no está resuelta. Levcopoulos y Gudmundsson muestran una 12-aproximación de una pseudo-triangulación que puede ser calculada con complejidad de $O(n^3)$ [GL07]. Los autores dan una aproximación de $O(\log n w(\text{MST}))$ de una pseudo-triangulación de peso mínimo en un tiempo de $O(n \log n)$ donde $w(\text{MST})$ es el peso del Árbol de Expansión Mínimo (*Minimum Spanning Tree, MST*) el cual es un subconjunto de la estructura obtenida.

En cuanto a las poligonizaciones, estudiamos el problema de obtener poligonizaciones de área mínima para un conjunto de puntos en el plano (*Minimum Area Polygonization, MAP*). Fekete ha demostrado que este problema es NP-duro [Fe00].

Debido a la complejidad de los problemas antes mencionados, aplicamos técnicas metaheurísticas para su resolución ya que permiten obtener soluciones que se aproximan a las óptimas.

En esta línea de investigación nos dedicamos al estudio, diseño y desarrollo de índices espacio-temporales, aplicables a diversos escenarios de movimiento (redes, espacios libres de obstáculos, etc.), considerando la geometría como una disciplina marco en la cual se formalizan aspectos propios de los problemas involucrados. En este contexto, nos proponemos el estudio de optimizaciones de estructuras geométricas que están relacionadas con las bases de datos ya mencionadas y al desarrollo de herramientas para la visualización de estructuras geométricas y aplicaciones vinculadas a las bases de datos espacio-temporales.

Líneas de investigación

La Geometría Computacional estudia problemas desde un punto de vista geométrico, dedicándose al diseño y análisis de algoritmos y/o estructuras geométricas adecuados para su resolución. Algunos de los problemas de

estudio la disciplina provienen de diversas áreas, por ejemplo en los sistemas de información geográfica, problemas de referidos a localización, planificación de rutas y búsqueda geométrica; en robótica, problemas de planificación de movimientos, de visualización y detección de colisiones, entre otros. En particular, los problemas de interés en la línea de investigación son aquellos vinculados al diseño de índices espacio-temporales para resolver en forma integral consultas espacio-temporales y su vinculación con la problemática de objetos móviles; y al tratamiento de problemas geométricos NP-duros. Para estos últimos, en la búsqueda de soluciones proponemos utilizar técnicas metaheurísticas, ya que proporcionan estrategias adecuadas para la obtención de soluciones aproximadas de múltiples problemas de optimización.

Como objetivos específicos de estudio podemos enumerar los siguientes:

i) Indexación espacio-temporal sobre objetos en movimiento para diversos escenarios. Desarrollo de las estructuras de almacenamiento, los algoritmos de consulta y la evaluación experimental, mostrando el desempeño de los distintos índices en aplicaciones de diferentes magnitudes respecto de la población de objetos en movimiento. Desarrollo de aplicaciones con herramientas de Geometría Computacional y Bases de Datos Espacio-Temporales. Desarrollo de una herramienta integral en el ámbito de la Salud para el seguimiento de focos epidémicos utilizando base de datos espacios-temporales.

ii) Estudio de configuraciones geométricas generales de puntos en el plano, que cumplan medidas de calidad como: peso, dilación, área, entre otras. Optimización de Triangulaciones, Pseudotriangulaciones y Poligonizaciones considerando medidas de calidad mínimas o máximas, aplicando técnicas metaheurísticas y diversas estrategias algorítmicas.

iii) Diseño y desarrollo de herramientas para la generación, visualización y manipulación de diferentes configuraciones geométricas de conjuntos de puntos en el plano.

Resultados Obtenidos

Para la resolución de problemas de optimización de Triangulaciones y Pseudo-triangulaciones de Peso Mínimo aplicamos las técnicas metaheurísticas: Optimización basada en Colonia de Hormigas (Ant Colony Optimization, ACO) y Recocido Simulado (Simulated Annealing, SA), técnicas determinísticas Voraces (Greedy) y Triangulación de Delaunay. Se llevó a cabo el estudio, adecuación y evaluación experimental de las técnicas metaheurísticas mencionadas para la búsqueda de triangulaciones y pseudo-triangulaciones que cumplan dicha propiedad. Los resultados obtenidos para los problemas MWT y MWPT utilizando la técnica metaheurística ACO están publicados en [DGLH11b] [GDLH11]. Actualmente se está realizando la evaluación experimental y análisis de los resultados obtenidos con la técnica Recocido Simulado. Además se está realizando una evaluación estadística para poder comparar ambas técnicas y poder determinar cuál es la mejor [DGLH11a] [DGLH11c].

Para el problema de optimización de poligonizaciones de puntos en el plano, se utilizaron técnicas metaheurísticas basadas en Colonia de Hormigas (Ant Colony Optimization, ACO) y técnicas Voraces (Greedy) para aproximar poligonizaciones de área mínima. Los resultados obtenidos se publicaron en [TGH11].

Por otra parte, se ha implementado una herramienta para la generación y visualización de triangulaciones pseudo-triangulaciones y poligonizaciones de conjuntos de puntos en el plano. La herramienta permite generar diferentes tipos de las estructuras geométricas ya mencionadas, realizar operaciones sobre ellas y describir sus propiedades más importantes, observar la morfología de cada instancia generada y el comportamiento de los algoritmos generadores de cada una de ellas, recuperar instancias de tales configuraciones geométricas a partir de resultados obtenidos de las generaciones experimentales alcanzados mediante técnicas metaheurísticas [DGPH10].

Se está desarrollando una aplicación en el ámbito de la Salud para el seguimiento de focos epidémicos utilizando base de datos espacios-temporales y herramientas de Geometría Computacional.

Actualmente, se continúa trabajando en los problemas de optimización mencionados considerando otras técnicas metaheurísticas, que puedan adecuarse para su resolución. Se continúa con estudios experimentales para determinar qué técnica ofrece mejores resultados a los problemas planteados.

Como trabajo futuro, específicamente para problemas geométricos, nos enfocamos en el estudio problemas de optimización de propiedades que pueden satisfacer las triangulaciones, pseudo-triangulaciones y poligonizaciones, tales como dilación, peso, área, entre otras. Se pretende estudiar y proponer otras técnicas para su resolución, para lo cual se realizará la evaluación y análisis de las distintas técnicas para determinar su funcionalidad e impacto en la comunidad científica. El análisis incluirá el correspondiente tratamiento estadístico, y comparativo con otros algoritmos, ya sean de tipo aproximado o de tipo exacto.

Formación de Recursos Humanos

La formación del grupo de trabajo en la Universidad Nacional de San Luis, se ha consolidado con actividades de cooperación mutua e intercambio recíproco de información científica, tecnología y desarrollo de nuevos conocimientos con investigadores locales y de otras universidades.

Entre las actividades desarrolladas, se mencionan:

- i) formación de recursos humanos plasmada en trabajos finales de Licenciatura en Ciencias de la Computación, Tesis de Maestría en Ciencias de la Computación y Tesis Doctorales en desarrollo, ii) direcciones de becas de investigación de Ciencia y Técnica-FCFMyN-UNSL y de CONICET, iii) realización y dirección de pasantías de investigación con docentes de otras universidades (UNNOBA), iv) actividades de formación académica, cursos de posgrado y de especialización,

v) actividades de divulgación científica, conferencias y publicaciones en congresos y revistas en el ámbito nacional e internacional.

Nos proponemos continuar con las actividades integradoras y relacionadas al presente proyecto, proponiendo actividades de formación académica, de formación de recursos humanos locales y de otras universidades nacionales, investigación, desarrollo, y otras actividades académico-científicas vinculantes.

Referencias

- [BFM97] Bäck T., Fogel D., Michalewicz Z. *Handbook of Evolutionary Computation*. IOP Publishing Ltd and Oxford University Press. 1997.
- [BKOS97] de Berg, M., van Kreveld, M., Overmars, M., Schwarzkopf O. *Computational Geometry: Algorithms and Applications*. Springer-Verlag, Heidelberg, 1997.
- [DGPH10] Dorzán M.G., Gagliardi, E. O., Palmero P.R., Hernández Peñalver, G. *Una herramienta para la generación y visualización de Triangulaciones y Pseudotriangulaciones*. XVI Congreso Argentino de Ciencias de la Computación 2010 (CACIC 2010), 2010
- [DGLH11a] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Approaches for MWT and MWPT Problems*. XVII Congreso Argentino de Ciencias de la Computación 2011 (CACIC 2011), 2011
- [DGLH11b] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Using ACO metaheuristic for MWT problem*. XXX International Conference of the Chilean Computer Science Society. ISBN 978-0-7695-4689-6. Chile. 2011
- [DGLH11c] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Metaheuristic approaches for MWT and MWPT Problems*. XIV Encuentros de Geometría Computacional. Publicado por Centre de Recerca Matemàtica (CRM) en sus series Documents. ISSN 2014-2323 (printed edition). ISSN 2014-2331 (electronic edition), pág: 79-82. 2011
- [Fe00] Fekete, S. P. *On simple polygonizations with optimal area*. Discrete and Computational Geometry, 23, pp. 73-110. 2000.
- [GDLH11] Gagliardi E. O., Dorzán M. G., Leguizamón M. G. y Hernández Peñalver. G.; *Approximations on Minimum Weight Pseudo-Triangulation problem using Ant Colony Optimization*. XXX International Conference of the Chilean Computer Science Society. ISBN 978-0-7695-4689-6. Chile. 2011
- [GL07] Gudmundsson J., Levcopoulos C.; *Minimum weight pseudo-triangulations*. Computational Geometry. Theory and applications. Elsevier Vol. 38-pages 139-153, 2007.
- [MR06] Mulzer W., Rote G. *Minimum weight triangulation is NP-hard*. In Proceedings of the 22nd Annual ACM Symposium on Computational Geometry. 2006.
- [MF04] Michalewicz Z., Fogel D., *How to Solve It: Modern Heuristics*, Springer, 2004.
- [TGH11] Taranilla, M. T.; Gagliardi, E. O.; Hernández Peñalver G. *“Approaching Minimum Area Polygonization”*. XVII Congreso Argentino de Ciencias de la Computación 2011 (CACIC 2011), 2011.