

Geometría Computacional y Bases de Datos

Maria Gisela Dorzán, Edilma Olinda Gagliardi, María Teresa Taranilla
Departamento de Informática, Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis, Argentina
{mgdorzan, oli, tarani}@unsl.edu.ar
Gregorio Hernández Peñalver
Departamento de Matemática Aplicada, Facultad de Informática
Universidad Politécnica de Madrid, España
gregorio@fi.upm.es

RESUMEN

La línea de investigación *Geometría Computacional y Bases de Datos* pertenece al proyecto *Tecnologías Avanzadas de Bases de Datos*, de la Universidad Nacional de San Luis. En esta línea de trabajo nos dedicamos al diseño e implementación de diversos tipos de estructuras de almacenamiento, tales como índices para el almacenamiento y consulta de datos de tipo espacio-temporales, como asimismo al estudio y diseño de estructuras geométricas y problemas relacionados a las mismas, mediante el análisis de propiedades que constituyen medidas de calidad que permiten estimar la bondad de las mismas.

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

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 donde se estudia el diseño y desarrollo de herramientas para administrar eficientemente sistemas de bases de datos no estructuradas. Este proyecto posee tres líneas de investigación orientadas al desarrollo de nuevos modelos para buscar y administrar la información en almacenamientos de este tipo, 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 particular, la

línea de investigación *Geometría Computacional y Bases de Datos*, perteneciente a dicho proyecto, viene desarrollándose desde el año 2002, 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.

Nos dedicamos al estudio y tratamiento de objetos y/o estructuras de tipo geométrico, que son de utilidad en diversos campos de aplicación, por ejemplo, robótica, visión artificial, computación gráfica, sistemas de información geográfica, computación móvil, diseño asistido por computadora, entre otras, y que se relacionan con las bases de datos antes mencionadas. Respecto de las temáticas de investigación, hemos vinculado las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas, debido a que en diversas aplicaciones dentro del campo de las Ciencias de la Computación se requiere la construcción y manejo de diferentes objetos geométricos, con propiedades deseables.

También, 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.

En particular, algunos de los problemas estudiados necesitan algoritmos eficientes para su resolución, pero dada su naturaleza NP-dura, utilizamos técnicas metaheurísticas para hallar soluciones aproximadas. En este trabajo, presentamos los tópicos más relevantes,

actualmente en estudio, con las propuestas más recientes y/o de interés.

1. INTRODUCCIÓN

Considerando la diversidad de los datos disponibles en forma digital, la evolución de las tecnologías de información y comunicación, y la variedad de aplicaciones existentes, podemos observar el surgimiento de repositorios o almacenamientos no estructurados de información. En este contexto, se consultan nuevos tipos de datos tales como datos geométricos, texto libre, imágenes, audio y video, donde ocurre que la información no se puede estructurar en forma tradicional. Incluso, cuando es posible una estructuración clásica, nuevas aplicaciones requieren acceder a la base de datos por cualquier campo requiriendo muchas veces hacer uso de herramientas no tradicionales, tales como las provistas por la Inteligencia Computacional.

Otras aplicaciones requieren guardar y consultar información histórica y actual, acerca de los cambios de forma y/o posición que tuvieron los objetos de estudio en diferentes escenarios a lo largo del tiempo, por lo que se requiere de modelos de bases de datos espacio-temporales para tales requerimientos. Por tanto, 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 problemas desde un punto de vista geométrico, dedicándose al diseño y análisis de algoritmos y/o estructuras geométricas adecuados para su resolución. En ocasiones, permite obtener soluciones más eficientes a problemas que no parecen geométricos, por lo que descubrir que los datos de un problema verifican propiedades geométricas sirve para poder aplicar alguna técnica algorítmica o alguna estructura especial, y así describir una solución más apropiada. En el ámbito de esta disciplina, existen problemas que o bien, son de

naturaleza NP-dura, o bien, son problemas para los cuales no se conocen ninguna solución eficiente. De todos modos, resulta de interés encontrar soluciones a tales problemas, aunque las mismas sean aproximadas a las óptimas, por medio de métodos de carácter heurístico. Algunos de los problemas conciernen a ciertas configuraciones geométricas, obtenidas a partir de un conjunto de puntos u objetos en el plano, y se busca optimizar ciertas propiedades que miden la calidad de dichas configuraciones. Para ello, proponemos la aplicación de técnicas metaheurísticas, ya que las mismas nos permiten encontrar soluciones cercanas a las óptimas. Una metaheurística es un proceso de generación iterativo que guía la búsqueda de soluciones combinando inteligentemente diferentes conceptos de diversos campos, dando un marco algorítmico general, que adaptado a un problema específico, puede ser aplicado en problemas de optimización con pocas modificaciones. 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].

La optimización de problemas relacionados a configuraciones geométricas tales como las triangulaciones, pseudotriangulaciones y poligonizaciones, son de interés debido a que son utilizadas en diversos campos, como por ejemplo, en visibilidad y vigilancia, detección de colisiones, problemas de rigidez, entre otros.

Algunos de los criterios de optimización consisten en minimizar o maximizar alguna propiedad de triangulaciones o pseudotriangulaciones. En particular, la triangulación de peso mínimo (*Minimum Weight Triangulation*, *MWT*) y la pseudotriangulación de peso mínimo (*Minimum Weight Pseudo-Triangulation*, *MWPT*) minimizan la suma de las longitud de las aristas proveyendo así una medida de calidad para determinar cuán “buena” es la estructura. La complejidad del cálculo de *MWT* fue uno de los problemas abiertos más estudiados en Geometría Computacional hasta

que, en 2006, Mulzer y Rote [MR06] demostraron que la construcción de MWT es un problema NP-duro. Por otra parte, la complejidad de la pseudotriangulación de peso mínimo aún no está resuelta. Levkopoulos y Gudmundsson [GL07] muestran una 12-aproximación de una pseudotriangulación que puede ser calculada con complejidad de $O(n^3)$. Los autores dan una aproximación de $O(\log n \cdot w(\text{MST}))$ de una pseudotriangulación 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.

Con respecto a las poligonizaciones, es interesante detectar aquellas que son óptimas respecto de ciertos criterios como por ejemplo, minimizar o maximizar el área o el perímetro del polígono obtenido. Fekete [Fe00] demostró que la optimización del área es un problema NP-completo. La minimización del perímetro es también un problema NP-completo dada su relación con la versión geométrica del Problema del Viajante [GJ79].

Dada la dificultad inherente de los problemas antes mencionados, los algoritmos aproximados aparecen como candidatos alternativos para la resolución de estos problemas ya que permiten obtener soluciones aproximadas a las óptimas.

Por lo expuesto, 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. Además, en este contexto, nos planteamos el estudio de optimizaciones de estructuras tales como triangulaciones, pseudotriangulaciones y poligonizaciones. Entre las principales medidas de calidad consideramos el peso, la dilación, el número de apuñalamiento, entre otras. También, nos dedicamos al estudio de problemas de minimización de la suma de longitudes, minimización del camino más largo entre sus vértices, optimización del

vector de ángulos de la triangulación, entre otros ejemplos.

2. LÍNEA DE INVESTIGACIÓN

La Geometría Computacional estudia técnicas, algoritmos y estructuras de datos adecuadas para la resolución de problemas geométricos. Estos problemas se encuentran en diversas áreas de las Ciencias de la Computación, tales como Computación Gráfica, Planificación de Movimientos, Visión Computacional, Robótica, Recuperación de la Información, Bases de Datos Espacio-Temporales, entre otras.

En particular, los problemas de interés en la línea de investigación son aquellos vinculados al diseño de índices espacio-temporales para resolver integralmente consultas espacio-temporales y su vinculación con la problemática de objetos móviles; y al tratamiento de problemas geométricos NP-duros. Para estos últimos, proponemos la búsqueda de soluciones utilizando técnicas metaheurísticas, que proporcionan estrategias adecuadas para la obtención de soluciones aproximadas de múltiples problemas de optimización.

Entre los tópicos de estudio de la línea, destacamos los siguientes:

i) Estudio de triangulaciones de puntos que cumplan ciertas medidas de calidad como son el peso, la dilación y el número de apuñalamiento. Optimización de triangulaciones de peso, dilación y número de apuñalamiento mínimos utilizando técnicas metaheurísticas. Estudio de problemas de visibilidad o iluminación en los poliedros-terreno. Resolución aproximada mediante metaheurísticas a los problemas de iluminación en triangulaciones planas. [MR06] [FLM08] [CS89] [M04].

ii) Optimización de rutas de vigilancia entre obstáculos poligonales mediante la aplicación de técnicas metaheurísticas. Resolución aproximada de diferentes variantes de problemas de vigilancia en polígonos, aplicando metaheurísticas. Metaheurísticas

para la optimización de poligonizaciones de puntos en el plano. [Fe00] [CNN93].

iii) Optimización de pseudotriangulaciones, que satisfagan propiedades como bajo peso, dilación pequeña, número de apuñalamiento pequeño, entre otras, mediante la aplicación de técnicas metaheurísticas. [PV96] [RSS06] [GL07].

iv) Indexación espacio-temporal sobre objetos en movimiento para diversos escenarios. Estudio de nuevas estrategias de ruteo en grafos geométricos para su aplicación a objetos móviles. Desarrollo de aplicaciones con herramientas de Geometría Computacional y Bases de Datos Espacio-Temporales. [CGG09] [CLGH10].

3. RESULTADOS OBTENIDOS Y VISIÓN DE FUTURO

Existen aplicaciones que requieren manejar objetos espacio-temporales, es decir, objetos cuya posición espacial o forma cambia en distintos instantes de tiempo. Por consiguiente, deberían considerarse métodos de acceso espacio-temporales, que sean capaces de soportar estos nuevos tipos de datos y de responder tipos de consultas apropiados, satisfaciendo de esta manera, las demandas de los usuarios. En general, TimeSlice, Intervalo, Eventos y Trayectoria son los tipos de consultas de interés. En la línea de investigación, estudiamos métodos de acceso espacio-temporal, que integralmente, permiten resolver estos cuatro tipos de consultas, sin aumentar la complejidad espacio-temporal, en diversos escenarios. Desarrollamos las estructuras de almacenamiento, los algoritmos de consulta y la evaluación experimental, mostrando el buen desempeño de los distintos índices en aplicaciones de diferentes magnitudes respecto de la población de objetos en movimiento [CGG09] [CLGH10].

Como antecedentes acerca de los problemas geométricos de complejidad NP-duros tratados con técnicas metaheurísticas, en el ámbito de la geometría computacional, que se resolvieron utilizando este tipo de métodos aproximados, podemos mencionar: Minimum Vertex Guard

[CH04], Maximum Hidden Vertex Set, [BCHM08], Descomposición de Minkowski [GLHT07]. En [DGLH09a] [DGLH09b], se muestran los resultados obtenidos para los problemas MWT y MWPT utilizando la técnica metaheurística Optimización basada en Colonias de Hormigas (Ant Colony Optimization, ACO) [DS04] donde además se describen los algoritmos ACO propuestos y el conjunto de instancias generadas para la evaluación experimental. Actualmente, se está trabajando en los problemas mencionados anteriormente, pero atacados con otras técnicas metaheurísticas, como por ejemplo Simulated Annealing y Algoritmos Genéticos. Se pretende realizar un estudio experimental para comprobar cuál es la técnica que brinda mejores resultados a los problemas planteados.

Como trabajo futuro, específicamente para problemas geométricos, están en estudio problemas de optimización de propiedades que pueden satisfacer las triangulaciones, pseudotriangulaciones y poligonizaciones, tales como dilación, apuñalamiento, peso, grado, área, entre otras. Por ello, consideraremos los siguientes problemas para las clases de estructuras geométricas mencionadas: construcciones de peso bajo, de dilación pequeña, con número de apuñalamiento pequeño, de bajo grado de vértices y de perímetro o área mínima o máxima.

Se propone el estudio y análisis de adecuación algunas técnicas de computación evolutiva [BFM97]; y a posteriori, el estudio y análisis de adecuación de técnicas de basadas en el paradigma de Inteligencia Colectiva (Swarm Intelligence) [KE01], entre otras [MF04].

La etapa experimental se realizará empleando casos de prueba de relevancia, mediante la generación y análisis pertinente de lotes de prueba, que sean de amplio espectro, y que permitan determinar cuáles técnicas permiten resolver con eficiencia los problemas propuestos. El análisis, la corrección y ajuste de los algoritmos propuestos se realizará a lo largo de desarrollo de trabajo en un proceso de refinamiento y perfección de las propuestas.

4. FORMACIÓN DE RECURSOS HUMANOS

Entre las actividades desarrolladas, se mencionan las siguientes: actividades de formación académica, actividades de divulgación científica en el ámbito nacional e internacional, 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, direcciones de becas de investigación (CyT-FCFMyN-UNSL y CONICET) y pasantías de investigación y docencia.

Las actividades se subvencionaron con fondos provenientes de Programa de Cooperación Interuniversitaria de la Agencia Española de Cooperación Iberoamericana (AECI); Fondo para Mejoramiento de la Calidad Universitaria (FOMEC); Proyecto AL2002-1010-2.43 / AL2003-1010-2.55 / AL2004-1010-2.53 / AL2005-PF-004 / AL2006-PF-013 / AL07-PAC-027 Geometría Computacional; Proyecto AL08-PAC-16 / AL09-PAC-12 Geometría Computacional, Algoritmos Aproximados y Bases de Datos subvencionado por la Universidad Politécnica de Madrid; Proyecto Fondo para Mejoramiento de la calidad Institucional (FOMEI) de la Universidad Nacional de San Luis; Proyecto Tecnologías avanzadas de Bases de Datos (22/F614) de la Universidad Nacional de San Luis; Aportes del Departamento de Informática Facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis; Subsidios de la Universidad Nacional de San Luis para traslados y pasantías de docentes, reglamentados por Ordenanzas N° 01/90-CS y N° 18/00-CS.

5. BIBLIOGRAFÍA

[BCHM08] Bajuelos A., Canales S., Hernández G., Martins A. *Estimating the Maximum Hidden Vertex Set in Polygons*. CG&A, Perugia. 2008.

[BFM97] Bäck T., Fogel D., Michalewicz Z. *Handbook of Evolutionary Computation*. IOP Publishing Ltd and Oxford University Press. 1997.

[CGG09] Carrasco F., Gagliardi E. O., García Sosa J. *I+3 R-Tree: un método de acceso espacio-temporal* (CACIC 2009), 2009

[CLGH10] Casanova C., Loyola R. Trabajo final de Licenciatura “Una Aplicación de Workforce Management con Herramientas de Geometría

Computacional y Base de Datos Espacio-Temporales”, bajo la dirección de MCs. E. O. Gagliardi y Dr. G. Hernández Peñalver.

[CH04] Canales S., Hernández Peñalver G. *Métodos Heurísticos en Problemas Geométricos: Visibilidad, Iluminación y Vigilancia*, Universidad Politécnica de Madrid. 2004.

[CNN93] Carlsson S., Nilsson B.J., Ntafos S. *Optimum guard covers and m-Watchmen Routes for restricted Polygons*, International Journal of computational Geometry and Applications, 3(1) 85-105, 1993.

[CS89] Cole R., Sharir M. *Visibility problems for polyhedral terrains*. Journal of Symbolic Computation, 7, pp. 11–30. 1989.

[DGLH09a] Dorzán, M., Gagliardi, E., Leguizamón, M., Hernández Peñalver, G. *Algoritmo ACO aplicado a la obtención aproximada de Triangulaciones de Peso Mínimo*, 32° CLEI, 2009.

[DGLH09b] Dorzán M. G., Gagliardi E. O., Leguizamón M. G., Hernández Peñalver G.. *Approximations on Minimum Weight Triangulations and Minimum Weight Pseudo-Triangulations using Ant Colony Optimization Metaheuristic*. I WEC. Jornadas Chilenas de Computación. 2009

[DS04] Dorigo M., Stützle T. *Ant Colony Optimization*. Massachusetts Institute of Technology. 2004.

[Fe00] Fekete, S. P. *On simple polygonizations with optimal area*. Discrete and Computational Geometry, 23, pp. 73-110. 2000.

[FLM08] Fekete S., Lübbecke M., Meijer H. *Minimizing the Stabbing Number of Matchings, Trees, and Triangulations*. Proceedings of the 15th ACM-SIAM Symposium on Discrete Algorithms. 2008.

[GL07] Gudmundsson J., Levcopoulos C.; *Minimum weight pseudo-triangulations*. Computational Geometry. Theory and applications. Elsevier Vol. 38-pages 139-153, 2007.

[GLHT07] Gagliardi E., Leguizamón M., Hernández Peñalver, G. Taranilla, M.T. *Algoritmo genético para la Descomposición de Minkowski de polígonos convexos*, XII Encuentro de Geometría Computacional, España. 2007.

[GJ79] Garey M. , Johnson, D. *Computers and Intractability: a guide of theory of NP-completeness*. Freeman, 1979.

[KE01] J. Kennedy, R. Eberhart. *Swarm Intelligence* (The Morgan Kaufmann Series in Artificial Intelligence). Morgan Kaufmann, 2001.

[M04] Mulzer V. *Minimum Dilation Triangulation for the Regular n-gon*. 2004

[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.

[PV96] Pocchiola M., Vegter G.. *Pseudo-triangulations: theory and applications*. Proceedings of the 12th Annual ACM Symposium on Computational Geometry: 1996.

[RSS06] Rote G., Santos F., Streinu I. *Pseudo-triangulations - a survey*. 2006.