

Geometría Computacional y Bases de Datos

Maria Gisela Dorzán, Susana Esquivel, Edilma Olinda Gagliardi,

Pablo Rafael Palmero, 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

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 nos dedicamos al diseño e implementación de estructuras y de índices adecuados para almacenar y resolver consultas de datos de tipo espacio-temporales, como asimismo al estudio 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: Geometría Computacional, Metaheurísticas, Bases de datos, Bases de datos espacio-temporales.

Contexto

El proyecto *Tecnologías Avanzadas de Bases de Datos* de la Universidad Nacional de San Luis, tiene como objetivo el estudio de bases de datos no convencionales donde se involucra 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.

Respecto de las temáticas de investigación, en la línea vinculamos el estudio de las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas, debido a que en diversos campos de aplicación es necesaria la construcción y utilización de objetos y estructuras geométricas que tengan propiedades deseables. En particular, algunos de los problemas relacionados con la optimización de estructuras geométricas son problemas NP-duros, por tanto utilizamos técnicas metaheurísticas para hallar soluciones aproximadas para ellos.

Por otra parte, existe la necesidad 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.

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 tópicos más relevantes, actualmente en estudio, junto con las propuestas más recientes de interés.

Introducción

La velocidad de crecimiento de los datos disponibles en forma digital, la variedad de aplicaciones existentes y la evolución de las tecnologías de información y comunicación han impulsado el surgimiento de almacenamientos no estructurados de información. Surge así, la necesidad consultar nuevos tipos de datos (texto libre, datos geométricos, imágenes, audio y video, entre otros) donde la información no se puede estructurar en forma tradicional. Incluso, cuando se puede utilizar una estructuración clásica, existen nuevas aplicaciones que requieren acceder a la base de datos por cualquier campo y es necesario utilizar de herramientas no tradicionales, como las provistas por la Inteligencia Computacional.

Asimismo, existen aplicaciones que necesitan 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. Para tales requerimientos son necesarios los modelos de bases de datos espacio-temporales.

Por tanto, es necesario contar con herramientas teóricas, que permitan modelar estos tipos de datos, realizar operaciones sobre ellos, definir lenguajes de consulta, analizar su expresividad, entre otras propiedades.

La Geometría Computacional es una disciplina dedicada al diseño de algoritmos y estructuras de datos para la solución de problemas que utilizan objetos geométricos. Una tarea fundamental en la disciplina es identificar conceptos, propiedades y técnicas para el desarrollo de algoritmos eficientes. Esto conduce al análisis y estudio de la complejidad combinatoria de estructuras geométricas, de estructuras de datos geométricas, de la complejidad de algoritmos, entre otras. Los métodos estudiados en esta área permiten el diseño y el análisis de algoritmos para la solución eficiente de numerosos problemas

geométricos que surgen en otras áreas de aplicación como astronomía, computación gráfica, sistemas de información geográfica, CAD/CAM, minería de datos, dibujo gráfico, imágenes médicas, metrología, modelismo molecular, robótica, tratamiento de señal, textil, tipografía, videojuegos, entre otras.

En el ámbito de esta disciplina, existen problemas NP-duros, o bien, son problemas para los cuales no se conoce 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 corresponden a configuraciones geométricas obtenidas a partir de un conjunto de puntos u objetos en el plano, y se busca optimizar ciertas propiedades que permiten medir la calidad de dichas configuraciones. Debido a esto, proponemos la aplicación de técnicas metaheurísticas, ya que las mismas nos permiten encontrar soluciones cercanas a las óptimas. Una metaheurística es un algoritmo de propósito general que generalmente puede ser aplicado en la resolución de cualquier problema de optimización. 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 y combina 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].

La optimización de problemas relacionados a configuraciones geométricas tales como las 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.

Los criterios de optimización consisten en minimizar o maximizar alguna propiedad de las estructuras geométricas. Con respecto a triangulaciones o pseudo-triangulaciones, 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 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 pseudo-triangulación de peso mínimo aún no está resuelta. Levcopoulos y Gudmundsson [GL07] muestran una 12-aproximación de una pseudo-triangulación que puede ser calculada con complejidad de $O(n^3)$. 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, se busca obtener aquellas que son óptimas respecto de criterios tales como el área o el perímetro del polígono obtenido. Fekete [Fe00] demostró que minimizar o maximizar el área es un problema NP-duro. La minimización del perímetro de una poligonización está estrechamente relacionada con la versión geométrica del Problema del Viajante y es también un problema NP-duro [GJ79].

Debido a la complejidad de los problemas antes mencionados, surgen los algoritmos aproximados como candidatos 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. En este contexto, nos planteamos el estudio de optimizaciones de estructuras tales como

triangulaciones, pseudo-triangulaciones y poligonizaciones. Entre las principales medidas de calidad consideramos el peso, la dilación, área, perímetro, entre otras.

También, proponemos el diseño y desarrollo de herramientas para la generación y visualización de triangulaciones y pseudo-triangulaciones de conjuntos de puntos en el plano, las cuales permiten generar diferentes tipos de dichas estructuras geométricas, realizar operaciones sobre ellas y describir sus propiedades más importantes. Dichas herramientas permiten observar la morfología de cada instancia generada y el comportamiento de los algoritmos generadores cada una de ellas. Además, es posible recuperar instancias de tales estructuras geométricas a partir de resultados obtenidos de las generaciones experimentales alcanzados mediante técnicas metaheurísticas [DGPH10].

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, por ejemplo en robótica (planificación de movimientos, problemas de visualización, y detección de colisiones), en los sistemas de información geográfica (localización y búsqueda geométrica y planificación de rutas), entre otros.

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 generales de estudio de la línea, destacamos los siguientes: i) Estudio de triangulaciones de puntos que cumplan ciertas medidas de calidad como peso y dilación.

Optimización de triangulaciones de peso y dilación mínimos utilizando técnicas metaheurísticas [MR06] [M04].

ii) Optimización de poligonizaciones de puntos en el plano. Resolución aproximada de diferentes variantes de problemas de vigilancia en polígonos. Optimización de rutas de vigilancia entre obstáculos poligonales mediante la aplicación de técnicas metaheurísticas [Fe00] [CNN93].

iii) Optimización de pseudo-triangulaciones, que satisfagan propiedades como bajo peso, dilación pequeña, entre otras, mediante la aplicación de técnicas metaheurísticas [PV96] [RSS06] [GL07].

iv) Diseño e implementación de herramientas para la generación y visualización de triangulaciones y pseudo-triangulaciones de conjuntos de puntos en el plano [DGPH10].

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 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 [CGG09] [CL10].

Resultados Obtenidos

Para la resolución de problemas de optimización de Triangulaciones y Pseudo-triangulaciones de Peso Mínimo se utilizaron las técnicas metaheurísticas: Optimización basada en Colonia de Hormigas (Ant Colony Optimization, ACO) y Recocido Simulado (Simulated Annealing, SA), y técnicas determinísticas Voraces (Greedy) y Triangulación de Delaunay. Se llevo 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. En [DGLH09a] [DGLH09b] [DGLH10a] y [DGLH10b], se muestran los resultados obtenidos para los problemas MWT y MWPT

utilizando la técnica metaheurística ACO. Actualmente se está realizando la evaluación experimental de los resultados obtenidos con la técnica Recocido Simulado [DGLH10c]. Además se está realizando una evaluación estadística para poder comparar ambas técnicas y poder determinar cuál es mejor.

Se está diseñando y desarrollando una herramienta de generación y visualización de triangulaciones y pseudo-triangulaciones de conjuntos de puntos en el plano. Además esta herramienta permite recuperar instancias de tales estructuras geométricas a partir de resultados obtenidos de las generaciones experimentales alcanzados mediante técnicas metaheurísticas de modo de poder observar su comportamiento [DGPH10].

Respecto al problema de optimización de poligonizaciones de puntos en el plano, estamos utilizando técnicas metaheurísticas basadas en Colonia de Hormigas (Ant Colony Optimization, ACO) y técnicas Voraces (Greedy) para obtener poligonizaciones de área mínima. Se está realizando la evaluación experimental con las técnicas mencionadas.

Actualmente, se continúa trabajando en los problemas de optimización mencionados considerando otras técnicas metaheurísticas, que puedan adecuarse para su resolución. Se está realizando en un estudio experimental para determinar la técnica que brinda 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

Se ha consolidado la formación del grupo de trabajo en la Universidad Nacional de San Luis, realizando 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 (CyT-FCFMyN-UNSL y CONICET), iii) pasantías de investigación y actividades de formación académica, iv) actividades de divulgación científica, tales como 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, investigación, desarrollo y experimentación, 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.
- [CGG09] Carrasco F., Gagliardi E. O., García Sosa J. *I+3 R-Tree: un método de acceso espacio-temporal* (CACIC 2009), 2009
- [CGG09] Carrasco F., Gagliardi E. O., García Sosa J. *I+3 R-Tree: un método de acceso espacio-temporal* (CACIC 2009), 2009
- [CL10] Casanova C., Loyola R., Dorzán, M.G., Gagliardi, E.O., Hernández Peñalver, G. Taranilla, M.T. “Una Aplicación de Workforce Managment con Herramientas de Geometría Computacional y Base de Datos Espacio-Temporales”, (CACIC 2010), 2010
- [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.
- [DGPH10] Dorzán M.G., Gagliardi, E. O., Palmero P.R., Hernández Peñalver, G. *Una herramienta para la generación y visualización de Triangulaciones y Pseudotriangulaciones*. (CACIC 2010). 2010
- [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
- [DGLH10a] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Globally optimal triangulations of minimum weight using Ant Colony Optimization metaheuristic*. Journal of Computer Science & Technology ISSN 1666-6038. Vol. 10 – N° 2. 2010
- [DGLH10b] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Triangulaciones y Pseudo-triangulaciones de Peso Mínimo: resolución aproximada con Simulated Annealing*. VII Jornadas de Matemática Discreta y Algorítmica. ISBN: 978-84-693-3063-0. 2010
- [DGLH10c] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Simulated Annealing aplicado a Triangulaciones y Pseudotriangulaciones de Peso Mínimo*. XVI Congreso Argentino de Ciencias de la Computación 2010 (CACIC 2010). ISBN 978-950-9474-49-9. pp: 62-71. 2010
- [DS04] Dorigo M., Stützle T. *Ant Colony Optimization*. Massachusetts Institute of Technology. 2004.
- [Fe00] Fekete, S. P. *On simple polygonalizations 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., Levkopoulos C.; *Minimum weight pseudo-triangulations*. Computational Geometry. Theory and applications. Elsevier Vol. 38-pages 139-153, 2007.
- [GJ79] Garey M., Johnson, D. *Computers and Intractability: a guide of theory of NP-completeness*. Freeman, 1979.
- [KE01] Kennedy, J. Eberhart R.. *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.