



Universidad Autónoma Metropolitana
Unidad Cuajimalpa

División de Ciencias Naturales e Ingeniería

Departamento de Matemáticas Aplicadas y Sistemas

Diseño automatizado de zonas geográficas mediante algoritmos mono y multiobjetivo

Proceso de selección para plaza de Profesor Visitante
Propuesta de investigación
Elaborada por: Dr. Alejandro Lara Caballero

A. ANTECEDENTES

El diseño de zonas es un problema de particionamiento que consiste en agrupar pequeñas áreas o unidades geográficas (UG) en regiones que satisfagan los requerimientos impuestos por un contexto específico. Existen diversas aplicaciones, en áreas de la ingeniería y ciencias sociales, como el diseño de zonas de ventas, distritos de mandos policial, servicios públicos, de emergencia, entre otros. Entre las características deseadas se puede incluir la formación de zonas conectadas, balanceadas con respecto a un criterio, con cierto tamaño y forma, entre otras.

El diseño de zonas se puede modelar con un problema de optimización combinatoria, con múltiples variables de decisión a optimizar de manera simultánea. En general, para plantear el problema se define $X = \{x_1, x_2, \dots, x_n\}$ como un conjunto de UG que se deben agrupar en k distritos o zonas. De igual manera se define Z_i como el conjunto de todas las áreas que corresponden a la i -ésima zona. Las restricciones del diseño varían según el problema considerado, pero se contemplan las siguientes restricciones comunes:

$$Z_i \neq \emptyset, \quad \forall i \in \{1, 2, \dots, k\} \quad (1)$$

$$Z_i \cap Z_j = \emptyset, \quad \forall i \neq j \quad (2)$$

$$\bigcup_{i=1}^k Z_i = X \quad (3)$$

La variedad de aplicaciones, sus características específicas, así como las limitaciones de tiempo inherentes a toda investigación, obliga a restringir esta propuesta a uno de los casos más populares: el diseño de zonas electorales. En este entorno la creación de zonas busca una adecuada representación democrática, a través del principio de mayoría relativa, y dificultar la manipulación de las demarcaciones con fines políticos. Entre los criterios empleados con frecuencia para el diseño de distritos electorales, se encuentran la creación de zonas conexas con equilibrio poblacional (aproximadamente el mismo número de habitantes), evitando formas retorcidas y alargadas (compacidad geométrica) [1]. De manera complementaria, también se pueden considerar criterios como la competitividad electoral, conformidad con límites políticos, integración de comunidades de interés, tiempo de traslado, entre otros. En la Tabla 1, se muestran algunas funciones comúnmente utilizadas en México para los criterios de compacidad y equilibrio poblacional. Es importante mencionar que, en la literatura revisada, existen más funciones que aquellas mostradas en la Tabla 1. Asimismo, su elección depende del contexto legal de cada país. Para estudios sobre distintas medidas de compacidad, se pueden consultar los trabajos de Saxon [2] y Niemi et al. [3].

La distritación asistida por computadora tiene sus orígenes en 1972 con el trabajo pionero de Nagel, quien explora técnicas exhaustivas para enumerar y tabular todas las posibles soluciones. Posteriormente, se demostró que el diseño de zonas es un problema en términos de complejidad computacional NP-Duro [4], y por lo tanto, normalmente se recomienda el uso de técnicas heurísticas para poder encontrar soluciones de buena calidad con un costo computacional razonable.

Criterio	Ecuación	Descripción
Equilibrio poblacional	$\sum_{i=1}^k \left(\frac{1 - \frac{P_i}{P_{med}}}{d_a} \right)$	P_{med} se refiere a la población media estatal, P_i a la población de la zona i , d_a es el porcentaje de desviación poblacional máximo aceptable y k el número de zonas.
Compacidad	$\frac{PC_{Z_i}}{4\sqrt{AC_{Z_i}}} - 1$	PC_{Z_i} es el perímetro de la zona Z_i y AC_{Z_i} su área.

Tabla 1 Criterios comunes del diseño de zonas

En los últimos 50 años, el diseño de zonas ha atraído la atención de gran número de investigadores que han propuesto una variedad de modelos y técnicas de optimización como búsqueda tabú [5], recocido simulado [6], algoritmos genéticos [7], optimización por enjambre de partículas [8], algoritmos evolutivos [9]. La mayoría de las estrategias ha abordado el problema mediante estrategias clásicas mono-objetivo, específicamente mediante una suma ponderada. También existe un creciente número de trabajos que aplican técnicas de optimización multi-objetivo como son derivaciones de algoritmo genético NSGA-II, SPEA-II y de recocido simulado [10] [11] [12] [13].

Tanto por su complejidad computacional, como su área de aplicabilidad, la creación de algoritmos y herramientas que permitan la creación eficiente de distritos electorales es un tema de particular interés. El uso de computadoras para asistir la creación de zonas promueve una mayor participación de los tomadores de decisiones como especialistas, académicos, autoridades legales de distintos países, educadores y también fomenta la innovación [14]. Actualmente las herramientas disponibles para llevar a cabo distritación mediante el uso de computadoras y técnicas heurísticas son escasas. Las aplicaciones comerciales, usualmente son integradas como módulos de Sistemas de Información Geográficas (SIG). En esta categoría se encuentran herramientas producidas por empresas de corte cartográfico como Caliper Corporation, Esri, Digital Engineering Corporation, Corona Solutions y Manifold Systems, entre otras. Algunos de los principales inconvenientes encontrados por un estudio independiente realizado por Altman et al. [15] en estos proyectos, consisten en considerar como único criterio a optimizar el equilibrio poblacional, no hay posibilidad de incluir criterios o fórmulas adicionales, los mecanismos heurísticos no se encuentran bien documentados y la calidad de las soluciones obtenidas es discutible.

En la comunidad académica, por su parte, existen algunos desarrollos de software que permiten la creación de planes de distritación automatizados mediante heurísticas como son BARD [14], Autoreldistrict [16] y Antimander [17]. En general, estas propuestas ofrecen una mayor flexibilidad en la creación de planes de distritación mediante heurísticas, la modificación de sus parámetros y las técnicas empleadas están bien documentadas. Sin embargo, también presentan algunas áreas de oportunidad como ofrecer un conjunto más diverso de técnicas heurísticas que resuelvan el problema desde un enfoque multiobjetivo, más allá de las técnicas tradicionales de suma ponderada y reduzcan la dependencia de herramientas externas como SIG y paquetes estadísticos para el preprocesamiento, procesamiento, visualización y análisis posterior de las soluciones encontradas.

La motivación del presente proyecto está en la contribución que podría representar la creación de una herramienta interactiva que permita generar escenarios robustos para el problema de diseño de zonas utilizando técnicas heurísticas mono y multi-objetivo, así como preprocesar, procesar, visualizar y analizar la calidad de las soluciones obtenidas.

B. HIPÓTESIS

Mediante el desarrollo de un software especializado que utilice técnicas heurísticas mono y multi-objetivo se pueden generar soluciones de buena calidad, contrastar estrategias y mejorar la toma de decisiones en el diseño de zonas electorales.

C. OBJETIVOS

Objetivo general

Desarrollar un software basado en técnicas metaheurísticas mono y multiobjetivo que asista el diseño automatizado de zonas electorales.

Objetivos específicos

1. Definir los criterios que se considerarán en este proyecto para el problema de diseño de zonas electorales
2. Adaptar e implementar enfoques de resolución mono y multi-objetivo representativos del estado del arte. Asimismo, probar su desempeño sobre instancias relevantes.
3. Diseñar e implementar un flujo de preprocesamiento y procesamiento eficiente y transparente para la manipulación de datos geográficos.
4. Crear una interfaz interactiva que permita visualizar, comparar y analizar la calidad de las soluciones obtenidas, empleando gráficas y métricas propias del campo de optimización.
5. Validar el software mediante un caso de estudio.

D. PRODUCTOS ESPERADOS

- 1.1. Creación de un software especializado que aborde áreas de oportunidad de los sistemas actuales para la creación de zonas electorales y sea accesible de usar.
- 1.2. Participación en la difusión de la herramienta en un evento relacionado con la aplicación de estrategias heurísticas para resolver problemas combinatorios o áreas afines. Es importante mencionar que debido a la contingencia de COVID 19, se buscarán eventos virtuales. Asimismo, se realizará según la disponibilidad de asignación de recursos.
- 1.3. Propuesta de un artículo en revista indizada, relacionada con la aplicación de técnicas de optimización a problemas de ingeniería o relacionadas al caso de estudio.

E. METODOLOGÍA

La metodología para el desarrollo del proyecto se divide en las siguientes etapas.

- a) Generar un conjunto de instancias que incluyan casos con un número pequeño, mediano y grande de unidades geográficas.
- b) Analizar y definir los criterios, su representación y elementos a considerar en el modelo. Asimismo, definir las técnicas heurísticas a utilizar.
- c) Implementar algoritmos de búsqueda local y poblacional, que emplearán una suma ponderada de objetivos para guiar el proceso de optimización. Los algoritmos obtenidos serán probados sobre las instancias generadas de acuerdo con lo planteado en el inciso a).

Los ponderadores utilizados incluirán diferentes combinaciones de valores normalizados, de tal forma que se pueda obtener una aproximación al frente de Pareto para cada instancia.

- d) Implementar algoritmos multi-objetivo y probarlos sobre el conjunto de instancias generado previamente.
- e) Proponer, diseñar e implementar diferentes módulos que permitan coadyuvar en las diferentes etapas del proceso de distribución y sean accesibles al usuario.
- f) Probar y evaluar mediante un caso de estudio la usabilidad de la herramienta, su desempeño y determinar las fortalezas y retos de la propuesta.

F. PROGRAMA DE ACTIVIDADES

1. Actualización del estado del arte.
 - 1.1.Revisión de las publicaciones más recientes relacionadas con el problema de interés.
 - 1.2.Revisión de las técnicas heurísticas mono y multiobjetivo empleadas.
2. Revisión de software relacionado.
 - 2.1.Revisión de las herramientas de software actuales.
 - 2.2.Proponer modificaciones que permitan mejoras y aborden áreas de oportunidad detectadas.
3. Selección de criterios y métricas relevantes
 - 3.1. Definir los criterios relevantes a incluir, así como su formulación.
4. Implementación de algoritmos basados en un enfoque mono-objetivo.
 - 4.1.Implementación, depuración y calibración de algoritmos basados en un enfoque mono-objetivo, empleando estrategias heurísticas significativas detectadas durante la actualización del estado del arte.
5. Implementación de algoritmos basados en un enfoque multi-objetivo
 - 5.1.Implementación, depuración y calibración de algoritmos basados en un enfoque multi-objetivo, empleando las estrategias heurísticas significativas encontradas durante la actualización del estado del arte.
6. Automatización del preprocesamiento de la información geográfica
 - 6.1.Generación de un conjunto de instancias relevantes de información disponible sobre entidades federativas de la República Mexicana y algunos estados extranjeros.
 - 6.2.Análisis, depuración y adecuación.
 - 6.3.Creación de un módulo para la generación de los datos geográficos necesarios para la parte de procesamiento.
7. Procesamiento de la información
 - 7.1.Creación de un módulo para administrar y gestionar los algoritmos implementados en el punto 4 y 5.
8. Visualización y análisis de la información
 - 8.1.Creación de un módulo para visualizar los escenarios y sus costos.

8.2. Creación de módulo para gráficas, estadísticas y métricas relevantes de optimización

9. Integración y validación

9.1. Integrar los módulos planteados en los incisos anteriores y realizar pruebas.

9.2. Validar la herramienta mediante un caso de estudio.

10. Difusión de resultados.

10.1. Presentación de los resultados en eventos especializados nacionales.

10.2. Propuesta de un artículo para someterlo en revista indizada.

El cronograma de actividades se presenta en la Tabla 2.

Actividad	Trimestre 20-O	Trimestre 21-I	Trimestre 21-P
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Tabla 2. -Cronograma de actividades

G. BIBLIOGRAFÍA

- [1] E. A. Rincón García y M. Á. Gutiérrez Andrade, «Redistricting by square cells,» *Lecture notes on artificial intelligence*, vol. 5845, pp. 669-679, 2009.
- [2] J. Saxon, «Reviving Legislative Avenues for Gerrymandering Reform with a flexible automated tool,» *Political Analysis*, pp. 1-58, 2020.
- [3] R. Nlemi, B. Grofman, C. Carlucci y T. Hofeller, «Measuring Compactness and the Role of a Compactness Standard in a Test for Partisan and Racial Gerrymandering,» *The Journal of Politics*, vol. 52, nº 4, pp. 1-14, 1990.
- [4] M. Altman, «Is Automation the Answer: The Computational Complexity of Automated Redistricting,» *Rutgers Computer Law Technology Journal*, vol. 23, pp. 81-141, 1997.
- [5] B. Bozkaya, E. Erkut y G. Laporte, «A tabu search heuristic and adaptive memory procedure for political districting,» *European Journal of Operational Research*, vol. 144, pp. 12-26, 2003.
- [6] E. A. Rincón García, M. Á. Gutiérrez Andrade, J. Ramírez Rodríguez, P. Lara Velázquez y S. G. de los Cobos Silva, «Applying simulated annealing to design compact zones,» *Fuzzy Economic Review*, vol. 15, nº 2, pp. 11-24, 2010.
- [7] F. Bacao, V. Lobo y M. Painho, «Applying genetic algorithms to zone design,» *Soft Computing*,

vol. 9, pp. 341-348, 2005.

- [8] E. A. Rincón García, M. Á. Gutiérrez Andrade, S. G. de los Cobos Silva, P. Lara Velázquez, R. A. Mora Gutiérrez y A. Ponsich, «A discrete particle swarm optimization algorithm for designing electoral zones,» de *Methods for Decision Making in an Uncertain Environment*, World scientific, 2012, pp. 174-187.
- [9] C. Chung-I, «A Knowledge-based Evolution Algorithm approach to political districting problem,» *Computer Physics Communications*, vol. 182, nº 1, pp. 209-212, 2011.
- [10] K. Deb, A. Pratap, S. Agrawal y T. Meyarivan, «A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II,» *IEEE Transactions on Evolutionary Computation*, vol. 6, nº 2, pp. 182-197, 2002.
- [11] A. Lara, M. Á. Gutiérrez y E. A. Rincón, «A Simulated Annealing-Based Multiobjective Optimization Algorithm for Political Districting,» *IEEE Latin America Transactions*, vol. 16, nº 6, pp. 1723-1731, 2018.
- [12] A. Lara, S. de los Cobos Silva, R. Mora-Gutiérrez, E. Rincón-García, M. Á. Gutiérrez-Andrade y P. Lara-Velázquez, «Multiobjective Genetic Algorithms for Reinforcing Equal Population in Congressional Districts,» *Mathematical Problems in Engineering*, 2019.
- [13] L. Vanneschi, R. Henriques y M. Castelli, «Multi-Objective Genetic Algorithm with Variable Neighbourhood Search for the Electoral Redistricting Problem,» *Swarm and Evolutionary Computation*, nº 36, 2017.
- [14] M. Altman y M. McDonald, «BARD: Better Automated Redistricting,» *Journal of Statistical Software*, vol. 42, nº 4, pp. 1-22, 2011.
- [15] M. Altman, M. McDonald y K. McDonald, «From Crayons to Computers: The Evolution of Computer Use in Redistricting,» *Social Science Computer Review*, vol. 3, nº 23, p. 334-346, 2005.
- [16] H. Levin y S. Friedler, «Automated Congressional Redistricting,» *ACM Journal of Experimental Algorithmics*, vol. 1, nº 1, pp. 1-24, 2019.
- [17] J. Simon y J. Lehman, «Antimander: open source detection of gerrymandering though multi-objective evolutionary algorithms,» de *GECCO '20: Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*, Cancún, México, 2020.