“Uso de computación evolutiva para hallar ubicaciones óptimas de una cadena de supermercados”

Autor/Autores: Añanca Arango, Pedro Christian. De Lama Ramírez, Diego Gonzalo. Díaz Villanueva, Julio Leonardo. Tello Rivera, Erika Elizabeth.

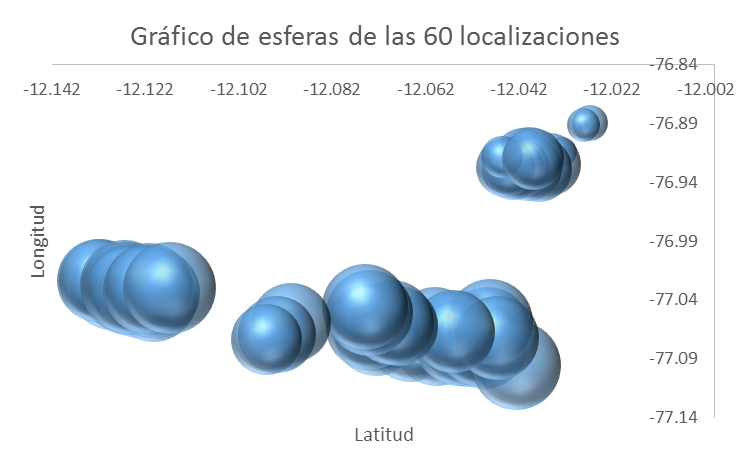
**Resumen-** Este proyecto busca las 10 mejores ubicaciones para nuevos supermercados basado en coordenadas geográficas y población de impacto. Para ello, se creó una versión de fitness que usa un dataframe con información geoespacial, calcula distancias entre ubicaciones, y acumula estos datos en un cromosoma. Para la población, se utilizó un porcentaje de la suma total de las localidades, dividido entre la población total de 60 localidades. Se ejecutó el algoritmo ajustando hiper parámetros como probabilidad de mutación, crossover y tasa de mutación, y se usaron dos tipos de algoritmos: mono objetivo y multi objetivo. Los resultados obtenidos fueron los siguientes ……

——————————◆——————————

1. **Introducción**

El problema plantea la búsqueda de las soluciones óptimas para optimizar 10 ubicaciones para supermercados. Para lo cual, se plantean 60 posibles ubicaciones distribuidas a lo largo de la provincia de Lima. Cada ubicación o punto, se cuenta con su latitud, longitud y población a 500 metros a la redonda.

Se ha elaborado un plot con la distribución de las localizaciones en función de su longitud y latitud.



Se tienen dos objetivos: maximizar la suma de la población que vive a 500 metros alrededor de los supermercados y maximizar la suma de distancias entre los supermercados escogidos.

Se propone emplear algoritmos genéticos mono-objetivo y multiobjetivo.

1. **Metodología**

* Describir la representación de estados/individuos elegida,
* Describir el comportamiento entrada/salida del (los) enfoques.
* Describir los operadores/funciones desarrolladas para solucionar el problema planteado. No describir métodos genéricos, sino detalles de cómo se aplicó/adaptó al problema
* Utilizar figuras para auxiliar la explicación
  1. **Primer enfoque (Algoritmo mono-objetivo)**

Tanto en el primer como en el segundo enfoque los individuos serán descritos como un cromosoma de longitud 10 cuyos alelos irán en un rango de 0 a 59. Este número indica la localización (total 60 candidatos). Por ello es necesario que no se repita ningún alelo en el cromosoma,

La obtención del fitness básicamente consistirá en el producto entre las sumas de distancias de las locaciones escogidas y la suma normalizada de las poblaciones.

* 1. **Segundo enfoque (Algoritmo multi-objetivo):**

La obtención del fitness básicamente consistirá en un array de dos valores. Como primer valor se obtendrá la suma de distancias de las locaciones escogidas y como segundo valor la suma normalizada de las poblaciones.

* 1. **Operadores para solucionar el problema planteado.**

La clase Individuo tiene operadores asociados al cruzamiento de los padres y la mutación de los individuos hijos:

* **crossover\_one\_point**: Se elige un punto aleatoriamente a lo largo del cromosoma (lista de localizaciones) y se intercambia los elementos de la lista del primer padre con indice menor igual al punto al segundo padre y los elementos de la lista del segundo padre mayores al punto elegido se intercambian al primer padre. Para adaptar a nuestro problema se parametrizar para no intercambiar los elementos de cada lista que se tuvieran en común y asi evitar localizaciones repetidas en un mismo cromosoma.
* **crossover\_uniform**: Cada alelo (elemento de la lista de localizaciones) se elige aleatoriamente de los dos progenitores con una probabilidad del 50%. Si es un alelo con valores en común en ambos padres no se hace el intercambio.
* **mutation\_flip:** Se elige un alelo del cromosoma y cambia su valor aleatoriamente en un rango específico del 0 al 59 excluyendo los valores en común de ambos padres.
* **mutation\_multiflip:** Se elige hasta un 50% de alelos como máximo y cambia sus valores aleatoriamente en un rango específico.

1. **Experimentación y Resultados**
   * **Setup experimental:**

* **Descripción de datos usados**

Para resolver el problema se nos proporcionó un mapa (1) de la ciudad en Google Earth con 60 locales candidatos donde la cadena de supermercados podría localizar sus supermercados. De ese mapa se podría extraer las coordenadas, pero optamos por usar un excel llamado ‘Candidatos\_supermercados’, el cual contiene una tabla con las localizaciones numeradas. Por cada localización candidata hay una coordenada asignada en formato longitud y latitud, la descripción de la dirección de dicha localización y además la población estimada que hay en 500 metros a la redonda de esta. El campo llave de esta tabla es el número de localización y es este dato que se ingresará en una lista de 10 localizaciones elegidas aleatoriamente, cada lista será un individuo de la población generada. Esta población es la que será el parámetro principal de los algoritmos genéticos Monoobjetivo y Multiobjetivo que optimizará la solución del problema.

Tabla de datos proporcionados con 10 primeras localizaciones

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

* **Descripción de métrica de evaluación**

Cómo métrica de evaluación en los algoritmos genéticos se debe plantear un fitness. Para este problema los componentes del fitness son:

Suma de distancias de las localizaciones

Para esto se generó previamente una matriz de 60 x 60, las cuales representan el número de localizaciones, que calcula las distancias cada dos localizaciones a partir de su latitud y longitud, obtenida del excel anteriormente descrito.

Para sumarizar las distancias de las 10 localizaciones candidatas a evaluar se sumariza todas las distancias entre las 10 localizaciones que podemos identificar en en el triángulo superior de la matriz de distancias. Este resultado luego se convierte en kilómetros.

* **Matriz de distancias**

Imagen de la pantalla de una computadora

Descripción generada automáticamente con confianza baja

* **Suma normalizada del tamaño de la población de las localizaciones**

Para el tamaño de población se asocia el número de localización que se encuentra en la lista candidata a evaluar con el número de fila de la tabla de datos proporcionada y se suma el valor del campo ‘Población 500m’ de las 10 localizaciones del individuo evaluado. Luego para normalizar esta variable, el resultado se divide entre la suma de las 60 localizaciones candidatas.

* **Métrica de evaluación del algoritmo genético Monoobjetivo**

Para determinar el mejor candidato el algoritmo Monoobjetivo primero se calcula el fitness de la mejor generación de la población como el producto de la suma de distancias y la suma normalizada del tamaño de la población. Luego se identifica el mayor valor de fitness del set de experimentos. Si hay 2 o más generaciones que tienen el mismo fitness entonces se identifica quien tiene mayor tamaño normalizado de la población y luego qué generación tiene mayor suma de distancias.

Métrica de evaluación del algoritmo genético Multiobjetivo

A diferencia del algoritmo genético Monoobjetivo, este algoritmo determina un conjunto de soluciones llamada Frontera de Pareto. Para determinar el mejor candidato se identifica qué experimento tiene mejor fitness del algoritmo Monoobjetivo y luego se plotea el valor de su suma de distancias y el valor de su tamaño normalizado de población junto con las soluciones de la Frontera de Pareto si revisa si la solución está dominada o no.

* **Descripción de experimentos hechos**

Para cada algoritmo se realizaron 20 experimentos con las siguientes características

Número de elementos por cada individuo: 10 (números del 0 al 59 que no se repitan)

Número de individuos de la población: 300

Número de generaciones: 50

Número de corridas: 3

|  |  |  |  |
| --- | --- | --- | --- |
| **Nº de experimento** | **Función de cruzamiento** | **Función de mutación** | **Tasa de mutación** |
| 0 | onepoint | flip | 0.15 |
| 1 | onepoint | multiflip | 0.15 |
| 2 | uniform | flip | 0.15 |
| 3 | uniform | multiflip | 0.15 |
| 4 | onepoint | flip | 0.25 |
| 5 | onepoint | multiflip | 0.25 |
| 6 | uniform | flip | 0.25 |
| 7 | uniform | multiflip | 0.25 |
| 8 | onepoint | flip | 0.5 |
| 9 | onepoint | multiflip | 0.5 |
| 10 | uniform | flip | 0.5 |
| 11 | uniform | multiflip | 0.5 |
| 12 | onepoint | flip | 0.75 |
| 13 | onepoint | multiflip | 0.75 |
| 14 | uniform | flip | 0.75 |
| 15 | uniform | multiflip | 0.75 |
| 16 | onepoint | flip | 1 |
| 17 | onepoint | multiflip | 1 |
| 18 | uniform | flip | 1 |
| 19 | uniform | multiflip | 1 |

* **Resultados**

**Resultados algoritmo mono-objetivo:**

Se presentan 20 resultados, en los que se presentan iteraciones para Crossover: uniform y onepoint, Mutation: flip y multiflip, Tasa de mutación: 0.15, 0.25, 0.50, 0.75 y 1.

Obteniendo que la combinación de parámetros que alcanza un mayor fitness es la de Onepoint, flip y un Pmut de 0.15.

Las 10 ubicaciones óptimas son los puntos 2, 11, 14, 15, 28, 29, 39, 50, 57, 60. Con un fitness de 92.58.

**Resultados algoritmo multiobjetivo:**

Se presentan 20 resultados, en los que se presentan iteraciones para Crossover: uniform y onepoint, Mutation: flip y multiflip, Tasa de mutación: 0.15, 0.25, 0.50, 0.75 y 1.

Obteniendo que la combinación de parámetros que alcanza un mayor fitness es la de Onepoint, multiflip y un Pmut de 0.75.

Para este algoritmo, se hallaron los valores máximos de suma de distancias y suma de población. Obteniendo que los valores máximos son: Suma de distancias = 0.2225 y Suma de población = 575.5784. Las 10 ubicaciones óptimas son los puntos: Los puntos óptimos son: 11, 14, 15, 28, 29, 38, 45, 50, 56 y 60.

* **Discusión**

Los experimentos deben ser planificados para poder caracterizar/comparar los enfoques desarrollados. Aquí algunas preguntas que deben responder los experimentos

* ¿El enfoque desarrollado resuelve siempre el problema?
* ¿Qué tan eficientemente lo resuelven?
* ¿Cuál es el desempeño comparado con otros modelos?
* ¿Cómo influyen los parámetros del enfoque en su desempeño?

1. **Conclusión**

* Dar las conclusiones principales con base a los resultados obtenidos y a lo que fue planteado en su hipótesis, ¿qué se puede decir del o los enfoques desarrollados y/o del problema abordado?

1. **Sugerencias de trabajos futuros**

* Para trabajos futuros, para el algoritmo multiobjetivo, se plantea hallar métodos para obtener las soluciones con resultados máximos (para cada objetivo).

1. **Link del repositorio del trabajo**

Github: <https://github.com/OSOMASTER123/Uso-de-computaci-n-evolutiva-para-hallar-ubicaciones-ptimas-de-una-cadena-de-supermercados-PUCP>

1. **Declaración de contribución de cada integrante**

Describir los aportes de cada integrante al proyecto

1. **Referencias**

[1] Mapa con las 60 loclizaciones https://drive.google.com/open?id=1w7n77ByWK6TrX74lOBTwID719LRj\_UQz&usp=sharing