

# Agrupamiento jerárquico balanceado para el problema de enrutamiento de flotas

Juan José Sapuyes Pino,, Juan Manuel Pajoy López, Juan Pablo Ortega Medina, and Sebastián Rendón Giraldo.

Universidad Nacional de Colombia, Medellín, Colombia.  
{jsapuyesp,jpajoyl,jportegame,serendongi}@unal.edu.co

**Abstract.** Este artículo propone un método de agrupamiento jerárquico geográfico para tratar de solucionar una parte del problema del enrutamiento de flotas, en este caso para laboratorios con tomas de muestras a domicilio. El algoritmo es un algoritmo jerárquico aglomerativo, que utiliza una filosofía bottom-up para el agrupamiento y de divide y vencerás para hallar los pares de grupos más cercanos.

**Keywords:** Clústering · Agrupamiento · Jerárquico.

## 1 Introducción

Existen laboratorios que prestan el servicio de toma de muestras en casa, un servicio bastante útil para personas que por alguna dificultad no pueden desplazarse hasta el lugar donde se toman las muestras, estas muestras se toman y son transportadas al laboratorio a bajas temperaturas, para poder preservarlas. Sin embargo, los laboratorios que prestan estos servicios se encuentran con un gran problema, ¿cuál es la ruta óptima para la persona que toma la muestra? Esto es un problema bastante recurrente para los laboratorios y es de vital importancia resolverlo, pues las muestras se encuentran en neveras portátiles, las cuales preservan las bajas temperaturas, pero no durante mucho tiempo. Estos recorridos se realizan en horas de la mañana, tomando como punto de partida el hogar del domiciliario  $P_0$ , este debe visitar los diferentes lugares designados para la recolección de muestras  $P_r$  y finalizar el recorrido en el laboratorio  $P_e$ . Para determinar las rutas optimas para los domiciliarios, se realiza un proceso de agrupamiento o clústering geográfico de los diferentes puntos de recolección  $P_r$  por zonas [1], para esto se tiene en cuenta la cantidad de domiciliarios disponibles  $K$ , los cuales se encargarán de tomar las muestras.

Este procedimiento no sólo beneficiará a los laboratorios que presten estos servicios representando menores costos logísticos y probablemente un mayor número de servicios diarios, sino también a los domiciliarios encargados de tomar las muestras, pues tendrán rutas más consistentes diariamente y recorrerán menores distancias.

## 2 Planteamiento del problema

Dado un conjunto de puntos de recolección de muestras  $P_R$  con coordenadas  $(x, y)$  y una cantidad de enfermeros domiciliarios  $K$  para un día de la semana, necesitamos encontrar  $K$  clusters de tamaño mínimo  $\lfloor \frac{R}{K} \rfloor$  y tamaño máximo  $\lceil \frac{R}{K} \rceil$ .

El problema es un subproblema del enrutamiento de flotas, en el cuál debemos encontrar los grupos geográficos que minimicen las distancias entre clientes, y asignar cada uno de estos grupos a un enfermero. El problema de hallar la ruta óptima dentro de cada grupo no se aborda en esta propuesta.

### 2.1 Técnicas propuestas por otros autores

Se han propuesto métodos de clústering geográfico para redes de recolección y entrega de paquetes, que buscan minimizar el número de vehículos de la flota que la empresa debe usar, utilizando algoritmos jerárquicos aglomerativos con una función de distancia que logren minimizar los tiempos (costo) entre clientes de un clúster [1].

Nazari, et al. [2], proponen un algoritmo *bottom-up* de clústering jerárquico con puntos superpuestos con complejidad  $O(n^2)$ , agrupando parejas en cada nivel, encontrando intersecciones entre clústers y creando un nuevo clúster como la unión de estos dos.

## 3 Método

Para la creación de los grupos, se va a utilizar un algoritmo aglomerativo, que implica que se van a calcular los puntos más cercanos de todo el conjunto de puntos, que se irán juntando poco a poco hasta formar grupos más grandes con la cantidad deseada. Esta distancia se va a calcular teniendo en cuenta el algoritmo de divide y vencerás para la selección de estos puntos, armando parejas de puntos lo más juntos posibles. Una vez se unan los puntos, se generará un punto medio, que hará parte de la nueva lista de puntos, para así repetir el proceso.

Al finalizar la selección de los clusters, se hará la asignación de los grupos a cada uno de los domiciliarios, que se hace de manera iterativa sobre los domiciliarios disponibles. Se comienza encontrando el clúster más cercano para todos los domiciliarios disponibles y en caso de desempate, se seleccionará el domiciliario más lejano (para que la varianza de las distancias domiciliario-clúster sea lo más baja posible).

### 3.1 Algoritmo

El pseudocódigo del algoritmo propuesto es el siguiente:

```
clusters_completos = []
clusters = puntos_entrega
```

```

clusters_actuales = tamaño de clusters
mientras clusters_actuales < numero_domiciliarios):
    # Función para calcular los puntos más cercanos por medio de divide y vencerás
    pares_cercanos = pares_cercanos(puntos_entrega)
    clusters = pares_cercanos['puntos_sobrantes']
    punto_medio: pares_cercanos['pares']
    si tamaño del cluster con centroide en punto_medio >= numero_domiciliarios:
        agregar a clusters_completos el punto_medio
    sino:
        agregar a clusters el punto_medio
    restar 1 a clusters_actuales

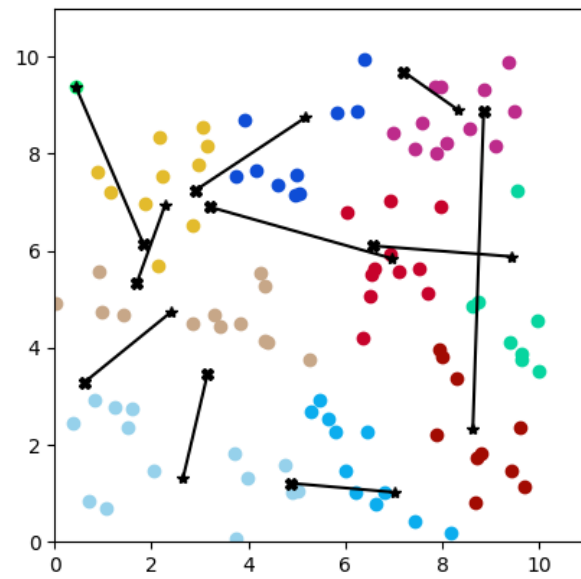
```

La complejidad de este algoritmo es de  $O(N^2 \log^2(N))$ , que es peor que el de otros métodos de agrupación, sin embargo, es más preciso que otros métodos como K-means [2] y creemos que esto compensa su peor eficiencia.

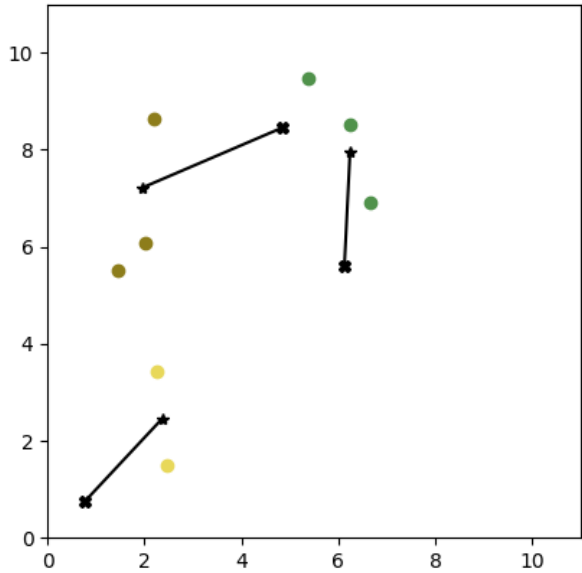
## 4 Resultados

El algoritmo es viable para un número de puntos de entrega no mayor a 100, lo cual es viable para empresas con alrededor de 10 domiciliarios.

**Nota:** las pruebas fueron realizadas en un MacBook Air M1 a 3,2 GHz.



**Fig. 1.** Resultados del algoritmo para 100 puntos de entrega y 10 domiciliarios. Tiempo de ejecución: 29,25 ms.



**Fig. 2.** Resultados del algoritmo para 8 puntos de entrega y 3 domiciliarios. Tiempo de ejecución: 109,20  $\mu s$

## 5 Conclusiones

El agrupamiento, en su mayoría genera resultados satisfactorios, sin embargo, es necesario una mayor investigación para crear una versión modificada capaz de crear clusters más balanceados.

El algoritmo propuesto presenta una manera más precisa de agrupamiento que otros algoritmos mencionados, y aunque su complejidad en tiempo sea mayor, puede ser útil en casos donde se busque un agrupamiento más preciso.

En este experimento se realizaron varios intentos de hacer una correcta normalización de la cantidad de los datos en los cluster, sin embargo estos no tuvieron éxito debido a las diferentes condiciones en las que se presentan algunos puntos y algunos casos extraños que requieren mayor investigación.

## Referencias

1. Jarrah, A., Bard, J.: Pickup and delivery network segmentation using contiguous geographic clustering. *Journal of the Operational Research Society* **62**(10) (2011). <https://doi.org/10.1057/jors.2010.123>
2. Nazari, Z., Nazari, M., Kang, D.: A bottom-up hierarchical clustering algorithm with intersection points. *International Journal of Innovative Computing, Information and Control* **15**(1) (2019). <https://doi.org/10.24507/ijicic.15.01.291>