

Selección de centros

Teoría de Algoritmos I (75.29 / 95.06)

Ing. Víctor Daniel Podberezski

✉ vpodberezski@fi.uba.ar

El problema de la selección de centros

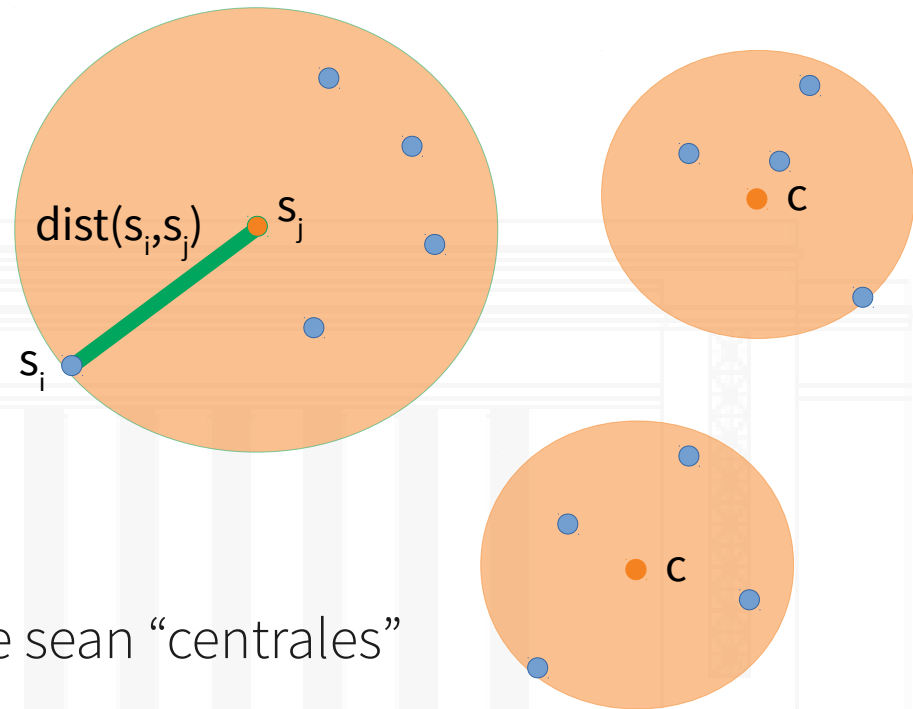
Sea

Un set S de n sitios

Una función “distancia”

Queremos

seleccionar k centros de tal forma que sean “centrales”



Función distancia

Función numérica

que mide el cuan cerca se encuentra un punto x de otro punto y

Debe satisfacer:

- $\text{dist}(x,x) = 0$
- $\text{dist}(x,y) = \text{dist}(y,x)$ (simetría)
- $\text{dist}(x,y) + \text{dist}(y,z) \geq \text{dist}(x,z)$ (desigualdad triangular)

Ejemplo: Distancia euclidiana

“Centralidad”

Sea C el set de centros.

Cada sitio “s” tiene una distancia a cada centro c de C .

Asignaremos a “s” a la órbita del centro c mas cercano.

$$\text{dist}(s, C) = \min_{c \in C} \text{dist}(s, c).$$

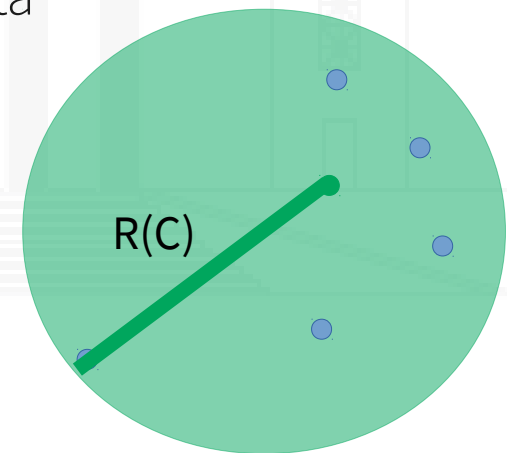
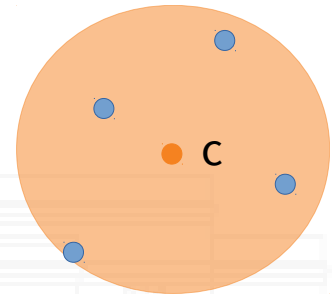
Definimos a r como el radio de cobertura.

Distancia máxima entre cada sitio y el centro al que órbita

$r(C)$ = radio de cobertura

Intentaremos seleccionar el set C de k centros

para minimizar el radio de cobertura



Tratando de construir un buen algoritmo

Supongamos que

conocemos que existe un set C^* de k centros con radio de cobertura $r(C^*) \leq r$

Cada sitio s debe tener un centro c^* en C^* que lo cubra

Podemos construir

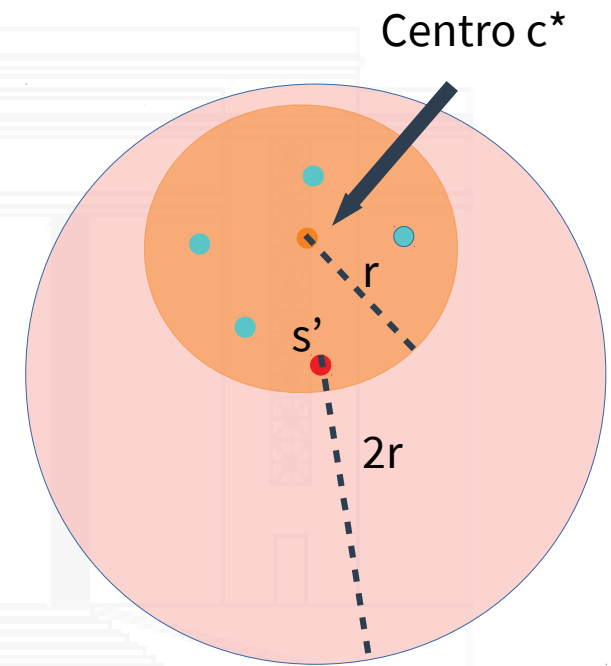
un set de k centros con radio de cobertura de como mucho $2r$

Seleccionamos un sitio s'

El sitio s' se encuentra a distancia máxima r de su centro c^*

Definimos s' como centro con cobertura $2r$

El nuevo centro s' contendrá a todos los sitios del centro c^*



Algoritmo propuesto

Sea S' la lista de sitios no asignados a un centro

$S' = S$

Sea $C = \emptyset$

Mientras $S \neq \emptyset$

 Seleccionar cualquier sitio $s \in S$ y agregarlo a C

 Borrar todos los sitios de S que tengan menor distancia a s de $2r$

Si $|C| \leq k$

 Retornar C como los centros seleccionados

Sino

 No hay k centros con radio de cobertura de al menos r

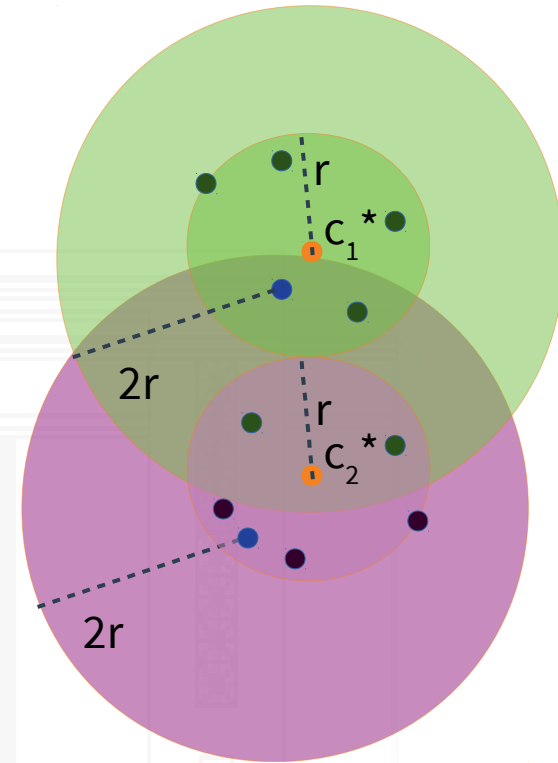
Análisis del algoritmo

Si el algoritmo retorna un set de k centros

$$r(c) \leq 2r$$

Si selecciona más de k centros

$r(C^*) > r$ contradiciendo la afirmación



¿Qué valor tendrá r en nuestro algoritmo?

Podemos realizar un algoritmo iterativo probando diferentes valores de r

Iniciamos con un $r = \text{distancia máxima entre 2 sitios} / 2$

Ejecutamos el algoritmo y comprobamos si existe resultado

Si existe puedo probar con un r más chico

Si no existe puedo probar con un r más grande

Puedo iterar modificando los radios (similar a búsqueda binaria) y aproximar al resultado tanto como considere oportuno.

El resultado será lo centros aproximados

Un algoritmo greedy sin presuponer el radio

Queremos resolver el problema

sin necesidad de suponer un radio de cobertura en la solución óptima

Implica un cambio pequeño en el algoritmo anterior

Seleccionamos siempre como próximo centro al punto disponible más lejano a los centros existentes

Si existe un centro a más de $2r$ de distancia de aquellos,

entonces el punto mas lejano debe formar parte de él

Algoritmo propuesto

Asumimos $k \leq |S|$ (sino definimos $C = S$)

Seleccionar cualquier sitio s y convertirlo en un centro $C = \{s\}$

Mientras $|C| < k$

 Seleccionar sitio $s \in S$ que maximice la distancia $\text{dist}(s, C)$

$C = C \cup \{s\}$

Returnar C como los sitios seleccionados

Análisis del algoritmo

Si

C^* es un set de centros optimos

C es el set que encuentra el algoritmo

Entonces

$$r(C) \leq 2r(C^*)$$

Por contradiccion

asumiremos $r(C^*) < \frac{1}{2} r(C)$

Por cada sitio c perteneciente a C , consideremos un circulo de radio $\frac{1}{2} r(C)$ a su alrededor

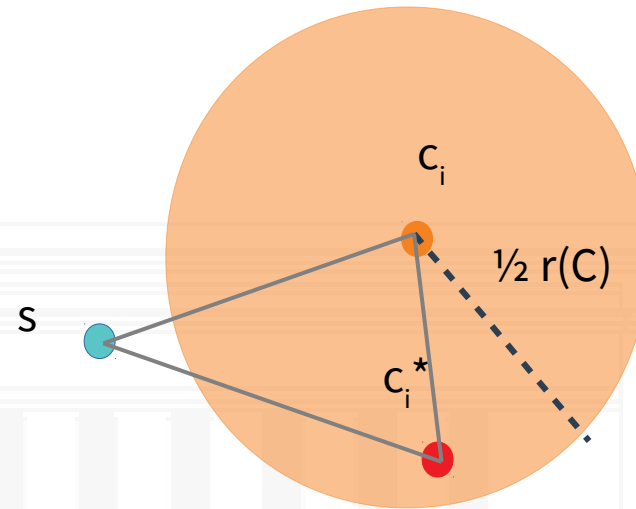
Llamaremos c_i a uno de ellos

Hay exactamente un único c^* dentro del circulo

Llamaremos c_i^* al que se encuentra dentro del circulo de c_i

Consideremos cualquier sitio s y su centro mas cercano de la solución óptima C^*

$$\begin{aligned} \text{dist}(s, C) &\leq \text{dist}(s, c_i) \leq \text{dist}(s, c_i^*) + \text{dist}(c_i^*, c_i) \leq 2r(C^*) \\ &\leq r(C^*) \text{ por que } c_i^* \text{ es su centro mas cercano} \end{aligned} \quad \longrightarrow \quad r(C) \leq 2r(C^*)$$



Conclusión

La solución presentada

Es un algoritmo de tipo greedy

Corresponde a una 2-aproximación

Del problema de selección de centros

Se ejecuta en tiempo polinomial



Presentación realizada en Julio de 2020