

# Detección de Comunidades en Redes Complejas

Aarón Rosas

Francisco Javier Gijón

# Contenido

- Redes complejas
- Comunidades
- Detección de comunidades
- Evaluación
- Implementación
- Wisper

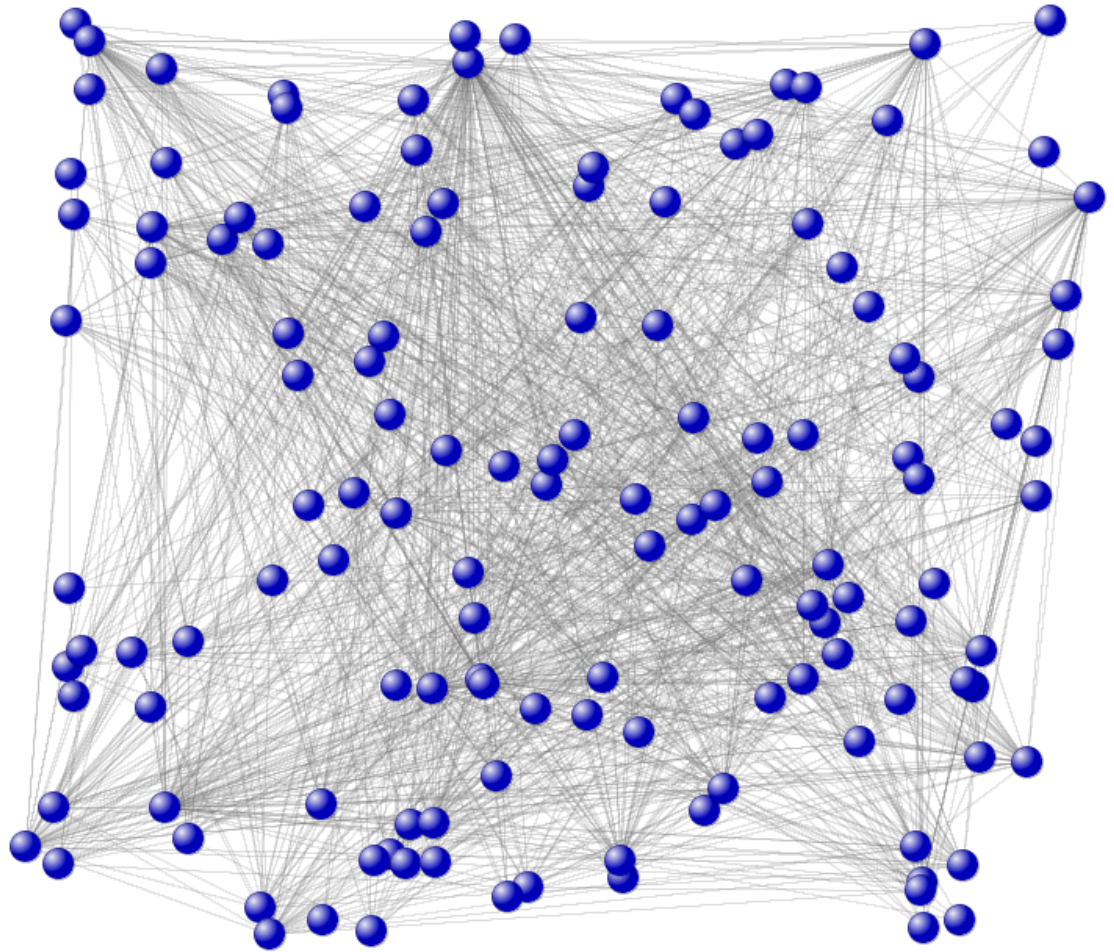
# Redes Complejas

- Modelan interacciones entre elementos
  - Redes de transporte
  - Redes biológicas
  - Redes sociales
  - Redes comerciales
  - Etc.



## Red de amigos de Francisco en Facebook

- 137 amigos
- 1336 relaciones de amistad



# Comunidades

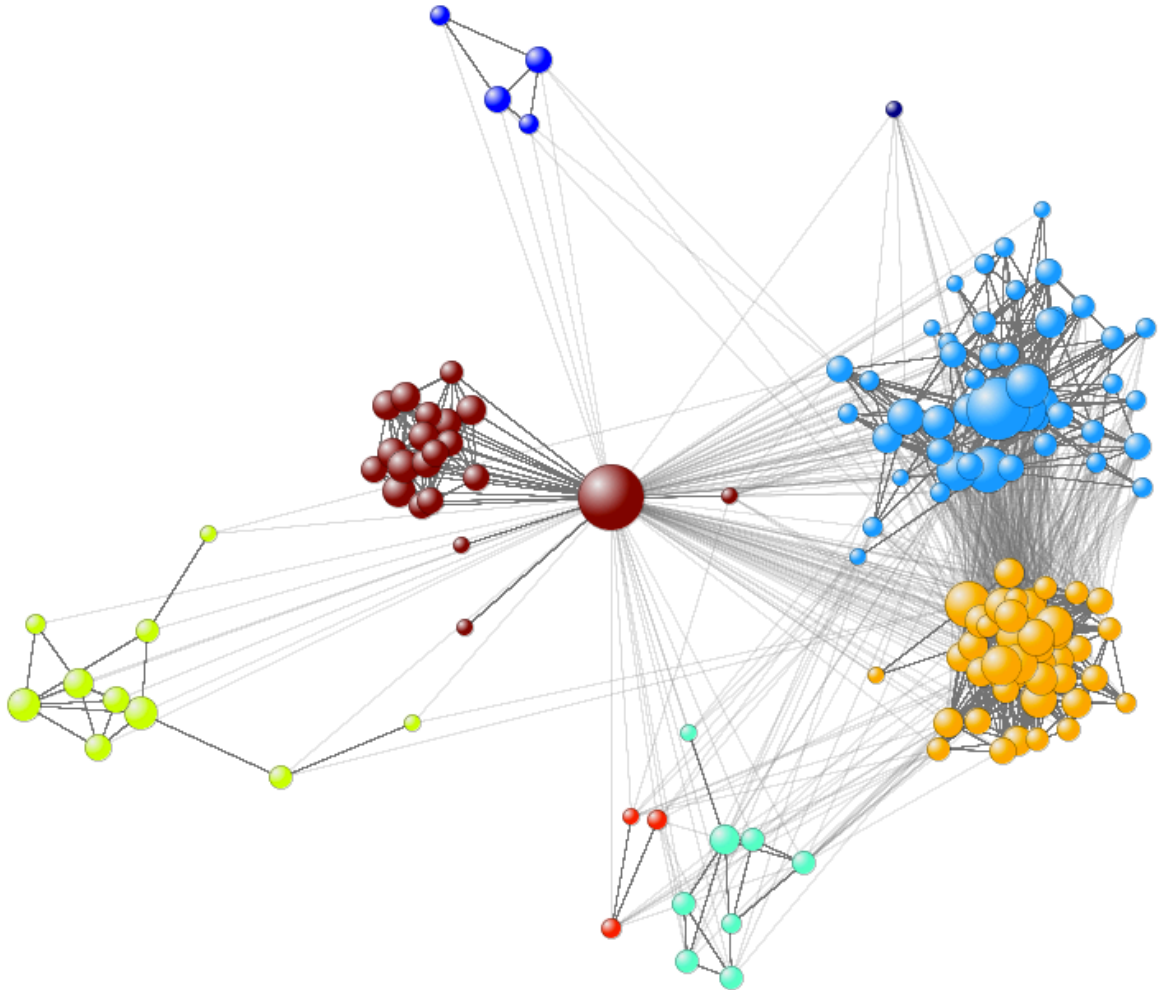
- Estructura comunitaria
- Sociología

*Conjunto de personas  
con intereses y actividades comunes*

- Presentan
  - Alto grado de cohesión interna
  - Alto grado de separación entre comunidades

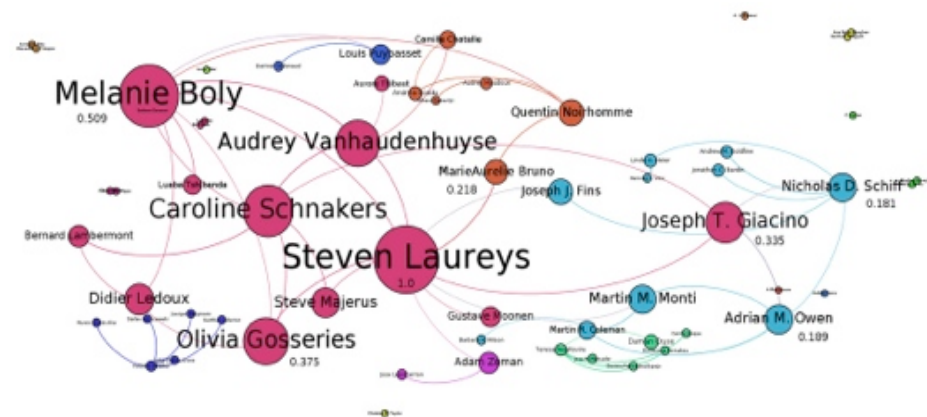
## Red de amigos de Francisco en Facebook

- 137 amigos
- 1336 relaciones de amistad



# Objetivo

- Dada una red, encontrar sus comunidades.
- Aplicaciones
  - Citas bibliográficas
  - Redes sociales
  - Procesos Biológicos



# Definición del problema

Red = Grafo

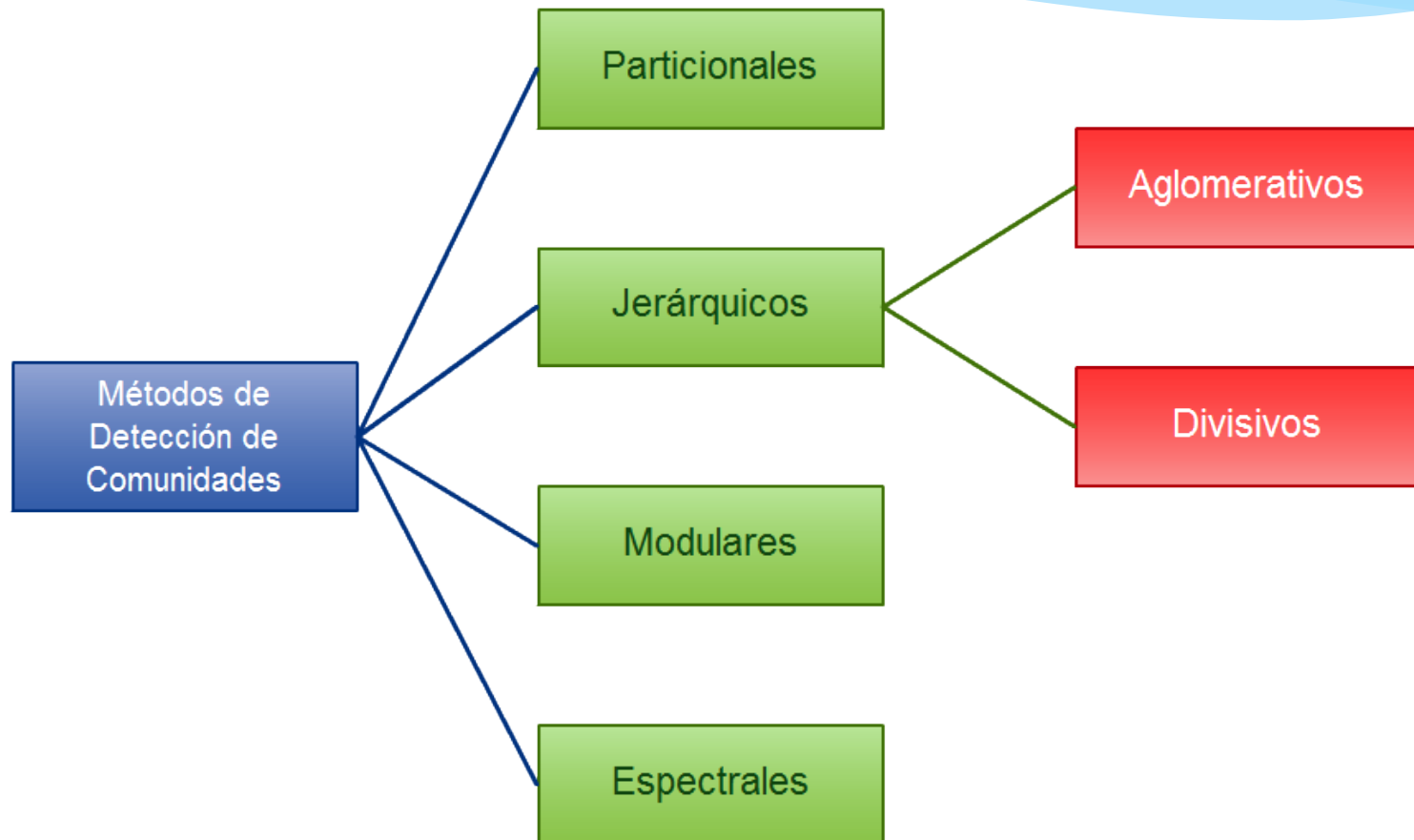
Elementos = Nodos

Interacciones = Enlaces

Sea  $G=(V,E)$  un grafo, encontrar la partición de los nodos de  $G$  en varios subconjuntos de forma que cada uno cumpla la noción intuitiva de comunidad.

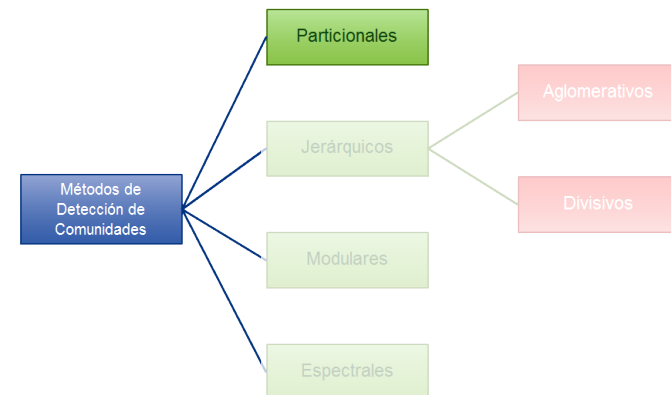
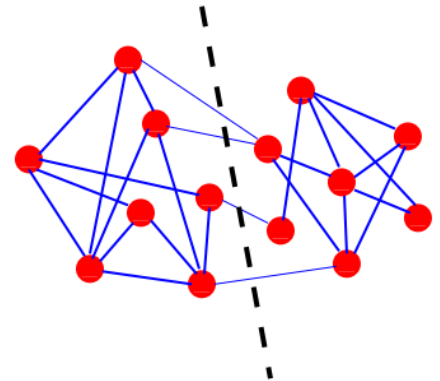


# Técnicas de detección



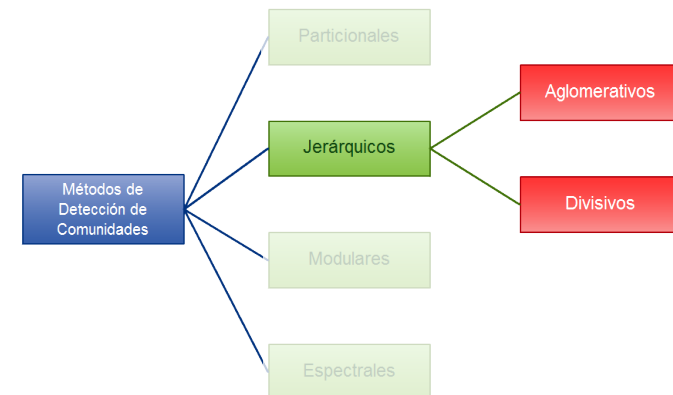
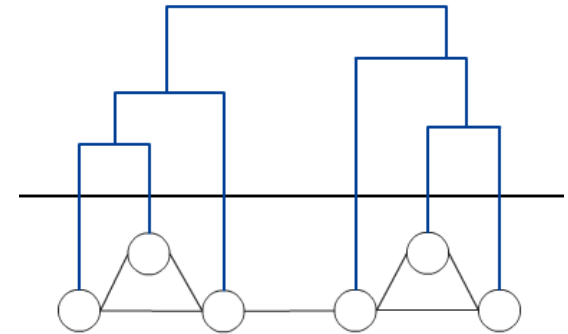
# Particionales

- \* Partir el grafo en k grupos
- \* Diferentes enfoques
  - \* Espacio métrico
    - \* Agrupan puntos del espacio métrico
    - \* KMeans
  - \* Corte de grafos
    - \* Minimizan función de corte
    - \* Kernighan-Lin



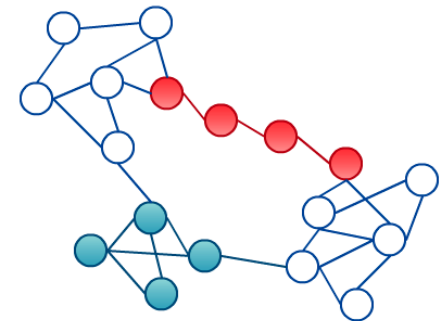
# Jerárquicos

- \* Construyen el dendrograma
- \* Se centran en los enlaces
- \* 2 enfoques
  - \* Aglomerativos
  - \* Divisivos

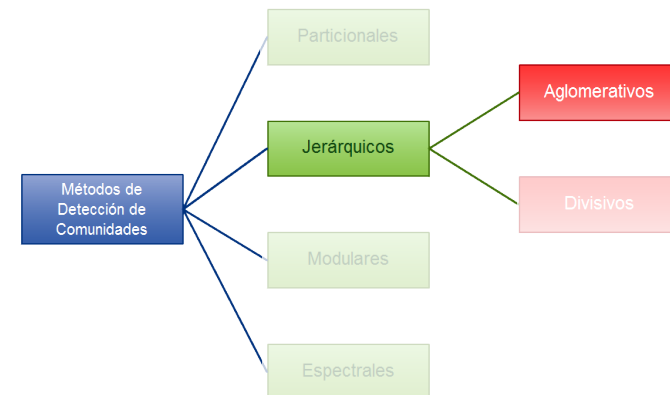


# Jerárquicos Aglomerativos

- \* *Bottom-up*
- \* Principio
  - \* fusionar clusters cercanos
- \* Variantes
  - \* Single-Link: distancia mínima
  - \* Complete-Link: distancia máxima
  - \* Average-Link: distancia media

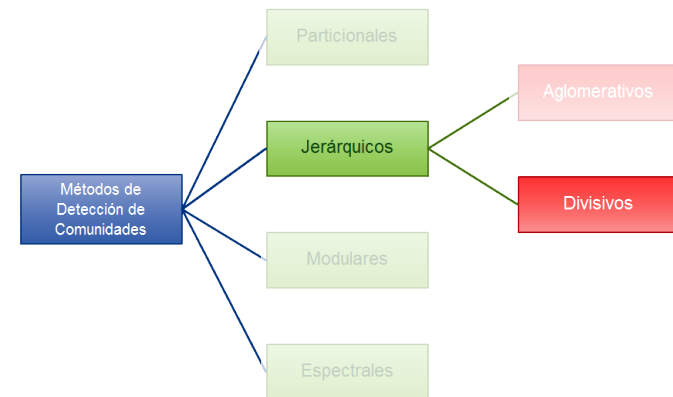
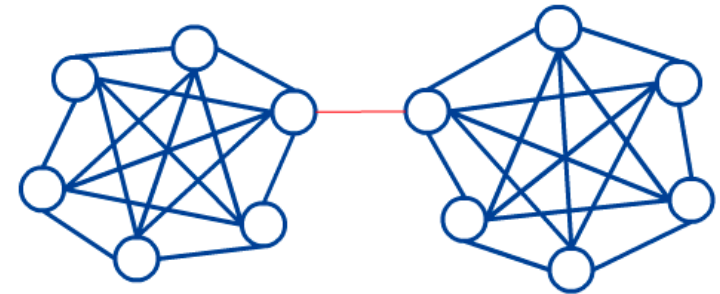


Encadenamiento



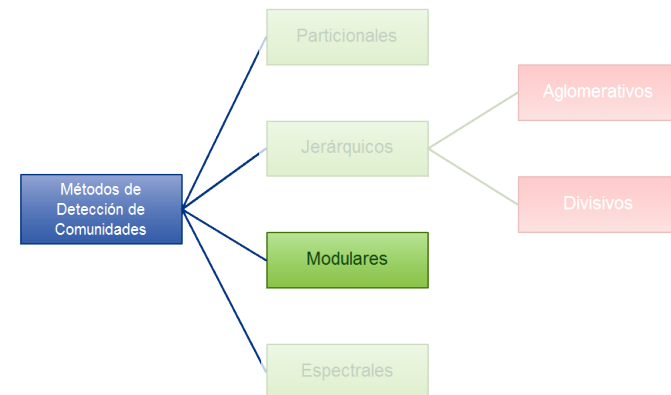
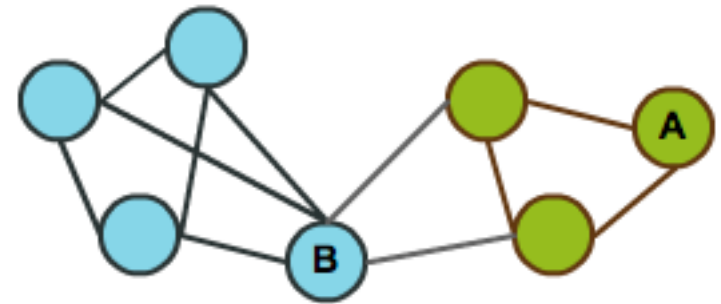
# Jerárquicos Divisivos

- \* *Top-down*
- \* Principio
  - \* Identificar enlaces puente
- \* Variantes
  - \* Newman-Girvan  
*link betweenness*
  - \* Radicchi  
*link clustering coefficient*



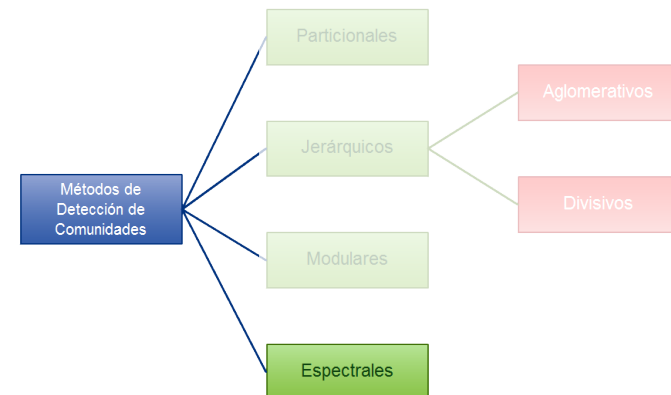
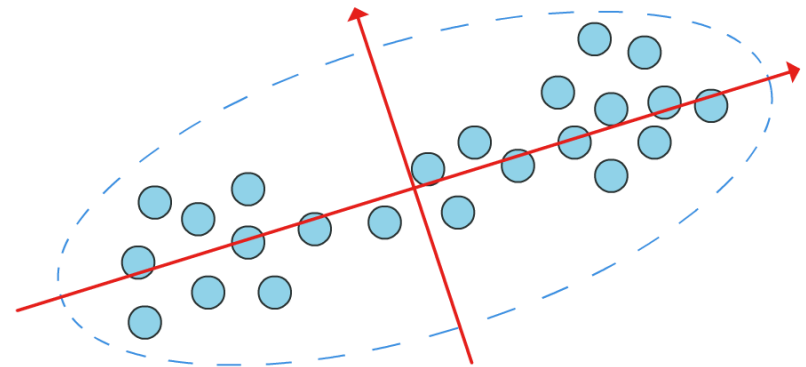
# Modulares

- \* Enfoque Greedy
- \* Optimizan Modularidad
- \* 2 versiones
  - \* Fast Greedy
  - \* Multi Step Greedy



# Espectrales (1/2)

- \* Particionamiento
- \* Funcionamiento general
  - \* Matriz Laplaciana
  - \*  $k$  vectores propios
  - \* Vectores  $\rightarrow$  Columnas
  - \* Agrupa puntos en  $k$  grupos



# Espectrales (2/2)

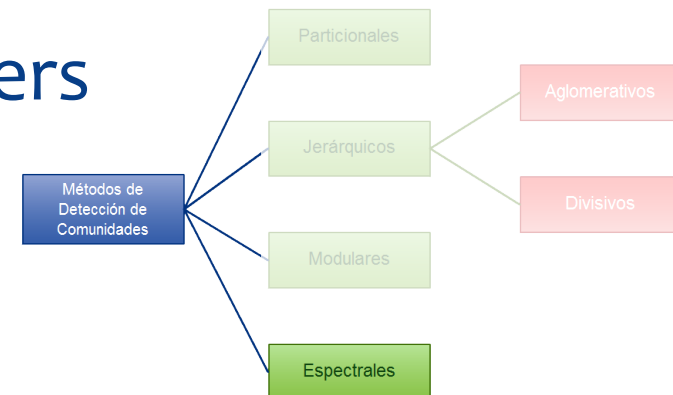
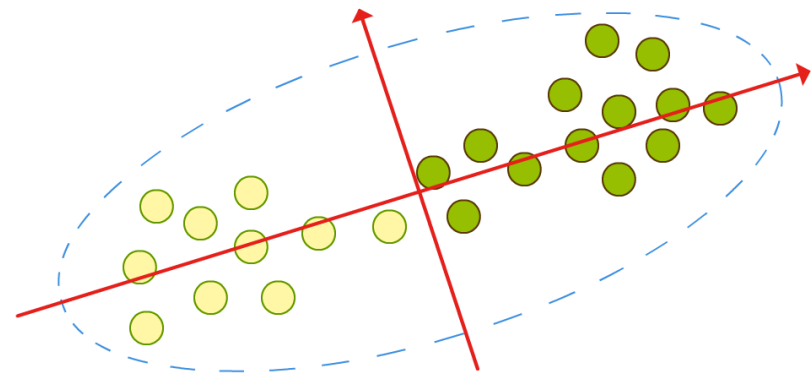
- \* Variantes

- \* Bisección Espectral

- \* UKmeans

- \* NJW

- \* Estimación número de clusters

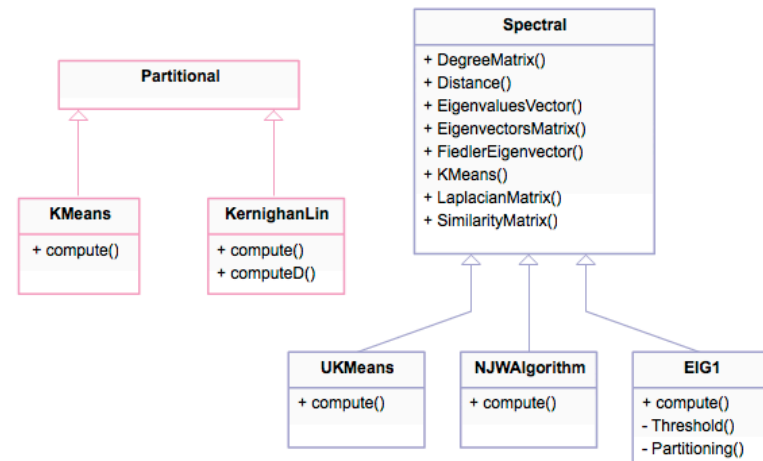
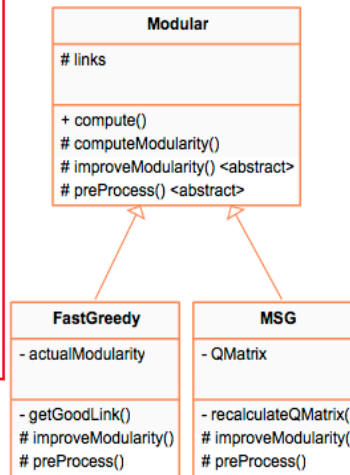
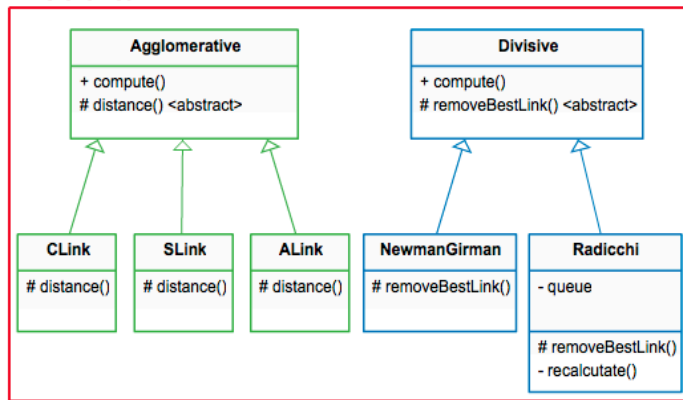




# Implementación (1/3)

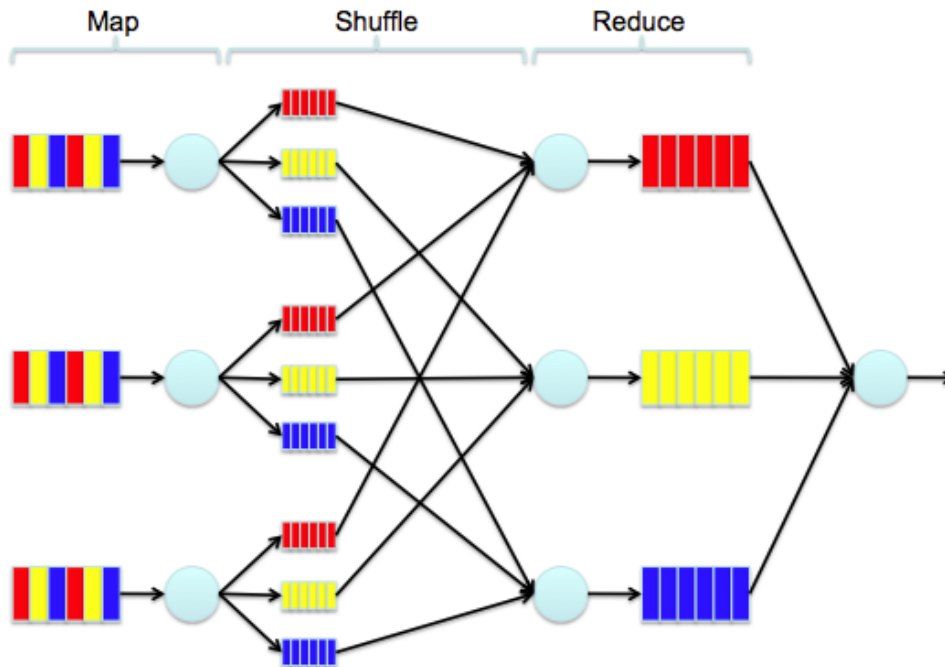
## \* Integrado en NOESIS

### Hierarchical



# Implementación (2/3)

## \* Paralelización: MapReduce



# Implementación (3/3)

- \* Código fuente

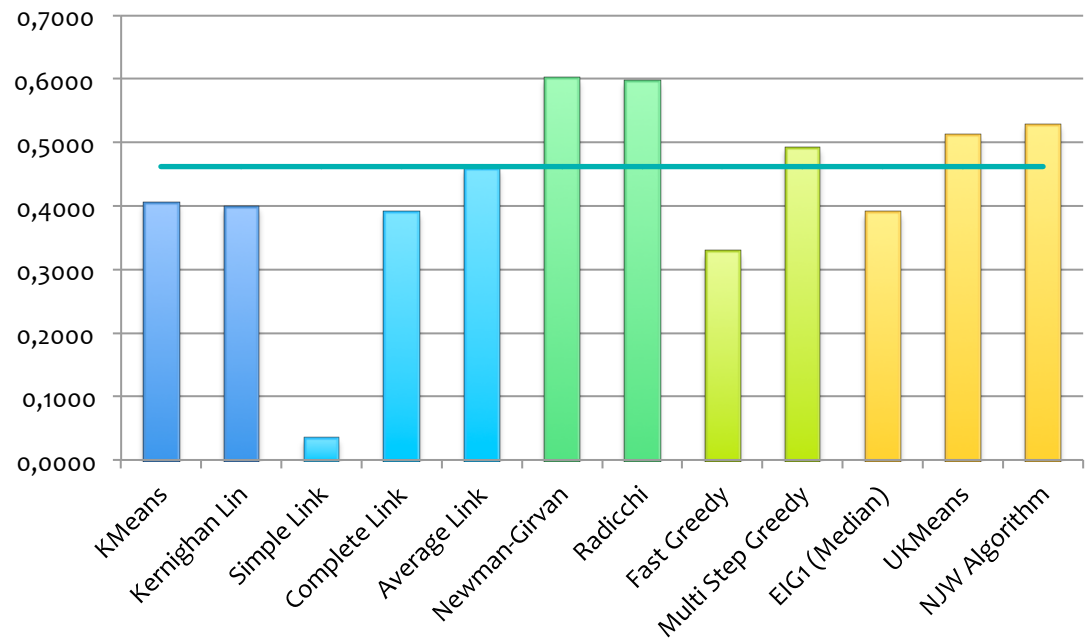
- \* <https://code.google.com/p/cdws/>

- \* Licencia BSD

# Evaluación (1/2)

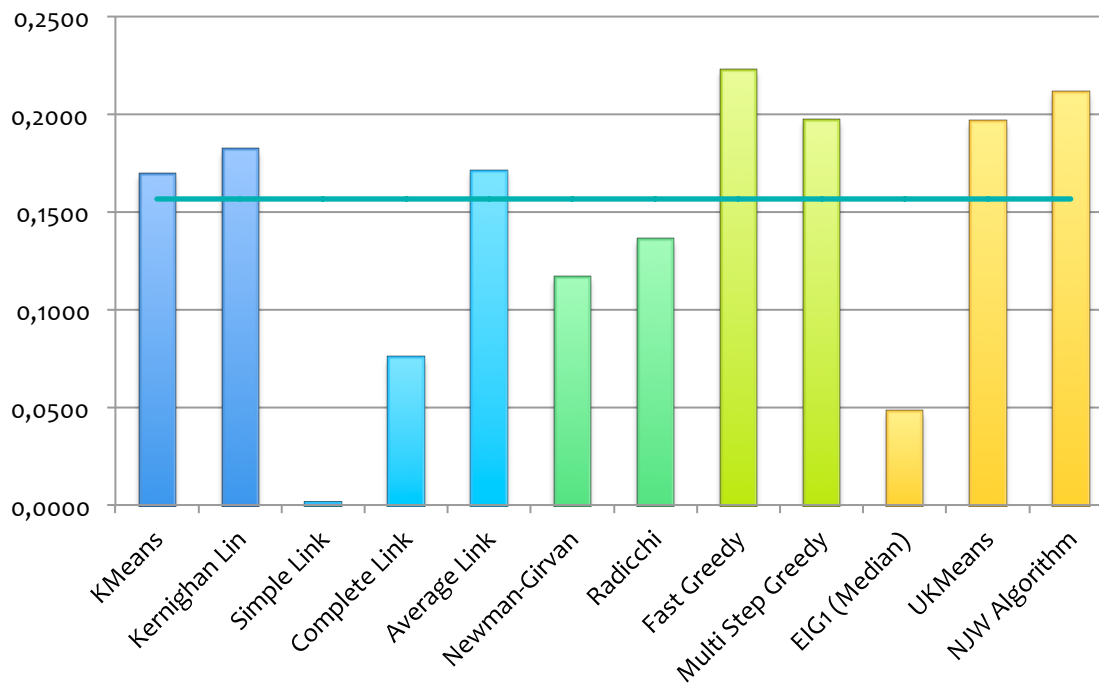
## \* Medidas

- \* Modularidad
- \* Cohesión
- \* Separación
- \* Silueta
- \* Cobertura



Modularidad: Red Futbol (115 nodos, 613 enlaces)

# Evaluación (2/2)



Algoritmo	Tiempo
KMeans	56
Kernighan Lin	815
Simple Link	393
Complete Link	933
Average Link	346
Newman-Girvan	33207
Radicchi	1467
Fast Greedy	11021
Multi Step Greedy	84166
EIG1 (Median)	224
UKMeans	117
NJW Algorithm	235

Red de Facebook: Aarón (270 nodos, 3277 enlaces)

# wisper.es (1/2)

- \* Aplicación a redes sociales
  - \* Facebook
  - \* Twitter
  - \* Etc.
- \* Determina grupos de usuarios según sus interacciones
- \* Visualización de contenidos más ordenada



# wisper.es (2/2)

- \* Trabajando con Twitter
- \* FOTO/VIDEO DEMO (como alternativa si falla Internet)

