30/04/2024

Analisis de estructuras en red (grafos)

1. Introducción

Redes: conjunto de objetos interconectados

Existen muchos tipos de redes:

* Redes de información:
* Redes de transporte y logística
* Redes sociales
* Redes biológicas
* Otras: redes financieras, redes comerciales, redes de autores, etc.
* También se usan en sistema informáticos de última generación

1. Aplicaciones del análisis de redes

Que se puede hacer con las redes? Ejemplos:

* Es probable que se propague un rumor en la red?
* Quien son las personas más influyentes?
* Es probable que un club de divida en dos grupos? Si es así, que nodos formarían parte?
* Qué aeropuertos tienen mayor riesgo de propagación de cierto virus? Qué destinos están menos expuestos

1. Conceptos

Red (o grafo): representación de conexiones entre elementos

Nodos: los elementos de la red

Enlaces (o aristas) enlaces entre nodos. Las redes pueden ser:

* Simetricales: relaciones bidireccionales (familiares)
* Asimétricas: es importante saber que nodo apunta a quien (cadena de alimentación de animales)
* Este concepto se llama direccionalidad: no dirigida (simétrica), dirigida (asimétrica)

Redes ponderadas: la relaciones tienen un valor numerico de ocurrencias (las veces que 2 personan han almorzado juntas|)

Redes con signo: la relación tiene un valor positivo o negativo

Enlaces con atributos diversos: relaciones diferentes

Multigrafos: unir dos nodos con múltiples relaciones a la vez

1. Grafos bipartidos

Grafos que estan compuestos de 2 partes. Ejemplo: fans y equipos de futbol. Los nodos de este grafo solo se pueden relacionar con los grafos del otro lado.

Se puede hacer una proyección L y R del grafo, obteniendo las relaciones entre los dos grupos del grupo

07/05/2024

Parte 2

1. Distancias

¿Cuanto de lejos esta el nodo A del nodo H?

Camino: secuencia de nodos conectados por un enlace.

Saltos: nodos por los que hay que pasar

Distancia entre dos nodos: la longitud del camino más corto entre ellos

Búsqueda por amplitud: procedimiento para calcular la distancia de un nodo hasta el resto

Diámetro: distancia máxima entre cualquier par de nodos del grafo

Excentricidad: de un nodo ‘n’ es la mayor distancia entre ‘n’ y todos sus nodos. Para cada nodo, busca cual es la mayor distancia para llegar al nodo mas alejado

El radio: mínima excentricidad. El nodo mas central, a cual le costaría llegar menos

Periferia: conjunto de nodos que tienen excentricidad igual al diámetro

Centro: excentricidad es igual al radio

1. Conectividad

Grafo no dirigido:

Los enlaces no tienen dirección. Se considera conectado si existe un camino para cualquier par de nodo.

* Componente conectado: los componentes tienen que estar conectados entre si y no tener conexión con otros nodos fuera del subconjunto

Grafo dirigido:

Se considera fuertemente conectado si hay un camino entre A -> B y B -> A

Se considera débilmente conectado si al remplezar los enlaces dirigidos con enlaces no dirigidos en grafo está conectado

1. Comunidades

Identificar subconjunto de nodos

Algoritmo de Lovain, busca la optimización de la modularidad del grafo

Modularidad:

* Mide la calidad de la partición de un comunidad
* Premisa: una buena partición tiene una partición similar dentro del mismo grupo, partición diferente de otro grupo

1. Consistencia

La consistencia es la capacidad de una red para mantener sus propiedad estructural cuando se enfrenta a fallos o ataques

* Tipo de ataques: eliminar nodos o enlaces
* Propiedades estructurales
* Ejemplos: un aeropuerto ciera

21/05/2024

Medidas de centralidad

1. Relevancia de un nodo

La importancia de un nodo depende del criterio que se tome

* Numero de amigos
* Fracción de los caminos más cortos que pasan por el nodo

Las medidas de centralidad sirven para:

* Nodos ifluyentes en una red social
* Nodos centrales en una epidemia

Estas medidas son múltiples:

\*

Grado de centralidad

Nodos con mas conexiones (numero de vecinos)

Redes no dirigidas: degree

Redes dirigidas: in-degree / out-degree

Intermediacion

Numero de caminos mas cortos que pasan por ese nodo (esta en el medio).

Este algoritmo es bastante costoso computacionalmente. Debido a ello, es posible coger tan solo una muestra en grafos con muchos nodos

También es posible trabajar con subconjutos para limitar los elementos

También es posible utilizar enlaces en lugar de nodos para medir la centralidad

Page rank

Forma de calcular la centralidad para casos concretos.

Trata de determinar que páginas web son las más importantes.

Trabaja con redes dirigidas para conocer como se distribuyen los hipervínculos

La idea del PageRank es que haciendo un recorrido aleatorio, se llegue a cierto nodo después de *k* pasos

Puede existir un problema si existen dos nodos que solo se apuntan entre ellos, abosorviendo todo el page rank. Es por ello que se creo un ‘factor de amortiguación’ llamado Scaled PageRank para evitar estos casos. El valor Alpha permite controlar la desviación de este algoritmo.

Comparativas de medidas de centralidad