

# BISOFT-28

## Estructuras de Datos 2

### Grafos II: Propiedades

- La interpretación de las diferentes propiedades de un grafo puede tener aplicaciones en el dominio del **problema**.
- La **alcanzabilidad** puede ser usada para detectar componentes fuertes en un grafo.
- Algoritmos para *web mining* toman en cuenta estas propiedades para explotar el significado y deducir nodos **autoritarios** y nodos **hubs**.

# Algoritmo de Floyd

- Encuentra **todos** los caminos más cortos en un grafo, iniciando en cualquier nodo.
- C: matriz de **costos**.
- A: matriz de costos de los caminos más **cortos**.
- n: número de nodos en el grafo.
- Eficiencia:  $O(n^3)$

# Algoritmo de Floyd (cont.)

$A \leftarrow C$

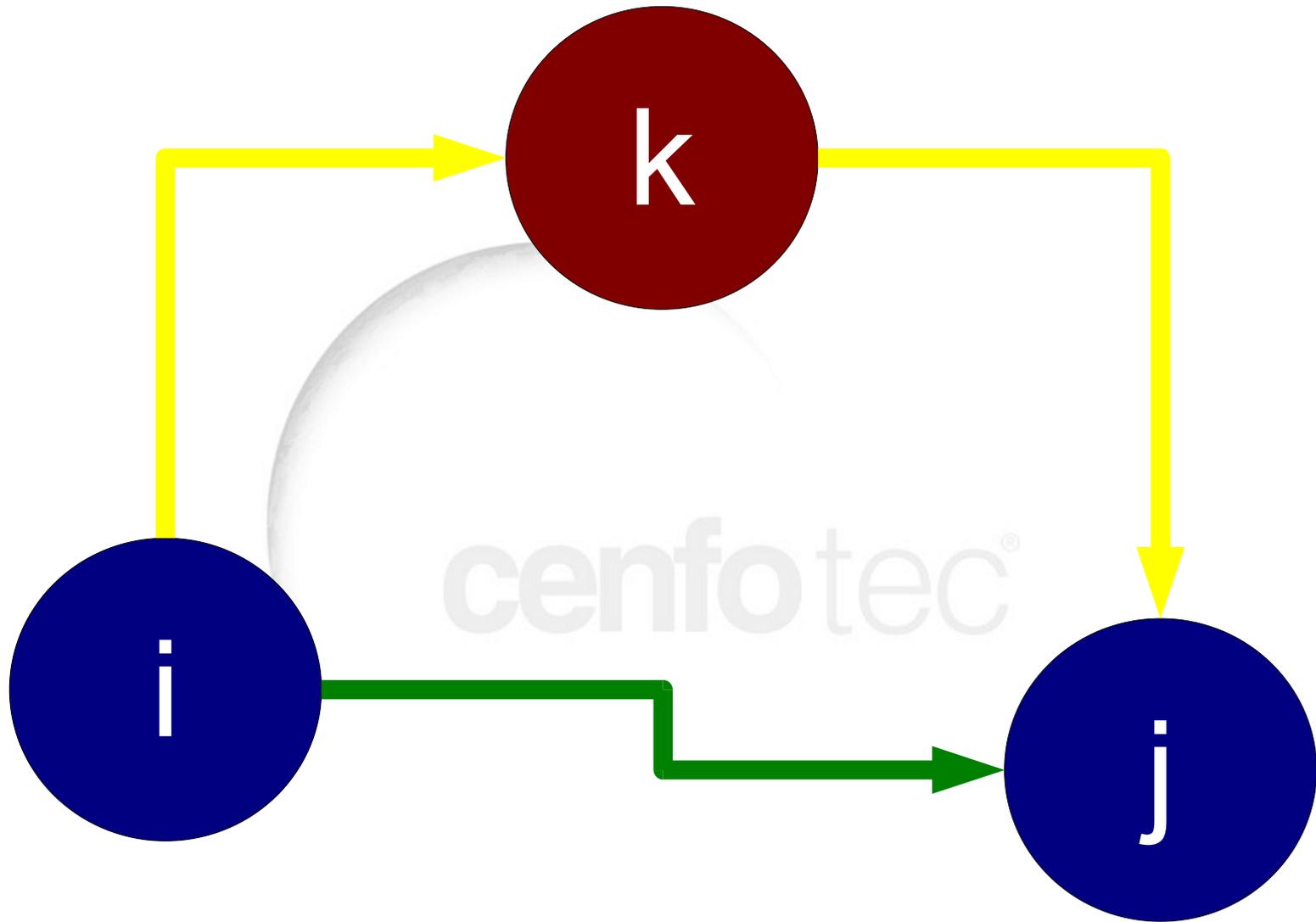
for  $k=1$  to  $n$

  for  $i=1$  to  $n$

    for  $j=1$  to  $n$

$A[i,j] \leftarrow \min(A[i,j], A[i,k]+A[k,j])$

# Algoritmo de Floyd (cont.)



# Algoritmo de Warshall

- Encuentra **todos** los nodos a los que se puede llegar desde cualquier nodo.
- M: matriz de **adyacencia**.
- D: matriz de **alcanzabilidad**.
- n: número de nodos en el grafo.
- Eficiencia:  $O(n^3)$

# Algoritmo de Warshall (cont.)

$D \leftarrow M$

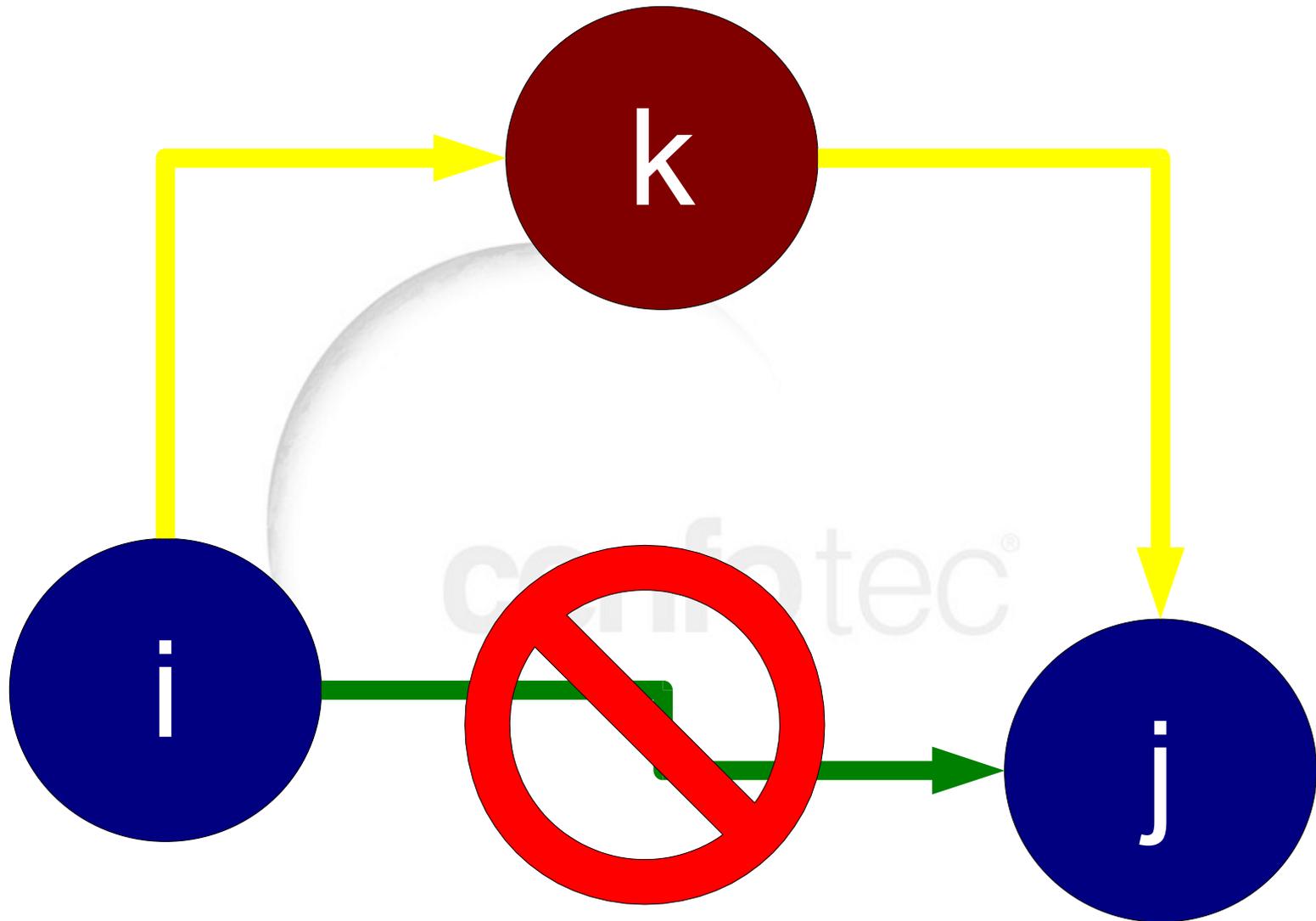
for  $k=1$  to  $n$

  for  $i=1$  to  $n$

    for  $j=1$  to  $n$

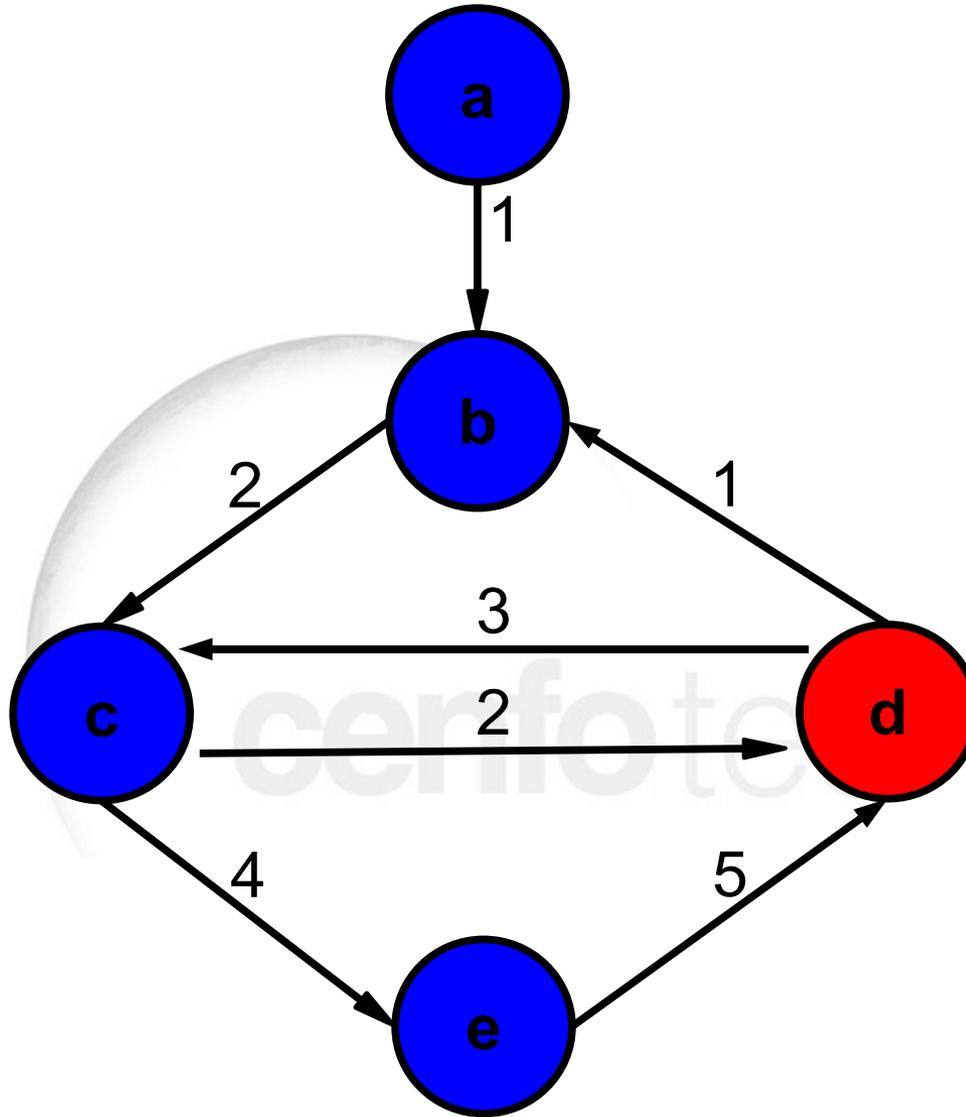
$D[i,j] \leftarrow D[i,j] \text{ OR } (D[i,k] \text{ AND } D[k,j])$

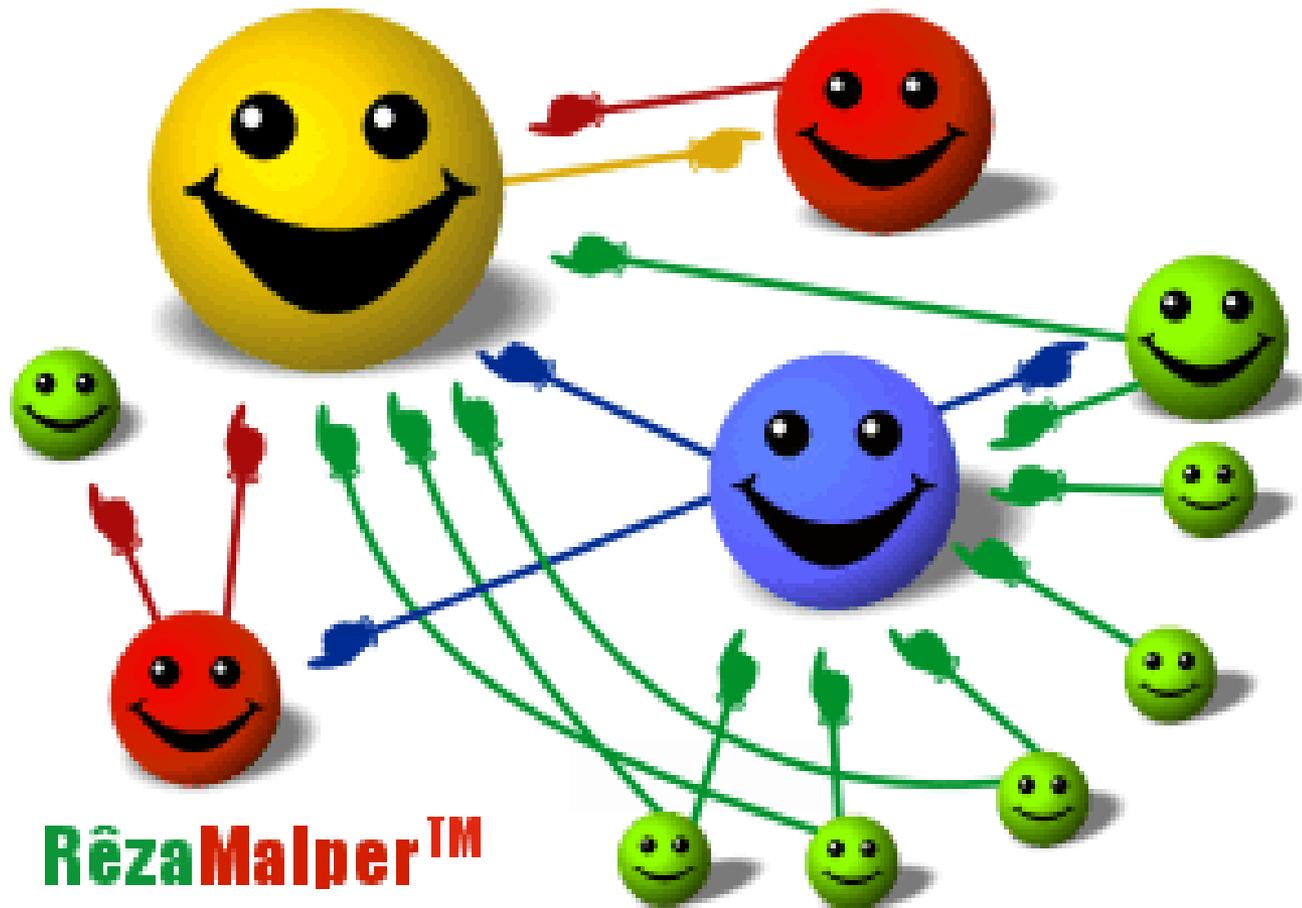
# Algoritmo de Warshall (cont.)



- Sea un **grafo**  $G=(V,A)$ .
- La **excentricidad** de  $v$  es:  
 *$\max\{\text{longitud mínima de un camino de } w \text{ a } v\}$*
- El **centro** de un grafo  $G$  es el nodo con menor excentricidad.
- El **diámetro** de un grafo es la mayor de las excentricidades.

# Ejemplo



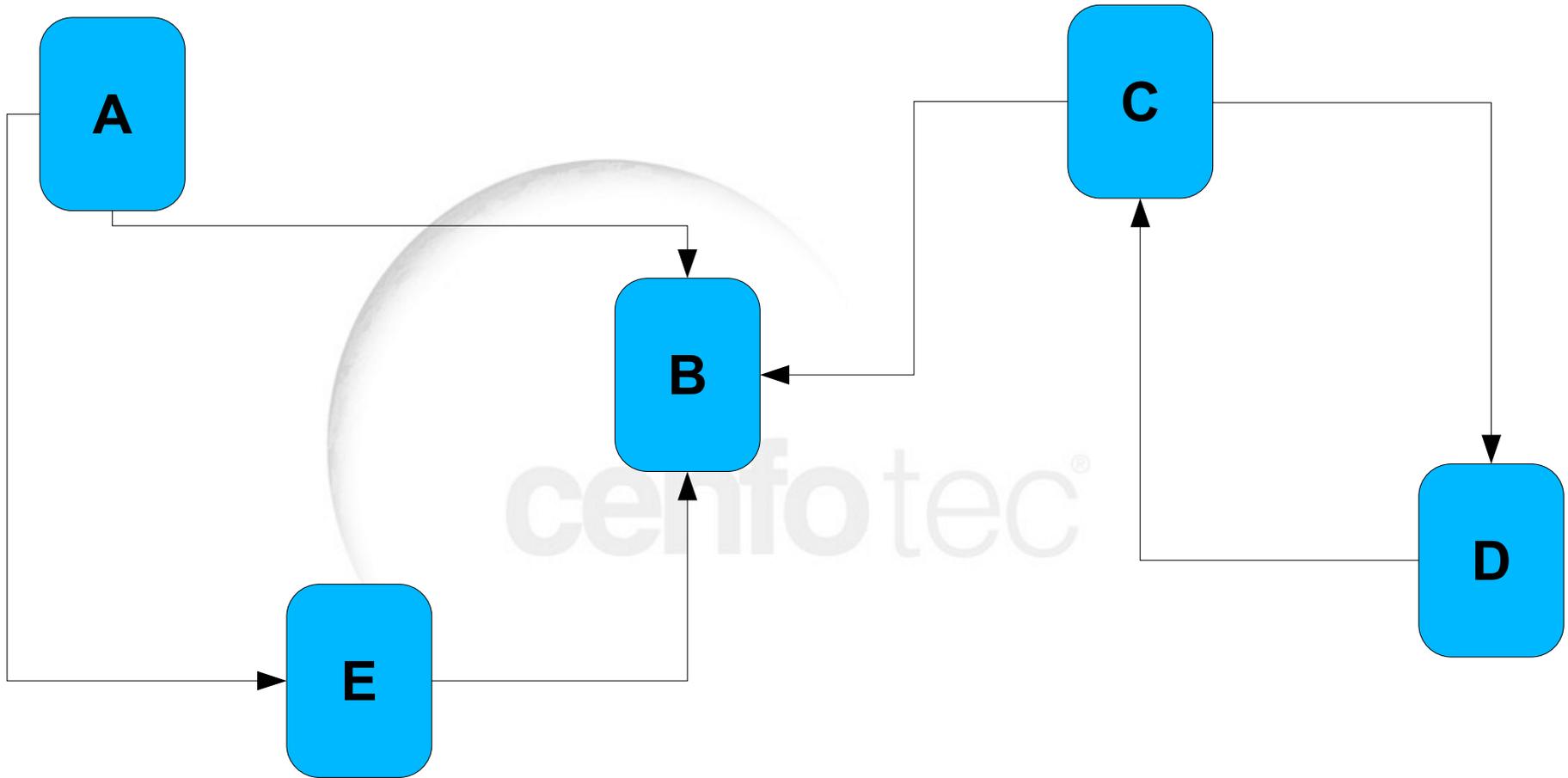


- Utilizado para analizar la **estructura** de la web.
- La web es un grafo **dirigido**.
- Determina **autoridades**: nodos mejor referenciados.
- Genera un *ranking* de todas las páginas de acuerdo a su autoridad.

$$PR(V) = (1-d) + d (PR(T1)/C(T1) + \dots + PR(Tn)/C(Tn))$$

- Las páginas  $T_i$  tienen al menos un arco directo hacia  $V$ .
- $d$  típicamente es 0.85
- El PR de todas las páginas inicialmente es 0.85

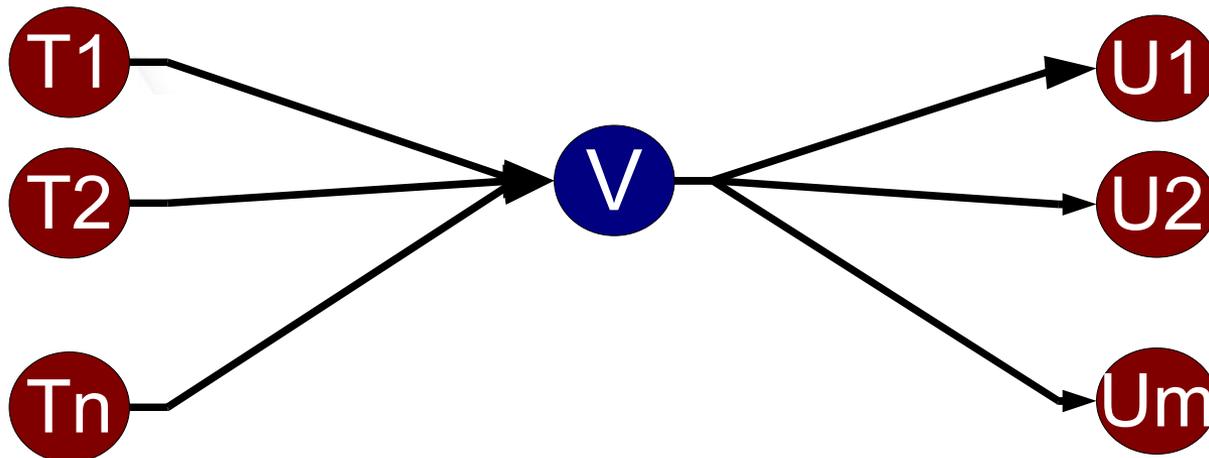
# Ejemplo



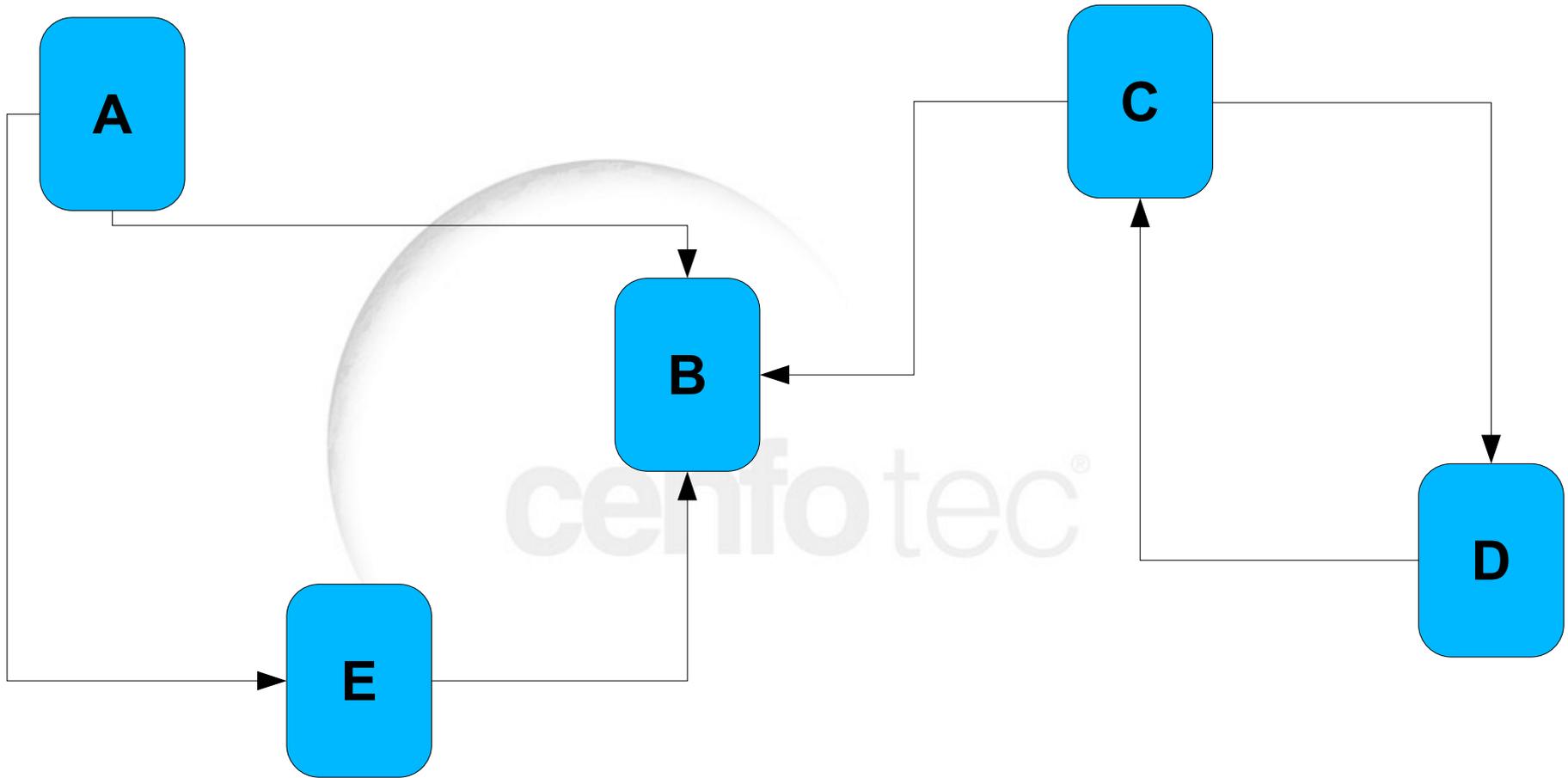
- Utilizado también para analizar la **estructura** de la web, pero clasifica los nodos en autoridades y *hubs*.
- **Autoridades**: nodos que contienen información de alta calidad.
- **Hubs**: listas comprehensivas de enlaces a autoridades.
- Cada nodo tiene un poco de **ambos** perfiles. Cada página tiene una medida de autoridad y una medida de *hub*.

# HITS (algoritmo)

1. Inicialmente, para toda página  $V$ , se tiene que:  
 $AT(V) = 1$   
 $HB(V) = 1$
2.  $AT(V) = HB(T1) + HB(T2) + \dots + HB(Tn)$
3.  $HB(V) = AT(U1) + AT(U2) + \dots + AT(Um)$
4. Repetir los pasos 2 y 3 hasta que no haya cambios sustanciales.

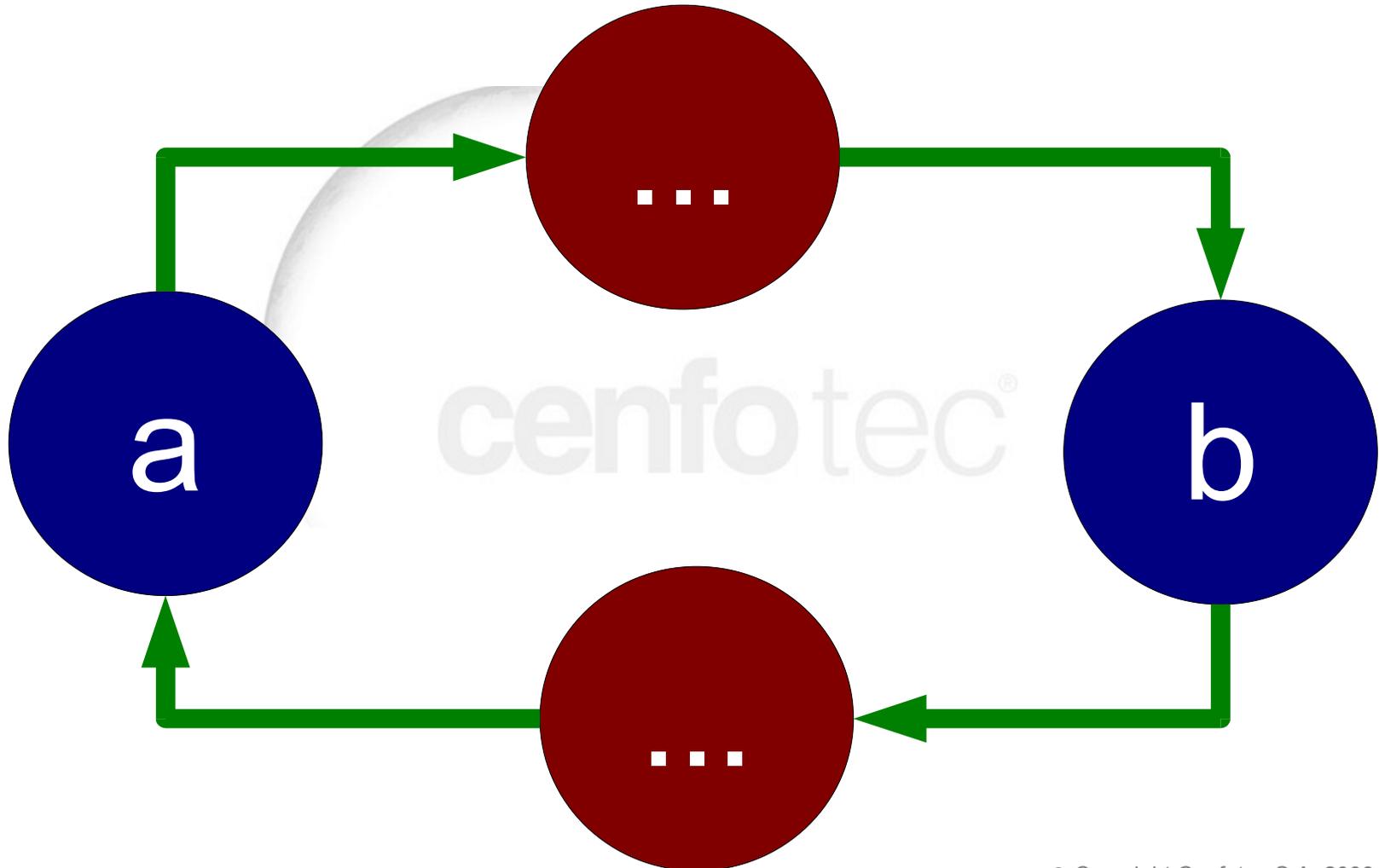


# Ejemplo



# Componentes fuertemente conexos

- Los nodos a y b están **fuertemente** conectados, si existe un camino de a hasta b y otro de b hasta a.



# Algoritmo

- Calcula el componente fuertemente **conexo** más grande al que pertenece el nodo  $x$ .
- Paso 1: calcular todos los nodos que son alcanzables desde  $x$ .
- Paso 2: calcular todos los nodos que pueden alcanzar a  $x$ .
- Hacer la intersección de los conjuntos de los pasos 1 y 2.

