Sistemas Complejos 2

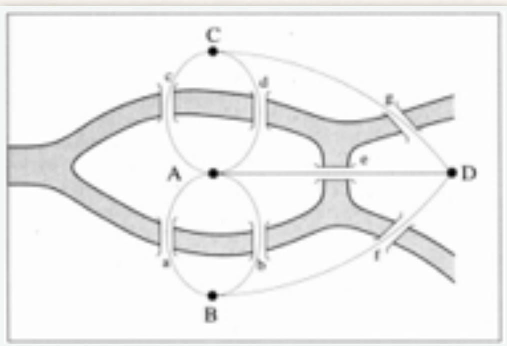
Redes complejas: 1. Elementos de Teoría de Gráficas

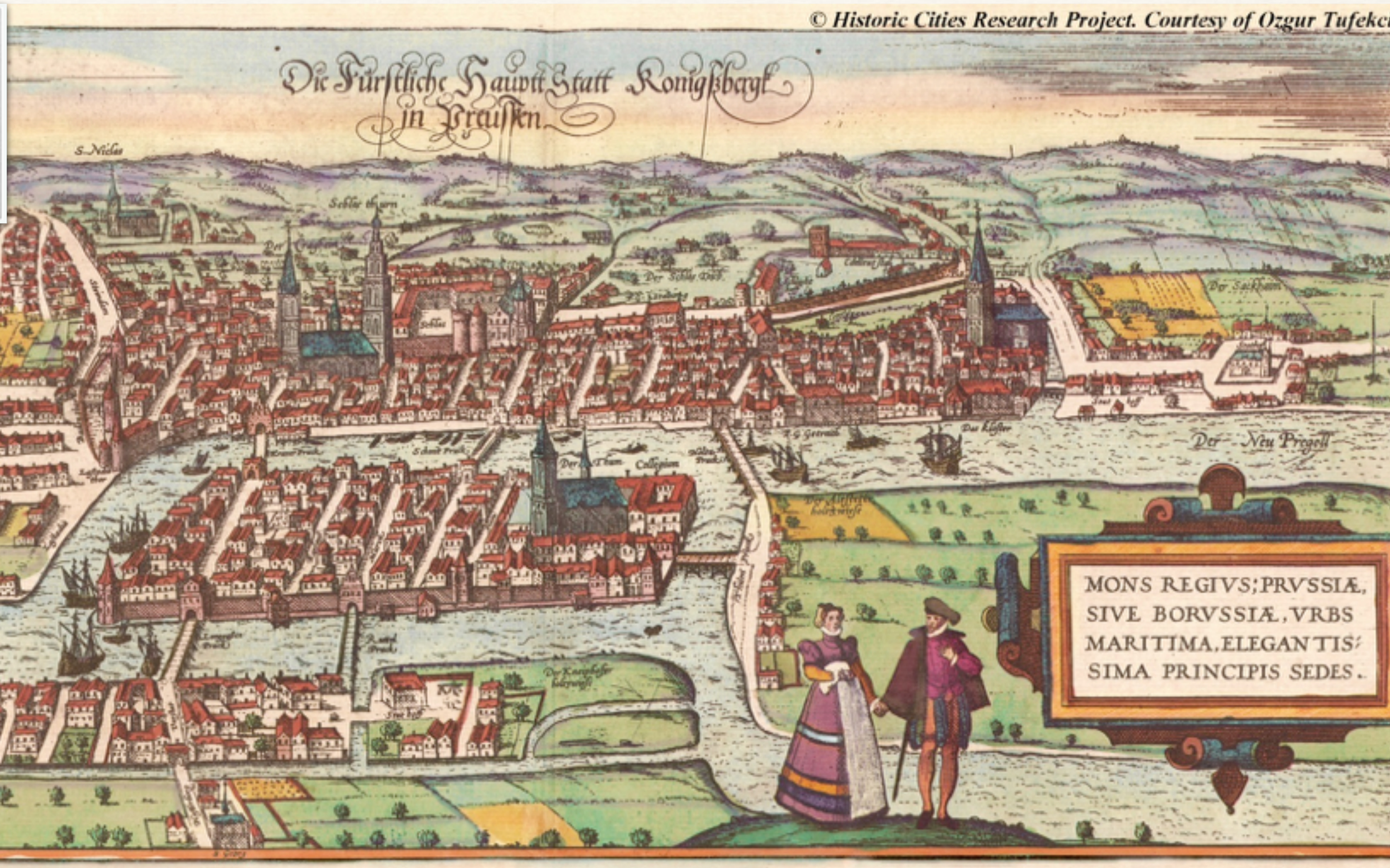
Dr. Felipe Contreras / UACM

Gráficas

- Se estudian desde el s.XVIII, con Euler y los puentes de Königsberg
- Es un campo muy activo de las matemáticas puras
- También es un área muy aplicada. Su aplicación a los sistemas complejos se le llama *redes*
- Aquí se han usado de diversas maneras, en especial, estadísticas. Aunque como herramienta, no la restringiremos a eso.

Königsberg y sus puentes



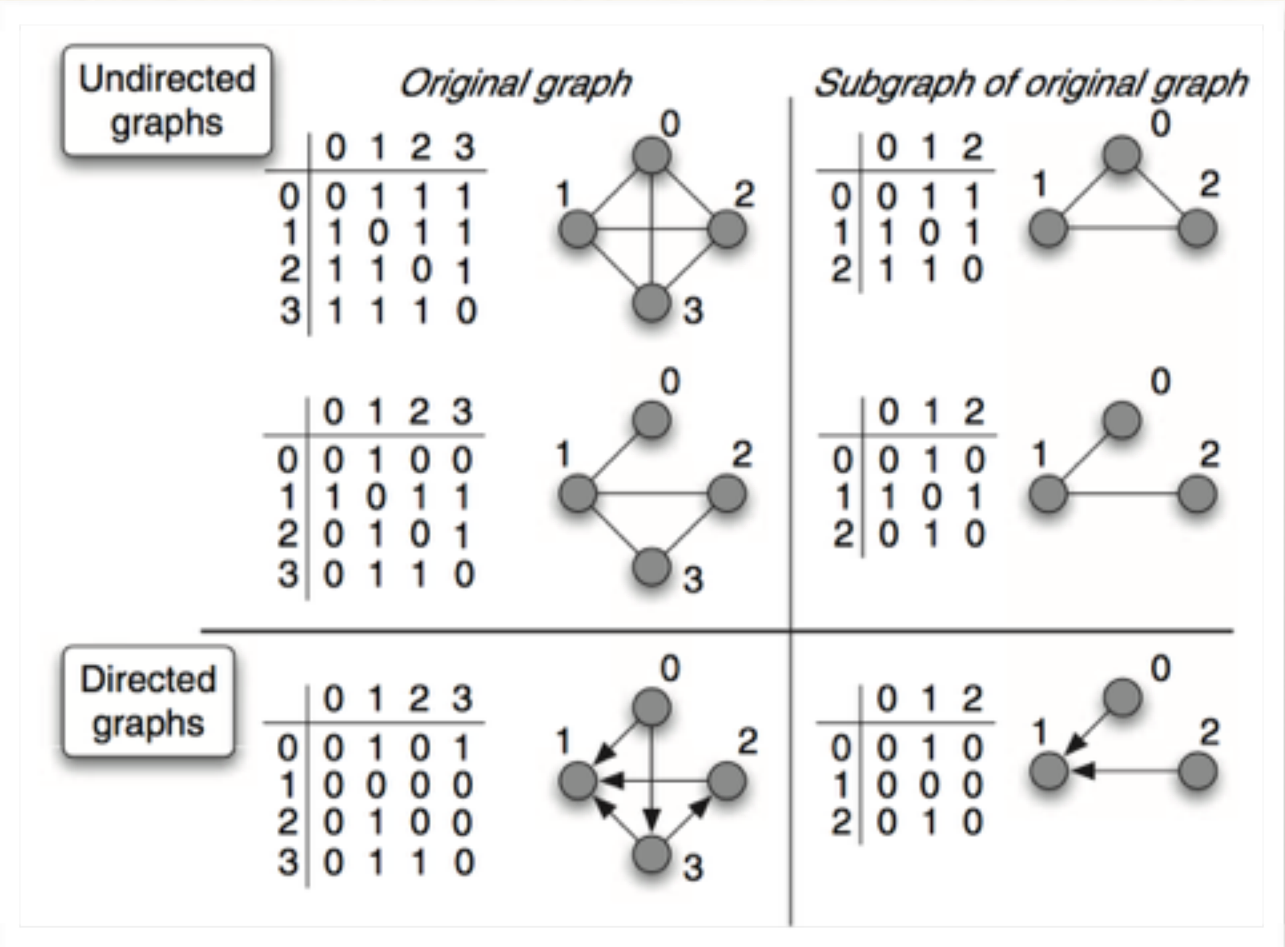


Problema: recorrer los 7 puentes exactamente una vez

Definiciones

- Una *gráfica* $G=(V, E)$ consiste del conjunto V de n elementos llamados *nodos* (o *vértices*), y el conjunto E de m parejas de nodos, llamadas *aristas*.
- Si la pareja (v_1, v_2) es un elemento de E, entonces los nodos v_1 y v_2 se dice que son *adyacentes* o *vecinos* uno del otro.
- Si nos interesa que $(v_1,v_2) \neq (v_2,v_1)$, entonces se dice que las aristas y G son *dirigidas*. A G también se le llama *digráfica*.
- G se puede representar como una *matriz de adyacencia* donde las entradas son

$$m_{i,j} = \begin{cases} 1 & \text{si existe la arista } (v_i,v_j) \\ 0 & \text{en caso contrario} \end{cases}$$



- En general (*gráfica simple*), no se permite que v sea vecino de sí mismo ni que entre dos nodos haya más de una arista (paralelas)
- ¿Hasta cuántas aristas *no dirigidas* puede tener una gráfica con n nodos?
- En España (y algunos mexicanos profanos) a las gráficas le llaman *grafos*.

- Una *subgráfica* $G'=(V', E')$ de G es tal que $V' \subseteq V$ y $E' \subseteq E$ *y si* (v_1', v_2') es arista de G', entonces v_1', v_2' son nodos de V'
- Si el número de aristas se parece al número de vértices a la alfa, se escribe: $|E| \sim |V|^{\alpha}$, entonces G se llama *rala* o *escasa* si $\alpha < 2$, y se llama *densa* si $|E| \sim |V|^2$.
- G también se puede representar como *lista de adyacencia.* Es decir, como la lista donde el elemento l_i es a su vez la lista de nodos $[v_1, v_2, \ldots]$ adyacentes a v_i.
- Si G es rala, o sea m es pequeña, comparada con el tamaño de su matriz de adyacencia, que es n^2; dicha matriz consiste en su mayoría de ceros; por lo que computacionalmente es más conveniente representar G como lista de adyacencia.
    - ¿de qué tamaño es una lista de adyacencia?

- Una *cadena* entre dos nodos a y b de G, es una subgráfica, cuyos nodos son $V'=\{a_i, i=1,2,\ldots,r+1 : a_1 = a, a_{r+1}=b\}$ y cuyas aristas son $E'=\{(a_i, a_{i+1}) \in E, i=1, 2, \ldots, r\}$. Al valor r se le llama la *longitud* de la cadena. (nota: al ser un conjunto, V' no tiene elementos repetidos)
- Se dice que G es *conexa* si para todo par de nodos, se puede encontrar una cadena entre ellos, de lo contrario se dice que G es *disconexa*.
- Un subconjunto S' del conjunto S es *maximal* bajo la propiedad p, si cumple p, pero $S' \cup \{e\}$ no cumple p, para cualquier $e \in S$.
- Una *componente conexa* $G'=(V', E')$ es una subgráfica de G, si V' y E' son maximales bajo la propiedad de que G' sea conexa.
- Así la *descomposición conexa* de G es la lista que consiste de sus componentes conexas.

  - ¿Cuál es la componente más grande de una gráfica conexa?

- De entre las cadenas entre los nodos *a* y *b*, la de menor longitud, aunque puedan haber varias, se llama *la ruta más corta*, y su longitud es llamada la distancia de *a* a *b*. Se escribe $d(a,b)$.
- Si *a* y *b* están en diferentes componentes conexas, se define $d(a,b)=\infty$.
- A la máxima distancia d_G entre cuales quiera dos nodos de una gráfica *G* se le llama *diámetro*.

  - ¿Cuánto es el diámetro de una gráfica disconexa?
  - $d(a,b)=d(b,a)$ si la gráfica es no dirigida. ¿Puedes dar un contra-ejemplo para cuando la gráfica es dirigida? ¿y si además es conexa?

- La *excentricidad* $ec(v)$ del nodo v es la máxima distancia $d(v,x)$, para todo nodo $x \neq v$.
- El *radio* de una gráfica rad_G es la mínima excentricidad entre sus nodos.
- La *distancia o longitud promedio se define como* $\langle l \rangle = \sum_{a,b}\{d(a,b)\}[n(n-1)]^{-1}$
- Si G' es la componente más grande y |G'|=Ω(n), a G' se le llama *componente gigante*.

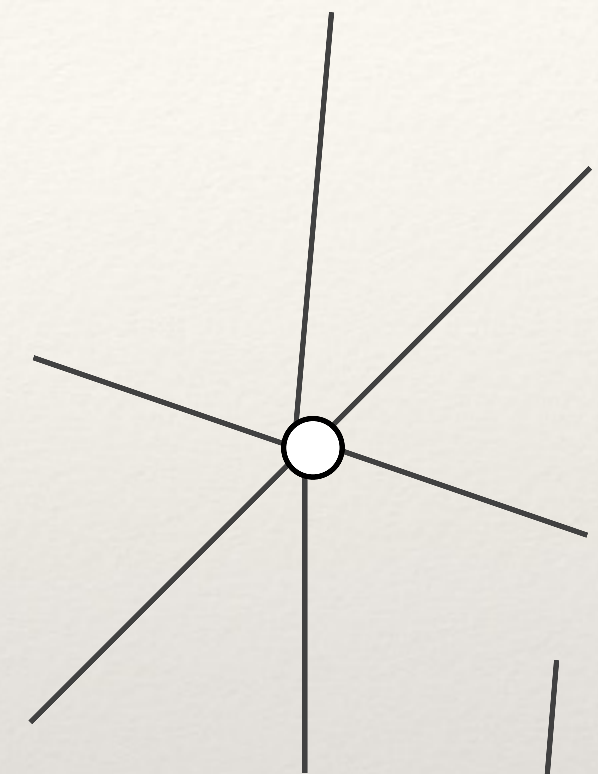
- demostrar $rad_G \leq d_G \leq 2rad_G$

Medidas de centralidad

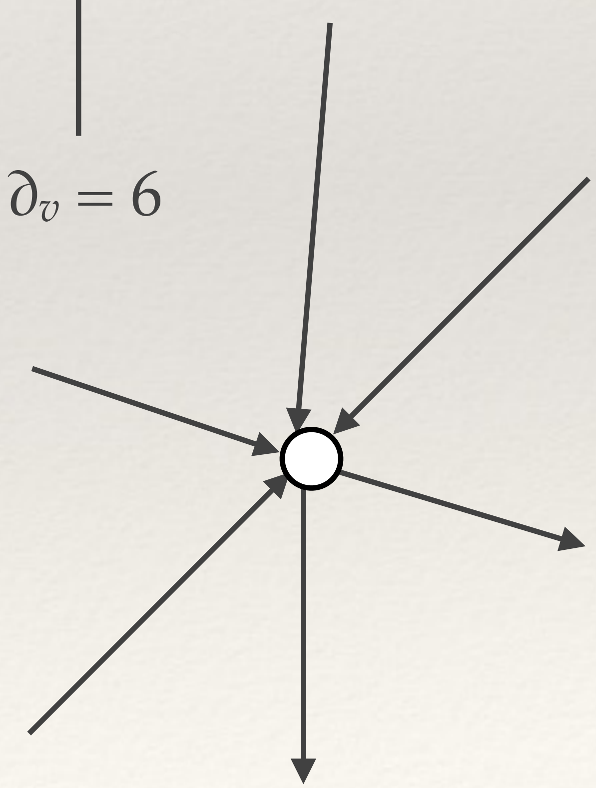
- El *grado* ∂_v del nodo v, es el número de aristas en las que participa.
  $\partial_v = |\{v \in e : e \in E\}|$

- Si G es dirigida, se distinguen los grados *exterior* ∂^{out}_v e *interior* ∂^{in}_v, como el número de aristas que tienen a v como su extremo final o inicial, respectivamente. Es decir:
  $\partial^{in}_v = |\{(a,v) \in E\}|$ y $\partial^{out}_v = |\{(v,a) \in E\}|$

- Se observa que: $\partial_v = \partial^{in}_v + \partial^{out}_v$



$\partial_v = 6$



$\partial^{in}_v = 4$, $\partial^{out}_v = 2$

- La *centralidad de cercanía* (*closeness*) es el promedio de las distancias de un nodo a los demás: $g_v = [\sum_{a \neq v}\{d(a,v)\}]^{-1}$
- Es más grande cuanto más distancias más cortas tenga con los demás nodos.
- La *centralidad de intermediación* (*betweenness*) dá el promedio de rutas más cortas que pasan por un nodo dado: Si $s_{a,b}$ es el número de rutas más cortas de a a b y $s_{a,b}(v)$ son aquellas que **además** pasan por v, entonces: $b_v = \sum_{a \neq v \neq b}\{s_{a,b}(v)\} / \sum_{a \neq b}\{s_{a,b}\}$

Agrupamiento

- Una *gráfica completa* es aquella en la que hay una arista entre cada par de nodos (par ordenado de nodos, en caso de ser dirigida)
- Un *clan* (*clique*) en una gráfica no dirigida *G*, es una subgráfica completa de *G*. Es decir, un subconjunto de sus nodos que están todos conectados entre sí.
- Las redes "naturales" tienden a formar clanes.
- El *coeficiente de agrupamiento*, mide el porcentaje de los vecinos de v que forman clanes con él, es decir, triángulos. Dicho de otra forma, es el porcentaje de los vecinos de v que son vecinos entre sí:
  $C(v) = e_v[\partial_v(\partial_v-1)/2]^{-1}$, e_v=número de aristas entre vecinos de v
- El *coeficiente de agrupamiento promedio* se define como $\langle C\rangle = \sum_v\{C(v)\}n^{-1}$
  - ¿Cuántas aristas tiene una gráfica completa?
  - ¿Qué valores pueden tomar: grado, centralidades de cercanía y de intermediación y el coeficiente de agrupamiento (promedio)? Dá ejemplos.