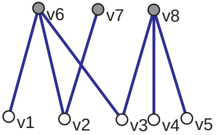
# 1.022 Introduction to Network Models

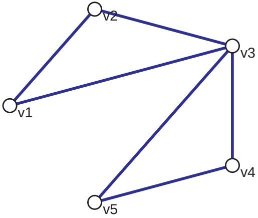
Amir Ajorlou

Laboratory for Information and Decision Systems
Institute for Data, Systems, and Society
Massachusetts Institute of Technology

Lecture 3

- A graph $G(V,E)$ is called bipartite when
  - $\Rightarrow$ $V$ can be partitioned in two disjoint sets, say $V_1$ and $V_2$; and
  - $\Rightarrow$ Each edge in $E$ has one endpoint in $V_1$, the other in $V_2$




- Useful to represent e.g., membership or affiliation networks
  - $\Rightarrow$ Nodes in $V_1$ could be people, nodes in $V_2$ clubs
  - $\Rightarrow$ Associated graph $G(V_1, E_1)$ joins members of same club

# Adjacency matrix

- Algebraic graph theory deals with matrix representations of graphs

  $\Rightarrow$ Leverage algebra to 'visualize' graphs as if being plotted

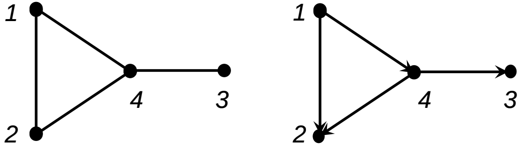
- Q: How can we capture the connectivity of $G(V, E)$ in a matrix?
- A: Binary, symmetric adjacency matrix $\mathbf{A} \in \{0,1\}^{|V| \times |V|}$, with entries

$$A_{ij} = \begin{cases} 1, & \text{if } (i,j) \in E \\ 0, & \text{otherwise} \end{cases} .$$

  $\Rightarrow$ Note that vertices are indexed with integers $1, \ldots, |V|$

- In words, $\mathbf{A}$ is one for those entries whose row-column indices denote vertices in $V$ joined by an edge in $E$, and is zero otherwise

- Examples for undirected graphs and digraphs




$$\mathbf{A}_u = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \quad \mathbf{A}_d = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

- If the graph is weighted, store the $(i, j)$ weight instead of 1

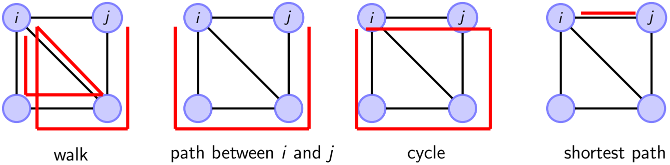
- Adjacency matrix useful to store graph structure.
  - $\Rightarrow$ Also, operations on $\mathbf{A}$ yield useful information about $G$
- Degrees: Row-wise sums give vertex degrees, i.e., $\sum_{j=1}^{|V|} A_{ij} = d_i$
- For digraphs $\mathbf{A}$ is not symmetric and row-, colum-wise sums differ

$$\sum_{j=1}^{|V|} A_{ij} = d_i^{out}, \qquad \sum_{i=1}^{|V|} A_{ij} = d_j^{in}$$

- Spectrum: $G$ is $d$-regular if and only if $\mathbf{1}$ is an eigenvector of $\mathbf{A}$, i.e.,

$$\mathbf{A}\mathbf{1} = d\mathbf{1}$$

# Walks, Paths, and Cycles




- Walks: Let $\mathbf{A}^r$ denote the $r$-th power of $\mathbf{A}$, with entries $A^r_{ij}$
- $[A^2]_{ij} := \sum_{k=1}^{n} A_{ik}A_{kj}$
- Corollary: $\mathrm{tr}(\mathbf{A}^2)/2 = |E|$ and $\mathrm{tr}(\mathbf{A}^3)/6 = \#\triangle$ in $G$

  $\Rightarrow$ You will prove this in your homework

# Incidence matrix

- A graph can be also represented by its $|V| \times |E|$ incidence matrix $\mathbf{B}$

  $\Rightarrow$ $\mathbf{B}$ is in general not a square matrix, unless $|V| = |E|$
- For undirected graphs, the entries of $\mathbf{B}$ are

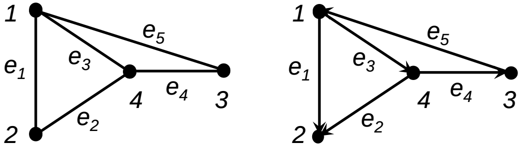
$$B_{ij} = \begin{cases} 1, & \text{if vertex } i \text{ incident to edge } j \\ 0, & \text{otherwise} \end{cases} .$$

- For digraphs we also encode the direction of the edge, namely

$$B_{ij} = \begin{cases} 1, & \text{if edge } j \text{ is } (k, i) \\ -1, & \text{if edge } j \text{ is } (i, k) \\ 0, & \text{otherwise} \end{cases} .$$

# Incidence matrix examples

- Examples for undirected graphs and digraphs

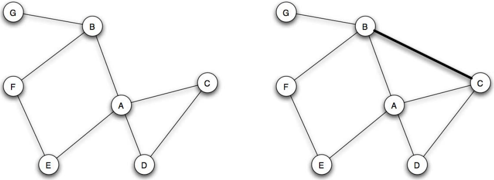



$$\mathbf{B}_u = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}, \quad \mathbf{B}_d = \begin{pmatrix} -1 & 0 & -1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & -1 & 1 & -1 & 0 \end{pmatrix}$$

- If the graph is weighted, modify nonzero entries accordingly

- Networks are rarely static structures $\Rightarrow$ Think about their evolution

  $\Rightarrow$ How are edges formed? $\Rightarrow$ Universal feature $\Rightarrow$ Triadic closure
- *If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends at some point in the future*




(a) *Before B-C edge forms.* (b) *After B-C edge forms.*

Easley, David, and Jon Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World.* Cambridge University Press, 2010. 

- Triadic closure is very natural $\Rightarrow$ Some reasons ...

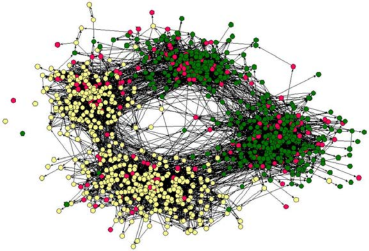
  $\Rightarrow$ Opportunity: B and C have a higher chance of meeting

  $\Rightarrow$ Trusting: B and C are predisposed to trusting each other

  $\Rightarrow$ Incentive: A might have incentive to make B and C friends

# Homophily

- We tend to be similar to our friends ⇒ Well known for long time
  - ⇒ Age, race, interests, beliefs, opinions, affluence, ...
- Contextual (as opposed to intrinsic) effect on network formation
  - ⇒ Contextual: Friends because we attend the same school
  - ⇒ Intrinsic: Friends because a common friend introduces us





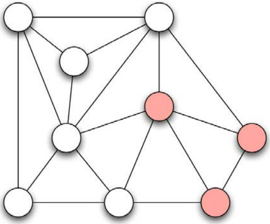
- In previous slide, B and C high chance of becoming friends
  - ⇒ Even if they are not aware of common knowledge of A

# Measuring homophily

- Is homophily present or is it an artifact of how the network is drawn?

  $\Rightarrow$ We need to formulate a precise mathematical measure

- Consider a small network of girls ($q = 3/9$) and boys ($p = 6/9$)



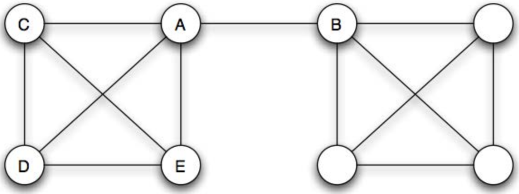
- If edges are agnostic to gender, portion of cross-gender edges is $2pq$

  $\Rightarrow$ Homophily Test: If the fraction of cross-gender edges is significantly less than 2pq, then there is evidence for homophily

  $\Rightarrow$ Cross-gender edges $5/18 < 8/18 = 2pq$ $\Rightarrow$ Mild homophily

# Hearing about a new job

- Mark Granovetter (1973) interviewed people that changed jobs
- Most heard about new jobs from acquaintances rather than close friends
  - ⇒ Explanation takes into account local properties and global structure

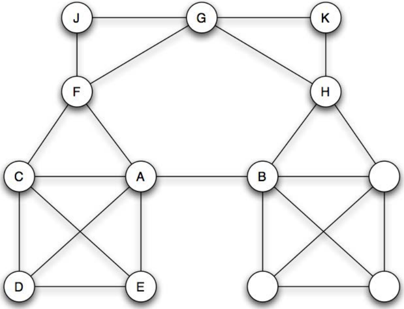





- $A$'s friends $E$, $C$, and $D$ form a tightly-knit group
- $B$ reaches to a different part of the network ⇒ New information
- Deleting $(A, B)$ disconnects the network ⇒ $(A, B)$ is a bridge
  - ⇒ But bridges are rare in real-world networks

# A social network closer to reality

- In real life, there are other multi-step paths joining $A$ and $B$
  - $\Rightarrow$ If $(A, B)$ is deleted, distance becomes more than 2 $\Rightarrow$ Local bridge
  - $\Rightarrow$ An edge is a local bridge when it is not part of a triangle

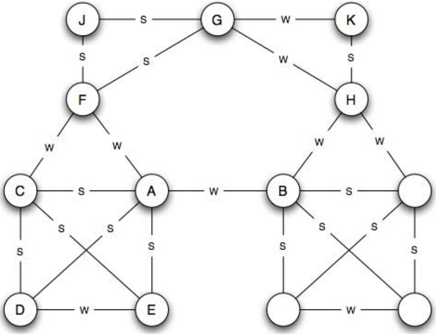



Easley, David, and Jon Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, 2010.


- Closely knit group of friends are eager to help
  - $\Rightarrow$ But have almost the same information as you

# Strong triadic closure

- How does overrepresentation of bridges relate to acquaintances?
- Consider two different levels of strength in the links of a social network
  - $\Rightarrow$ Strong ties correspond to friends, weak ties to acquaintances



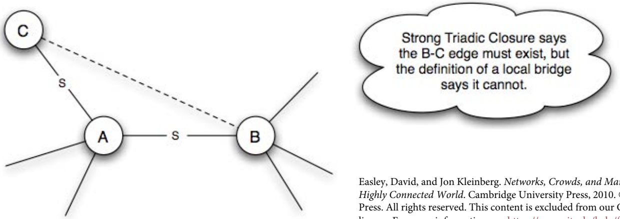



- $A$ violates the Strong Triadic Closure if it has strong ties to two other nodes $B$ and $C$, and there is no edge at all (strong or weak) between $B$ and $C$

# Local bridges and weak ties

- Tie strength ⇒ Local/interpersonal feature
- Bridge property ⇒ Global/structural feature
- How do these two features relate in light of the strong triadic closure?

- *If A satisfies the strong triadic closure and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie*






- Acquaintances are natural sources of new information

  ⇒ Strict modeling assumptions, first-order conclusions, testable