Social Network Analysis

SIG 2 Fan Zheng (Jayden)

CI 5330

Outline

1. Share
2. Introduction
3. Break-----10 min
4. Activity 2: Tools
5. Debrief & General Discussions

1. Share

Do you have any interesting or impressive experiences or stories about using social network or social network analysis?

http://elitedaily.com/life/culture/the-power-of-social-media-10-years-10-amazing-stories-made-possible-by-facebook/

2. Introduction

2.1 Basics

What is social network analysis?

2.1 Basics

“Social network analysis (SNA) is the process of investigating social structures through the use of network and graph theories. It characterizes networked structures in terms of nodes (individual actors, people, or things within the network) and the ties or edges (relationships or interactions) that connect them.”

----------From Wikipedia

2.1 Basics

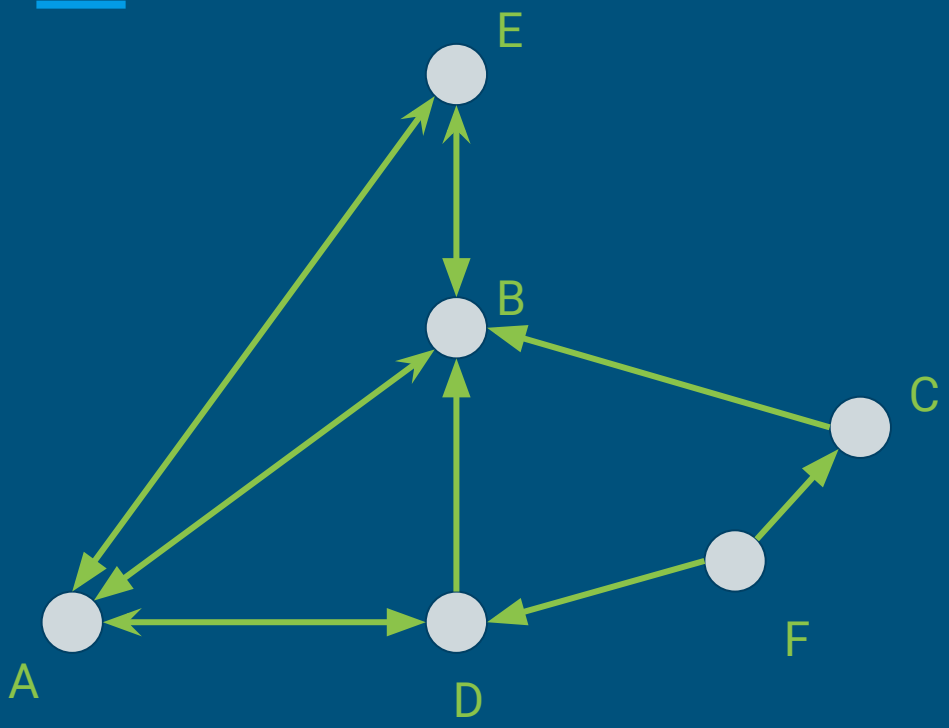
“Using social network techniques, these data can be used to indicate characteristics of positions held in a network and characteristics of the network structure. Positions in a network reveal who controls, facilitates, or inhibits the flow of information, and who has similar information needs or uses. Network structures reveal how information flows around the whole environment.” (p 17, *An Approach and Technique for the Study of Social Network Analysis: Information Exchange*, Caroline Haythornthwaite)

2.2. Activity 1: Discussions about readings

- What insights could SNA offer in the educational setting? And how (e.g., with which measures)?

- What other thoughts about the readings?

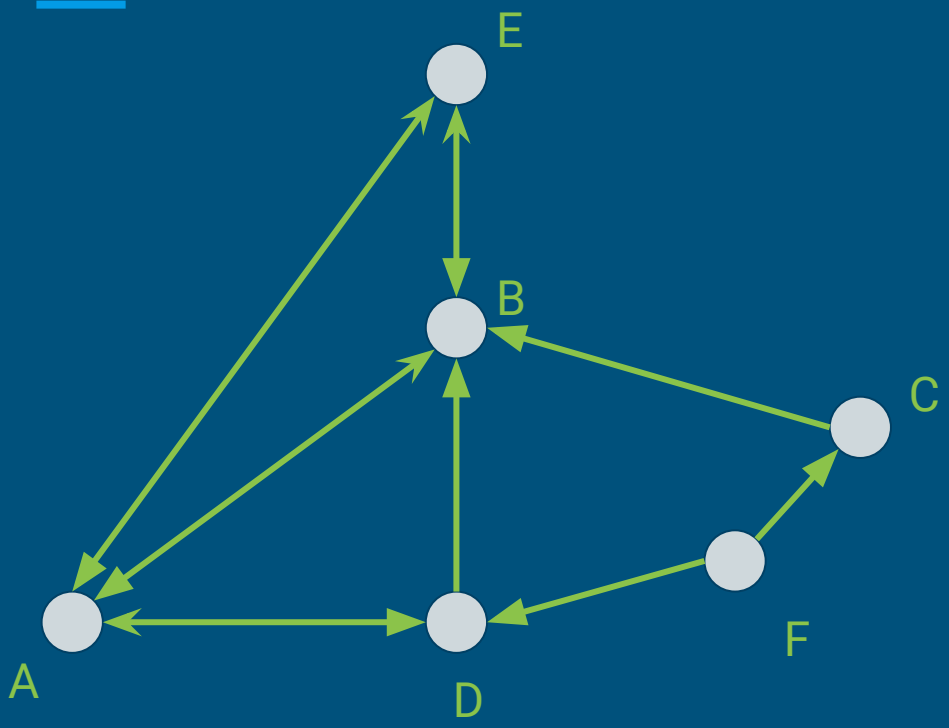
2.2 Terminology



1. Node: individual actors, people, or things within the network. Sometimes called a vertex.
2. Tie: a relationship between two nodes. Sometimes called a link, sometimes an edge.

- Directed Edge - A directed edge is an ordered pair of nodes that can be represented graphically as an arrow drawn between the nodes.
- Undirected Edge - An undirected edge disregards any sense of direction and treats both nodes interchangeably.

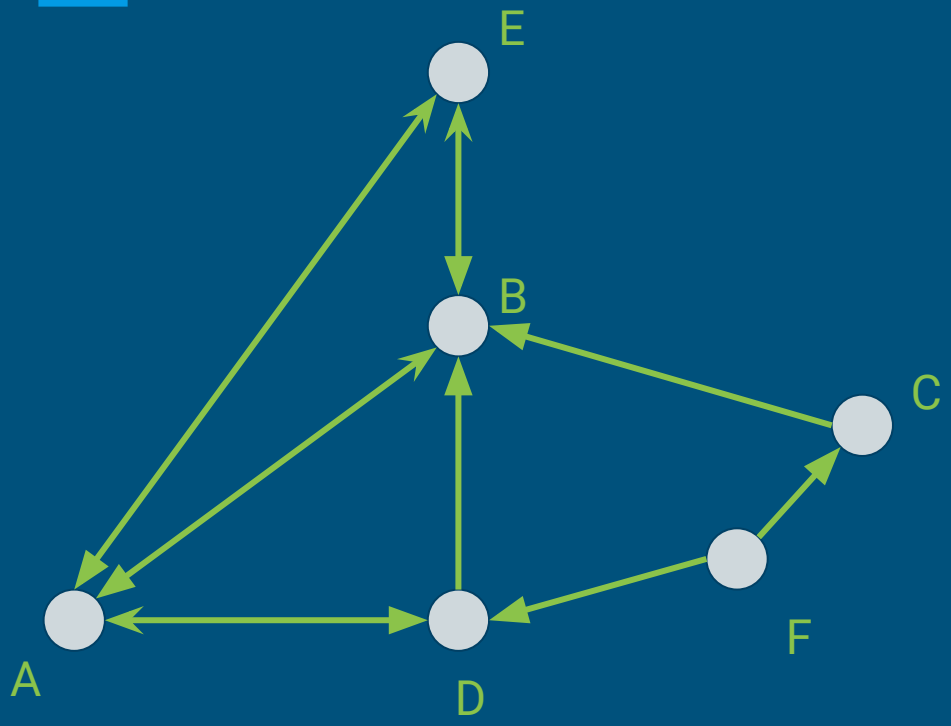
2.2 Terminology



1. The complete set of nodes and edges is often called the social graph, or simply graph.
2. Degree: The number of ties a node has is its degree, which can be distinguished between in-degree and out-degree.

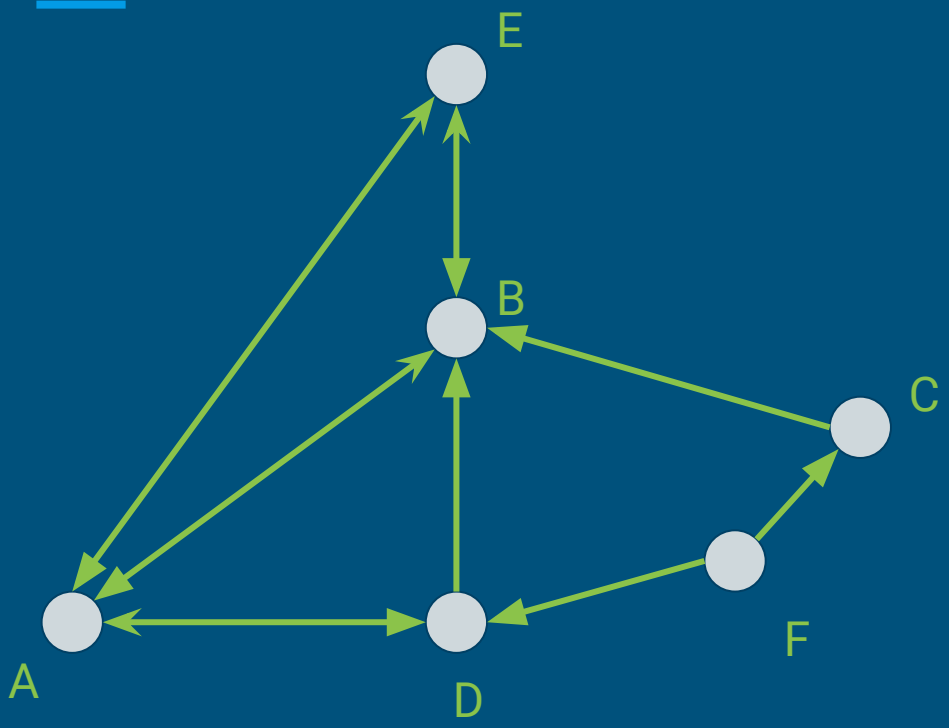
   Eg: Node B has an in-degree of 4. Node E has a out-degree of 2.
3. Distance: The distance between two nodes is defined as the number of edges along the shortest path connecting them.

2.2 Terminology



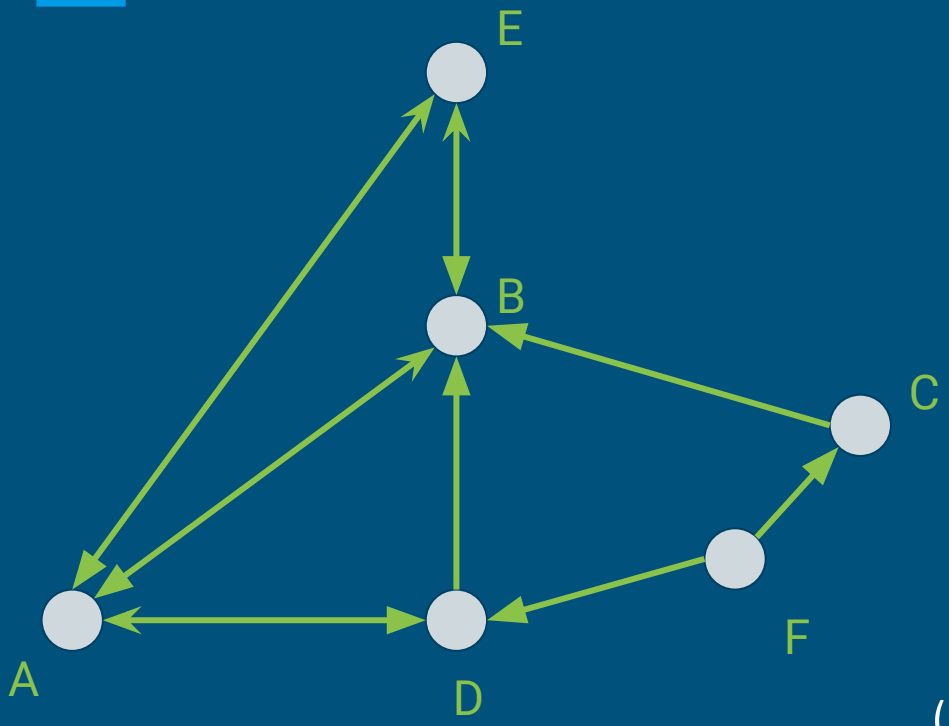
1. Centrality (Centrality measures address the question: "Who is the most important or central person in this network?" There are many answers to this question, depending on what we mean by importance.)

2.2 Terminology



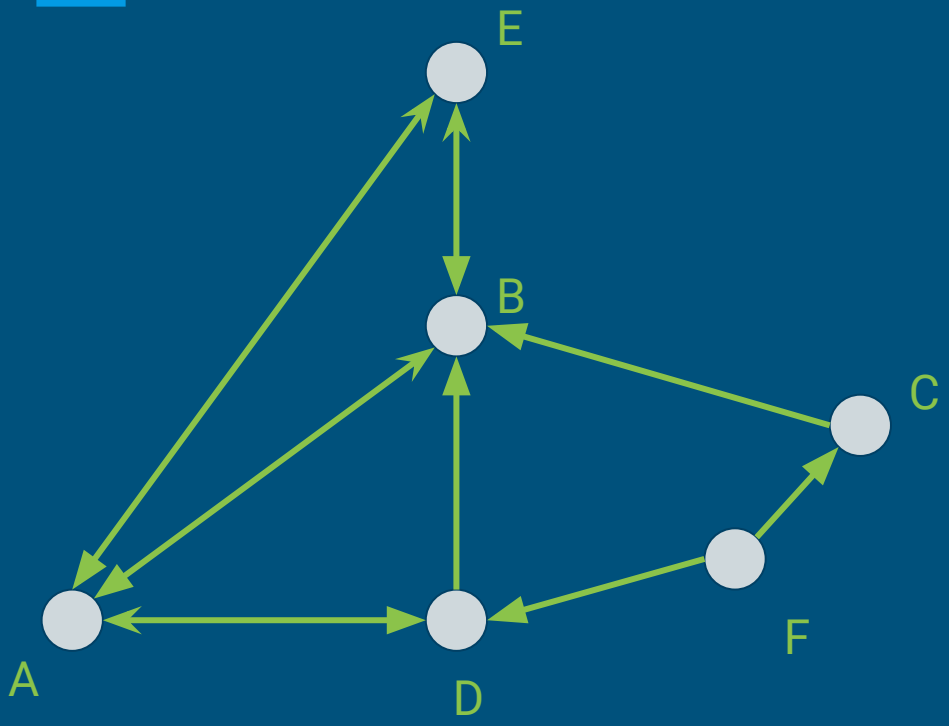
- Degree centrality(The nodes with higher degree is more central)
- Closeness centrality(the more central a node is the lower its total distance to all other nodes.)
- Betweenness centrality(Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes.)

2.2 Terminology



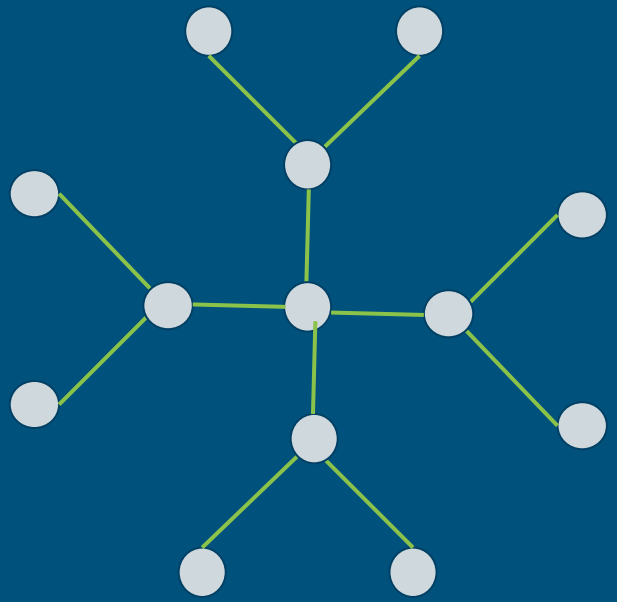
- PageRank centrality: There are three distinct factors that determine the PageRank of a node: (i) the number of links it receives, (ii) the link propensity of the linkers, and (iii) the centrality of the linkers. The first factor is not surprising: the more links a node attracts, the more important it is perceived. Reasonably, the value of the endorsement depreciates proportionally to the number of links given out by the endorsing node: links coming from parsimonious nodes are worthier than those emanated by spendthrift ones. Finally, not all nodes are created equal: links from important vertices are more valuable than those from obscure ones. (http://www.sci.unich.it/~francesc/teaching/network/pagerank)

2.2 Terminology



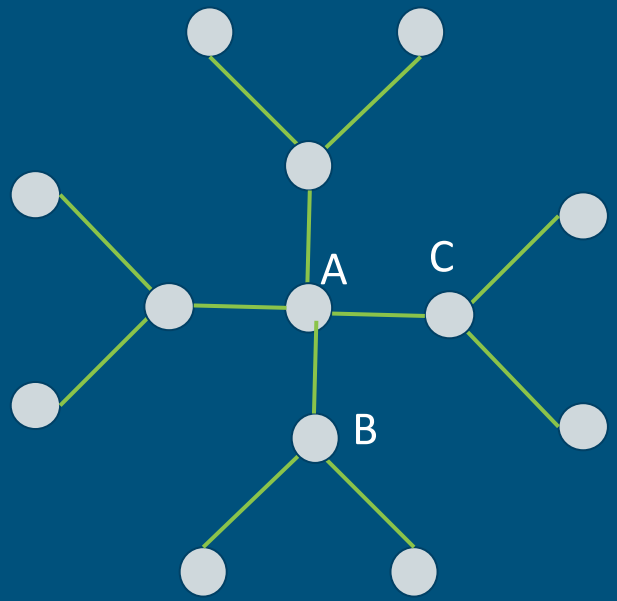
- Eigenvector centrality: Eigenvector centrality not only counts the number of nodes each node is connected to, but also weights these nodes according to their centrality. Eigenvector centrality is a measure of the quality of the nodes a node is connected to, rather than the quantity of connections of a node.(http://djjr-courses.wikidot.com/soc180:eigenvector-centrality)

2.2 Terminology



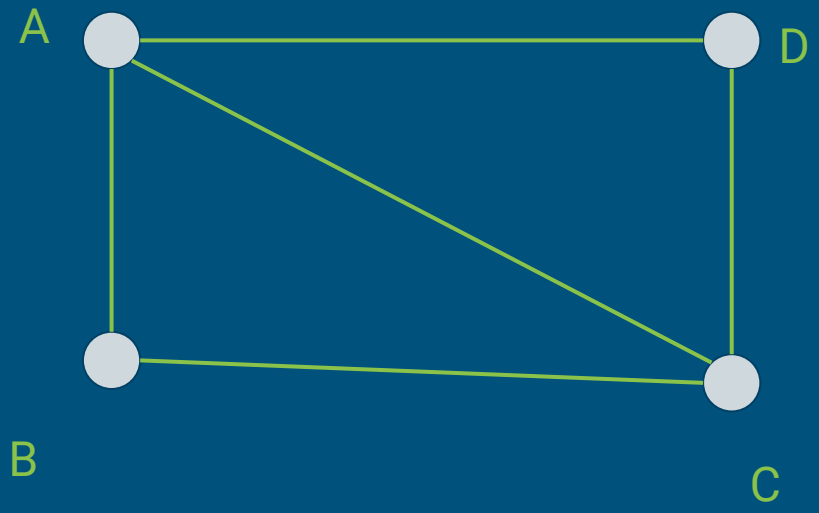
1. Centralization: Centralization measures the extent to which a set of actors are organized around a central point.It refers not to the relative prominence of points, but to the overall cohesion or integration of the graph. Measures of centralization can tell us whether a graph is organized around its most central points.

2.2 Terminology



In such a network, information that needs to pass from Node B to Node C can go through one intermediary. If the network had been arranged in a circle instead, there would have been many more intermediaries in the information distribution process,with an attendant possibility of adding error to the information exchange process.

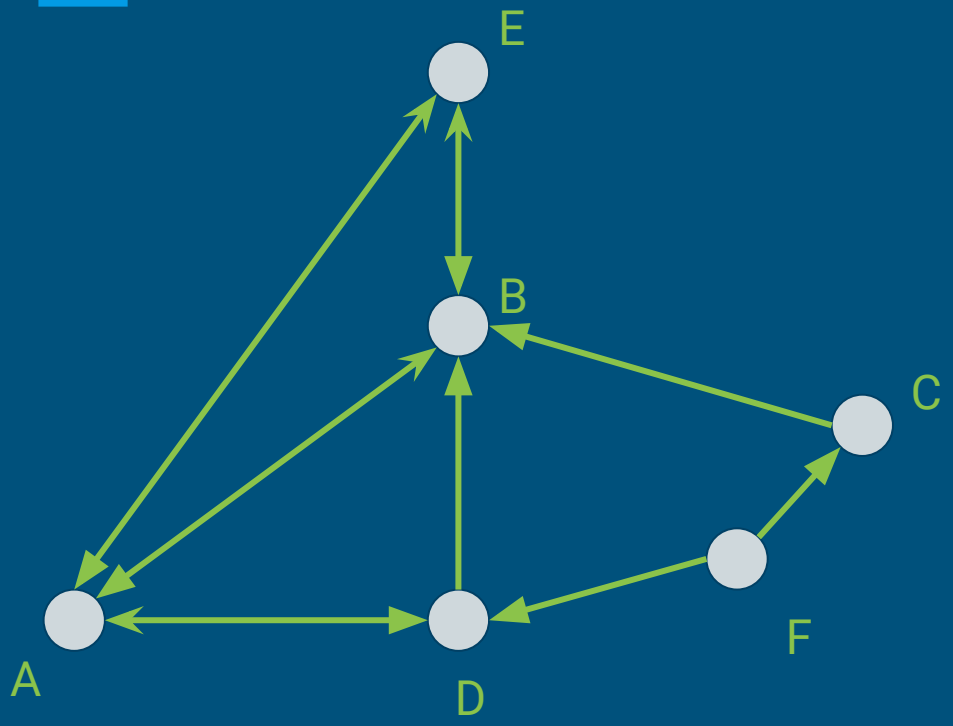
2.2 Terminology



D=5/6

1. Density: A network's density is the ratio of the number of edges in the network over the total number of possible edges between all pairs of nodes. Density is useful in comparing networks against each other.

2.2 Terminology



A new algorithm for clique-detection in a graph is introduced. The method rests on the so-called “decomposition of a graph into a chain of subgraphs”

http://www.sciencedirect.com/science/article/pii/0771050X75900352

2.3 Tools

- Gephi: visualization and basic network metrics
- iGraph: for Programming assignments
- NetLogo: modeling network dynamics
- NetworkX (Python): open source, extensive functionality
- Netlytic: cloud-based text and social networks analyzer
- NodeXL (Windows only): SNA integrated into Excel
- Pajek (Windows only)
- SNA and igraph packages in R
- SNA in Python
- SoNIA: social network image animator specialized for longitudinal analysis of networks
- UCINet (Windows only): sociology-focused functionality
- SNAPP: Social networks adapting pedagogical practice

2.4 Applications

1. Business use SNA to analyze and improve communication flow in their organization, or with their networks of partners and customers.
2. Law enforcement agencies(and the army) use SNA to identify criminal and terrorist networks from traces of communication that they collect; and then identify key players in these networks. https://www.harmari.com/2013/02/social-network-analysis-for-law-enforcement/
3. Social Network Sites like Facebook use basic elements of SNA to identify and recommend potential friends based on friends-of-friends.

2.4 Applications

http://www.slideshare.net/welrifai/social-network-analysis-big-data-telecommunications-and-more

4. BREAK

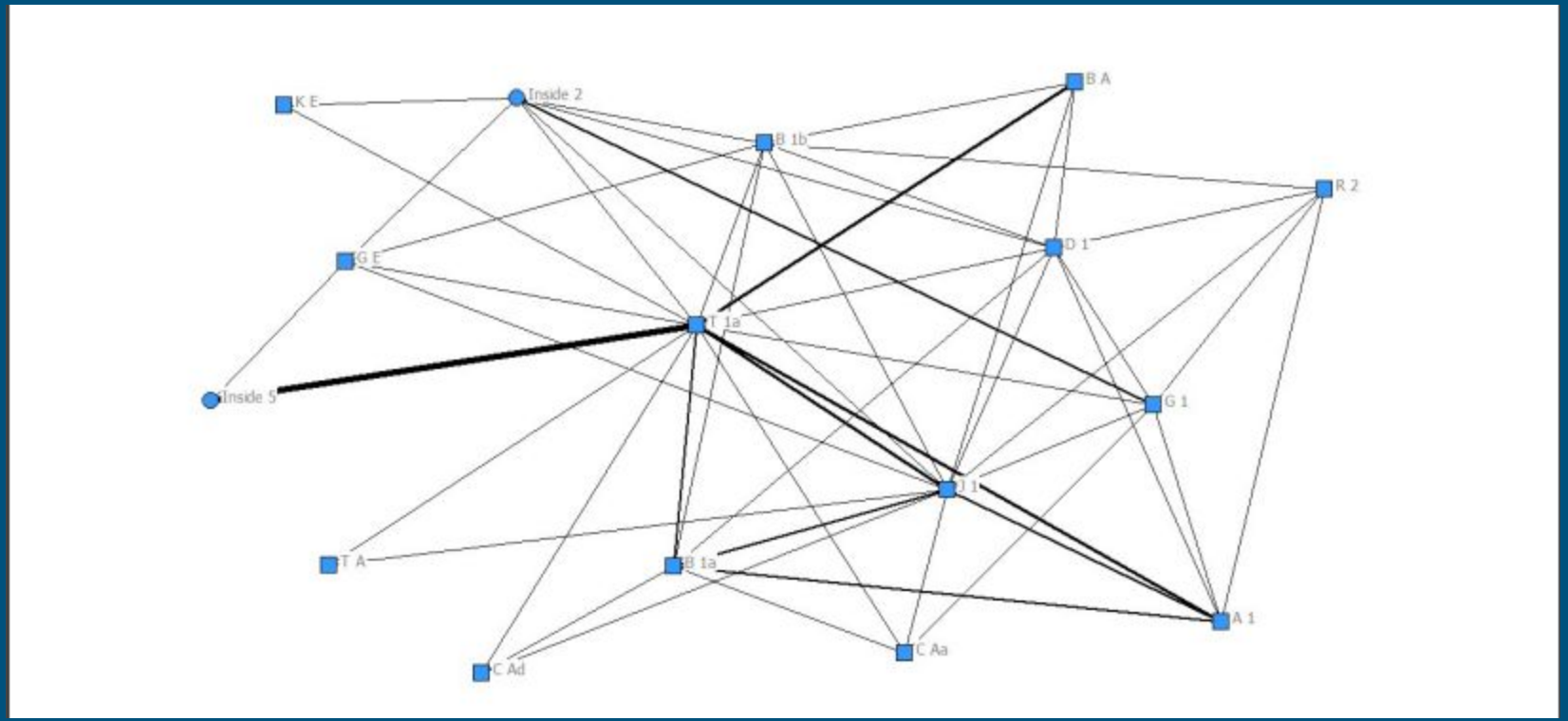
10 MINUTES BREAK!!!!!!!!!!!!!!!!!!!!

2.5 Example

Examining graduate committee faculty compositions- A social network analysis example

http://www.uky.edu/~kdbrad2/Kate.pdf

2.5 Example



5. Activity 2: TOOLS

Social Network Analysis tools: Gephi, Netlytic, and Condor

3 Groups

15 minutes

6. Debrief & General Discussions

- Tool features you like or dislike?
- One interesting finding to share?
- Potential use of the tool(s) you see?

Additional resources

- Coursera MOOC on Social Network Analysis: https://www.coursera.org/course/sna
- Blog: http://sna433.weebly.com/blog
- Videos: https://www.youtube.com/watch?v=VjOVhWfh6iI

Wrap up

1. Basics of social network analysis
2. Terminologies in social network analysis: Node, Tie, Graph. Degree, Distance, Centrality, Centralization, Density, and clique-detection.
3. Tools: Gephi, Netlytic, and Condor
4. Applications
5. Examples

Thank you !