Outline

- Network Science
  - A “Random” History

- Network Analysis
  - Network Topological Analysis: Random, Scale-Free, and Small-world Networks
  - Node level analysis
  - Link Analysis

- Network Visualization

- Network-based Business Intelligence Application
Network Science

Network science is an interdisciplinary academic field which studies complex networks such as information networks, biological networks, cognitive and semantic networks, and social networks. It draws on theories and methods including (Wiki)

- **Graph theory** from mathematics, e.g., Small-world
- **Statistical mechanics** from physics, e.g., Rich get richer,
- **Data mining** and **information visualization** from computer science,
- **Inferential modeling** from statistics, e.g., Collaborative filtering
- **Social structure** from sociology, e.g., weak tie, structural holes

Network science can be defined as "the study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena."
A “Random” History: Math, Psychology, Sociology…

- The study of networks has emerged in diverse disciplines as a means of analyzing complex relational data.

- Network science has its root in **Graph Theory**.
  - *Seven Bridges of Königsberg* written by *Leonhard Euler* in 1736.
  - Vertices, Edges, Nodes, Links,
  - a branch of mathematics that studies the properties of pairwise relations in a network structure

- **Social Network Analysis**
  - *Jacob Moreno*, a psychologist, developed the **Sociogram** and to “precisely describe the interpersonal structure of a group”.
  - Jacob’s experiment is the first to use Social Network Analysis and was published in the New York Times (April 3, 1933, page 17).
Jacob Moreno’s experiment on Friendship Network

EMOTIONS MAPPED BY NEW GEOGRAPHY
Charts Seek to Portray the Psychological Currents of Human Relationships.

New York Times
April 3, 1933
Now…

Complex Networks in the Real World

<table>
<thead>
<tr>
<th></th>
<th>Nodes</th>
<th>Links</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social network</td>
<td>People</td>
<td>Friendship, kinship, collaboration</td>
</tr>
<tr>
<td>Inter-organizational</td>
<td>Companies</td>
<td>Strategic alliance, buyer-seller relation, joint venture</td>
</tr>
<tr>
<td>network</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Citation network</td>
<td>Documents/authors</td>
<td>Citations</td>
</tr>
<tr>
<td>Internet</td>
<td>Routers/computers</td>
<td>Wire, cable</td>
</tr>
<tr>
<td>WWW</td>
<td>Web pages</td>
<td>hyperlink</td>
</tr>
<tr>
<td>Biochemical network</td>
<td>Genes/proteins</td>
<td>Regulatory effect</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>
Examples of Real-World Complex Networks

A collaboration network of physicists (size < 1K)
Source: (Newman & Girvan, 2004)

The Internet (size > 150K),
Source: Lumeta Corp.,
The Internet Mapping Project
Network Analysis: Topology Analysis

Network Topology Analysis takes a macro perspective to study the physical properties of network structures. Network topological measures include:

- **Size**,
- **Density**,  
- **Average Degree**,  
- **Average Path Length**: on average, the number of steps it takes to get from one member of the network to another.  
- **Diameter**  
- **Clustering Coefficient**: a measure of an "all-my-friends-know-each-other" property; small-world feature

\[
CC(i) = \frac{2E_i}{k_i(k_i-1)} \\
CC = \sum_{i=1}^{k_i} ClusteringCoeff(i)
\]

- \(k_i = C_d(i) = \# \text{ of neighbors of node } i\)
- \(E_i = \# \text{ of links actually exist between } k_i \text{ nodes}\)
Topology Analysis: Three Topology Models

- Random Network
- Erdős–Rényi Random Graph model
  - used for generating random graphs in which edges are set between nodes with equal probabilities
Topology Analysis: Three Topology Models

- Small-World Network
- Watts-Strogatz Small World model
  - used for generating graphs with small-world properties
  - large clustering coefficient
Topology Analysis: Three Topology Models

- Scale-Free Network
- Barabási–Albert (BA) Preferential Attachment model
  - A network model used to demonstrate a preferential attachment or a "rich-get-richer" effect.
  - An edge is most likely to attach to nodes with higher degrees.
  - Power-law degree distribution
## Network Analysis: Topology Analysis

<table>
<thead>
<tr>
<th>Topology</th>
<th>Average Path Length (L)</th>
<th>Clustering Coefficient (CC)</th>
<th>Degree Distribution (P(k))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Graph</td>
<td>( L_{\text{rand}} \approx \frac{\ln N}{\ln \langle k \rangle} )</td>
<td>( CC_{\text{rand}} = \frac{\langle k \rangle}{N} )</td>
<td>Poisson Dist.: ( P(k) \approx e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} )</td>
</tr>
<tr>
<td>Small World (Watts &amp; Strogatz, 1998)</td>
<td>( L_{\text{sw}} \leq L_{\text{rand}} )</td>
<td>( CC_{\text{sw}} \gg CC_{\text{rand}} )</td>
<td>Similar to random graph</td>
</tr>
<tr>
<td>Scale-Free network</td>
<td>( L_{\text{SF}} \leq L_{\text{rand}} )</td>
<td></td>
<td>Power-law Distribution: ( P(k) \sim k^{-\gamma} )</td>
</tr>
</tbody>
</table>

\( \langle k \rangle \): Average degree
Network Scientists

• Paul Erdős (Random graph model)
• Duncan Watts (Small-World model)
• A.-L. Barabási (Scale-Free model); “Linked”
• Mark Newman (SW and SF models)
Network Analysis: Node-level Analysis

- Node Centrality can be viewed as a measure of influence or importance of nodes in a network.

- Degree
  - the number of links that a node possesses in a network. In a directed network, one must differentiate between in-links and out-links by calculating in-degree and out-degree.

- Betweenness
  - the number of shortest paths in a network that traverse through that node.

- Closeness
  - the average distance that each node is from all other nodes in the network
Example: Centrality Measures of Bin Laden in a Global Terrorist Network

- The changes in the degree, betweenness and closeness of the node bin Laden from 1989 to 2002
Findings and Possible Explanations

The changes described in the above figure show that

- From 1994 to 1996, bin Laden’s betweenness decreased a lot and then increased until 2001
  - In 1994, The Saudi government revoked his citizenship and expelled him from the country
  - In 1995, he then went to Khartoum, Sudan, but under U.S. pressure was expelled Again
  - In 1996, bin Laden returned to Afghanistan established camps and refuge there

- From 1998 to 1999, there is another sharp decrease in betweenness
  - After 1998 bombings of the United States embassies around world, President Bill Clinton ordered a freeze on assets linked to bin Laden
  - Since then, bin Laden was officially listed as one of the FBI Ten Most Wanted Fugitives and FBI Most Wanted Terrorists
  - In August 1998, the U.S. military launched an assassination but failed to harm bin Laden but killed 19 other people
  - In 1999, United States convinced the United Nations to impose sanctions against Afghanistan in an attempt to force the Taliban to extradite him
Network Analysis: Link Analysis

- Link analysis focuses on the prediction of link formations between a pair of nodes based on various network factors. Its applications include:
  - Finance: Insurance fraud detections
  - E-commerce: recommendation systems, e.g., Amazon
  - Internet Search Engine: Google PageRank
  - Law Enforcement: Crime link predictions
Network Visualization: Expert Partition of the Collaboration Network

Weapons of massive destruction

Terrorism in Europe

Not well-defined group

Legal perspective of terrorism

Criminal justice

An international terrorism conf.

Rand Corp.

Historical and policy perspective of terrorism
Network-based Business Applications

- Facebook: People you may know