Representations of Network Data cont.

Matrix representation and visual layout

Information Visualization MOOC

Unit 6 – “With Whom”: Networks

Overview and Terminology

Reference:

http://ivmooc.cns.iu.edu
Network Analysis Examples and Goals

• Natural Networks: Neuronal, cell signaling, food webs
• Social Networks: Friendship, business, communication, collaboration networks
• Technological: Water networks, power grid, Internet, WWW

Importance of network thinking:
• Food webs might completely disassemble if just one species goes extinct.
• Weak ties and brokers are extremely important in professional networks.

Relevant Research Disciplines

The study of networks has a long tradition in
• graph theory and discrete mathematics
• sociology
• communication research
• bibliometrics/scientometrics
• webometrics/cybermetrics
• biology
• physics

Today, it is conducted in mathematics, statistics, physics, social network analysis, economics, information science, and computer science, etc.
Representations of Network Data

Matrices

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.015</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>250.5</td>
<td>0</td>
<td>-280</td>
<td>33.32</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

Adjacency Matrix

Lists of nodes & links

* Vertices 3
  1 "Doc1" 0.0 0.0 0.0 ic Green bc Brown
  2 "Doc2" 0.0 0.0 0.0 ic Green bc Brown
  3 "Doc3" 0.0 0.0 0.0 ic Green bc Brown
  *Arcs
  1 2 3 c Green
  2 3 5 c Black
  *Edges
  1 3 4 c Green

Network layout

When to use what kind of representation?

Representations of Network Data cont.

Matrix representation and visual layout

<table>
<thead>
<tr>
<th>Undirected Graphs</th>
<th>Original graph</th>
<th>Subgraph of original graph</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully connected</td>
<td>(a) 0 1 2 3 to 2 0 1 0</td>
<td>(b) 0 1 2 0 1 0</td>
</tr>
<tr>
<td></td>
<td>1 1 0 1 1</td>
<td>1 1 1 0 0</td>
</tr>
<tr>
<td></td>
<td>2 1 1 0 1</td>
<td>1 1 0 0 0</td>
</tr>
<tr>
<td></td>
<td>3 1 1 0 1</td>
<td>0 1 0 0 0</td>
</tr>
<tr>
<td>Partially Connected</td>
<td>(c) 0 1 2 3</td>
<td>(d) 0 1 2 0 1 0</td>
</tr>
<tr>
<td></td>
<td>0 1 0 1</td>
<td>1 1 0 1 0 0</td>
</tr>
<tr>
<td></td>
<td>1 1 0 1</td>
<td>1 0 1 0 1 0</td>
</tr>
<tr>
<td></td>
<td>2 0 1 0 1</td>
<td>0 0 1 0 0 1</td>
</tr>
<tr>
<td></td>
<td>3 1 1 0 1</td>
<td>2 0 0 0 0 0</td>
</tr>
<tr>
<td>Directed Graphs</td>
<td>(e) 0 1 2 3</td>
<td>(f) 0 1 2 0 1 0</td>
</tr>
<tr>
<td>Partially Connected</td>
<td>0 1 0 1 0</td>
<td>1 0 0 0 0 0</td>
</tr>
<tr>
<td></td>
<td>1 0 1 0 1</td>
<td>2 0 0 0 0 0</td>
</tr>
<tr>
<td></td>
<td>2 0 0 1 0</td>
<td>0 0 1 0 0 0</td>
</tr>
<tr>
<td></td>
<td>3 0 1 0 0</td>
<td>1 0 0 0 0 0</td>
</tr>
</tbody>
</table>
Terminology

Network (or graph) is composed of nodes (or vertices) and links (or edges).

Nodes can be
• **Isolated**: Unconnected to other nodes.
• **Labeled**: Have labels attributes—e.g., weights.

Edges can be
• **Undirected** (symmetric) or **directed** (non-symmetric).
• **Labeled**: Have labels attributes—e.g., weights.
• **Signed**: Be positive and negative (friend/foe, trust/distrust).

Terminology

Networks can be
• **Labeled**: Network contains labels (weights, attributes) on nodes and/or edges.
• **Temporal**: For each node/edge we know the time when it appeared in the network.
• **Undirected**: Relations between pairs of nodes are symmetric or **directed** (also called **digraph**) with directed links.
• **Multigraph**: Network has multiple edges between a pair of nodes.
• **Bipartite**: Nodes can be divided into two disjoint sets U, V such that every link connects a node in U to one in V.
• **Signed**: Edges can be positive and negative (friend/foe, trust/distrust).
Network Measurements

- Node and Link Properties
- Network Properties
- Statistical Properties
- Network Types

Reference:

Node and Link Properties

- **Isolated node**: Not connected to any other node.
- **Degree of a node**: Number of links connected to it.
- **Betweenness centrality of a node**: Number of shortest paths between pairs of nodes that pass through a given node.
- **Betweenness centrality of a link**: Number of shortest paths among all possible node pairs that pass through a given link.
- **Shortest path length**: Lowest number of links to be traversed to get from nodes $i$ to $j$. 
Network Properties

**Number of**
- Nodes
- Isolated nodes
- Edges
- Self-loops

**Diameter:** Longest of all shortest paths among all possible node pairs in a network—i.e., #links to be traversed to interconnect the most distant node pairs.

**Density:** Ratio of the number of edges in the network to the square of the total number of nodes.

Network Properties cont.

**Number of**
- *Strongly connected components:* There is a directed path from each node in the network to every other node.

- *Weakly connected components:* Maximal subgraph in which all pairs of vertices are reachable from one another—disregards link directions.
Network Properties cont.

- **Clustering coefficient**: Measures the average probability that two neighbors of the node $i$ are also connected.

![Diagram of network components]

Statistical Properties

- **Node degree distribution** $P(k)$ of an undirected graph is defined as the probability that any randomly chosen node has degree $k$.

Power law distribution $P(k)$ for a scale free network

![Graph with power law distribution]

- Very many nodes with very few links
- Some nodes with a few links
- Few hubs with large number of links
Network Properties and Network Types

- **Average clustering coefficient** measures the average probability that two neighbors of the node $i$ are also connected.

- **Average path length**: Average number of steps along the shortest paths for all possible pairs of network nodes.

Adopted from Watts and Strogatz, 1998

Network Types

**Small World**: Road maps, food chains, electric power grids, networks of brain neurons, telephone call graphs, and social influence networks.

**Scale-free**: World-Wide Web, the Internet, social networks, airline.