**Social Network Analysis**

Social network analysis uses networks and graph theory to get better insight into complex social structures.

**Examples of Social Network Analysis Use Cases**

**Transportation network**: To find different modes of transportation in a region and their connections. From this you can ask:

* What is the best way to get from point A to point B? Is there a bus route that gets me to point B in the shortest time? What if I take a taxi part of the way and then use a bus route?

**Search Engine Optimization:** To find which webpages link to other webpages. From this you can ask:

* Where does my webpage stand among other webpages? To which sites is it being linked? Among which webpages would I like to increase my connections?

**Social Media:** To find how users are connected on social media platforms? From this you can ask:

* Who are the most highly connected users in a social media platform? How are groups formed on social media and what are the connections between those groups?

Both of these use cases are utilized in AdTech to help advertisers understand their audience and the shortest path to reaching new audiences.

**Social Network Analysis Software**

**Gephi:** GUI interface; does not require coding

**NetworkX:** Python package

**Data Inputs for Building Network Graphs**

**Nodes:** Each vertex or node represents a person in the network

**Edges:** Edges connect the vertices of pairs of people who have connected in the network

Diagram

Description automatically generated

**Weight:** The size of a vertex encodes the number of times in which people connect in a given time period

**Signed network:** A graph that represents a positive or negative connection between each node pairs.

**Multi-Graph network:** A network in which node pairs have multiple edges

**Bi-Partite Graph:** A graph with two sets of nodes in which every single edge connects a node from one set to the other. i.e. a network of a team who has at least 1 fan in common

**Further Analysis for Visualizing Network Graphs**

**Directed or Undirected Edges:** Edges can either be directed i.e. person 1 reached out to person 2, or undirected i.e person 1 and person 2 connected (but we don’t know who reached out to who).

**Centrality Measures:** Metrics that quantify the influence of a node in the network. In an undirected network, this measures the degree of a node i.e. how many edges a node has. For a directed network, there are three different degree measures:

1. **Degree:** How many in-degree and out-degree edges a node has.
   1. **In-Degree:** Incoming connections to a particular node.
   2. **Out-degree:** Outgoing connections from a particular node to others.
2. **Closeness:** How well connected a node is to all other nodes in the network. How many nodes does one particular node have to go through to reach another particular node. Example: In a hierarchical organization, how many people would I have to go through to get a message to the CEO?
3. **Betweenness:** The number of shortest paths the node is included in relative to the total number of shortest paths. Example: In a hierarchical organization, someone who has direct access to the CEO but also the most connections to individuals in the organization would have the highest betweenness centrality.

**Degree Distribution:** The number of unique connections a specific user has within the network; essentially a popularity metric showing a user that connects with many others in the network.

Analysis: Is number of followers and degree distribution correlated?

**Weighted Degree Distribution:** The number of unique connections and the weight of those connection (how many times 2 people have connected) a specific user has within the network.

**Network Modularity:** How easily the network can be split into smaller, more well-connected, communities.

“It is generally thought that a high value for the modularity means that the network has a complex underlying community structure. These communities can sometimes have signiﬁcant meaning in the network.” - <https://medium.com/@mehulved1503/network-analysis-an-application-case-the-simpsons-season-4-6dec557fa4e0>

**Analysis:**

* Top Co-Appearances
* In-Network Co-appearances vs out-network
* Top 20 people based on PageRank score