SOCIAL NETWORKS

February 17 2015

Lecture # 2

Topics to be covered:

1) Graph/Network Theory
   (Linear Algebra, Graph Laplacian & Ranking)
2) Data Science
   (Ranks, Clusters: Cliques, Clans, Clubs)
3) Game Theory
   (Signaling Games, Privacy, Security & Ethics)
4) Data Sciences
   (Big Data, Statistical Inference)
5) Distributed Computing
   (Infrastructure, Implementation)
6) Internet Economics
   (Auction, Pricing, Payment System)

Graph Theory:

a) Combinatorial Structure

b) Algebraic Structure ~ Spectral Properties

c) Probabilistic Structure ~
Random Graphs & Their Evolution.

Social Interactions (Pair-wise)

- Graph Theory (Interaction Choices)
- Game Theory (Strategic Choices)

We wish to model Strategic Interactions among Rational Agents.

Ingredients:

$V \equiv$ Set of Actors

$E \subseteq V \times V \equiv$ Set of Links

$S_v \equiv$ Strategy space $v \in V$

$u_v : \prod_{v \in V} S_v \to \mathbb{R}_+ \equiv$ Pay-off functions of $v \in V$

$\left(V, E, S_v\big|_{v \in V}, u_v\big|_{v \in V}\right)$

determine

A Social Network and how its agents behave.

Defn GRAPHS (Networks)

A graph $G = (V, E)$ consists of a set of vertices V together with a set of edges $E \subseteq V \times V$.

$\Rightarrow$ A mathematical object describing an irreflexive, symmetric binary relation on a discrete set, which need not be finite.

Example: Friendship.

IRREFLEXIVE : One is not his own friend.
$\langle v, v \rangle \notin E$ (no self-loop)

SYMMETRIC : One is a friend to a friend.
$\langle v, w \rangle \in E \Leftrightarrow \langle w, v \rangle \in E$

NON TRANSITIVE : One is not necessarily a friend to a friend's friend.
$\langle u, v \rangle \in E \wedge \langle v, w \rangle \in E \not\Rightarrow \langle u, w \rangle \in E$.

Friendship relation in a Social Network can be described by an undirected graph.

- An edge $e = (u, v) \in E$ (where $E \subseteq V \times V$) is described by the unordered pair of vertices (players), which serve as edge's end-points.

NOTE BOOK

- Two vertices ~~are~~ u and v are adjacent if $\exists_e\ e=(u,v)$ connecting u and v.

- Notation: $|V| = n = \#$ vertices
  $|E| = m = \#$ edges.

$$m \le \binom{n}{2} = \frac{n(n-1)}{2} \quad \begin{cases} n = \#\text{ ways to choose } u \\ n-1 = \#\text{ ways to choose } v \\ \text{identifying edge} \\ \quad (u,v) \equiv (v,u)\ [\text{Symm}] \end{cases}$$

STRICT GRAPH:

No self-loop $\quad (u,u) \notin E \quad \forall_u$

No multi-edge $\quad e_1 = (u_1, v_1) \neq e_2 (u_2, v_2)$

$$\Rightarrow (u_1 = u_2) \Rightarrow (v_1 \neq v_2)$$
$$\wedge\ (u_1 = v_2) \Rightarrow (v_1 \neq u_2)$$
$$\forall e_1, e_2$$

Two vertices u, and v are <u>adjacent</u>

$$\exists_e\ e = (u,v)$$

Two edges e and f are <u>incident</u>

$$\exists_u\ e = (u,v) \wedge f = (u,w)$$

NOTEBOOK

$$d(v) = |\{u \mid (u,v) \in E\}| = \text{Degree.}$$

The number of vertices adjacent to a given vertex v is called the degree of the vertex.

$$\sum_{v \in V} d(v) = 2|E| = 2m$$

Average degree of the graph

$$\bar{d} = \frac{\sum_{v \in V} d(v)}{V} = \frac{2m}{n}$$

Density = Ratio of the number of edges to the number of possible total

$$= \frac{|E|}{\binom{n}{2}} = \frac{2m}{n(n-1)} = \frac{\sum d(v)}{n(n-1)} = \frac{\bar{d}}{n-1}$$

Density = 1 $\Rightarrow$ Graph is $\boxed{\text{complete.}}$

$$\bar{d} = (n-1) \qquad \forall v \;\; d(v) \leq n-1$$

$$\Rightarrow \forall_v \; d(v) = n-1$$

A graph is complete if all of its vertices are adjacent to all others.

- If a social network has many "well-connected individuals" then the network is "dense," since

  density $= \bar{d}/n-1$

  $\Rightarrow$ Finding and connecting to "well-connected" subnetworks is beneficial.

  CLIQUE CLUB

  CLAN.

Distance between two individuals:

Path: A sequence of adjacent distinct vertices

$$v_0, v_1, \ldots, v_n$$

$$\forall_i \quad (v_i, v_{i+1}) \in E, \qquad 0 \le i < n$$

$$\forall_{i,j} \quad v_i \neq v_j \qquad 0 \le i \neq j \le n$$

in which no vertex occurs more than once.

Path $\le$ Trail $\le$ Walk

Vertices are distinct — Edges are distinct — Neither.

- The (geodesic)-distance between two vertices is the length of the shortest path connecting them.

  $d(u,v)$ = Geodesic distance between u and v.

◇ The maximal geodesic-distance in a graph is its <u>diameter</u>, $\mathcal{D}(G)$

$$\forall_{u,v} \quad d(u,v) \leq \mathcal{D}(G).$$

$$G_1(V_1, E_1) \subseteq G(V,E) \quad (\text{Subgraph})$$
$$\text{iff} \quad V_1 \subseteq V \wedge E_1 \subseteq E.$$

◇ A subgraph of a graph G is a graph whose vertices and edges are contained in G.

$$G_1 \subseteq G, \quad |V_1| = k, \quad |E_1| = \binom{k}{2}$$

◇ A clique is a maximal complete subgraph.

◇ The subgraph $G(S)$ of a graph G induced by the set of nodes S is defined as the maximal subgraph of G that has vertex set S.

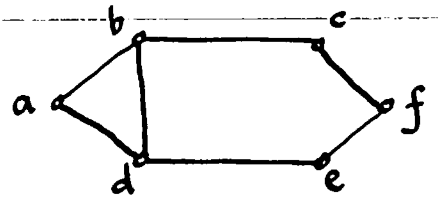
COHESIVE SUBNETWORKS

K-clique: A k-clique S of a graph is a maximal set of nodes in which
$\forall u, v \in S \quad d(u,v) \leq k$.
1. Clique = Clique.

K-Club: A k-club is a subset S of nodes such that in the subgraph $G(S)$ induced by S, the diameter is k or less.
$\mathcal{D}(G(S)) = k$

K-Clan $\equiv$ k-Clique $\cap$ k-club.

A k-clan is a k-clique in which all pairs of vertices are at distances less than or equal to k,
even when the geodesics/paths are restricted to involve only members of S.



{a, b, d} = 1-clique
{a, b, c, d, e} = 2-clique.
Note d(c, e) = 2
(through f).
{b, c, d, e, f} = 2-clique.
{a, b, c, d} ≠ 2-clique
(not maximal)

{b, c, d, e, f} = 2-clan
{a, b, c, d, e} ≠ 2-clan

⇒ k-clique $\not\supseteq$ k-clan.

{a, b, c, d} = 2-club
{a, b, c, d, e} ≠ 2-club

⇒ k-clique ≠ k-club

$$\text{k-clique} \rightarrow \text{k-clan} \leftarrow \text{k-club}$$

LS-Sets:

Let H be a set of nodes in $G=(V,E)$ and let K be a proper subset of H.

Let $\alpha(K)$ denote the number of edges linking members of K to $V \setminus K$.

H = An LS-set of G if

$$\forall_{K \subsetneq H} \quad \alpha(K) > \alpha(H).$$

⇒ Members of H have more ties with the "insiders" (other members of H) than "outsiders" (members of $V \setminus H$).

◇◇◇