AMRITA
VISHWA VIDYAPEETHAM
DEEMED TO BE UNIVERSITY

19CSE337 Social Networking Security

Lecture 8

Topics to Discuss

- Betweenness Centrality

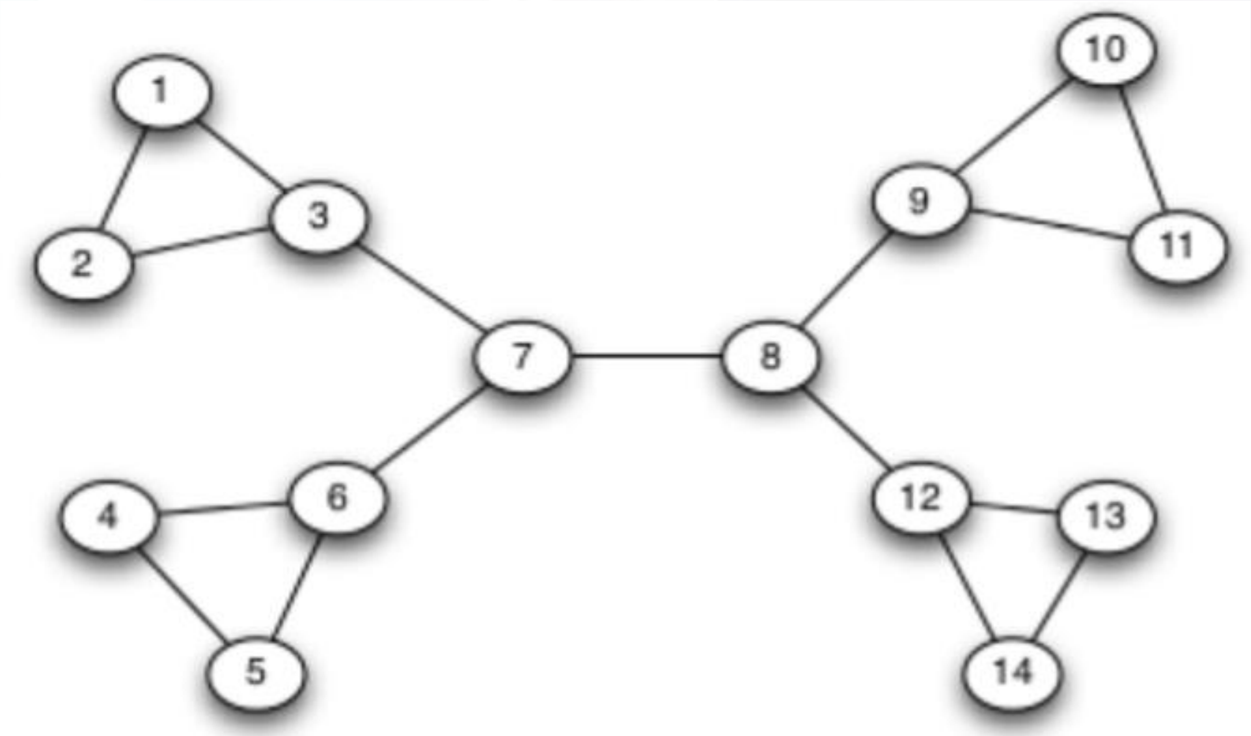
Betweenness Centrality

- Betweenness centrality measures the number of times a node lies on the shortest path between other nodes.
- It is a measure of how often a node appears in the shortest path connecting two other nodes.
- Nodes with a high betweenness centrality score are the ones that most frequently act as ‘bridges’ between other nodes.
- They form the shortest pathways of communication within the network.
- Usually this would indicate important gatekeepers of information between groups.

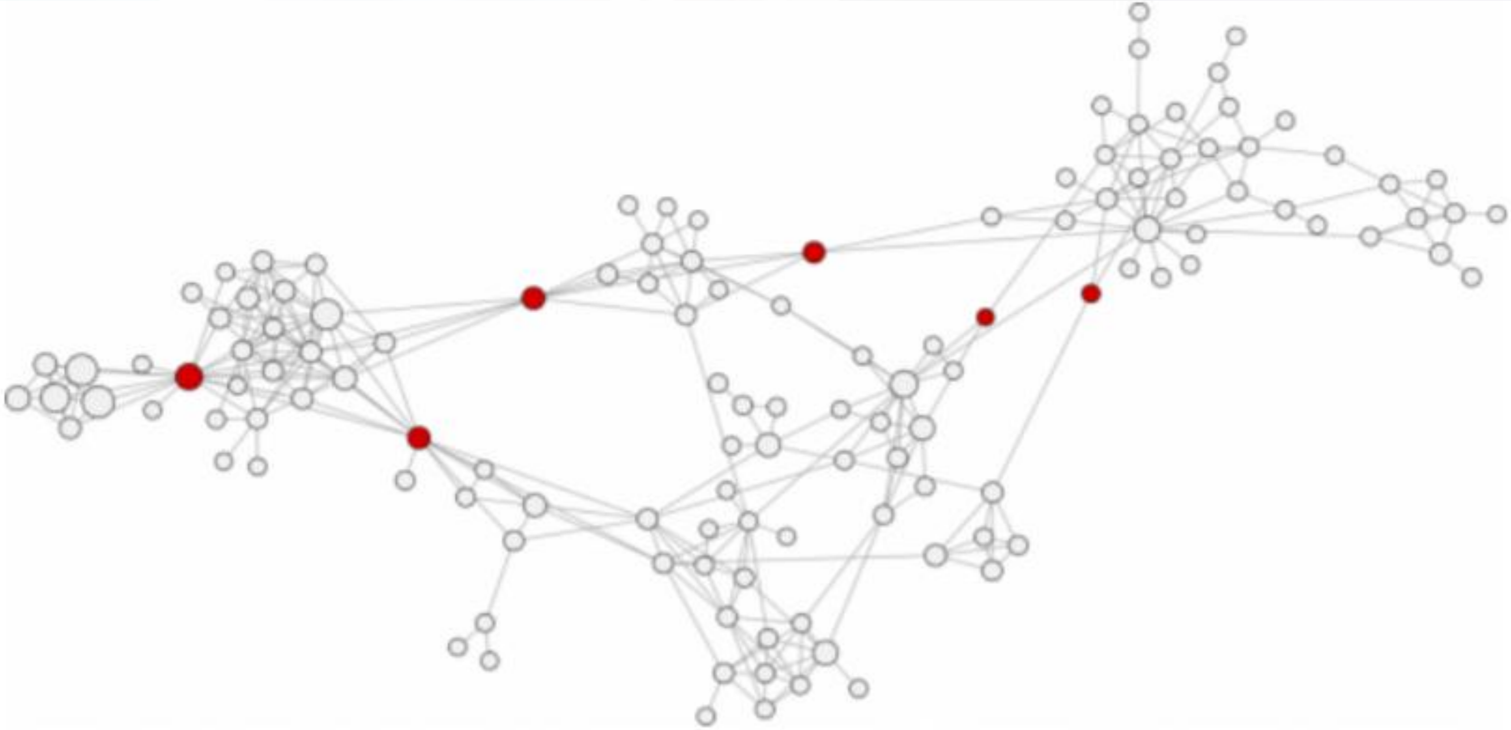
Betweenness Centrality

- Useful in finding the individuals who influence the flow around a system.
- A high betweenness count could indicate someone holds authority over disparate clusters in a network, or just that they are on the periphery of both clusters.

Betweenness Centrality

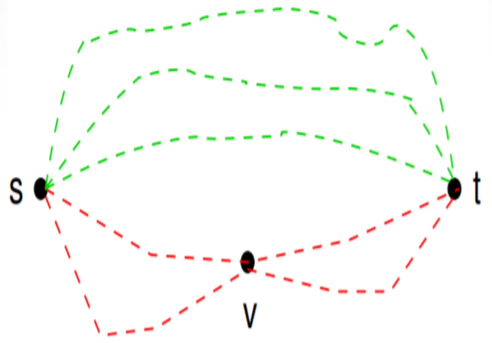


Betweenness Centrality



Betweenness Computation

- The betweenness centrality of a node V is defined as the proportion of shortest paths between all pairs of nodes that go through V.
- Consider a node v and two other nodes s and t.
- Each shortest path between s and t shown in green doesn't pass through node v.
- Each shortest path between s and t shown in red passes through node v.



Betweenness Computation

- Any shortest path between nodes s and t will be called an s-t shortest path.
- Let σ_{st} denote the number of all s-t shortest paths.
- Let $\sigma_{st}(v)$ denote the number of all s-t shortest paths that pass through node v.
- Consider the ratio $(\sigma_{st}(v) / \sigma_{st})$:
  - This gives the fraction of s-t shortest paths passing through v.
  - The larger the ratio, the more important v is with respect to the pair of nodes s and t.
  - To properly measure the importance of a node v, we need to consider all pairs of nodes (not involving v).

Betweenness Computation

- The betweenness centrality of a node v, denoted by β(v), is defined by

$$\beta(v) = \sum_{\substack{s\ ,\ t \\ s \neq v\ ,\ t \neq v}} \left[\frac{\sigma_{st}(v)}{\sigma_{st}} \right]$$

- Note: For two nodes x and y, if β(x) > β(y), then x is more central than y.

Betweenness Computation

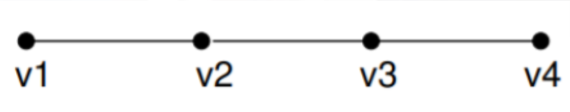
- Suppose we want to compute $\beta(v)$ for some node v. The formula suggests the following steps.
  - Set $\beta(v) = 0$.
  - For each pair of nodes s and t such that $s \neq v$ and $t \neq v$,
    - Compute σ_{st} and $\sigma_{st}(v)$.
    - Set $\beta(v) = \beta(v) + \sigma_{st}(v)/\sigma_{st}$.
  - Output $\beta(v)$.

Betweenness Computation

- Note: Here, there is only one path between any pair of nodes. (So, that path is also the shortest path.)

- Consider the computation of β(v2) first.
- The s-t pairs to be considered are: (v1, v3), (v1, v4) and (v3, v4).
- For the pair (v1, v3):
  - The number of shortest paths between v1 and v3 is 1; thus, $\sigma_{v1,v3} = 1$.
  - The (only) path between v1 and v3 passes through v2; thus, $\sigma_{v1,v3}(v2) = 1$.
  - So, the ratio $\sigma_{v1,v3}(v2)/\sigma_{v1,v3} = 1$.

Betweenness Computation



- In a similar manner, $\sigma_{v1,v4} = 1$, and $\sigma_{v1,v4}(v2)=1$.
- So, for the pair (v1, v4), the ratio $\sigma_{v1,v4}(v2)/\sigma_{v1,v4} = 1$.

- For the pair (v3,v4), $\sigma_{v3,v4} = 1$, and $\sigma_{v3,v4}(v2)=0$.
- So, for the pair (v3, v4), the ratio $\sigma_{v3,v4}(v2)/\sigma_{v3,v4} = 0$.

Betweenness Computation

- Therefore,
- $\beta(v2) = 1$ (for the pair (v1, v3))
  $+ 1$ (for the pair (v1, v4))
  $+ 0$ (for the pair (v3, v4))
  $= 2$.
- Repeat the procedure to compute $\beta(v1)$, $\beta(v3)$, $\beta(v4)$.
- The values are $\beta(v1)=0$, $\beta(v3)=2$, $\beta(v4)=0$ respectively.
- In the above example, nodes v2 and v3 are most influential nodes.

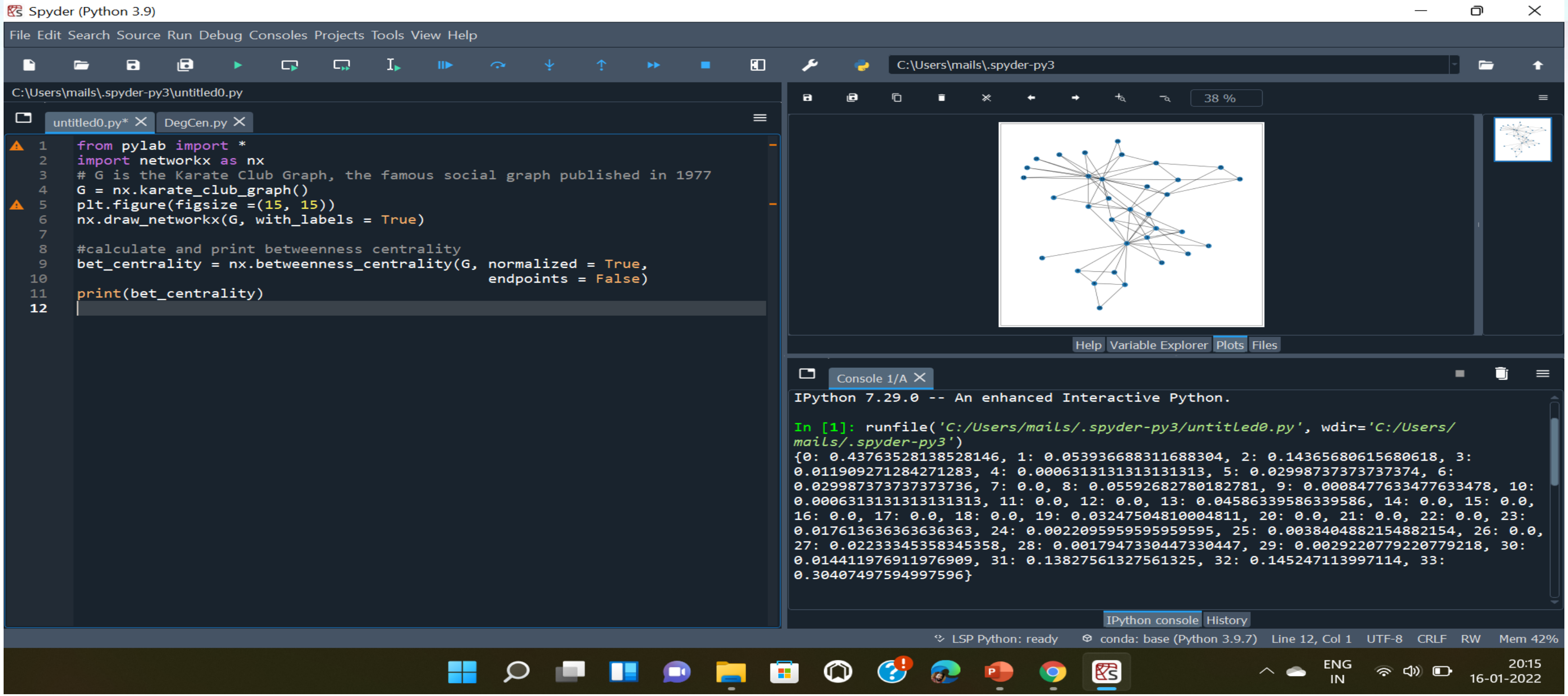
Betweenness Computation

- Graphs with larger number of nodes will tend to have higher centrality than graphs with smaller number of nodes.
- Because in large graphs, there are more nodes, s and t, to choose from to compute the centrality of the nodes.
- So, if we want to compare betweenness centrality across networks, it's useful to normalize.

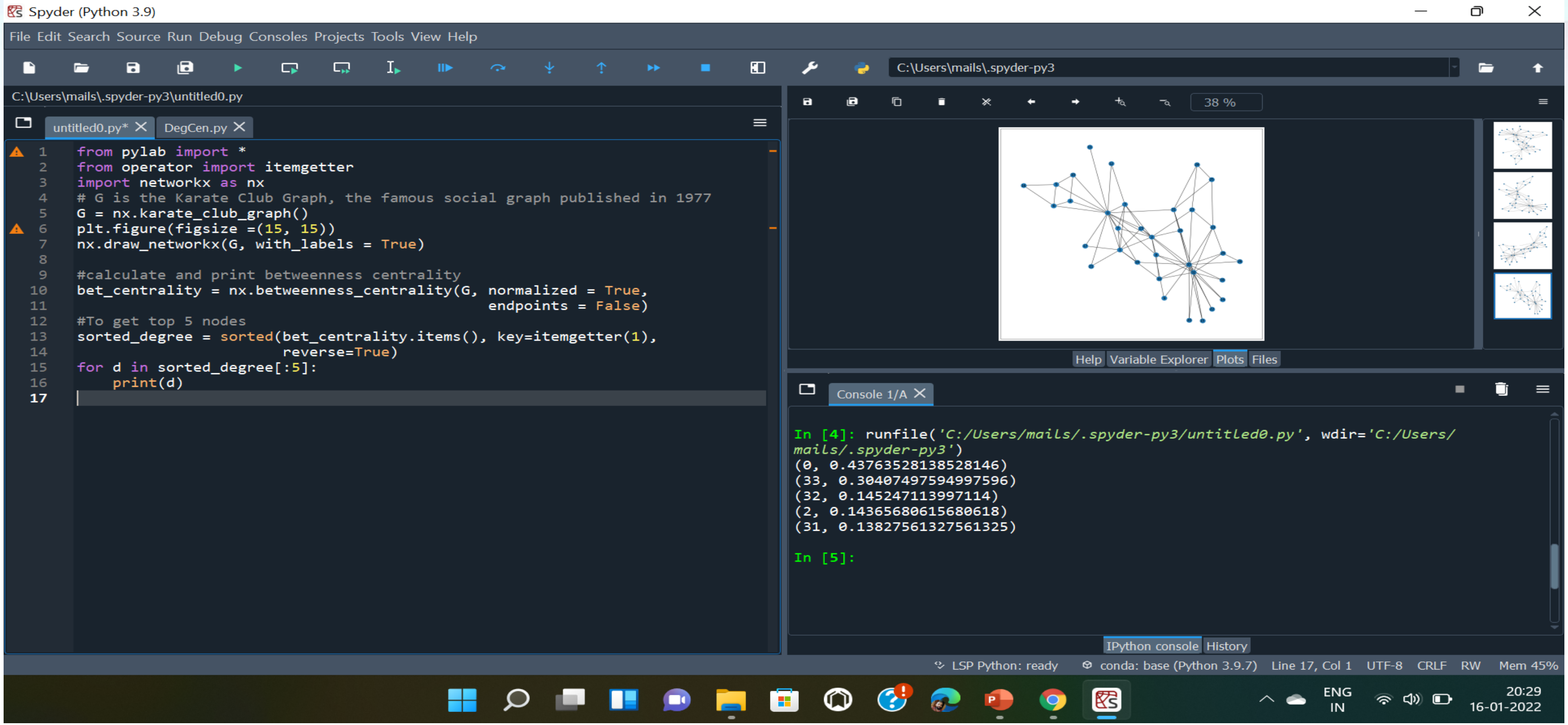
Betweenness Computation

- Normalization can be done by dividing the betweenness centrality of node v by the number of possible pairs of nodes in the network, excluding node v.
- It is different for directed and undirected graphs.
  - For undirected graphs, [β(v)/((N-1)(N-2)/2)]
  - For directed graphs, [β(v)/(N-1)(N-2)] (Since you have twice the number of pairs because for any pair s-t, you could have a path from t-s also).
- So, in our example, the normalized value for v2 is [β(v2)/((N-1)(N-2)/2)]=2/3=0.67

Betweenness Centrality using NetworkX



Betweenness Centrality using NetworkX



Thanks...........