Aghil Sabu
2018EET2865

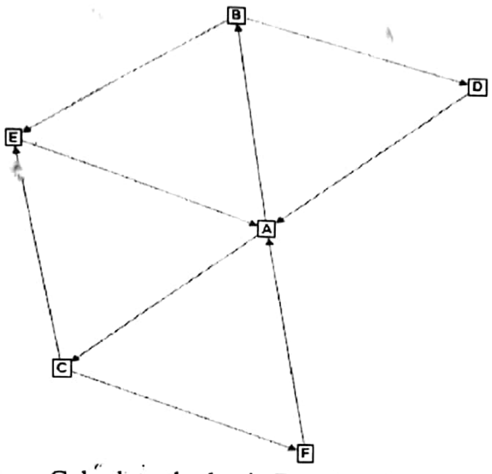
Social Network Analysis

Minor 2 – October 5, 2018

Total Marks: 25 (1 hour)

NAME:

$\frac{10}{10} + \frac{14}{15} = \frac{24}{25}$



Q1. a) Consider the graph above. Calculate the basic Page Rank values of the nodes after the 1st and 2nd iteration of the Basic Page Rank algorithm.

	A	B	C	D	E	F
1st iteration	$\frac{1}{6}\times 3 = 1/2$	1/12	1/12	1/12	1/6	1/12
2nd iteration	$\frac{1}{12}+\frac{1}{12}+\frac{1}{6}=\frac{1}{3}$	1/4	1/4	1/24	1/12	1/24

[6 marks]

initially all nodes have probability = 1/6

6

b) Give 2 disadvantages of the Basic/Scaled Page Rank algorithm. *[2 marks]*

i) If a particular section of the networks has only links which are pointed towards it from the rest of the network, and there are no outlinks from that section to the rest of the network, then Page rank will get accumulated on that section of the network.

2) If two sections of network are not at all connected, then also the results obtained wont be that accurate

2

3) more prone to scams,

4) not effective in case of dynamic pages

Q2. You have to choose between 2 mobiles A and B. Due to reputation of the brands probability A being better is 3/5. Before buying you read reviews comparing the two. The probability of review recommending the right product is 3/4. You decide to buy mobile B after reading a recommendation favoring B. Using information cascade theory determine the probability that you bought the wrong mobile. *[3 marks]*

C → correct mobile

W → wrong mobile.

R → review)

$$P(W/R) = \frac{P(R/W)\cdot P(W)}{P(R)} = \frac{P(R/W)\cdot P(W)}{P(R/W)\cdot P(W) + P(R/C)\cdot P(C)}$$

3

$$= \frac{\frac{1}{4}\times\frac{3}{5}}{\frac{1}{4}\times\frac{3}{5}+\frac{3}{4}\times\frac{2}{5}} = \frac{3}{3+6} = \frac{3}{9} = \frac{1}{3}$$

Q3. Assume that individuals are connected to 4 other people in a tree structure. What is the maximum *contagion probability* that a disease can have so that it does not turn into an epidemic? Why? *[2 marks]*

The disease will turn into an epidemic,

if, (Probability of disease) × (no of connections) > 1

$P \times 4 > 1$

1.5 $P > 0.25$

Hence the maximum Probability is 0.25

Mention Reproductive probability

b) How can a SIS epidemic model be represented by a SIR epidemic model? *[2 marks]*

In SIR epidemic model, once an infected node is recovered, it won't be considered for further analysis. It is treated like, it won't catch the disease again.

In SIS epidemic model, some nodes are infected and all other nodes are in succeptible state. once an infected node is recovered it goes to succeptible state. Due to this, in SIS epidemic model, the disease die much slower as compared to SIR epidemic model.

We can represent SIS epidemic model by using SIR epidemic model, by adding the recovered nodes back to the network as succeptible nodes

Time expanded network

1.5

Q4. In the basic "six degrees of separation" question, one asks whether most pairs of people in the world are connected by a path of at most six edges in the social network, where an edge joins any two people who know each other on a first-name basis. Now let's consider a variation on this question. For each person in the world, we ask them to rank the 30 people they know best, in descending order of how well they know them. (Let's suppose for purposes of this question that each person is able to think of 30 people to list.) We then construct two different social networks:

(a) The "close-friend" network: from each person we create a directed edge only to their ten closest friends on the list.

(b) The "distant-friend" network: from each person we create a directed edge only to the ten people listed in positions 21 through 30 on their list.

Let's think about how the small-world phenomenon might differ in these two networks. In particular, let C be the average number of people that a person can reach in six steps in the close-friend network, and let D be the average number of people that a person can reach in six steps in the distant-friend network (taking the average over all people in the world). **When researchers have done empirical studies to compare these two types of networks (the exact details often differ from one study to another), they tend to find that one of C or D is consistently larger than the other. Which of the two quantities, C or D, do you expect to be larger? Give a brief explanation for your answer.** *[5 marks]*

Here D will be larger than C.

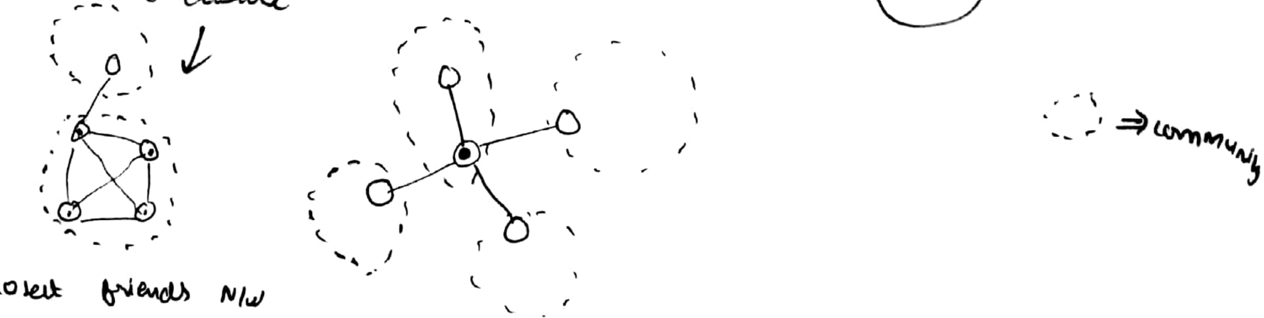
~~[illegible]~~ C is the average number of people one can reach ~~[illegible]~~ to closest friends network. ~~[illegible]~~ The closest friends list most probably contains people of the same community. but the distant friends list will connect many communities and hence can reach many persons quickly. but in case of the closest friends network, it will be difficult for one person inside it to reach a person in a different communities.

Hence the distant friends network can reach ~~[illegible]~~ more people as compared to closest friends network

Hence D will be larger than C

5

"triadic closure"



closest friends N/w

(less coverage)

distant friends N/w

(more coverage)

Q5. Suppose that some researchers studying educational institutions decide to collect data to address the following two questions.

(a) As a function of k, what fraction of IIT classes have k students enrolled?
(b) As a function of k, what fraction of 3rd-grade government school classrooms in New Delhi have k pupils?

Which one of these would you expect to more closely follow a power-law distribution as a function of k? Give a brief explanation for your answer, using some of the ideas about power-law distributions developed in class. *[5 marks]*

I think (a) will power law distribution more closely as compared to (b)

A certain decision ~~[illegible]~~ is said to follow power law, when that decision ~~[illegible]~~ is made depending on ~~[illegible]~~ others decision.

In this case IIT's have huge reputation among people and everyone wants join in an IIT, but in case of a 3rd grade government ~~school~~ class, the decision is made mostly based on individual preferences (of child/family). but Here also decision of others affect to some extent but not as much compared to that in case of IIT's. In case of IIT's, students choose IIT college or course mainly based on what the other top rankers did, on based on the placements, both of which are highly inter related.

Hence, (b) will be more random as compared to (a)
hence (a) follows power law distribution more closely.

Not about joining IIT, but more about courses within IIT being popular.

5

2018EET2865
Aghil Sabu

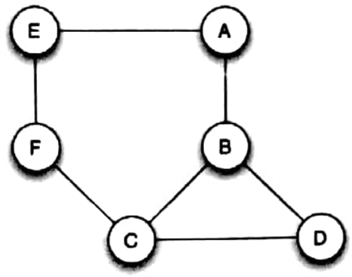
Social Network Analysis

Minor 1 -August 25, 2018

Total Marks: 25 (1 hour)

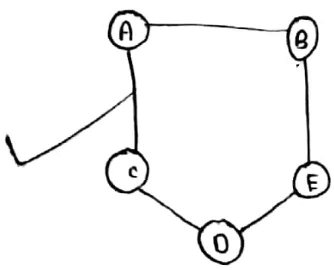
19/25

1. We say that a node X is *pivotal* for a pair of distinct nodes Y and Z if X lies on every shortest path between Y and Z (and X is not equal to either Y or Z). In the figure, B is pivotal for paths A—C and A—D, E is pivotal for A—F , D is pivotal for none.



8/10 + 11/15

(a) Give an example of a graph in which every node is pivotal for at least one pair of nodes. Explain your answer. *[2 marks]*



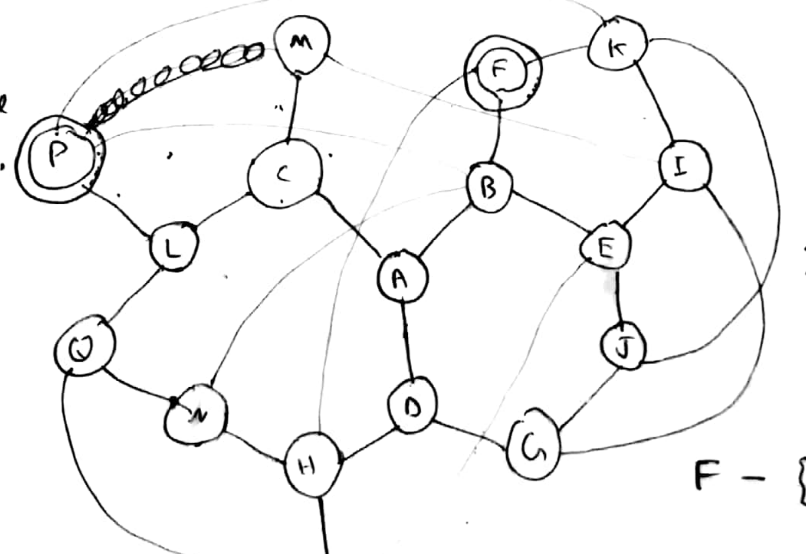
A Pivotal for B-C
B Pivotal for AE
C Pivotal for A-D
D Pivotal for C-E
E Pivotal for B-D

2

(b) Give an example of a graph in which every node is pivotal for at least two different pairs of nodes. Explain your answer. *[2 marks]*

0

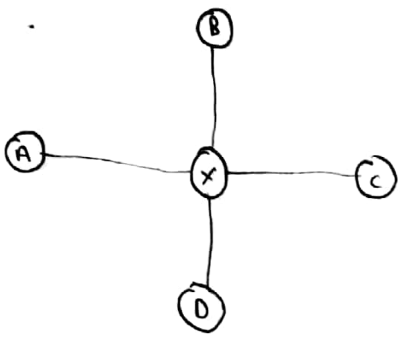
Here each node has atleast 3, connections, and connections are given such that each node is pivotal to atleast two different pairs of nodes. This is achieved by giving atleast one connection to a far point.



A – {B, P} {C, D}
B – {A, E}, {F, E}
C – {A, M}, {L, M}
D – {A, G}, {A, H}
E – {B, G}, {B, I}
~~F – {K, B}, {H, B}~~
F – {H, K}

(c) Give an example of a graph having at least four nodes in which there is a single node X that is pivotal for every pair of nodes (not counting pairs that include X). Explain your answer.

[2 marks]

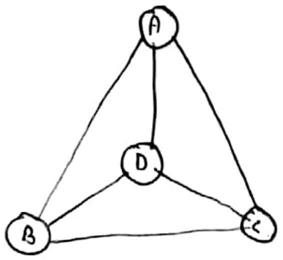


x is pivotal to A - C , A - D, A - B,
B - C, B - D, C - D

2

(d) Give an example of a graph having at least four nodes in which NO node is pivotal.

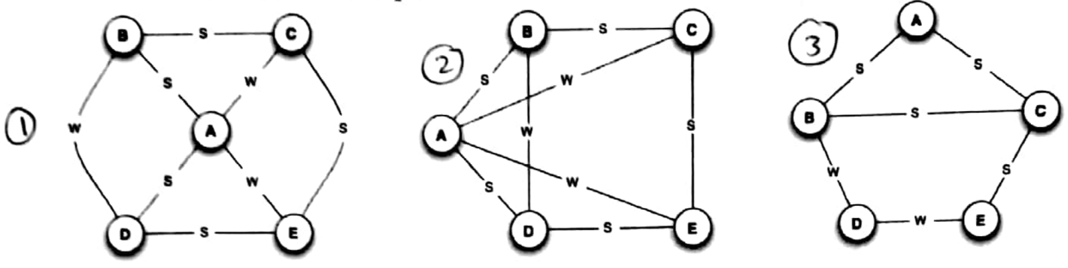
[1 mark]



A is pivotal to none
B is pivotal to none
C is pivotal to none
D is pivotal to now

1

2. In the social networks depicted in the figures below, with each edge labeled as either a strong or weak tie, which nodes satisfy the Strong Triadic Closure Property, and which do not? Provide an explanation for your answer. *[3 marks]*



3

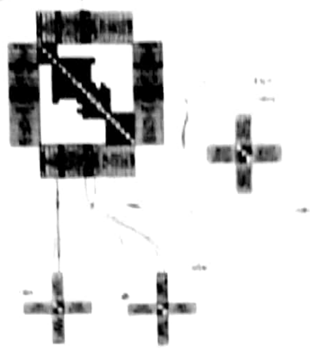
According to strong Triadic Closure property, if a node has strong ties with two other nodes then those two other nodes will atleast have a weak tie between

1 here nodes C and E doesn't satisfy strong Triadic closure Property, as in case of C, it has strong ties with B and E, but there is no tie between nodes B & E. In case of node E, it has strong ties with nodes C and D, but C & D doesn't has any ties between them.

2 here nodes C, E, doesn't satisfy strong triadic closure propery. as in cas of C, it has strong ties with B & E, but B & E has no ties between them. In case ab E, it has strong ties with C & D, but C & D has no ties between them

3 here C has strong ties with A, B & E, but there is no tie between A, E and B, E. Hence C doesn't satisfy strong triadic closure property.

3. Answer the following questions for the visualizations below: *[3 marks]*



i) Name the visualization

bar graph visualization. X

ii) State 1 advantage of this visualization method

~~less space require~~

better utilization of space compared to other methods

iii) State 1 disadvantage of this visualization method

difficult to implement

2/3

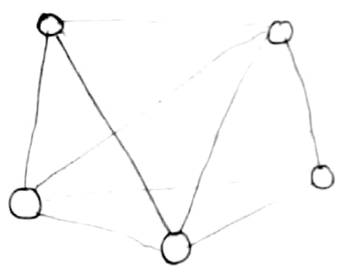
4. What is minor of a graph? How is it used in Planarity Testing? *[2 marks]*

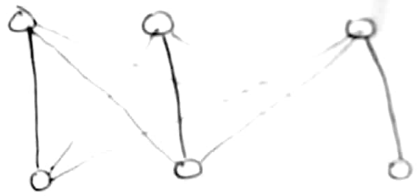
A graph can be called as a minor of a particular graph, if we can obtain that graph by removing some of the vertices or edges of the main graph.

If any minor of a graph is non planar, then we can say that the graph is non planar. for example, consider the following two graphs below.

K_5 → non planar

$K_{3,3}$ → non planar



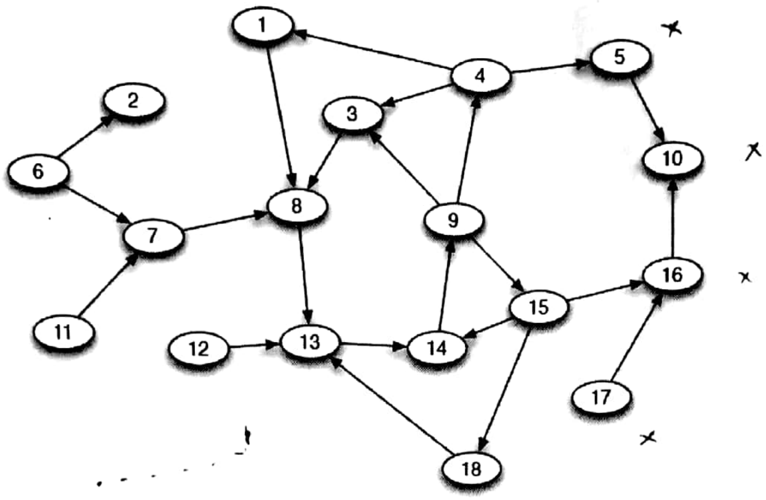


hence K_5 and $K_{3,3}$ are examples of non planar graphs.

Hence we can say that, any graph which has K_5 and $K_{3,3}$ as a minor is also non planar

2/2

Q5.



Assume that the graph above shows the Bow-Tie structure. *[5x2=10 marks]*

(a) How many nodes are there in the Giant SCC component?

9 nodes ✓

(b) Which nodes belong to the set OUT?

5, 16, 10 ✓

(c) Which nodes belong to Tendrils from IN nodes?

~~11~~, ~~6~~, 2

(d) Give an example of an edge that can be deleted to increase the size of IN.

edge 4→1, or edge 15→18 ✓...

7/10

(e) How can we form a Tube in this graph?

17 gives input directly to 16, which is an out node.