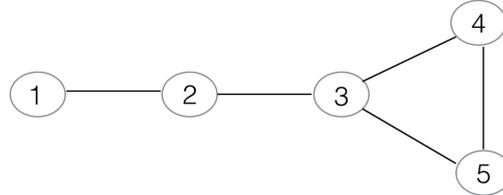


COMPUTER SCIENCE TRIPOS Part IA – 2017 – Paper 3

8 Machine Learning and Real-world Data (SHT)

Consider the following graph of users in a social network.



- (a) Give definitions of the following quantities, and state their values for nodes 1, 2 and 3:
- (i) Degree [1 mark]
 - (ii) Clustering Coefficient [2 marks]
 - (iii) Betweenness Centrality [2 marks]
- (b) (i) Which property of nodes in a network do clustering coefficient and betweenness centrality attempt to model? Give examples from naturally occurring networks. [4 marks]
- (ii) In the network depicted above, do the betweenness centrality values you calculated correspond to your intuition about the function of nodes 1, 2, and 3? If not, which factors of the formula are responsible for the divergence? [3 marks]
- (c) Suppose the link between 3 and 4 is removed and a link between 2 and 4 is added. How does this affect the betweenness centrality of 2 and 3, and why? [3 marks]
- (d) Some pairs of nodes in the new network from part (c) now have more than one shortest path that connects them. Identify these paths, and explain how their existence affects the calculation of betweenness centrality. [2 marks]
- (e) Professor Miller claims that it is possible for a particular pair of nodes in a graph to be connected by a number of shortest paths which is exponential with respect to the number of nodes in the graph. Give an example where this is the case, or prove that it cannot be the case. [3 marks]