Question 1

Question 1.a

Power law is a distribution of networks that is more common in reality, especially when the scale of network is very large. Different from The Erdos-Renyi Model, which rarely gives extreme events and follows Poisson distribution, Power law distribution exhibits long-tail phenomenon. In Power law distribution, the possibility of seeing a few nodes with high degree is much more common than Erdos-Renyi Model. For example, if we regard each website as nodes and link between them as edges (the direction of edge from A to B is 1 if a link on A directs to B), there will be a few websites, such as Google.com, that have much more website directed to them than the average websites.

Question 1.b

SIR model of epidemics is a model that is used to predict how disease spread in a region. The node is people in this region, and the link is how one person spread the disease to other people in the region. One person got infected, there is a period when this person could transmit the disease to others with a possibility of p. Different from other models, a person in SIR model is assumed to be resistant to the disease after he/she already had the infection but recovered. In the reality, the spread of chickenpox could be modeled using SIR model, since if one person got infected by chickenpox but recovered, he/she becomes resistant to chickenpox in the future.

Question 2

Question 2.a

The maximum flow is 11.
flow(s-a) is 6, flow (a-c) is 4, flow (a-d) is 2
flow(s-b) is 5, flow (b-c) is 2, flow (b-d) is 3
flow (c-t) is 6, flow (d-t) is 5

The minimum cut is 11.
Edges: E(a-c), E(a-d), E(b-c), E(b-d)

Question 2.b

The maximum flow is the maximum possible value of flow from node s to node t without exceeding the capacity of every edge. The minimum cut is the minimum sum of weights of edges that, if removed, would disconnect the node s and node t. From the definition of maximum flow and minimum cut, it is clear that the minimum cut is the threshold of possible flow from the source to the sink. Since the flow must be equal or smaller than the threshold, the maximum flow is always less than or equal to the minimum cut.

Question 3

Question 3.a

Yes. The PageRank of each node depends on its predecessor, or how many other nodes point to this one and the importance of those nodes. In figure 2, page C and page E have links to page D, and page D has link to page E. Therefore, we could write down the adjacency matrix for this directed graph as below:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \; W = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \end{bmatrix} \; W^T = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{4} \\ 1 & 0 & 0 & 0 & \frac{1}{4} \\ 0 & 1 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 1 & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Therefore $r = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{4} \\ 1 & 0 & 0 & 0 & \frac{1}{4} \\ 0 & 1 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 1 & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} r$

In the algebraic form:

$$r_A = \frac{1}{4} * r_E$$
$$r_B = r_A + \frac{1}{4} * r_E$$
$$r_C = r_B + \frac{1}{4} * r_E$$
$$r_D = r_C + \frac{1}{4} * r_E$$
$$r_E = r_D$$

Apparently, D and E should have the same PageRank.

Question 3.b

Page A.

From the equations derived in Question 3.a, it is obvious that the PageRank of Page A is the lowest.

Question 3.c

Page D and E.

From the equations derived in Question 3.a, the PageRank of Page A is the lowest, and the PageRank of Page B is larger than that of Page A, and the PageRank of Page C is larger than that of Page B, and the PageRank of Page D is larger than that of Page C. Since the PageRank of Page D is the same as that of Page E, Page D and E have the highest page rank.