6.439 Problem Set 3

Xudong Sun

4.1

(a)

The adjacency matrix of the citation network $G_{citation}$ is given by:

$$A_{ij} = \begin{cases} 1, & if\ i\ cites\ j; \\ 0, & otherwise \end{cases}$$

For the co-citation network, $G_{co-citation}$, the node $i\ and\ j$ are connected by an edge if they are both cited by another node k in network A:

$$A_{ki} = A_{kj} = 1$$

And the weight of the edge in co-citation network, $G_{co-citation}$, with adjacency matrix, C, is given by the number of co-citations. Therefore, the weighted edge is:

$$C_{ij} = \sum_{k=1}^{N} A_{ki}A_{kj} = \sum_{k=1}^{N} (A^T)_{ik}A_{kj}$$

The adjacent matrix C can be written as:

$$C = A^T A$$

(b)

For the bibliographic coupling network, $G_{bibliographic}$, the node $i\ and\ j$ are connected by an edge if they both cite another node k in network A:

$$A_{ik} = A_{jk} = 1$$

And the weight of the edge in co-citation network, $G_{bibliographic}$, with adjacency matrix, B, is given by the number of co-citations. Therefore, the weighted edge is:

$$B_{ij} = \sum_{k=1}^{N} A_{ik}A_{jk} = \sum_{k=1}^{N} A_{ik}(A^T)_{kj}$$

The adjacent matrix C can be written as:

$$C = AA^T$$

(c)

A high weight in co-citation network means the two papers are related in their subject content. While bibliographic coupling network can also indicate such relation, it basically means that the two papers are based on similar research background, but their content can also differ in the future. Additionally, the value of co-citation can change, which reflects the changing in scientists' perceptions. While the bibliographic coupling is fixed. In sum, I think co-citation network is a better measure.