

Citation Network

Hervé Guillon

11 May 2020

Contents

Citation Network	1
------------------	---

Citation Network

The citation network was extracted from article metadata which stores identifiers for the cited articles and citing article (i.e. the article to which the metadata are attached to). A total of 29,900 citations were found between 4,603 unique articles of the English corpus. The resulting bibliographic network is defined by its $4,603 \times 4,603$ -adjacency matrix B and is filtered to remove edges between node external to the collected corpus. The citation network between countries, B_C is obtained by:

$$B_C = C^T \times B \times C$$

where C is a $4,603 \times 23$ -matrix containing the output of the machine learning probabilistic predictions of the location of study between 23 possible country of study. In consequence, B_C is a weighted directed adjacency matrix. Similar reasoning leads to the citation network between specific research topics, method topics and water budget topics by changing the matrix C in the previous equation for the LDA probabilistic predictions of topics.