The Impact of Citation and Co-Citation Analysis on the Relevancy of Keywords in Digital Libraries

STU FIIT

Co-citations

Primary goals
- Keyword extraction in digital libraries
- Three sources of text for extraction methods:
    - Fulltext
    - Citations context
    - Co-citations

E   F   G

A   B   C   D

CCI   --> Co-citation weight
CCPI --> Co-citation proximity index

Evaluation
- Method 1: Fulltext
- Method 2: Fulltext + Citations
- Method 3: Fulltext + Citations + Co-citations

$$P(M1) < P(M2) < P(M3)$$

Important features
- Citations highlight different aspects of articles
- Co-cited articles tend to be similar
- Using citations' context of any size produces better keywords
- It is possible to highlight all relevant aspects of an article only with citations
- The number of highlighted aspects starts to converge with the increasing count of analyzed citation's contexts

Hypothesis:
Using co-citations produces higher precision of the keyword extraction method compared to the method using only citations, especially when the article has only a small quantity of citations.

Author: Máté Vangel   Contact: mate.vangel@gmail.com   Supervisor: Róbert Móro

PeWe@FIIT
personalized web group

ACL Anthology Network
The Association Of Computational Linguistics Anthology Network

Annota