A recommendation system based on hierarchical clustering of an article-level citation network

The scholarly literature is expanding at a rate that necessitates intelligent algorithms for search and navigation. For the most part, the problem of delivering scholarly articles has been solved. If one knows the title of an article, locating it requires little effort and, paywalls permitting, acquiring a digital copy has become trivial. However, the navigational aspect of scientific search – finding relevant, influential articles that one does not know exist – is in its early development.

In this paper, we introduce Eigenfactor Recommends – a citation-based method for improving scholarly navigation. The algorithm uses the hierarchical structure of scientific knowledge, making possible multiple scales of relevance for different users. We implement the method and generate more than 300 million recommendations from more than 35 million articles from various bibliographic databases including the AMiner dataset. We find little overlap with co-citation, another well-known citation recommender, which indicates potential complementarity.

In an online A-B comparison using SSRN, we find that our approach performs as well as co-citation, but this new approach offers much larger recommendation coverage. We make the code and recommendations freely available at babel.eigenfactor.org and provide an API for others to use for implementing and comparing the recommendations on their own platforms.