Abstract

The explosive growth and variation of digital repository renders the searching task non-tedious for a researcher. In such a provision, research paper recommendation system is advocated. In the designing of research paper recommendation system, proper representation of research articles as well as user profile, capturing users dynamic notion, appropriate ranking formalization and an effective recommending technique are essential tasks. To achieve this goal, several research paper recommendation frameworks have been proposed in recent literature. However, the reported recommendation systems have several drawbacks in regard of representing articles, capturing users' dynamic interest in timely manner, alleviating cold start problem. In addition, in most of the reported literature, an effective ranking strategy has not been implied properly or considered at all. Whereas, a proper ranking enhance the recommendation quality. This research addresses these gap and aims to design a multi-criteria based research paper recommendation system.

The first objective in this thesis is to identify and extract new features, such as keyword diversification, text complexity, citation analysis over time, and scientific quality measurement to represent a research article. The proposed features help to incorporate variation in recommendation, analyze text complexity of papers, decide relevancy of papers, and measure the quality of papers. At the same time, proposed features are also useful to define user characteristics as well.

The second objective is to select relevant and non-redundant features from the feature set comprising with proposed features and traditional features of a paper. To optimize features in unsupervised domain, a two-stage unsupervised feature selection strategy for mixed dataset has been proposed. The proposed approach addresses the drawbacks of existing unsupervised feature selection techniques and enhance the performance of recommendation system. Moreover, it is scalable to data size and applicable to any type of data.

Another important module of recommendation system is representing user profile and capturing users' need dynamically. However, the existing solutions do not capture user's intention that is necessity at the moment to make the recommendation perfect. In addition, they merely consider user's long term and short term interest both to analysis the trend of users notion. An approach with deep sequential topic analysis has been proposed in this thesis to overcome the existing limitations.

Finally, to recommend relevant articles to the researcher a multi-criteria based personalized research paper recommendation system has been developed. For a given search query, the proposed system efficiently provide recommendation for old users as well as new users. further, a feature engineering technique is utilized to unfold valuable insights of papers through multiple hidden features. These features are used as a context of users as well as multiple criteria for ranking papers. Additionally, the system predicts a user's intention beyond the user's preference to capture the dynamic notion of a user. Finally, a novel ranking strategy is proposed to retrieve personalized and the most important papers.

Keywords: Research paper recommendation system, Feature engineering, Users intention, Topic modeling, Ranking, Multi-criteria analysis, Personalization.