



INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Social Network Analysis in Information Retrieval System

Lalit Gehlod*

* Department of Computer Engineering, Institute of Engineering and Technology, Indore

Abstract

A social network is a social structure made up of a set of actors (such as individuals or organizations) and the dyadic ties between these actors. The social network perspective provides a clear way of analyzing the structure of whole social entities. Social network analysis (SNA) is the methodical analysis of social networks. Social network analysis views social relationships in terms of network theory, consisting of *nodes* (representing individual actors within the network) and *ties* (which represent relationships between the individuals, such as friendship, kinship, organizational position, sexual relationships, etc.) we show how. Social network analysis (SNA) can be used in information retrieval system.

Keywords: . Social network analysis (SNA), Information Retrieval

Introduction

Information retrieval is the process of gathering information resources relevant to an information need from a huge collection of information resources. Request for information can be based on metadata or on full-text indexing. An information retrieval process begins when a user enters a query into the system. Queries are formal statements of information needs, for example search strings in web search engines. In information retrieval a query does not uniquely identify a single object in the collection. Instead, several objects may match the query, perhaps with different degrees of relevancy. An object is an entity that is represented by information in a database. User queries are matched against the database information. Depending on the application the data objects may be, for example, text documents, images, audio, mind maps or videos. Often the documents themselves are not kept or stored directly in the IR system, but are instead represented in the system by document surrogates or metadata. Most IR systems compute a numeric score on how well each object in the database matches the query, and rank the objects according to this value. The top ranking objects are then shown to the user. The process may then be iterated if the user wishes to refine the query [1]. A social network is a social structure made up of a set of actors (such as individuals or organizations) and the dyadic ties between these actors. The social network perspective provides a clear way of analyzing the structure of whole social entities [1].

Social network analysis (SNA) is the methodical analysis of social networks. Social network analysis views social relationships in terms of network theory, consisting of *nodes* (representing individual actors within the network) and *ties* (which represent relationships between the individuals, such as friendship, kinship, organizational position, sexual relationships, etc.) [3, 5, 7, 6]

A major limitation of most existing information retrieval models and systems is that the retrieval decision is made based solely on the query and document collection information about the actual user and search context is largely ignored [2]. The proposed research is aimed to investigate the effects of social network analysis to enhance information retrieval systems. As social network of a person has a significant impact on his or her information acquisition. Therefore the research proposes the application of available social network data in the context of information retrieval systems. Therefore the research aims at the application of available social network data in the context of information retrieval systems. An outline of the research design for the exploration of meaningful sources for social network extraction and the impact of meaningful social network analysis methods and measures in the context of information retrieval systems

The proposed research is aimed to investigate one or more Factors having impact on Information retrieval systems using Social network analysis (SNA). It is

further proposed to establish a model to validate the investigated results by experimentation and/or by comparison with the similar published work. It is also proposed to develop a computer program to study the effects of the parameters analytically.

A Brief Review Of The Work Already Done In The Field

Many researchers have worked the improvement of information retrieval systems. Information Retrieval research explores a wide range of questions, all somehow connected to the goal of helping a person and useful information in response to a need. The challenge over almost fifty years of work is improving the accuracy of search systems. The unspoken assumption is that if searchers are given more accurate systems, they will be more successful in finding answers [1].

Shen, Xuehua [2] reveals in his paper the major limitation of most existing retrieval models and systems is that the retrieval decision is made based solely on the query and document collection; information about the actual user and search context is largely ignored. In this paper, we study how to exploit implicit feedback information, including previous queries and click through information, to improve retrieval accuracy in an interactive information retrieval setting. We propose several context sensitive retrieval algorithms based on statistical language models to combine the preceding queries and clicked document summaries with the current query for better ranking of documents. We use the TREC AP data to create a test collection with search context information, and quantitatively evaluate our models using this test set. Experiment results show that using implicit feedback, especially the clicked document summaries, can improve retrieval performance substantially.

Mislove, Alan [3] in his paper presents a large-scale measurement study and analysis of the structure of multiple online social networks. They examine data gathered from four popular online social networks: Flickr, YouTube, Live Journal, and Orkut. They crawled through publicly accessible user links on each site, obtaining a large portion of each social network's graph. Their data set contains over 11.3

million users and 328 million links. They believe that this is the first study to examine multiple online social networks at scale. Their results confirm the power-law, small-world, and scale free properties of online social networks. They observe that the in-degree of user nodes tends to match the out-degree; that the networks contain a densely connected core of high-degree nodes; and that this core links small groups of strongly clustered, low-degree nodes at the fringes of the network. Finally, they discuss the implications of these structural properties for the design of social network based systems.

Bonchi, Francesco [6] has explained how Social network analysis has gained significant attention in recent years, largely due to the success of online social networking and media-sharing sites, and the consequent availability of a wealth of social network data. In spite of the growing interest, however, there is little understanding of the potential business applications of mining social networks. While there is a large body of research on different problems and methods for social network mining, there is a gap between the techniques developed by the research community and their deployment in real-world applications. Therefore the potential business impact of these techniques is still largely unexplored. In his article he uses a business process classification framework to put the research topics in a business context and provide an overview of what we consider key problems and techniques in social network analysis and mining from the perspective of business applications. In particular, he discusses data acquisition and preparation, trust, expertise, community structure, network dynamics, and information propagation. In each case they present a brief overview of the problem, describe state-of-the-art approaches, discuss business application examples, and map each of the topics to a business process classification framework. In addition, they provide insights on prospective business applications, challenges, and future research directions. The main contribution of their article was to provide a state-of-the-art overview of current techniques while providing a critical perspective on business applications of social network analysis and mining.

Noteworthy Contributions In The Field Of Proposed Work

Many researchers have worked / working in the field Information retrieval and application of social network analysis (SNA).

Hsiao-Hsuan Lu [3] Social networking sites have becoming important platforms for users to establish the relationships between each other. As time goes by, the links between people will form the So-called “Strong Links”. For those users, information provided by the friends with strong link is considered as more interesting and useful. Most of recent search engines are designed based on only measuring the similarity between keywords and articles. However, the social relations between authors of articles and searcher have not been taken into account in recent research. Therefore, in order to improve the performance of recent search engines, they include the measurement of social relationships in search engine and expect the search quality can be improved. In this study, they collected the data from Face book to calculate the social relationship. About the content, the data will be processed by using CKIP (Chinese word net) and TF-IDF. Finally, they combine key-word frequency and social relations as a value, which is called the Social Ranking value. The value will be used as the key to rank the search results. In this paper, they also demonstrated a real example to explain the proposed methodology as well as a system interface.

Yang, Christopher C. [6] in his work shows Social media is actively utilized by extremists to spread out their ideologies. While the Internet provides a platform for any users around the world to share their opinions, some opinions in social media can be related to the national security and threatening to others. Given the large volume and exponential growing rate of messages on the social media platforms, it is impossible to analyze the messages by manual effort. An effective way to identify the threat through social media is detecting the influential users automatically. By identifying the influential users, we can determine the impact and the neighborhood of these users. In his work, he develops weights to incorporate message content similarity and response immediacy to measure the degree of influence

[http:// www.ijesrt.com](http://www.ijesrt.com)© *International Journal of Engineering Sciences & Research Technology*

between any two users on a social networking site and integrate the weights with the typical link analysis techniques. In his experiment, he investigates the impact of weights and the basic algorithms (iterative or prestige) on the user influence ranking. The experiment is conducted on the Dark Web forum provided in the ISI-KDD Challenge. The result shows that the weights make substantial impact on the ranking results, especially on the in-degree algorithm.

Proposed Methodology

The work plan is very important criteria to carry out the research work. The following is the methodology or work plan for achieving the above objectives is illustrated in Figure1. At first, the theoretical foundations for the usage of online information sources for the extraction of social networks are outlined. Secondly, a model for a social network enhanced information retrieval system was developed. Therefore an evaluation of methods that are applicable to information retrieval system has been conducted. On the basis of the chosen methods a model for the relevance computation of a document within an information retrieval system was developed. At third, the model was implemented as part of the search engine/information retrieval system. The fourth step is an evaluation whether and how the relevance ranking based on centrality measures has an impact on the relevance ranking of the documents of the information retrieval system



Figure1: Methodology / Work Plan

Conclusion

Information retrieval systems using Social network analysis (SNA). The work may include:

- a. Social Network Extraction.
- b. Social Network Analysis on extracted Social Network to determine one or more parameters of our interest.
- c. Ranking document based on one of these factors.

- d. Apply document ranking as extra parameter in existing Information Retrieval System to enhance relevance ranking of documents

References

1. Fang, Hui, Tao Tao, and ChengXiang Zhai. "A formal study of information retrieval heuristics." In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 49-56. ACM, 2004.
2. Kelly, A., & Kittle Autry, M. (2011). A humanistic approach to the study of social media: Combining social network analysis & case study research. *Proceedings of the 29th ACM International Conference on Design of Communication*, 257-260
3. Willinger, Walter, Reza Rejaie, Mojtaba Torkjazi, Masoud Valafar, and Mauro Maggioni. "Research on online social networks: time to face the real challenges." *ACM SIGMETRICS Performance Evaluation Review* 37, no. 3 (2010): 49-54.
4. Teevan, Jaime, Daniel Ramage, and Meredith Ringel Morris. "# TwitterSearch: a comparison of microblog search and web search." In *Proceedings of the fourth ACM international conference on Web search and data mining*, pp. 35-44. ACM, 2011.
5. Allan, James, Ben Carterette, and Joshua Lewis. "When will information retrieval be good enough?." In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 433-440. ACM, 2005
6. Shen, Xuehua, Bin Tan, and ChengXiang Zhai. "Context-sensitive information retrieval using implicit feedback." In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 43-50. ACM, 2005.

Author Bibliography

