Joeran Beel, Bela Gipp, and Jan Olaf Stiller. Information Retrieval on Mind Maps – What could it be good for? In *Proceedings of the 5th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom'09)*, pages 1–4, Washington (USA), November 2009. IEEE. ISBN 978-963-9799-76-9. Downloaded from www.docear.org

Information retrieval on mind maps – what could it be good for?

Joeran Beel
Otto-von-Guericke University
Computer Science/ITI/VLBA-Lab
Magdeburg, Germany
SciPlore.org
beel@sciplore.org

Bela Gipp
Otto-von-Guericke University
Computer Science/ITI/VLBA-Lab
Magdeburg, Germany
SciPlore.org
gipp@sciplore.org

Jan-Olaf Stiller
University of Wolfenbüttel
Karl Scharfenberg Faculty
Wolfenbüttel, Germany
SciPlore.org
stiller@sciplore.org

***act*—Mind maps are used by millions of people. In this paper we present how information retrieval on mind maps could be used to enhance expert search, document summarization, keyword based search engines, document recommender systems and determining word relatedness. For instance, words in a mind map could be used for creating a skill profile of the mind maps' author and hence enhance expert search. This paper is a research-in-progress paper which means no research results are presented but only ideas.**

Keywords-data mining, information retrieval, mind maps, expert seach, document clustering, document classification

I. Introduction

Mind maps were originally invented by Tony Buzan in the 1970s [1] and are nowadays used by millions of people for brainstorming, note taking, project planning, decision making, and document drafting. Many software tools exist to support the creation of mind maps [2, 3]. The probably most popular ones are *MindManager* with about 1.5 million users [4] and *FreeMind* with about 150,000 downloads a month [5]. Hundreds of books and research articles were published about how to create mind maps and about evaluating mind maps' effectiveness, for instance, in the field of education [6-9].

However, to our knowledge, no research exists whether information extracted from mind maps could be used for enhancing other applications. We believe, it can and present our ideas in this paper. Each of the next sections deals with enhancing a particular application, namely

- Expert search
- Document summarization
- Keyword based search engines
- (Document) Recommender Systems
- Determining word relatedness

After presenting the ideas, the concept of information retrieval on mind maps in general is discussed as well as future research.

II. Expert Search

Finding the right experts in a big company is a difficult endeavor. In first attempts, databases were used and employees could enter their skills manually [10, 11]. In the last decade much research has been performed on automatically creating skill profiles. The probably most promising approach is analyzing documents. For instance, if a researcher has published many documents containing the word 'mind map', she probably has some expertise in the field of mind mapping. Typical documents being analyzed are emails, visited websites, scholarly articles and documents published in a company's intranet [12-16]. Mind maps have not been used so far.

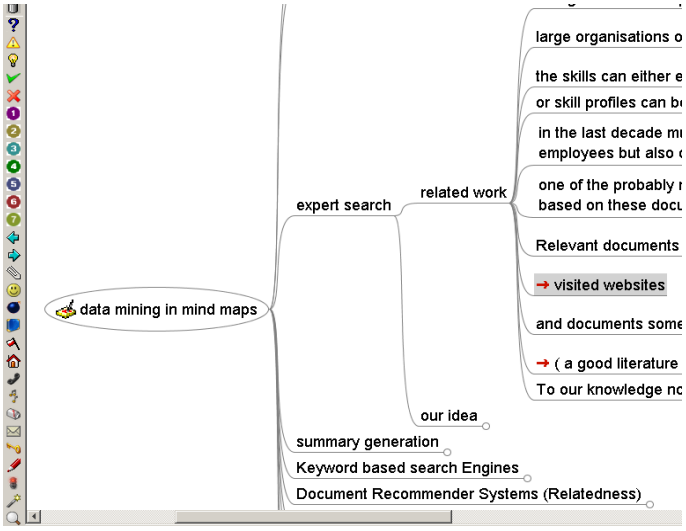

Figure 1. A mind map (early draft of this paper)

A mind map (see Figure 1) seems well suited for creating a skill profile of its author. The words in a mind map should specify quite well the author's expertise. In addition, nodes can contain notes and links which could also be analyzed. In contrast to text documents, a mind map seems likely to contain less stop and other irrelevant words. This should facilitate the creation of skill profiles.

III. Document Summarization

Search engines usually display summarized data for each search result. This could be the document's title, URL, or a short extract of the document's text. Academic search engines additionally display data such as the author, publishing date or the abstract (see Figure 2). This does not always deliver satisfying results. In Figure 2, for instance, the extract is not very informative, it equals basically the title. Alternatively to

text extracts, some researchers attempted to automatically create abstracts [17-19] or summarizing documents based on user generated data such as hyperlinks [20], social annotations [21] and annotated bibliographies [22].

Using **Mind Maps** To Teach Social Problems Analysis.
AR Peterson, PJ Snyder - 1998 - eric.ed.gov
... ED424882 - Using **Mind Maps** To Teach Social Problems Analysis. ... ERIC #: ED424882. Title: Using **Mind Maps** To Teach Social Problems Analysis. ...
Cited by 10 - Related articles - Cached - Import into BibTeX

Figure 2. Example of summary data on Google Scholar

Mind maps could be used to complement summarization of documents. Most mind mapping tools allow to link nodes in the mind map with documents on the user's hard drive or to link a node to a webpage. The node's text, and the text of parent nodes, could be seen as a summary for the linked document. Figure 3 illustrates this approach: The node with the red arrow and gray background links to the PDF file of the scholarly article 'Are your citations clean?' [23]. This article deals with problems of citation analysis. In this example the node which is linking to the PDF and its parent nodes summarize the article's content well:

```
Citation Analysis -> Problems -> Technical Problems ->
problem: different authors with the same name
```

Certainly, one occurrence in a mind map would not be sufficient for a thoroughly summary. But if several users would link a document in several mind maps, this could add up to a descent summary, highlighting what readers found most relevant in the document.

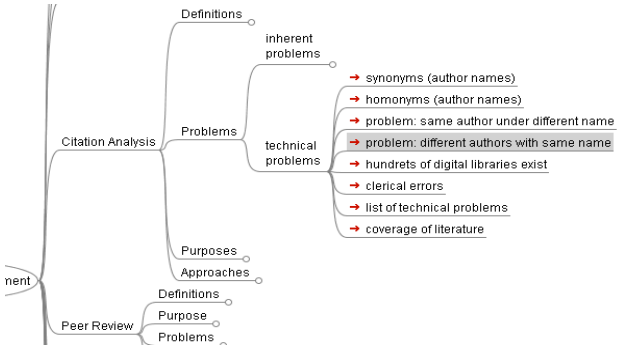

Figure 3. Mind maps as document summary and for determining word relatedness

IV. Keyword Based Search Engines

When searching for documents, usually a keyword is entered and the search engine returns those documents containing the keyword. Various algorithms exist to calculate how relevant a document is for a certain keyword search (e.g. tf-idf and BM25(f)), but usually only words contained in the documents are considered. Only few approaches consider words of 'neighbored' documents additionally [24]. Considering neighbored documents means, document A could be found for a keyword search even if document A does not contain the keyword, but document B, which is linking to document A. Usually this kind of link analysis is applied to scholarly literature and websites. However, it seems likely that the same concept could be applied to mind maps.

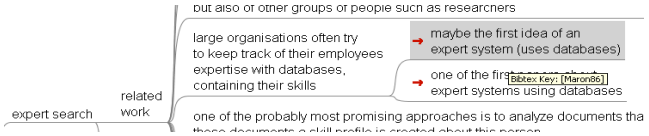

Figure 4. Enhance keyword based search engines

If a mind map links to a document, the words of the linking (and parental) node could be assigned to the linked document. Figure 4 illustrates this: The mind map contains a node called 'expert search' and child nodes link to documents related to expert search (those with the red arrows). However, many of these documents do not contain the term 'expert search', but other expressions such as 'expert finder', 'expertise management' or 'skill management'. If search engines would analyze mind maps and treat them as 'neighbored' documents, recall in document retrieval could be increased.

V. (Document) Recommender Systems

One common recommendation approach is to recommend those items which are related to items a user likes (item based recommendations). For scholarly literature and websites, relatedness often is determined via citation analysis and hyperlink analysis respectively. The same concept could be applied to mind maps.

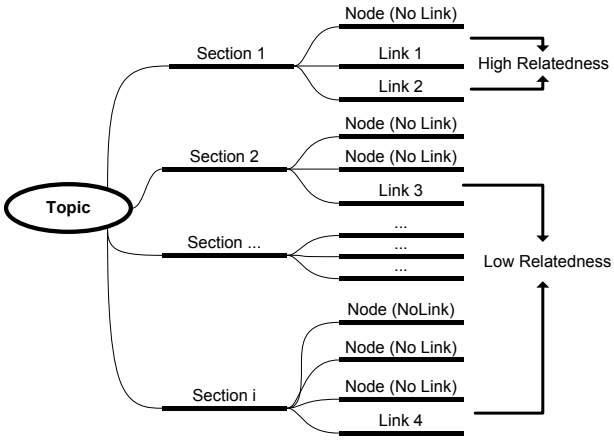

Figure 5. Expected Link Relatedness (Illustration)

The basic idea of what we call 'Mind Map Citation Analysis': when two documents A and B are linked by a mind map, document B could be recommended to those users liking document A. This concept could be enhanced with common citation analysis approaches. For instance, if two documents are linked in high proximity, their relatedness can be expected to be higher than two documents linked in lower proximity [25, 26]. Figure 5 illustrates this concept: Link 1 and 2 are in direct proximity. Therefore, the linked documents can be expected to be highly related. Between link 3 and 4 is a higher distance, so their relatedness is likely to be lower.

VI. Determining Word Relatedness

Knowing how words are related is important for many applications. For instance, search engines want to determine synonyms [27-29] and offer search query recommendations [30-32]; social tagging systems often recommend related tags to their users [33-35]; and, among others, for web 2.0

applications, (semi) automatic generation of ontologies is desirable [36-38]. Again, it seems likely that information retrieval on mind maps could help enhancing these applications.

A mind map is a graph and nodes are in hierarchical order. As such the nodes' terms are in direct relation to each other. For instance, in Figure 3: Based on the mind map a search query recommender could recommend the terms 'problems' or 'definition' to someone searching for 'citation analysis' in order to specify his search. Or, if a person is searching for 'citation analysis', then 'peer review' might be an interesting search term to broaden the search.

VII. Discussion and Further Research

In this paper we presented how data of mind maps could be used to enhance expert search, document summarization, keyword based search engines, document recommender systems and determining word relatedness. The presented ideas are not yet supported by research and it could turn out that data of mind maps is not able to enhance the mentioned applications. In addition, two more challenges exist. First, it is unknown if a sufficient number of people create mind maps and if they are willing to share their data. Second, the robustness of data seems critical. All platforms analyzing data of users do have to cope with spam and fraud as soon as they become successful. There is no reason to assume that this would be different if information retrieval on mind maps became successful.

As part of the *SciPlore.org* project we will further research information retrieval on mind maps. Recently we developed a special mind mapping software focusing on researchers needs [39]. This software will help to gather and analyze mind maps in order to see if the here presented ideas may be realized.

References

[1] Toni Buzan. *Making the Most of your Mind*. Pan Books, 1977.

[2] Open Directory Project DMOZ. Open Directory - Reference: Knowledge Management: Knowledge Creation: Mind Mapping: Software. Website, Juli 2009. URL http://www.dmoz.org/Reference/Knowledge_Management/Knowledge_Creation/Mind_Mapping/Software//.

[3] Mind-Mapping.org. Software for mindmapping and information organisation. Website, Juli 2009. URL http://www.mind-mapping.org/mind-mapping-software/35?selectedCategories[]=mind%20maps&selectedOSes[]=all%20operating%20systems&pastOrPresent[]=current&datePicker2=&filterData=Show+selected+items.

[4] MindJet. MindJet: About MindJet. Website, Juli 2009. URL http://www.mindjet.com/about/.

[5] SourceForge. SourceForge.net: Project Statistics for FreeMind. Website, 2008. URL http://sourceforge.net/project/stats/detail.php?group_id=7118&ugn=freemind&type=prdownload&mode=year&year=2008&package_ihttp://sourceforge.net/project/stats/detail.php?group_id=7118&ugn=freemind&type=prdownload&mode=year&year=2008&package_id=0.

[6] Sanford B. Frumker. *Mind Map: Your Guide to Prosperity and Fulfillment*. Health Associates, 1993.

[7] Anthony J. Mento, Patrick Martinelli, and Raymond M. Jones. Mind Mapping in Executive Education: Applications and Outcomes. *The Journal of Management Development*, 4(4), 1999.

[8] Claire A. Michelini. Mind Map: A New Way to Teach Patients and Staff. *Home Healthcare Nurse*, 18(5):318–322, 2000.

[9] Steve Morris and Jane Smith. *Understanding mind maps in a week*. Hodder & Stoughton, 1998.

[10] T.H. Davenport and L. Prusak. Working knowledge: How organizations manage what they know. *Ubiquity*, 1(24), 2000.

[11] ME Maron, S. Curry, and P. Thompson. An inductive search system: Theory, design, and implementation. *IEEE Transactions on Systems, Man and Cybernetics*, 16(1):21–28, 1986.

[12] C.S. Campbell, P.P. Maglio, A. Cozzi, and B. Dom. Expertise identification using email communications. In *Proceedings of the twelfth international conference on Information and knowledge management*, pages 528–531. ACM New York, NY, USA, 2003.

[13] N. Craswell, D. Hawking, A.M. Vercoustre, and P. Wilkins. P@noptic expert: Searching for experts not just for documents. In *Ausweb Poster Proceedings, Queensland, Australia*, 2001.

[14] B. Dom, I. Eiron, A. Cozzi, and Y. Zhang. Graph-based ranking algorithms for e-mail expertise analysis. In *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, pages 42–48. ACM New York, NY, USA, 2003.

[15] M. Maybury, R. D'Amore, and D. House. Expert finding for collaborative virtual environments. 2001.

[16] J. Wang, Z. Chen, L. Tao, W.Y. Ma, and L. Wenyin. Ranking user's relevance to a topic through link analysis on web logs. In *Proceedings of the 4th international workshop on Web information and data management*, pages 49–54. ACM New York, NY, USA, 2002.

[17] HP Luhn. The automatic creation of literature abstracts. *Advances in Automatic Text Summarization. MIT Press, Cambridge, MA*, pages 58–63, 1956.

[18] CD Paice. The automatic generation of literature abstracts: an approach based on the identification of self-indicating phrases. In *Proceedings of the 3rd annual ACM conference on Research and development in information retrieval*, pages 172–191. Butterworth & Co. Kent, UK, UK, 1980.

[19] CD Paice. Constructing literature abstracts by computer: techniques and prospects. *Information Processing and Management: an International Journal*, 26(1):171–186, 1990.

[20] J.Y. Delort, B. Bouchon-Meunier, and M. Rifqi. Enhanced web document summarization using hyperlinks. In *Proceedings of the fourteenth ACM conference on Hypertext and hypermedia*, pages 208–215. ACM New York, NY, USA, 2003.

[21] H. Zhang, Z.C.W. Ma, and Q. Cai. A study for documents summarization based on personal annotation. In *Proceedings of the HLT-NAACL 03 on Text summarization workshop-Volume 5*, pages 41–48. Association for Computational Linguistics Morristown, NJ, USA, 2003.

[22] M.Y. Kan, J.L. Klavans, and K.R. McKeown. Using the annotated bibliography as a resource for indicative summarization. *Arxiv preprint cs/0206007*, 2002.

[23] D. Lee, K. Jaewoo, M. Prasenjit, L. Giles, and O. Byung-Won. Are your citations clean? *Communications of the ACM*, 50:33–38, 2007.

[24] X. Qi and B.D. Davison. Classifiers without borders: incorporating fielded text from neighboring web pages. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 643–650. ACM New York, NY, USA, 2008.

[25] Bela Gipp, Jöran Beel, and Christian Hentschel. Scienstein: A Research Paper Recommender System. In *International Conference on Emerging Trends in Computing (ICETiC'09)*, pages 309–315. IEEE, January 2009. Available on http://www.sciplore.org.

[26] Bela Gipp and Jöran Beel. Citation Proximity Analysis (CPA) - A new approach for identifying related work based on Co-Citation Analysis. In *Proceedings of the 12th International Conference on Scientometrics and Informetrics (ISSI'09)*, pages 571–575. International Society for Scientometrics and Informetrics, July 2009. Available on http://www.sciplore.org.

[27] V.D. Blondel, A. Gajardo, M. Heymans, P. Senellart, and P. Van Dooren. A measure of similarity between graph vertices: Applications to synonym extraction and web searching. *Siam Review*, 46(4):647–666, 2004.

[28] M. Clements, A.P. de Vries, and M.J.T. Reinders. Detecting synonyms in social tagging systems to improve content retrieval. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 739–740. ACM New York, NY, USA, 2008.

[29] D. Sanchez and A. Moreno. Automatic discovery of synonyms and lexicalizations from the Web. *Artificial intelligence research and development*, 131:205–212, 2005.

[30] R. Baeza-Yates, C. Hurtado, and M. Mendoza. Query recommendation using query logs in search engines. *EDBT Workshops*, 3268:588–596, 2004.

[31] O.R. Zaiane and A. Strilets. Finding similar queries to satisfy searches based on query traces. *Lecture notes in computer science*, pages 207–216, 2002.

[32] Z. Zhang and O. Nasraoui. Mining search engine query logs for query recommendation. In *Proceedings of the 15th international conference on World Wide Web*, pages 1039–1040. ACM New York, NY, USA, 2006.

[33] R. Jaschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme. Tag recommendations in folksonomies. *Lecture Notes In Computer Science*, 4702:506, 2007.

[34] L.B. Marinho and L. Schmidt-Thieme. Collaborative tag recommendations. *Data Analysis, Machine Learning and Applications*, pages 533–540, 2008.

[35] Y. Song, Z. Zhuang, H. Li, Q. Zhao, J. Li, W.C. Lee, and C.L. Giles. Real-time automatic tag recommendation. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 515–522. ACM New York, NY, USA, 2008.

[36] C. Blaschke and A. Valencia. Automatic ontology construction from the literature. *Genome Informatics Series*, pages 201–213, 2002.

[37] B. Fortuna, M. Grobelnik, and D. Mladenic. System for semi-automatic ontology construction. *Demo at ESWC*, pages 11–14, 2006.

[38] J.U. Kietz, A. Maedche, and R. Volz. A method for semi-automatic ontology acquisition from a corporate intranet. In *EKAW-2000 Workshop "Ontologies and text"*, 2000.

[39] Jöran Beel, Bela Gipp, and Christoph Müller. 'SciPlore MindMapping' – A Tool for Creating Mind Maps Combined with PDF and Reference Management. 2010. to appear.