

{tag}

{/tag}

International Journal of Computer Applications
© 2014 by IJCA Journal

Volume 96 - Number 19

Year of Publication: 2014

Authors:

S. S. Sonawane

P. A. Kulkarni

10.5120/16899-6972

{bibtex}pxc3896972.bib{/bibtex}

Abstract

A common and standard approach to model text document is bag-of-words. This model is suitable for capturing word frequency, however structural and semantic information is ignored. Graph representation is mathematical constructs and can model relationship and structural information effectively. A text can appropriately represented as Graph using vertex as feature term and edge relation can be significant relation between the feature terms. Text representation using Graph model provides computations related to various operations like term weight, ranking which is helpful in many applications in information retrieval. This paper presents a systematic survey of existing work on Graph based representation of text and also focused on Graph based analysis of text document for different operations in information retrieval. In this process taxonomy of Graph based representation and analysis of text document is derived and result of different methods of Graph based text representation and analysis are discussed. The survey results shows that Graph based representation is appropriate way of representing text document and improved result of analysis over traditional model for different text applications.

ences

Refer

- Jae-Yong Chang and Il-Min Kim Analysis and Evaluation of Current Graph-Based Text Mining Researches. *Advanced Science and Technology Letters* Vol. 42, 2013, pp. 100-103.
- Hassan S. , Mihalcea R. , Banea C. , Random-Walk Term Weighting for Improved Text Classification . *IEEE International Conference on Semantic Computing, ICSC-2007*, 2007.
- H. Balinsky, A. Balinsky, and S. Simske, Document Sentences as a Small World, in *Proc. of IEEE SMC 2011*, pp. 9-12,
- Wei Jin and Rohini Srihari, Graph-based text representation and knowledge discovery. In *proceedings of the SAC conference, 2007*, pp 807-811.
- Faguo Zhou, Fan Zhang and Bingru Yang. Graph-based text representation model and its realization. In *Natural Language Proceeding and knowledge Engineering (NLP-KE)*, 2010, pp 1-8.
- Francois Rousseau, Michalis Vazigiannis, Graph-of-word and TW-IDF: New Approach to Ad Hoc IR. *Proceedings of the 22nd ACM international conference on Conference on information and knowledge management 2013*, pp. 59-68.
- Bordag, S. , Heyer, G. , Quasthoff, U. Small worlds of concepts and other principles of semantic search . In T. Bhme, G. Heyer, H. Unger (Eds.), *IICS*, 2003, lecture notes in computer science Vol. 2877, pp. 10-19.
- i Cancho, R. F. , Capocci, A. , Caldarelli, G. Spectral methods cluster words of the same class in a syntactic dependency network. *International Journal of Bifurcation and Chaos*, 2007, 17(7), pp. 2453-2463.
- Chris Biemann, Unsupervised Part-of-Speech Tagging Employing Efficient Graph Clustering *Proceedings of the COLING/ACL 2006 Student Research Workshop*, July 2006, pp. 7-12.
- Dorogovtsev, S. N. , Mendes, J. F. F. Language as an evolving word web. *Proceedings of The Royal Society of London. Series B, Biological Sciences* 268(1485), pp. 2603-2606.
- J. Wu, Z. Xuan, and D. Pan Enhancing Text Representation for Classification Tasks with Semantic Graph Structures. *International Journal of Innovative Computing, Information Control*, Vol. 7, No. 5(B), 2011, pp. 2689-2698.
- Rada Mihalcea and Paul Tarau. TextRank: Bringing order into texts. *Association for Computational Linguistics EMNLP'04*, pp. 404-411.
- Antoon Bronselear, Gabreilla Pasi, An approach to graph-based analysis of textual document , 8th European Society for Fuzzy Logic and Technology, *Proceedings EUSFLAT 2013*, pp. 634-641.
- Lakshmi Ramachandran , Edward F. Gehringer, Determining Degree of Relevance of Reviews Using a Graph-Based Text Representation. *Proceedings of the 2011 IEEE 23rd International Conference on Tools with Artificial Intelligence*, 2011, pp. 442-445.
- Galitsky, B. , Ilvovsky, D. , Kuznetsov, S. O. , and Strok, F. , Matching sets of parse trees for answering multi-sentence questions, *Proc. Recent Advances in Natural Language Processing (RANLP 2013)*, Bulgaria, 2013, pp. 285-294.
- Steyvers, M. , Tenenbaum, J. B. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 2005, Pp. 41-78.
- Svetlana Hensman, Construction of conceptual graph representation of texts. *Proceedings of the Student Research Workshop at HLT-NAACL 2004*, p. 49-54.

- Sajgalk, M. , Barla, M. , Bielikov, M. From ambiguous words to key-concept extraction . In Proceedings of 10th International Workshop on Text-based Information Retrieval at DEXA 2013,IEEE, 2013, pp. 63??67.
- Kozareva, Z. , Riloff, E. , Hovy, E. . Semantic class learning from the web with hyponym pattern linkage graphs. In Proceedings of ACL-08: HLT, Ohio: Association for Computational Linguistics,2008, pp. 1048??1056.
- Motter, A. E et al Topology of the conceptual network of language. Phy. Rev. E. Stat. Nonlin. Soft Matter Phys. , 65, 2002.
- Sigman, M. , Cecchi, G. A. The global organization of the WordNet lexicon. Proceedings of the National Academy of Sciences of the USA,2002, 99, pp. 1742??1747.
- Daisuke Kobayashi, Tomohiro Yoshikawa and Takashi Furuhashi, Visualization and Analytical Support of Questionnaire Free-Texts Data based on HK Graph with Concepts of Words, IEEE International Conference on Fuzzy Systems June 2011, pp. 27??30.
- Rio Blanco, Christina Lioma Graph-based term weighting for information retrieval. Information retrieval, 15(1), February 2012, pp 54??92.
- Gunes Erkan and Dragomir R. Radev. LexRank: Graph based centrality as salience in text summarization. Journal of Artificial Intelligence Research,Volume 22 issue 1, 2004, pp. 457??479.
- Kjetil Valle, Pinar Ozturk Graph-based Representations for Text Classification. India-NorwayWorkshop on Web Concepts and Technologies, October 3rd 2011.
- Jiang, F. Coenen, R. Sanderson and M. Zito, Text classification using graph mining-based feature extraction. Research and Development in Intelligent Systems XXVI, Springer, 2010, pp. 21??34.
- Gamon, M. Graph based text representation for novelty detection In: Proceedings of TextGraphs: the First Workshop on Graph Based Methods for Natural Language Processing, New York City, Association for Computational Linguistics, 2006, pp. 17??24.

Index Terms

Computer Science

Applied Mathematics

Keywords

Information Retrieval Graph Theory Natural Language Processing.

