International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 2 Issue 12, December - 2013

Neural Network Based Multi-Agent Semantic Web Content Mining

Prof. P. A. Khodke
Dept. of Computer Science & Engg.
PRMCEAM, Badnera Amravati University,

Miss. Ashwini H. Bhuskat
PRMCEAM, Badnera
Amravati University,

Abstract:

As with the rapid increase of the huge amount of online information, there is a strong demand for Information retrieving from the web which helps to discover some useful knowledge from Web documents. Single agent is unable to retrieve the information from multiple searching tools. Multi-agent neural network system is an effective solution to large scale Web mining. In this paper, Multi-agent neural network system is applied to the Semantic Web content mining. This paper proposes Multi-agent neural network based framework for mining contents of semantic web, which would provide query relevant knowledge using clustering technique.

Keywords: Multi-agent Systems, Hierarchical Clustering, neural network, web mining, content mining.

1. Introduction

Due to the rapid increase of the huge amount of online information, there is a strong demand for Web text mining which helps people to discover some useful knowledge from Web documents. There are number of search tools available for retrieving the information, information agent called search tools used to build database indices for web information. But these approaches often causes unacceptable access delays under highly competitive access situations as in search tools. As the amount of information stored in the database indices increases, it leads to the unacceptable access delay problem. To overcome this problem meta search services is used.

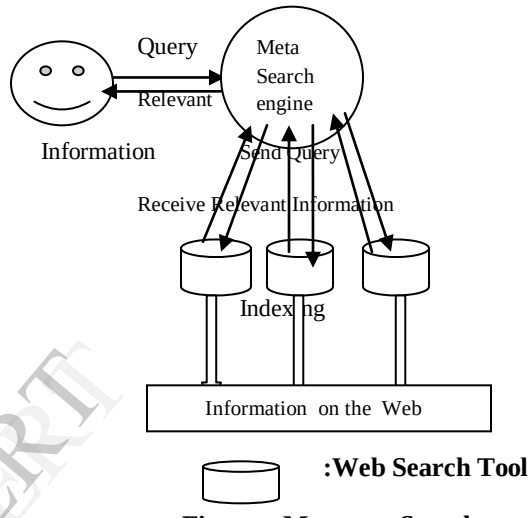

Fig. Meta Search Service

Meta search service retrieve the information from multiple search tools. But this technique also have a problem, i.e. multiple source problem means if there exist multiple search tool then the question arises to which source the given query be submitted to retrieve the information. To solve this question neural network is used. In this, the internal neural network mechanism discovers the search tools from which the associated information for the query is retrieved.

The single agent approach is impractical for large scale information using the Multi-Agent approach. In this IR system number of agents collaboratively retrieves the desired information from the distributed web search tools. Multi agent neural network system is solution to the large web text mining. A multi-agent system is one in which a number of agents cooperates and interact with each other in a complex and distributed environment.

Web mining [11] can be defined as mining of the World Wide Web (WWW) to find useful knowledge about user behavior, content, and

structure of the web. Web Content Mining focuses on extracting knowledge from the contents or their descriptions. It involves techniques for summarizing, classification and clustering of the web contents. It can provide useful and interesting patterns about user needs and contribution behavior. In this paper, the neural network is applied using multi-agent for semantic web content mining. In the present era of WWW, the user is more interested in getting useful, relevant and knowledge oriented contents from the WWW. The paradigm is shifting from demand of information to demand for knowledge. Web content mining when applied on semantic web contents can lead to discovery of knowledge that could be provided to end users to better serve their requirements.

2. Problem Definition

As web mining is complex process to obtain required information from internet, Web content mining process extract the knowledge from their description using the clustering technique. It can provide useful and interesting patterns about user needs and contribution behavior. Users are in need of systems to help them cope with the large amount of information available on the Web. Examples of such systems include Web search engines, meta-search engines, multi-agent systems and information filtering systems. Paper focuses on Multi-agent neural network based framework for mining contents of semantic web, which would provide query relevant knowledge using clustering technique. As the single agent information is unable to retrieve the information from multiple search tools on the web, multi-agent information approach using neural network for semantic web content mining have been used.

3. Related Work

Choi, Y.S. and Yoo, S.I. [1] et al. proposed multi agent learning approach to retrieve the web information using neural network. James A. Freeman, David M. Skapura, [5] uses the neural network for single agent but this have the problem of convergence as the number of search tools increases, the single agent information comes to have difficulty in training its neural network. Such problem is called 'bounded rationality' for single agent approach. This approach is impractical or inefficient for the multiple search tools, so the multi agent hierarchically organized multi-agent approach is proposed. In this system, information agents interacting with each other learn collaboratively about their environment from user's feedback so that it can retrieve the desired information effectively from the distributed Web search tools. In this system, each information agent can dynamically join or leave the collaborative organization and the information sources are subject to asynchronous changes of their themes, contents, and structures.

Web mining [3][11] can be defined as mining of the WorldWideWeb (WWW)to find useful knowledge about user behavior, content, and structure of the web. C.Dimou, A.Batzios, A.L.Symeonidis and P.A.Mitkas [2][4] et al. proposed Agent based Focused crawlers (ABFC), the intelligent miners which selectively seek web pages that are most relevant to input information.

Singh A [13] et al. proposed agent based framework for semantic web content mining which provide classification and clustering of the web content. This paper includes Interface agent, collection agent, content mining agent, Descriptive, semantic metadata agent, clustering agent and ontology database. Interface agent receive query given by user and passes it on to the collection agent explore query, ontology database for the meaning of phrase. Content mining agent work in coordination with indexes to extract knowledge from it. Clustering agent work on the table maintained by the Descriptive and semantic metadata agent. In this paper Clustering agent makes use of Hierarchical clustering algorithm for this purpose. The comprehensiveness of the ontology database can ensure context based information retrieval.

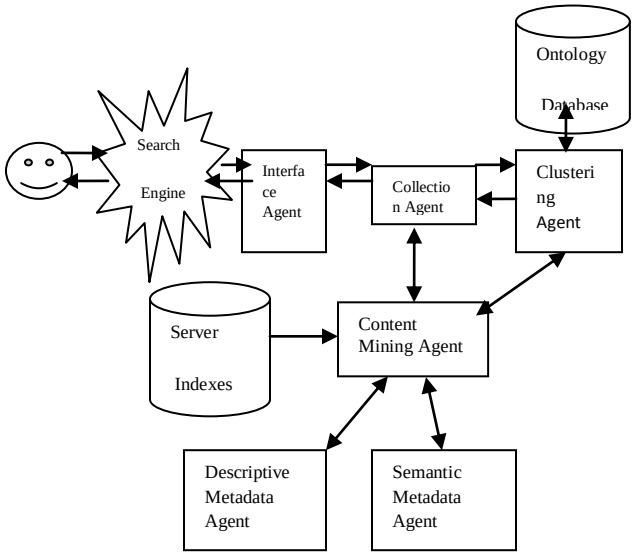

Fig. High level view of Semantic Web Mining System

Singh A., Juneja D. and Sharma A.K., et al.[14] uses Ontology to retrieve the information from web. Here ontology plays an important role in enhancing the efficiency of existing agent based focused crawlers [14]. Ontology is defined as a well organized knowledge scheme that represents high level background knowledge with concepts and relations [14] [15]. Ontology based crawling [10] eliminates simple keyword based crawling method as it introduces semantics/context for improving crawl efficiency in which a keyword is being searched . This work proposes an Ontology Driven Agent Based Focused Crawler that attempts to improve the efficiency of exiting ABFCs by introducing semantics in which a keyword is searched. ABFC follows the context-based approach that analyzes content of web page thereby reducing the redundant information and hence deduces the relevant information from a page. Tremendous growth of web-sites and text and multimedia contents on the WWW (World Wide Web) has lead to demand of a strategy which could provide knowledge from the vast data scattered over different servers and also could make useful predictions for otherwise uncertain user behavior.

Shu Bo and Kak Subhash[12] et al. proposed meta search engine-Anvish has learning neural network to classify and organize the search results. Anvish is further improved by expanding the current binary network into continuous network so that the Web pages can be ranked quantitatively.

4. Conclusion & Future Work

This work gives an overview of Multi-agent neural network system which is applied to the Semantic Web content mining with the aim to provide context based knowledge oriented results to the user. Web mining using neural network will lead to improved performance, reduced network traffic, and better results. However, implementation of this work is still under progress and is left as future work.

5. REFERENCES

[1] Choi, Y.S. and Yoo, S.I. Multi-agent Learning Approach to WWW Information Retrieval using Neural Network, in Proceedings of 1999 ACM International Conference on Intelligent User Interfaces.

[2] C.Dimou, A.Batzios, A.L.Symeonidis and P.A.Mitkas, 'A Multi-agent framework for spiders traversing the semantic web',Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence.

[3]Eirinaki M. & Vazirgiannis M., 'Web Mining for Web Personalization'. Published in ACM Transactions on Internet Technology, Vol.3, No. 1, February 2003, pp. 1-27.

[4] F. Buccafurri, G. Lax, D. Rosaci and D. Ursino, 'Dealing with semantic heterogeneity for improving web usage', Data Knowl. Eng. 58 (3) (2006), pp. 436–465.

[5]James A. Freeman, David M. Skapura, Neural Networks Algorithms, Applications, and Programming Techniques in COMPUTATION AND NEURAL SYSTEMS SERIES, Christof Koch, California Institute of Technology, 1991

[6]Karayannidis N. & Sellis T., 'Hierarchical Clustering for OLAP: The CUBE File Approach'. Published in The VLDB Journal — The International Journal on Very Large Data Bases, Vol. 17, Issue 4, July 2008.

[7] Kosala R. & Blockeel H., 'Web Mining Research: A Survey'. Published in ACM SIGKDD, Vol. 2, Issue 1, July 2000

[8]Nikola K. Kasabov, Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering, The MIT Press Cambridge, Massachusetts London, England,1998.

[9]Lean Yu, Shouyang Wang, and Lai Kin Keung, A Multi-Agent Neural Network System for Web Text Mining, Information Science Reference, *(*2008) *162-183*

[10]S.Ganesh, M. Jayaraj, V.Kalyan & G. Aghila ,'Ontology based Web Crawler', Published in proceedings of the international Conference on Information Technology Coding & Computation, 2004

[11] Sharma K., Shrivastava G. & Kumar V., 'Web Mining: Today and Tommorrow'. In Proceedings of the IEEE 3rd International Conference on Electronics Computer Technology, 2011.

[12]Shu Bo and Kak Subhash, A neural network-based intelligent metasearch engine, Information Sciences 120 (1999) 1-11

[13]Singh A, Agent Based Framework for Semantic Web Content Mining, International Journal of Advancements in Technology 3(April 12)

[14]Singh A., Juneja D. and Sharma A.K., 'Design of Ontology-Driven Agent based Focused Crawlers'. In proceedings of 3rd International Conference on Intelligent Systems & Networks (IISN-2009),Organized by Institute of Science and Technology, Klawad, 14 -16 Feb 2009, pp. 178-181. Available online in ECONOMICS OF NETWORKS ABSTRACTS, Volume 2, No. 8: Jan 25, 2010.

[15]Singh A., Juneja D., Sharma A.K., 'General Design Structure of Ontological Databases in Semantic Web'. Published in International Journal of Engineering, Science & Technology, Vol. 2, Issue 5, pp. 1227-1232, 2010

[16]Zhan L. & Zhijing L., 'Web Mining based on Multi-Agents'. Published in proceedings of Fifth International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'03), 2003.