WEB INTELLIGENCE TECHNOLOGY AND APPLICATIONS

Author: D.Hari Prashanth

Abstract-As the rapid growth of internet and web related technologies continue to take over the entire world in almost all the fields of information technology like data mining, information security, e-based applications etc., web intelligence is the particular field of research and development to explore the fundamental roles as well as practical applications of artificial intelligence which is slowly developing throughout the world in order to create next generation of products, services and frameworks based on internet.

Index Terms-Agent based systems, semantic web, social intelligence design, web farming, web information retrieval, web mining, and web spiders.   .

----

1.   Introduction:

The web intelligence technology is the foremost important aspects of the application of artificial intelligence in the field of web based applications. The future of web networks, and also development of extraordinary devices and machines were completely based on artificial intelligence technology. This will be one of the technology milestone reached by human knowledge. The future applications of artificial intelligence in web technology are semantic web, social intelligence design; web farming, web information retrieval, web mining and web spiders etc., the applications which were stated were not limited to the above but also in building of high end web security systems, anti-hacking security, business markets and many more. Here the various present and future applications of web intelligence technology were detailed below.

Author name: D.Hari Prashanth currently pursuing Masters in Software Engineering MS(SE) At VIT UNIVERSITY, Vellore, INDIA. Ph. +91-9789640904 e-mail: hprashanth610@gmail.com

2.   WEB INTELLIGENCE TECHNOLOGY DESIGN AND DEVELOPMENTS:

Web intelligence technology is combination of web analytics which examines how the visitors view and interact with the site's pages and features thedevelopment of this technique started from the introduction of new internet based services like online banking, online shopping, e-commerce, etc., the further enhanced development of web intelligence technology by high speed networks, system security, secure transactions will enable safer means of using the features offered by web and also can earn the peoples reputation which contributes   to   global   market.

Figure1 Evolution of Web technology

3. PRESENT TECHNOLOGY IN WEB SYSTEMS AND IT'S SECURITY:

At present the web developers have somehow managed to design high-tech web systems which enable the users to share, retrieve the confidential data and also customer related services like net-banking, online buying services etc., but some people who can really misuse the services could disturb the normal functioning and inter-dependency of both users and service providers and also the internet security systems as a whole like hacking, piracy, sharing of information between terrorists organization which cannot be detected by the international security systems spots out that there are still more techniques to be developed to design a powerful web security systems. The web intelligence systems needs to be further developed in order to offer full-fledged resistant to hacking, piracy etc. the spam control and detecting the location of the spam, anti-spam systems can be developed to deal with this situations.

4. WEB INTELLIGENCE SYSTEMS:

High frequency signal transfer systems, automatic detection systems for finding system malfunction in certain aspects. Even high-end flaw determining or failure of the functioning of certain parts of aircraft or submarine or a ship which sends a warning signal automatically within seconds of malfunction. Various security systems involved in opening and closing of security doors in banks, laser detection system etc. were few of the intelligence systems used in present.

5. WEB FARMING:

Web farming also known as the server farming or server cluster is the collection of computer servers usually maintained by an enterprise to accomplish server needs far beyond the capability of one machine. Server farms often have backup servers, which can take over the function of the primary servers in the event of the server power failure. Server farms are typically collocated with the network switches and/or routers which enable the communication between the different parts of the cluster and the users of the cluster. The computer routes, power supplies and related electronics which are typically mounted on 19-inch racks in a server room or data centre. Server farms are commonly used for the cluster computing. Many modern super computers comprise giant server farms of high speed processors connected by either gigabit Ethernet or custom inter connects such as infiniband or myrinet server farms are most increasingly being used instead of or in addition to mainframe computers.
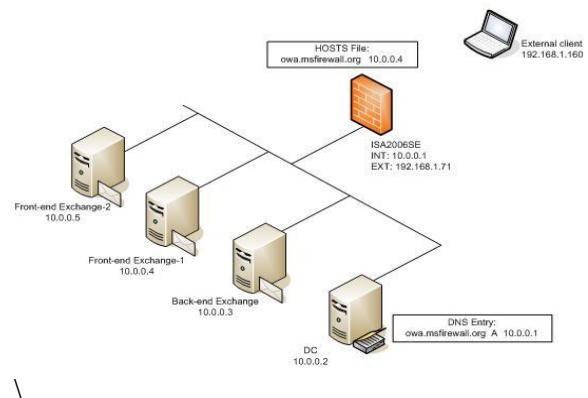

\

Figure2 Web Farming Methodology

6. WEB SPIDERS:

When most people talk about internet search engines, they really mean World Wide Web search engines. Before the web became the most visible part of the internet, there were already search engines in place to help people find information onthe net. Programs like "goper" and "archie" kept indexes of the files stored on servers connected to the internet, and dramatically reduced the amount of time required to find programs and documents. In the late 1980's, getting serious value from the internet meant knowing how to use gopher, archie, venoria and the rest.

Before the search engine can tell us where a file or document is, it must be found. To find the information on the hundreds of millions of web pages that exist, a search engine that employs special software robots, called 'Spiders' to build list of words found on the web sites. When a spider is building its list it is called as Web crawling. In order to maintain the useful lists of words, a search engine's spiders have to look at lots of pages. There are certain starting points in which how a web spider travels over web; the usual starting point is the lists of heavily used servers and very popular pages. The spider will begin with a popular site, indexing the words on its pages and following every link found within a site. In this way, the spidering system quickly begins to travel, spreading out across the most widely used portions of web. These web spiders can crawl over 100 pages per second, generating around 600 kilobytes of data each second.

7. WEB MINING:

Web mining is the application of the data mining techniques to discover patterns from the web. According to the analysis targets, web mining can be divided into three categories Web usage mining, Web content mining and Web structure mining.

8. WEB USAGE MINING:

Web usage mining is the process of extracting useful information from server logs i.e., users history. Web usage mining is the process offinding out what users are looking for internet. Some users might be looking at only textual data whereas some others might be interested in multimedia data.

9. WEB STRUCTURE MINING:

Web structure mining is the process of graph theory to analyse the node and connection structure of the web site. According to the type of web structural data, web structure mining can be divided into two kinds:

10. EXTRACTION PATTERNS FROM HYPERLINKS IN THE WEB:

A hyperlink is nothing but a structural component that connects the web page to a different location. The content present in the web will be displayed with respect to the user's search the extraction of information is purely based on the words or sentence typed by the user during search module.

11. MINING THE DOCUMENT STRUCTURE:

Analysis of a tree like structure of pages structures to describe HTML or XML tag usage. Mining of the document structure is partially related to web and partially based on the nature of the document presented by the user and displaying results related to structure.

12. WEB CONTENT MINING:

Mining, extraction and integration of useful data, information and knowledge from web page contents, the heterogeneity and the lack of structure that permeates much of ever expanding information sources on the world wide web, such as hypertext documents, makes automated discovery, organization, and search and indexing tools of internet and the world wide web such as Lycos, Alta Vista, Web Crawler, ALIWEB, MetaCrawler and others provide some comfort to the users, but they do not provide any structural information nor categorize, filter, or, interpret document. In recent years these factors have

prompted researchers to develop more intelligent tools for information retrieval, such as intelligent web agents, as well as to extend database and mining techniques to provide higher level of organisation for semi-structured data available on the web.

13. SECURITY THREATS TO WEB SYSTEMS:

Even the developments of the web technologies reached greater heights and also developing day by day there also grows the security threats to the web intelligence systems day by day. Advance the technology is, advance the security threats too. This can be simply understood by taking an example of the communication method of various terrorist organisations all over the world, their ability to contact, send and receive information between each other which can't be detected by the world security systems. And also the criminal organisations know how to hack sensitive information belonging to the government sometimes, which demonstrates the weakness of the security systems at present.

Another example, if we take into account the social networking such as Facebook at the beginning we all able to feel comfortable to connect with each other all over the globe. But now a days many fake profiles in Facebook and also insecure money transfer of buying few facilities in it continues to threaten the users all over the world. There isn't enough security against the Facebookhackers andfake profiles in it. These cases can also be considered as the weakness in web securities.

There are also certain devices like signal jammer, spam intrusion devices, and also advances signal hacking devices like drone hacker system etc. the hacker can eventually hack the drone aircraft(unmanned spy plane) and changes the command sent to it by its head. These are very cheap devices to build and operate but the losses caused by these devices were equally high if it was used as an act of criminal offence.

14. AGENT BASED SYSTEMS:

An agent-based model (ABM) (also sometimes related to the term multi-agent system or multi-agent simulation) is a class of computational models for simulating the actions and interactions of autonomous agents (both individual and collective entities such as organizations or groups) with a view to assessing their effects on the system as a whole. It combines elements of game theory, complex systems, emergence, computational sociology, multi-agent systems, and evolutionary programming. Monte Carlo Methods are used to introduce randomness. ABMs are also called individual-based models. A review of recent literature on individual-based models, agent-based models and multi-agent systems shows that ABMs are used on non-computing related scientific domains including Life Sciences, Ecological Sciences and Social Sciences.

The models simulate the simultaneous operations and interactions of multiple agents, in an attempt to re-create and predict the appearance of complex phenomena. The process is one of emergence from the lower (micro) level of systems to a higher (macro) level. As such, a key notion is that simple behavioural rules generate complex behaviour. This principle, known as K.I.S.S. ("Keep it simple and short") is extensively adopted in the modelling community. Another central tenet is that the whole is greater than the sum of the parts. Individual agents are typically characterized as bounded rational, presumed to be acting in what they perceive as their own interests, such as reproduction, economic benefit, or social status, using heuristics or simple decision-making rules. ABM agents may experience "learning", adaptation, and reproduction.

15. SOCIAL INTELLIGENCE DESIGN:

Web intelligence reflects the view that eventually we will build a totally new kind of collective intelligence on the Web computing infrastructure. To reach this goal we must solve several major problems. For example, embedding Web computing into our everyday lives and society

poses a more difficult problem than engineers might think. Because new technologies often consume financial resources without providing a comparable benefit, we must pay close attention to the social aspects of intelligence and how Web computing can augment knowledge processes, an attitude that underlies social intelligence design. Computer-supported collaborative work takes a similar approach, focusing on well-structured, goal-oriented groups. Social intelligence design, on the other hand, highlights collective knowledge processes in informal, loosely coupled groups. It thus focuses not only on technological development for Web intelligence but also on the design and analysis of a social framework for embedding Web intelligence into everyday life.
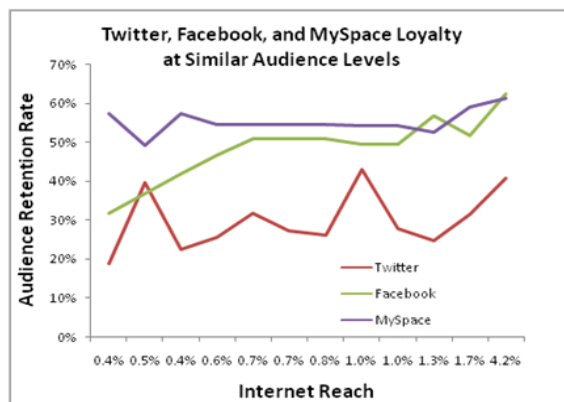

Figure3 Levels of users using Social networking sites Blue: FB, Green: Twitter, Red: MySpace

16. CONCLUSION:

Thereby the artificial intelligence technology in the field of web has reached greater heights at present and also there should be precautions to illegal usage of precious resources in web. The web intelligence technology is developing day by day and also trying its best to offer quality services to the users all over the world. Even super computers similar to the Human brain can also be developed within few years by 2020 it is expected to be deployed.

Details about Author:
D.HariPrashanth,
M.S.SoftwareEngg,
VIT UNIVERSITY, Vellore.
INDIA.

Residence address:
105,RoyanKuttai Street,
Big Kanchipuram,
Tamil nadu-631502.
Email: hprashanth610@gmail.com
Ph: +91-9789640904.