 ISSN : 2229-4333(Print) | ISSN : 0976-8491(Online)

A Survey of Agent-based Personalized Semantic information retrieval

[1]Dr. M.Thangaraj, [2]M. Chamundeeswari
[1]Dept. of Computer Science, Madurai Kamarj University, Palkalai Nagar, Madurai, India
[2]Dept. of Computer Science, V.V.V College for Women(A), Virudhunagar, India

Abstract
The Aim of personalized search is to provide users with information tailored to their individual contexts. Web personalization is the process of customizing a Web site to the needs of specific users, taking advantage of the knowledge acquired from the analysis of the user's navigational behavior (usage data) in correlation with other information collected in the Web context, namely, structure, content and user profile data. Due to the explosive growth of the Web, the domain of Web personalization has gained great momentum both in the research and commercial areas. Current Web search engines are built to serve all users, independent of the special needs of any individual user. Personalization of Web search is to carry out retrieval for each user incorporating his/her interests. The Semantic Web and Multi-Agent are effective means for constructing information retrieval systems.Despite a great deal of research, a number of challenges still exist before making Semantic Web and agent-based computing a widely accepted in information retrieval practice. The goal of this survey is study of the main concepts, existing methods, and practices of this area.

Key Words and Phrases
Multi-Agent, Web personalization, user profiling, Genetic algorithm, neural network.

I. Introduction
The explosive growth of documents in the Web makes it difficult to determine the most relevant documents for a particular user, given a general query. Recent search engines rank pages by combining traditional information retrieval techniques based on page content, such as the word vector space [63, 71] with link analysis techniques based on the hypertext structure of the Web [10, 25]. Traditional search engine has dealt with searching information on the web to a large extent, but it also has some problems at present [43]. The web information has enlarged from quantity to types, showing the trend of exponential growth, so the search engine cannot index all the pages; The web information has changed dynamically, so the search engine cannot be sure to update in time; Traditional search engine cannot meet the increasing need day by day that people want personal service for information retrieve; Search engine requires hardware owning more storage capacities, even hundreds of GB, and more servers. Besides the above stated problem a recent research has shown that only 13% of search engines show personalization characteristics [18]. Hence web personalization is one of the promising approaches to tackle this problem by adapting the content and structure of websites to the needs of the users by taking advantage of the knowledge acquired from the analysis of the users' access behaviors. Nowadays, various forms of digital contents like documents and web pages have been growing exponentially. In face of the overwhelming information volumes, people are struggling with information overload rather than its shortage. In order to handle multitudinous digital contents, information retrieval and related theories and technologies for the acquisition, management, and application of digital contents have risen as an important issue. But, it is well known that existing information retrieval systems based entirely on keywords have serious limitations and has led to the following problems: (1) Semantics in users' queries and documents cannot be extracted based on keywords. Most of the existing retrieval systems use keywords search and directory search. And the users' queries and the information in internet identified through keywords did not always meet the users' requirements. Because of synonym and polysemy in human language, information retrieval through keywords always left out other information with similar semantics; (2)Traditional retrieval system is lacking interaction with the users. According to some researches, the information of users' behavior can improve the rate of retrieval precision and retrieval recall. Even though some search engines record a large number of user behavior information through the log files, but they do not effectively use the information to establish the feedback mechanism to guide information retrieval and communicate with users. (3) Poor sort out search results. The results most search engines return are lack of precision. The users' search behaviors often bring a lot of spam. According to the evaluation of experts, the rate of relevant results the current major search engines return is less than 45%. At the same time, because the search matching algorithms are not ideal, so the search engines sort search results too rough. These deficiencies have restricted the development of information retrieval. The rest of the paper is organized as follows: Section II explain the architecture and function of Information retrieval based on multi-agent model related in this area. Then the semantic search directions are presented in Section III. Section IV provides a tour around the most well known applications of web personalization both at a research and a commercial level. Finally the conclusions are made in Section V.

II. Information Retrieval System Model and Multi-agent
Multi-agent system is composed of many different agents with different functionality. Multi-agent technology can be applied to the research of information retrieval. The combination of information retrieval and Multi-agent technology has the following features: (1) Adaptability: Based on the information of users' behaviors in internet, Agent can discover the users' interest, reason the user's needs and establish personalized documentation for each user; (2) Initiative: Agent can initiatively retrieve the corresponding information based on users' demand, and even can monitor the changes of information sources; (3) Collaborative: Agents can share the information with other Agents. For example, a user's Agent can access to a lot of useful information from other users' Agents that have the same data about users' interest . The information retrieval system (FIG. 1) designed to utilizes RDF and Multi-agent technology as the basis to transform users' queries and documents in database. The semantic pattern as triples, have Subject, Property and Object, so as to process, recognize , and match semantic. Users can perform semantic-based and Multi-agent-based information retrieval in the following process: First of all, users should submit their queries to User Agent, User Agent analyzed and determined user's characters about retrieval, and the query record will be stored in User Database, and the query will be transmitted to Extract Agent. Extract Agent will extract the semantic patterns in

queries which can represent actual users' requirement. The next step was finished by Semantic Matching Agent, it will complete matching user's semantics with document semantic which stored in document semantic database, and the results of the feedback to User Agent. User Agent will display search results based on user's character information in User Database. Based on the users' different requirements, Information Gathering Agent can select different types of web robot to collect information in the internet and monitor the robot. The documents gathered by Information Gathering Agent will be analyzed and extracted semantics by Semantic Extract Agent, and the semantics in the document will be stored in Document Database.

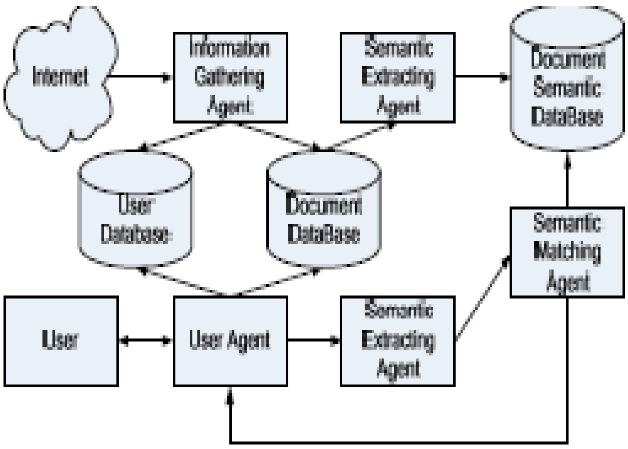

Fig. 1: Information Retrieval System Model.

III. Research Directions:

The ultimate goal of any user-adaptive system is to provide users with what they need without them asking for it explicitly [37] . Personalization is a central technology used in these system. In the context of the Web, personalization implies the delivery of dynamic content, such as textual elements, links, advertisement, product recommendations, etc., that are tailored to needs or interests of a particular user or a segment of users. Fig. 2. gives general structure of Web Personalization.

A. Using Mining Techniques

Traditional approaches to personalization using Mining have included content-based, collaborative, and rule-based filtering systems.[ ] Each of these approaches is distinguished by the specific type of data collected to construct user profiles, and by the specific type of algorithmic approach used to provide personalized content. Generally, the process

Fig. 2: General Structure of Personalization

of personalization consists of a data collection phase in which the information pertaining to user interests is obtained and a learning phase in which user profiles are constructed from the data collected. Learning from data can be classified into memory based (also known as lazy) learning and model based (or eager) learning depending on whether the learning is done online while the system is performing the personalization tasks or offline using training data. Standard user-based collaborative filtering and most content based filtering systems that use lazy learning algorithms are examples of the memory-based approach to personalization, while item-based and other collaborative filtering approaches that learn models prior to deployment are examples of model-based personalization systems. Memory based systems simply memorize all the data and generalize from it at the time of generating recommendations. They are therefore more susceptible to scalability issues. Model-based approaches, that perform the computationally expensive learning phase offline, generally tend to scale better than memory based systems during the online deployment stage. On the other hand, as more data is collected, memory based systems are generally better at adapting to changes in user interests compared to model based techniques in which model must either be incremental or be rebuilt in order to account for the new data. These advantages and shortcomings have led to an extensive body of research and practice comprised of a variety of personalization or recommender systems that generally fall into the aforementioned categories.

1. Rule-Based Personalization Systems

Rule-based filtering systems [44] rely on manually or automatically generated decision rules that are used to recommend items to users. Many existing e-commerce Web sites that employ personalization or recommendation technologies use manual rule-based systems. Such systems allow Web site administrators to specify rules, often based on demographic, psychographic, or other personal characteristics of users. In some cases, the rules may be highly domain dependent and reflect particular business objectives of the Web site. The rules are used to affect the content served to a user whose profile satisfies one or more rule conditions. Like most rule-based systems, this type of personalization relies heavily on knowledge engineering by system designers to construct a rule base in accordance to the specific characteristics of the domain or market research. The user profiles are generally obtained through explicit interactions with users. Some research has focused on machine learning techniques for classifying users into one of several categories based on their demographic attributes, and therefore, automatically derive decision rules that can be used for personalization.

The primary drawbacks of rule-based filtering techniques, in addition to the usual knowledge engineering bottleneck problem, emanate from the methods used for the generation of user profiles. The input is usually the subjective description of users or their interests by the users themselves, and thus is prone to bias. Furthermore, the profiles are often static, and thus the system performance degrades over time as the profiles age.

2. Content-Based Filtering Systems

In Content-based filtering systems, a user profile represent the content descriptions of items in which that user has previously expressed interest. The content descriptions of items are represented by a set of features or attributes that characterize that item. The recommendation generation task in such systems usually involves the comparison of extracted features from unseen or unrated items with content descriptions in the user profile.

Items that are considered sufficiently similar to the user profile are recommended to the user.

In most content-based filtering systems, particularly those used on the Web and in e-commerce applications, the content descriptions are textual features extracted from Web pages or product descriptions. As such, these systems often rely on well-known document modeling techniques with roots in information retrieval[63] and information filtering research. Both user profiles, as well as, items themselves, as represented as weighted term vectors (e.g., based on TF.IDF term-weighting model [63]) Predictions of user interest in a particular item can be derived based on the computation of vector similarities (e.g., using the Cosine similarity measure) or using probabilistic approaches such as Bayesian classification. Furthermore, in contrast with approaches based on collaborative filtering, the profiles are individual in nature, built only from features associated with items previously seen or rated by the active user.

Examples of early personalized agents using this approach include Letizia [27], NewsWeeder [28] , Personal WebWatcher [35], InfoFinder [26], Syskill and Webert [55] and the na¨ıve Bayes nearest neighbor approach used by Schwab et al.A survey of the commonly used text-learning techniques in the context of content-based filtering can be found in [33].

The primary drawback of content-based filtering systems is their tendency to over- specialize the item selection since profiles are solely based on the user's previous rating of items. User studies have shown that users find online recommenders most useful when they recommend unexpected items [65] suggesting that using content similarity alone may result in missing important "pragmatic" relationships among Web objects such as their common or complementary utility in the context of a particular task. Furthermore, content-based filtering requires that items can be represented effectively using extracted textual features which are not always practically given the heterogeneous nature of Web data.

3. Contextual Search

A new approach for the search named Just-in-Time IR (JITIR) [58] where the information system proactively suggests information based on a person's working context. Basically, the system continuously monitors the user's interaction with the software, such as typing in a word processor or surfing with Internet browsers, in a non-intrusive manner, automatically identifying their information needs and retrieving useful documents without requiring any action by the user. The retrieval process can exploit a variety of data sources, i.e., any number of pre-indexed databases of documents, such as e-mails or commercial databases of articles.

The JITIR approach combines the alerting approach of Google Alert, with personalization based on the events inside the user's local working context. Alerting pushes information related to predefined sets of topics toward the user regardless of his current activity, usually requiring a sudden change of user attention. By means of a dynamic user profile kept updated according to changes of the local working contexts, JITIR provides the information tailored to the current user activity.

4. Personalization Based on Search Histories

User queries are undoubtedly an important source in recognizing the information needs and personalizing the human-computer interaction. A search engine is able to access and process all this information in a non-invasive way, i.e., without installing external proxy servers or client desktop bots, therefore it can tailor the query results based on the previous requests and interests [30]. Simple log-in forms and cookies can be employed in order to identify the user and the related click streams data instead of complex heuristics based on IPs, last access times or user agents data, which cannot be considered entirely accurate. Approaches based on search history can be organized in two groups. Offline approaches exploit history information in a distinct preprocessing step, usually analyzing relationships between queries and documents visited by users. Online approaches capture these data as soon as they are available, affecting user models and providing personalized results taking into consideration the last interactions of the user. Even though the latter approaches provide updated suggestions, an offline approach can implement more complex algorithms because there are usually less urgent time constraints than in an online one.

5. Personalization Based on Rich Representations of User Needs

This section presents three prototypes of personalized search systems based on complex representations of user needs constructed using explicit feedback: ifWeb, Wifs and InfoWeb.

ifWeb [4] is a user model-based intelligent agent capable of supporting the user in Web navigation, retrieval, and filtering of documents taking into account specific information needs expressed by the user with keywords, free-text descriptions, and Web document examples. The ifWeb system exploits semantic networks in order to create the user profile. The ifWeb prototype also performs autonomous focused crawling collecting and classifying interesting documents. The user profile is updated and refined by explicit relevance feedback provided by the user: ifWeb presents a collection of documents to the user (usually no more than ten for each feedback session), who then explicitly selects the ones that meet his needs. Then, ifWeb autonomously extracts the information necessary to update the user profile from the documents on which the user expressed some positive feedback.

The Wifs system described is capable of filtering HTML or text documents retrieved by the search engine ALTAVISTA3 in response to a query input by the user. This system evaluates and reorders page links returned by the search engine, taking into account the user model of the user who typed in the query. The user can provide feedback on the viewed documents, and the system uses that feedback to update the user model accordingly. In short, the user model consists of a frame whose slots contain terms (topics), each one associated with other terms (co-keywords) which form a simple semantic network. Slot terms, that is, the topics, must be selected from those contained in a Terms Data Base (TDB), created a priori by experts who select the terms deemed most relevant for the pertinent domain. The filtering system is based on a content-based approach, where the documents retrieved by ALTAVISTA are assessed solely according to their contents. The document modeling is not based on `traditional IR techniques, such as the Vector Space Model, due to the high variability of Web information sources.

InfoWeb [40] A further approach to personalization is taken by InfoWeb , an interactive system developed for adaptive content-based retrieval of documents belonging to Web digital libraries The distinctive characteristic of InfoWeb is its mechanism for the creation and management of a stereotype knowledge base, and its use for user modeling. A stereotype contains the vector representation of the most significant document belonging to a specific category of users, initially defined by a domain expert. The system helps the domain expert build the stereotypes through a k-means clustering technique , which is applied to the whole document collection in an off-line phase. InfoWeb uses the stereotypes exclusively for the construction of the initial user model. The user's profile evolves over time in accordance to the user's information needs, formulated through queries, using an explicit relevance feedback algorithm that allows the user to provide an assessment of the documents retrieved by the system.

6. Adaptive Result Clustering

Several Web search engines organize results into folders by grouping pages about the same topics together, for example CLUSTY and KARTOO. The former is based on the VIV´ISIMO clustering engine that arranges results in the style of folders and subfolders. In addition to the traditional HTML layout, the meta search engine KARTOO organizes the returned resources on a graphic interactive map. When the user moves the pointer over those resources, a brief description of the site appears. The size of the icons corresponds to the relevance of the site to the given query. In the Web domain, clustering is usually performed after the retrieval of the Query results, therefore the whole process must be fast enough to be computed interactively, while the user waits for results. For this reason, the clustering algorithms usually take document snippets instead of whole documents as a representation of page contents. Since, unlike classification, clustering does not require pre-defined categories, the number and the organization of the clusters should be chosen so that the user can navigate easily through them. Finally, clustering should provide concise and accurate cluster descriptions that allow the user to find the most useful ones, even in case of polysemous or misleading queries.

Hyperlink-Based Personalization : Based on one of the enhanced versions of the PageRank algorithm[16] PROS that provides personalized ranking of Web pages according to user profiles built automatically, using user bookmarks or frequently-visited page sets. In short, the PageRank (PR) is a vote assigned to a page A collected from all the pages T1..Tn on theWeb that point to it. It represents the importance of the page pointed to, where a link to a page counts as a vote of support. The two algorithms, HubFinder and HubRank use the Web link structure to find topic-related pages and to rank the Web pages needed to build the user profile for the Personalized PageRank algorithm. The pages judged more interesting are collected and the expanded sets are built automatically, using bookmarks and the most visited pages. The process does not require explicit activity by the user.

7. Collaborative Filtering (CF)

Most collaborative filtering methods using data mining techniques fall into the following categories: Memory-based algorithm , Model-based algorithm and Hybrid recommenders.

8. Memory-Based Collaborative Filtering

Memory-based CF algorithms use the entire or a sample of the user-item database to generate a prediction. Every user is part of a group of people with similar interests. By identifying the so-called neighbors of a new user (or active user), a prediction of preferences on new items for him or her can be produced. The neighborhood-based CF algorithm, a prevalent memory-based CF algorithm, uses the following steps: calculate the similarity or weight, Wi,j , which reflects distance, correlation, or weight, between two users or two items, i and j; produce a prediction for the active user by taking the weighted average of all the ratings of the user or item on a certain item or user, or using a simple weighted average [61]. When the task is to generate a top-N recommendation, we need to find k most similar users or items (nearest neighbors) after computing the similarities, then aggregate the neighbors to get the top-N most frequent items as the recommendation. For item-based CF algorithms, the basic idea of the similarity computation between item i and item j is first to work on the users who have rated both of these items and then to apply a similarity computation to determine the similarity, Wi, j , between the two co-rated items of the users .For a user-based CF algorithm, first calculate the similarity, Wu,v, between the users u and v who have both rated the same items. Commonly used methods to compute similarly or weight between users or items are: Correlation-based and vector cosine-based similarities.

9. Model-based Collaborative Filtering

The design and development of models (such as machine learning, data mining algorithms) can allow the system to learn to recognize complex patterns based on the training data, and then make intelligent predictions for the collaborative filtering tasks for test data or real-world data, based on the learned models. Model-based CF algorithms [87] , such as Bayesian models, clustering models, and dependency networks, have been investigated to solve the shortcomings of memory-based CF algorithms. Usually, classification algorithms can be used as CF models if the user ratings are categorical, and regression models and SVD methods and be used for numerical ratings.

10. Hybrid Collaborative Filtering Techniques

Hybrid CF systems combine CF with other recommendation techniques (typically with content-based systems) to make predictions or recommendations. Content-based recommender systems make recommendations by analyzing the content of textual information, such as documents, URLs, news messages, web logs, item descriptions, and profiles about users' tastes, preferences, and needs, and finding regularities in the content [55]. Many elements contribute to the importance of the textual content, such as observed browsing features of the words or pages (e.g., term frequency and inverse document frequency), and similarity between items a user liked in the past[73]. A content-based recommender then uses heuristic methods or classification algorithms to make recommendations [53]. Content-based techniques have the start-up problem, in which they must have enough information to build a reliable classifier. Also, they are limited by the features explicitly associated with the objects they recommend (sometimes these features are hard to extract), while collaborative filtering can make recommendations without any descriptive data. Also, content-based techniques have the overspecialization problem, that is, they can only recommend items that score highly against a user's profile or his/her rating history [3,17] .

B. Neural Network Based Information Retrieval

[72] The single-agent approach where one information agent learns about all search tools may be inefficient and impractical for the large-scale IR environment that has quite a number of search tools. A neural net agent sends a given query to its directly accessible search tool or neighboring neural net agent (what we call a cooperator) and then receives the information relevant to that query. A neural net agent is defined by the 6-tuple $\alpha = < QB, IM, RF, SG, LM, QS >$. Each component of the tuple is defined as : Query Broadcaster (QB) broadcasts a given query to all cooperators of its neural net agent in order to receive all information relevant to that query from them. Information Merger (IM) merges the information submitted by the cooperators of its neural net agent and then presents it to the provider of query. Relevance Feedback (RF) receives the user's judgment for the information presented by the IM for a given query q and then

ISSN : 2229-4333(Print) | ISSN : 0976-8491(Online)

generates a binary vector representation $c_q=(c_q1, c_q2, \ldots, c_{qm})$ where if S is an ordered set of all cooperators, S = <A1, A2, …, AM>, of the neural net agent that the RF belongs to, then for i=1,2,.m , C_{iq} = {1 if information submitted by A_i is judged to relevant to the given query or 0 otherwise} Signature Generator (SG) transforms a query expressed as a character string into its signature. The signature of a certain character string is a binary vector representation of fixed size that is generated by applying a hashing function to that character string. Learning Mechanism (LM) is used to learn from user's relevance feedback and to recall when retrieving information. Each agent has its Learning Mechanism in the form of the neural network associative memory. Back propagation Neural Network (BPN) is adopted to take advantage of its learning and generalization properties. Query Sender (QS) sends a given query selectively to the cooperators of its neural net agent according to the output of the BPN recall phase. The training cost of neural net agent is one of the important factors for the finding the efficiency of the system and it may be heavily affected by the number of cooperators that a neural net agent learns about as well as the number of training queries. Further more, if new cooperators were added into the existing system, their neural net agent should be retrained for all coopertors with all training queries only to lean about those new cooperators. To overcome these difficulties , the multiagent approach may be used , in which IR knowledge is distributed over a number of neural net agents that are hierarchically organized. A multi-agent IR system M = <A,S,R> is hierarchically organized if M is represented as a balanced tree where the elements of A are interior nodes and the elements of S are leaf nodes

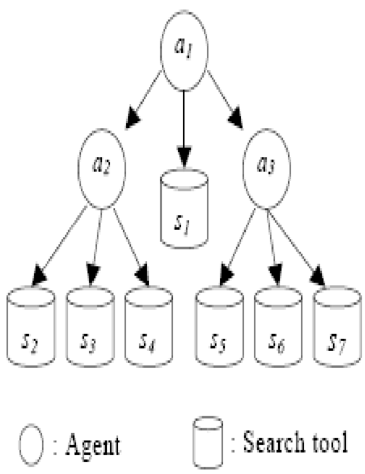

Fig. 3: Hierarchical organization of multi-agent IR system

For a neural net agent to be trained, all of its cooperators that are not a search tool should be trained beforehand. If not so, the neural net agent can be trained with the inaccurate information submitted by some cooperators. Therefore, the neural net agents of hierarchically organized multi-agent IR system should be trained in the post-order traverse order of the tree representing that IR system. Each neural net agent in a multi-agent IR system sends a given query to some of its cooperators according to its trained BPN and then presents the information submitted by them to the provider of that query by the information retrieval procedure of neural net agent as:

Step 1: SG transforms a given query q into its signature sq.

Step 2: LM (BPN) activated by sq produces an output vector oq by its recall phase.

Step 3: QS selects cooperators based on oq and sends q to the selected cooperators.

Step 4: IM merges all information submitted by the cooperators in step 3, and presents the information to the provider of q.

In step 2 and 3, a neural net agent locates its cooperators that is expected to give the desired information for a given query using the IR knowledge stored as the link-weight matrices of BPN, and sends that query to those cooperators to retrieve information. Multi-agent IR system provides a method for retrieving the desired information effectively from the distributed Web. Hierarchical organization of Multi-agent IR system makes it possible to scale up the number of information agents and information sources without radically incurring additional training cost because of the limited effects to the hierarchical agent organization. So, we need agent that can dynamically join or leave the collaborative organization and the information sources are subject to asynchronous changes of their themes, contents, and structures and also looking into an automated method to extract training queries from information sources.

C. Information Retrieval Using Genetic Algorithms

Genetic algorithms may be used in cases where the search space is very complex, and hard to understand, no mathematical analysis is available, classical search methods fail to offer an answer. The most important benefits of using Genetic Algorithms are that they can handle many constraint types and objectives and they are able to discover good solutions rapidly for difficult high-dimensional problems. The agent technology and genetic algorithms are enabling individuals and business to take advantage of the new and powerful medium of the World Wide Web. A framework for clever information retrieving may be seen like a multi-agent system that has specific components, each one with its goal. In the most cases, the information must be retrieved from the internet, a space where an amount of information must be searched for newer results. Referring to the clever retrieving framework of information in Internet (Fig 4), it is possible to observe the presence of mobile agents making the research activity starting from a set of input information, seeking for links of such information and evaluating them through some clever evaluation algorithms as the fuzzy logic, mainly used for the searching for web pages similar to the entry ones, given by the user.

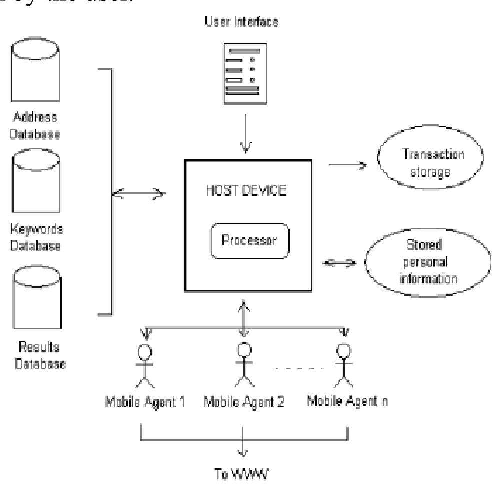

Fig.4 :

The mobile agents using genetic algorithms have, as their main purpose, the reduction of the useless information transfer, which yields the decrease of the net traffic. The "Processor" is the application responsible to produce and to classify the URLs database and to manage the interfacing and the execution programs. The task of the "user interface" is to initialize the requests for extraction made by the user and to show the achieved results. Transaction storage is meant to register all the transaction made by the users. Base on these records a user profile may be done.

The advantage of using genetic algorithm in this system is contributed to decrease the time response for obtaining the final results. Fuzzy c-means clustering algorithm may be used to find the documents through the links of elements of the current set(set of input documents representing the pool of the current solution) are compared with the input objects for similitude through the fuzzy analysis algorithm and choose best one. The fuzzy c-means algorithm is very similar to the k-means algorithm:

- choose a number of clusters;
- assign randomly to each point coefficients for being in the clusters.
- repeat until the algorithm has converged

When the documents in the pool of interest contain several links, this approach can be very slow because in order to choose the best elements of the current generation, all the documents belonging to it and the ones pointed by it have to be evaluated. In the multi-agent platform used for the clever search presented above, the mobile agents are sent in the sites where the useful documents are stored and implemented the evaluations taking back only the results. The genetic evaluation algorithms apply the "temporal locality" and "spatial locality" principles. This last one means that all the explorations are implemented in the environments close to server where the father-document is located in the same server or the local network; the "temporal locality" instead refers to the conservation of the elements according to the foreseen results and to the application to a subset of such elements of the mutation operator. The mobile agents are sent to more than one site at the same time; they execute the documents evaluation in parallel on remote servers and only the results are sent to the home server. A scheme of a possible implementation is described in the block diagram fig. 5

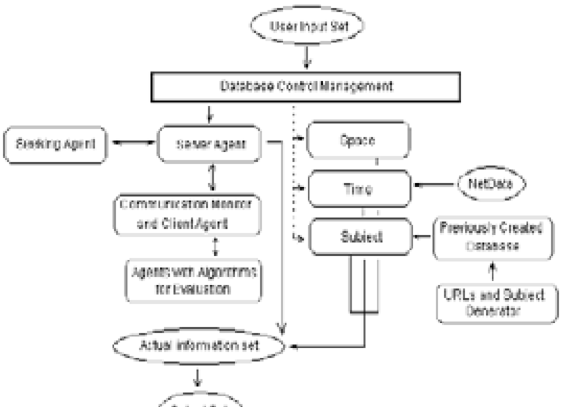

Fig. 5 : Dynamic implementation

Description of the dynamic implementation: Database Control Management – coordinates the static and dynamic application of mobile agents systems. It is responsible for the management of database. Server agent – is the application executed in the local server for the coordination of several mobile agents placed in the remote servers, and for signaling the best documents suitable for the output. Agents with Algorithms for Evaluation – is the application having the evaluation algorithms sent to the remote server, which will feed back only the results of such evaluation Communication Monitor and Client Agent allow the communication between the Server Agent resident in the local host and the several Agents with Algorithms for Evaluation resident in the remote sites. Subject – is the application that implements the mutation; Seeking Agent – is the application used to seek for the documents in the "best ones" set received from the Server Agent. Space - is the application responsible for the mutation performances; Time – is an application referring to the temporal locality principle that takes part in the mutation process.

To evaluate the mobile agents' positive contribution in a generic research algorithm it is needed to consider the static implementation of the same project presented in the initial fig., where the "Database Control Management" has been replaced by the "Database Control Tasks" that keeps only the tasks linked to the static components of the first one.

Agent technology and genetic algorithms represent powerful tools in the management of virtual organizations. Agents representing different entities such as manufacturers, sup-pliers, service providers, brokers, and other partners, can take advantage of new opportunities and changing circumstances in markets and organize themselves into virtual organizations or enterprises to achieve temporary objectives. Genetic algorithms offer fast solutions that may be included in the agent software, making them become very efficient to return the best results in a very short time.

Agent technology can reduce the costs of trading and thus, increase market efficiency and profitability, trading volumes, as well as the speed of trading. Agents can enable the move from traditional brick and mortar companies to intelligent and ubiquitous business.

D. A Hybrid Approach using ontology

A hybrid approach [42] is a personalized web search that uses ontology to represent the context of users' need, dynamic user's profile updated on time and recommendations received from similar users collectively. Empirical analysis reveals that the proposed way of including ontology, dynamic user profile and collaborative filtering together improves the accuracy of the Information retrieval. Ontology has been used for helping people interact with computers. It can offer good semantics to help Information Retrieval (IR) processes by matching users' needs.

The main key issues of the hybrid approach are : (1) A novel and simple way to get the context of query with least user's involvement using ontology and expansion of query to improve precision and recall; (2) Time based automatic user profile updating with user's changing behavior that helps in finding appropriate context of his/her need; and (3) Recommendation received from similar users are used in User profile in terms of weight of the query.

In response to a query , the user background knowledge is discovered from the world knowledge base and users browsing behaviour [70]. In [24] much ontology based personalized search systems with different approaches such as personal / reference ontology have been reported. Personalization can be performed , by constructing a user term weights matrix analogous to user-item matrix in memory based collaborative filtering algorithms and then applied traditional collaborative filtering predictive algorithms to predict a term weight in each user profile. In [54] researches proposed semantic similarity matching between web resources ontology and user context ontology.

ISSN : 2229-4333(Print) | ISSN : 0976-8491(Online)

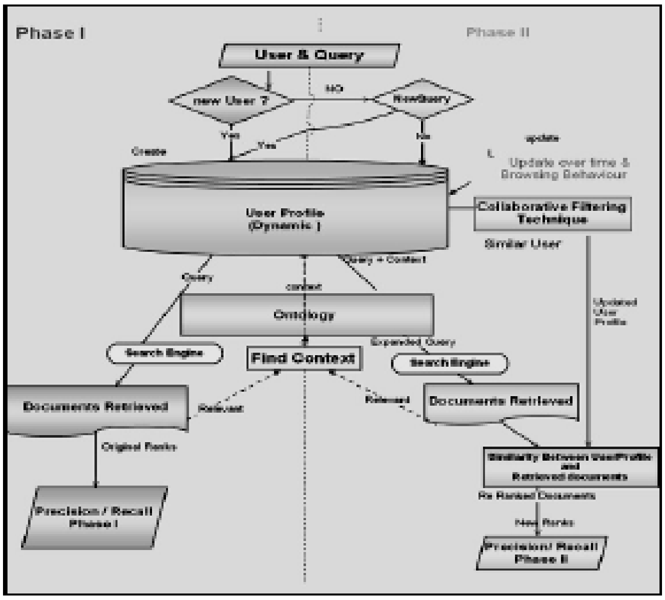

Fig.6: Hybrid approach using Ontology

Fig. 6 shows the overview of the two phases approach. The first phase includes the standard information retrieval using a search engine. The second phase uses the documents retrieved in first phase and steps forward using three modules namely, (i) Ontology for query expansion, (ii) Dynamic user profile and (iii) Collaborative filtering, as described below.

Ontology for query expansion: - the goal is to identify the accurate user context to personalize search results by re-organizing the results returned from a search engine for a given query. Context is the description of a user's aim/ need for information retrieving. In this research, context is extracted from Ontology in terms of concepts. Ontology is used to identify topics that might be of interest to a specific Web user. To get the appropriate context of query topic, use the WordNet[1] (http://wordnet.princeton.edu) [1]Ontology and retrieve appropriate context using algorithm as:

Input: Query Topic and a set of relevant search results
Output: Expanded Query
CON = {C1, ..., Cn}, contexts obtained from ontology for query q
R = {d1, ..., dn}, search results for query q
For each di R do
maxFreq= 0;
For each Cj CON do
Calculate freq[di,Cj] =
term_Frequency(di,Cj);
if freq[di,Cj] _ maxFreq then
{ Context (q) = Cj;
maxFreq = freq[di,Cj] }
end
end
expandedQuery= q+ Context(q);
Context searching algorithm

In Dynamic User Profile ,context is implicitly defined through the ontology based user profiles, which are updated over time to reflect changes in user interests/ needs by using the user's past search history. User Profile is defined as (Ui,{QueryTopicj, Contextj, Weightj }), where 1 < = i <= N, N denotes number of users and 1 <= j <= M, M denotes number of query topics for the user. Query topic is defined as set of query terms. Context Tj is context of jth Query topic extracted using algorithm described above. Weightj is weight given to the jth query topic for the user. When the user poses the similar query topic, the context retrieved from ontology is added in query topic in order to expand the query. Subsequently, we send expanded query topic to search engine for retrieving documents. The system also monitors the user's browsing history and updates the user's profile whenever the user's relevant retrieved page changes, in terms of recent context of query terms. Whilst user's older interest / context gets deviated over a period of time (threshold defined – 7 days) and it is also updated in his profile.

Collaborative Filtering

Initially, a cluster of users is generated based on similarity between queries posed by active users and documents found relevant by other similar users using collaborative filtering. From the cluster, first N users in the neighborhood are selected based on similarity measured by cosine similarity formula. It is found from experiments that value of N is best to be 5. Then, user's profile is updated in order to assign the weight for a query with the average weight of the same query in N similar users' profile from cluster. After retrieving the results from search engine, similarity between the user's updated profile P and the feature vector of the ith web page in search results (denoted as di) is computed using cosine similarity formula. Based on the values obtained, the re-ranking is done using re-ranking algorithm.

The proposed approach aims to effectively personalize search results according to each user's information need by accurately identifying the user context, updating user profile timely, recommending documents according to similar users and by reorganizing the information satisfying the needs. An overall functionality of the proposed approach is, a new user poses a query that goes to a search engine, results are retrieved and relevant documents are marked by noting down the user's clicks to the pages or asking the user through feedback. Query topic and documents both are preprocessed. by removing Stop Words (semantically non relevant terms) followed by stemming. Porter stemming 8 is used to reduce words to their stems. The TF-IDF (Term Frequency-Inverse Document Frequency) weighting scheme is used to find the weight of each term in the document. Subsequently, using WordNet context of the query for the active user is extracted and stored in user profile. Now, an old user poses an old query topic i.e query topic used earlier or similar; context from the user profile is retrieved using context searching algorithm and query is expanded with context. Subsequently, the user profile is updated with weight of terms. Simultaneously, we obtain mean weight of terms from similar n users using the collaborative filtering algorithm . Henceforth, the expanded query goes to the search engine; documents are retrieved and re-ranked using re-ranking algorithm and results are presented to user.

IV. Tools and Standards

It is obvious that personalizing the Web experience for users by addressing individual needs and preferences is a challenging task for the Web industry. Web-based applications (e.g., portals, e-commerce sites, e-learning environments, etc.) can improve their performance by using attractive new tools such as dynamic recommendations based on individual characteristics and recorded navigational history. This section provides a tour around the most well known applications of Web personalization both at a research and a commercial level.

Letizia [49] : one of the first intelligent agents, assists Web search and offers personalized lists of URLs close to the page being read using personal state, history and preferences (contents of current and visited pages) More, specifically , the agent automates a browsing strategy consisting of a best-first search augmented by heuristics inferring user interest from her behavior. WebWatcher [2] comprises a "tour guide" Web agent that highlights hyperlinks in pages based on the declared interests and path traversal pattern of the current user, as well as previous similar users. WebWatcher incorporates three learning approaches: (a) learning from previous tours, (b) learning from the hypertext structure and (c) Combination of the first two approaches A recommendation system that assists Web search and personalizes the results of a query based on personal history and preferences (contents and ratings of visited pages) is Fab [6] By combining both collaborative and content-based techniques, it succeeds to eliminate many of the weaknesses found in each approach. Humos/Wifs [1] has two components, the Hydrid User Modeling Subsystem and the Web-oriented Information Filtering Subsystem, assisting Web search and personalizing the results of a query based on an internal representation of user interests (inferred by the system through a dialogue). Another agent that learns users' preferences by looking at their visit records and then provides them with updated information about the Website is SiteHelper [41]. Personal WebWatcher [35] is a "personal" agent, inspired basically by WebWatcher, that assists Web browsing and highlights useful links from the current page using personal history. Manber [31] presents Yahoo! Personalization experience. Yahoo! was one of the scale. This work studies three examples of personalization Yahoo! Companion, Inside Yahoo! Search and My Yahoo! application, which were introduced in July 1996. Cingil [11] describe the need for interoperability when mining the Web and how the various standards can be used for achieving personalization. Furthermore, he establishes an architecture for providing Web servers with automatically generated, machine processable, dynamic user profiles, while conforming to user's privacy preference Mobasher [34] describe a general architecture for automatic Web personalization using Web usage mining techniques. WebPersonalizer is an advanced system aiming. At mining web log files to discover knowledge for the production of personalized recommendations for the current user based on her similarities with previous users. These user preferences are automatically learned from Web usage data, as wel as keeping them updated. The pre-processing steps outlined in Cooley [13] are used to convert the server logs into server sessions. The system recommend pages from clusters that closely match the current session. For personalizing a site according to the requirements of each user, Spiliopoulou [69] describes a process based on discovering and analyzing user navigational patterns. Mining these patterns, we can gain insight into a Web site's usage and optimality with respect to its current user population. Usage patterns extracted from Web data have been applied to a wide range of applications. WebSIFT [14,15] is a website information filter system that combines users, content and structure information about a website. The information filter automatically identifies the discovered patterns that have a high degree of subjective interestingness. Web Utilization Miner - WUM [67,68,69] specifies, discovers and visualized interesting navigation patters. In WUM the concept of navigation patterns includes both the sequence of events that Satisfies the expert's constraints and the routed connecting those events. Another Web usage miner designed for ecommerce applications is, in which a navigation pattern is a sequence of events satisfying the constraints posed by an expert who can specify, in a powerful mining language, which patterns have potential interest. IndexFinder [50] is a Web management assistant, a system that can process massive amounts of data about site usage and suggest useful adaptations to the Web master. This assistant develops adaptive Websites that semi-automatically improve their organization and presentation by learning from visitor access patterns. Adaptive Websites are defined in Perkowitz and Etziomi [49,4,47] . Finally, Rossi et al. [51] introduce an interesting approach based on the Object-Oriented Hypermedia Design Method (OOHDM). They build Web application models as object-oriented views of conceptual models and then refine the views according to users' profiles or preferences to specify personalization. In this context, the linking topology or the contents of individual nodes can be basically personalized. Extensible Markup Language (XML)b a simple very flexible text format originally designed to meet the challenges of large-scale electronic publishing. XML plays an increasingly important role in the exchange of a wide variety of data on the Web and the XML Query Languagec can be used for extracting data from XML documents.

Resource Description Framework (RDF)d is a foundation for processing metadata and constitutes a recommendation of W3C. It provides interoperability between applications that exchange machine- understandable information on the Web and its syntax can use XML. RDF applications include resource discovery, content description/relationships, knowledge sharing and exchange, Web pages' intellectual property rights, users' privacy preferences, Websites' privacy policies, and so forth.

Platform for Privacy Preferences (P3P)e was developed by the W3C in 1999 and comprises a standard that provides a simple and automated way for users to gain more control over their personal information when visiting Websites. Personal profiling is a form of web site visitors surveillance and leads to a number of ethical considerations.

Open Profiling Standard (OPS) is a proposed standard by Netscape that enables Web personalization. Ti allows users to keep profile records on their hard drives, which can be accessed by authorized Web servers. The users have access to these records and can control the presented information. These records can replace cookies and manual online registration. The OPS has been examined by the W3C, and its key ideas have been incorporated into P3P.

Customer Profile exchange (CPX)g is an open standard for facilitating the privacy-enabled interchange of customer information across disparate enterprise applications and systems. It integrates online /offline customer data in an XML-based data model for use within various enterprise applications both on and off the Web, resulting in a networked, customer-focused environment. The CPEX working group intends to develop open-source reference implementation

Personalized Information Description Language (PIDL)h aims at facilitating personalization of online information by providing enhanced interoperability between applications. PIDL provides a common framework for applications to progressively process original contents and append personalized versions in a compact format. It supports the personalization of different media (e.g., plain text, structured text, graphics,etc.), multiple personalization methods (such as filtering, sorting replacing etc.) and different delivery methods (for example SMTP, HTTP, IP-multicasting , etc.) .It created a unified frame work for services to both personalize and disseminate information. Using PIDL, services can describe the content and personalization methods used access methods.

V. Conclusion

Summarizing, in this study we explored the different faces of personalization. We traced back its roots and ancestors, and followed its progress. We provided detailed descriptions of personalization process and presented an overview of the interesting research initiatives and representative commercial tools. Then, we introduced and discussed several open research issues recommendations for solutions.

The future of Semantic Web is that the information can be processed by computer and support the service of intelligent web. The ultimate search technology in future should be personalized and intelligent completely. The third generation of search engine is personalized and intelligent preliminarily. The search engine should be close to users' need and simulate human's intelligence. As there is a lot of information in Semantic Web, multi-agent system based on Ontology is more accurate and efficient than the existing agent systems. But there is still some inefficiency: First of all, the model also needs to adopt appropriate ontology mapping algorithms to map the heterogeneous ontology to ontology base. Secondly, we need to design a semantic matching algorithm to calculate the matching degree about the retrieval results and retrieval requests which needs further studies.

References

[1] Amborsini,L.Cirillo,V, Micarelli,A(1997). "A Hybrid architecture for user-adapted information filtering on the World Wide Web". Preceding of the 6th International Conference on User Modeling (UM'97) (pp. 59-1) Springer-verlag.

[2] Armstrong, R.Joachims, D.Freitage, D. Mictchell, T(1995). "Web Watcher: A learning Proceedings of the AI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environments", Stanford, California, USA (pp. 6-13).

[3] G. Adomavicius, A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-theart and possible extensions," IEEE Transactions on Knowledgeand Data Engineering, vol. 17, no. 6, pp. 734–749, 2005

[4] Asnicar, F.A., Tasso, C.: "if Web: a prototype of user model-based intelligent agent for document filtering and navigation in the world wide web". In: Proceedings of Workshop Adaptive Systems and User Modeling on the World Wide Web (UM97), Sardinia, Italy (1997) 3–12

[5] G. Adomavicius, A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-theart and possible extensions," IEEE Transactions on Knowledgeand Data Engineering, vol. 17, no. 6, pp. 734–749, 2005

[6] Balabanovic, M & Shohan, Y(1977). "Content-based Collaboraive Recommendation". Communciation of the ACM , 40(3), 66-72

[7] C. Basu, H. Hirsh, W. Cohen, "Recommendation as classification: using social and content-based information in recommendation," in Proceedings of the 15th National Conference on Artificial Intelligence (AAAI '98), pp. 714–720, Madison,Wis, USA, July 1998

[8] J. Breese, D. Heckerman, C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," in Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI '98), 1998.

[9] Belkin, N., Croft, B.: "Information filtering and information retrieval". Communications of ACM 35(12) (2001) 29–37

[10] Brin, S., Page, L, "The anatomy of a large-scale hypertextual Web search engine". Computer Networks 30 (1998) 107–117

[11] Buchner, M., Baumgarten, M., Anand, S., Mulveena,m., Hughes,J.(1999). "Navigation pattern discovery from Internet data". WEBKDD'99. San Diego, California.

[12] Cingil , L.Dogac,A., Azgin,A. (2000). "A broader approach to personalization", Communication of the ACM , 43(8),136-141.

[13] Cooley, R. Mobasher, B., Srivastava,J.(1997). "Grouping Web Page references into tansactiion for mining World Wide Web Browsing patterns". Knowledge and Data Engineering Workshop, Newport Beach, CA(pp2-9).

[14] Cooley, R. Mobasher, B., Srivastava,J.(1999a). "Data preparation for minig World Web browsing patterns". Knowledge and Information System, I(1),5-32.

[15] Cooley,R. Tan,p., Srivastava,J. (1999b). "WebSIFT: The Web Site information filer systems". WEBKDD, San Diego,California. [Online] Available : http://www.acm.org./sigkdd/proceeding/wenkdd99/papers/papper11-vooley.ps

[16] Cooley,R.Tan,P., Srivastava, J. (2000). "Discovering of interesting usage patterns from Wen data". InM. Spiliopoulou(Ed.)NCS/LNAI Series. Springer-verlag.

[17] Chirita, P.A., Olmedilla, D., Nejdl, W.: Pros, "A personalized ranking platform for web search", In: 3rd International Conference Adaptive Hypermedia and Adaptive Web Based Systems (AH 2004). Volume 3137 of Lecture Notes in Computer Science, Eindhoven, The Netherlands, Springer (aug 004) 34–43

[18] M. K. Condiff, D. D. Lewis, D. Madigan, C. Posse, "Bayesian mixed-effects models for recommender systems," in Proceedings of ACM SIGIR Workshop of Recommender Systems: Algorithm and Evaluation, 1999.

[19] M. Eirinaki, M. Vazirgiannis, "Web mining for Web personalization", ACM Transactions on Internet Technology, 3(1), 2003, pp. 1–27.

[20] X. Fu, J. Budzik, K. J. Hammond, "Mining navigation history for recommendation," in Proceedings of the International Conference on Intelligent User Interfaces (IUI '00), pp. 106–112, 2000.

[21] Faloutsos, C. "Access Methods for Text". ACM Computing Surveys 17, 1 (1985), 49-74.[11] Werbos, P. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. PhD thesis, Harvard, Cambridge, MA, 1974.

[22] Floarea NASTASE, Radu –Ioan MOGOS, "An-agent based framework of E-commerce IR management using Genetic alg". The Bucharest Academy of Economic Studies, Romanianastasef@ase.ro, modos.radu@gmail.com

[23] Gauch, S., M. Speretta, et al. (2007). "User profiles for personalized information access." Lecture Notes in Computer Science4321: 54.

[24] Gentili, G., Micarelli, A., Sciarrone, F, Infoweb: An adaptive information filtering system for the cultural heritage domain". Applied Artificial Intelligence 17(8-9) (2003) 715–744.

[25] Jeh, G., Widom, J, "Scaling personalized web search". In: WWW '03: Proceedings of the 12th international conference on WorldWideWeb, New York, NY, USA, ACM Press (2003) 271–279

[26] Kleinberg, J, "Authoritative sources in a hyperlinked environment". Journal of the ACM 46 (1999) 604–632

[27] Krulwich, B., Burkey, C, "Learning user information interests through extraction of semantically significant phrases". In: Proceedings of the AAAI Spring Symposium on Machine Learning in Information Access, Stanford, California (March 1996)

[27] Lieberman, H.: Letizia, "An agent that assists web browsing".

In: Proceedings of the 14th International Joint Conference inArtificial Intelligence (IJCAI'95), Montreal, Quebec, Canada (August 1995) 924–929

[28] Lang, K, "Newsweeder: Learning to filter netnews". In: Proceedings of the 12th International Conference on Machine Learning, Tahoe City, California (July 1995) 331–339

[29] Lieberman, H, "Letizia: An agent thatassists web browsing". In: Proceedings of the 14th International Joint Conference in Artificial Intelligence (IJCAI'95), Montreal, Quebec, Canada (August 1995) 924–929

[30] Lawrence, S, "Context in web search". IEEE Data Eng. Bull. 23(3) (2000) 25–32

[31] Manber,U., Patel , A., Robison,J.(2000), "Expreience wirh personalization on Yahoo!, Communication of the ACM, 43(8),35-39.

[32] Micarelli, A., Sciarrone, F., Marinilli, M, "Web document modeling". In Brusilovsky, P., Kobsa, A., Nejdl, W., eds.: The Adaptive Web: Methods and Strategies of Web Personalization. Volume 4321 of Lecture Notes in Computer Science. Springer-Verlag, Berlin Heidelberg New York (2007)

[33] Mladenic, D, "Text-learning and related intelligent agents: A survey". IEEE Intelligent Systems 14(4) (July/August 1999) 44–54

[34] Mobasher, B. Cooley, R. Srivastava , J. (2000b), "Automatic personalization based on Web usage mining". Communications of the ACM, 43(8), 142-151.

[35] Mladenic, D, "Personal web watcher: Implementation and design". Technical Report IJSDP- 7472, Department of Intelligent Systems, J. Stefan Institute, Slovenia (1996)

[36] Mladenic, D, "Text-learning and related intelligent agents: A survey". IEEE Intelligent Systems 14(4) (July/August 1999) 44–54

[37] Mulvenna, M., Anand, S.S., B¨uchner, A.G, "Personalization on the net using web mining", Communication of ACM 43(8) (2000) 122–125

[38] MacQueen, J.B, "Some methods for classification and analysis of multivariate observations", In: Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability. Volume 1., University of California Press (1967) 281–297

[39] Micarelli, A., Gasparetti, F, "Adaptive focused crawling". In Brusilovsky,

[40] Micarelli, A., Sciarrone, F, "Anatomy andempirical evaluation of an adaptive web-based information filtering system". User Modeling and User-Adapted Interaction 14(2-3) (2004) 159–200

[41] Ngu, DWu, X.(1997), "SiteHelper. A localized agent that helps incremental exploration of the worldwide Web". Proceeding of the 6th World Wide Web Conference, Santra Clara ,CA

[42] Namita Mittal , Richi Nayak , MC Govil., KC Jain, "A Hybrid Approach of Personalized Web Information Retrieval"

[43] Ouyang Liubo, Li Xueyong, Li Guohui,Wamg Xin, "A Survey of Web Spiders Searching Strategies of Topicspecific Search Engine", Computer Engineering, vol. 13, pp. 32-33, 46, 2004

[44] Pazzani, M, "A framework for collaborative, content-based and demographic filtering". Artificial Intelligence Review 13(5-6) (1999) 393–408

[45] Pazzani, M., Billsus, D, "Learning and revising user profiles: The identification of interesting web sites". Machine Learning 27 (1997) 313–331

[46] M. J. Pazzani, "A framework for collaborative, content-basedand demographic filtering," Artificial Intelligence Review, vol.13, no. 5-6, pp. 393–408, 1999.

[47] Perkowitz., Etzioni, O. (1998), "Aaptive Web Sites: An AI challenge". Proceeding of the 15th International Joint Conference on the Artificial Intelligence.

[48] Perkowitz., Etzioni , O. (1999), "Aaptive Web Sites: Conceptual clusing mining". Proceeding of the 16th International Joint Conference on the Artificial Intelligence

[49] Perkowitz., Etzioni, O. (2000a), "Aaptive Web Sites", Communication of the ACM, 43(8) 152-158.

[50] Perkowitz., Etzioni , O. (2000b), "Towards adaptive Web Sites: Conceptual framework and case study". Artificial Intelligence, 118(1-2),245-275.

[51] Perkowitz., Etzioni , O. (1999), "adaptive Web Sites: Conceptual framework and case study". Artificial Intelligence, 118(1-2),245-275.

[52] J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", Morgan Kaufmann, San Francisco, Calif, USA, 1988.

[53] M. Pazzani, D. Billsus, "Learning and reveising user profiles: the identification of interesting web sites," MachineLearning, vol. 27, no. 3, pp. 313–331, 1997.

[54] Pan, X., et al (2007),"Context based Adaptive Personalized web Search for Improving Information Retrieval Effectiveness" IEEE pages 5427-5430

[55] Pazzani, M, "A framework for collaborative, content-based and demographic filtering". Artificial Intelligence Review 13(5-6) (1999) 393–408

[56] Rossi, G.Schwabe, D., Guimaraces, R.(2001), "Designing prersonalized Wen application". Proceeding of the WWW10, Hong Kong. 275-284.

[57] Rhodes, B.J, "Just-In-Time Information Retrieval", PhD thesis,MIT Media Laboratory, Cambridge, MA (May 2000) [Online] Available : http://citeseer.ist.psu.edu/rhodes00justtime.html.

[58] Rhodes, B.J, "Just-In-Time Information Retrieval". PhD thesis,MIT Media Laboratory, Cambridge, MA (May 2000)

[59] P. Resnick, H. R. Varian, "Recommender systems," Communications of the ACM, vol. 40, no. 3, pp. 56–58, 1997.

[60] Spiliopoulou,M., Pohlc,C.(2000), "Data Mining for measuring and improving the success of the Web sites", Data Mining and Knowledge Discovery, Special Issue on Electronic Commerce.

[61] Salton, G., McGill, M, "An Introduction to Modern Information Retrieval". McGraw-Hill, New York, NY (1983)

[62] Rivassanen, J, "understanding and applying data model to first order markov chains". Univ. of british coloumbia,Canada,technical report,Univ. of British Columbia,Canada, Canada (2003)

[63] Salton, G., McGill, M, "Introduction to Modern Information Retrieval". McGraw- Hill, New York, NY (1983)

[64] Schwab, I., Kobsa, A., Koychev, I, "Learning about users from observation", In: Adaptive User Interfaces: Papers from the 2000 AAAI Spring Symposium, Menlo Park, CA, AAAI Press (2000)

[65] Sinha, R., Swearingen, K, "Comparing recommendaions made by online systems and friends". In: Proceedings of Delos-NSF Workshop on Personalisation and Recommender Systems in Digital Libraries. (June 2001)

ISSN : 2229-4333(Print) | ISSN : 0976-8491(Online)

[66] B. M. Sarwar, G. Karypis, J. A. Konstan, J. Riedl, "Itembased collaborative filtering recommendation algorithms," in Proceedings of the 10th International Conference on World Wide Web (WWW '01), pp. 285–295, May 2001.

[67] Spiliopoulou,M., Faulstich, L.(1998). "WUM-A Web utilization miner". EDBT Workshop Web- DB98, Valencia, Spain

[68] Spiliopoulou, M., Faul stich, L., Wilkler. K.(1999a). "A data miner analyzing the navigational nehavior of Web users". Proceeding of the Work-shop on Machine Learning in User Modeling of the ACA199. Chania , Greece.

[69] Spiliopoulou, M., Pohel, C. (2000). Data Mining for measuring and improving the success of Web sites. Data Mining and Knowledge Discovery , Special Issue on Electronic Commerce.

[70] Tao, X., Y. Li, et al. (2008). "An Ontology-Based Framework for Knowledge Retrieval". In IEEE/WIC/ACM International Conference on Web Intelligence, pages 510--517, 2008

[71] Van Rijsbergen, C.: Information Retrieval. Butterworths, London (1979) Second edition.

[72] Yong S. Choi, Suk I. Yoo, Jaeho Lee, "Neural Network Based Multi-agent Information Retrieval System" Department of Computer Science Seoul National University Shilim-dong, Kwanak-ku Seoul, 151-742, Korea

[73] T. Zhu, R. Greiner, G. Haubl, "Learning a model of a web user's interests," in Proceedings of the 9th International Conference on User Modeling (UM '03), vol. 2702, pp. 65–75,Johnstown, Pa, USA, June 2003

[74] RDF. Resource Description Framework. "Model and syntax Specification. [Online] Availabe: http://www.w3.org/TR/REC-rdf-syntax/

[75] P3P. "Platform for Privacy Preferences Project". [Online] Available: http://www.w3.org/P3P

[76] CPEX. "Customer Profile Exchange". [Online] Aavilable: http://www.cpexchange.org/

[77] PIDL. "Personalized Information Description Language". [Online] Available: http://www.w3.org/TR/NOTE-PIDL