Implementation of an Intelligent Agent for Web Page Personalization and Link Prediction

Pallavi Atrey[1], Prof. Nilima Jichkar[2]

M.Tech Student, Dept. of Computer Science & Engineering, Y.C.C.E, Nagpur, India[1]

Professor, Dept. of Computer Science & Engineering, Y.C.C.E, Nagpur, India[2]

ABSTRACT: Due to the transformation from industrial age to knowledge era, the field of Web Personalization has increased exponentially. Web Personalization is relatively new and challenging field for web content delivery. Right from e-mail, e-trading, e-learning, buying and selling over internet, online shopping and internet forum to social networking based websites like facebook and twitter, all these sources of social networking site and the internet world use recommendation system for providing customized services to their loyal and authorized or registered users. People are being addicted to the usage of internet. The recommendation system learns from a customer's behavior and recommends a product in which user may be interested. Web personalization and recommendation system plays a major role in building a long lasting relationship with loyal users of websites. Nowadays internet seems to be one of the basic needs of human being. Without a help system, the user may spend lots of time on the web finding the information they are interested in. The more people will use it, more the demand of web personalization and recommendation systems increases. Thus there is need to develop such an automated system which is more enhanced than the traditional web page personalization and link prediction system, which will work more efficiently and can able to find out which user is interested in what activity, thought, or interested categories and based on their behavior recommend the web pages relevant to the web users. The paper proposes an intelligent system for effective web page personalization and link prediction. The experiment indicates to provide better accuracy and reluctance to the web users.

KEYWORDS: Web Personalization, recommendation, link prediction, intelligent system.

I. INTRODUCTION

Every day, the World Wide Web provides various kinds of web recommendations which are made available to users that include images, audio, video, query suggestion and web page itself. For an intelligent web system, web-page recommendation system plays a crucial role. The large amount of information is available on the World Wide Web (WWW). This information keeps growing exponentially as every day new information or web page is added. Thus it forms huge repository of information. But this information is randomly available and do not form an integrated structure or schema. Therefore it becomes difficult for the users to access relevant information effectively. Hence it became more challenging to provide highly relevant information to users with different needs. In recent years, web-page recommendation has proved to be valuable means of helping Web users by providing useful and effective recommendation or suggestions. The working process of all recommendation systems are like they learn users behavior, judge them and evaluate what users would like to view in future. For this the core techniques to be implemented are learning and prediction model [1]. In particular, these systems suggests interesting items from a large set of items based on the knowledge obtained from a valid Web user and his behavior in using the web sites. Based on the user's current Web navigation behavior, this Web-site recommendation can automatically recommend Web-sites that are most interesting to that particular user.

A. Working of Web page personalization:

Personalization also called as customization consists of tailoring a service or a product to accommodate specific individuals, sometimes tied to groups or segments of individuals. A wide variety of organization use personalization to improve customer satisfaction, digital sales conversion, marketing results, branding and improved website metrics as well as for advertising. It is the key element in social media and recommender systems. Web page personalization is becoming very demotic. They show links to related web page, related image or colloquial images at th different websites. When user types a particular query in the search engine, the user indirectly sends request to the web server that means session is created for that user. As the session is started, during this session when user browses a website the list of page related to the query entered in the browser is stored as a session data for that user. This sequence can be organized and stored as web sessions denoted as S =d1, d2, d3, where di =page ID of the i^{th} visited page [1]. Web pages

can be personalized based on the characteristics like interests of a particular user, social categories he is searching in, context to the query entered in the browser, actions like click on button, intent like making a purchase for example online shopping or checking the status of an entity, or any other parameter that can be identified and associated with an individual, therefore providing them with a tailored use experience.

Recommendations system's main aim is to learn from historic data of the current web user as well as the other user who is visiting the same websites as the current user is using. It works by analyzing the historic data of the users visiting the same websites. Recommendation system decides the domain of the current user from the historic information and then predicts the pages according to the current session. The current visited Web-page is referred to as state and k-previously visited pages denote the previous k states. And the Web-pages that will be visited in the next navigation step can be predicted. The performance of the Web-page personalization systems depends on the size of the training datasets. The bigger the training dataset size is, the higher the prediction accuracy is. This training datasets are obtained from the web access sequences by the web usage data.

The objective of the Web Personalization system is to provide users with the information they want or need, without expecting from them to ask for that information from an explicit means. Therefore, the requirement to fulfill users' needs of what they want can be obtained by predicting user needs in order to improve the usability and user retention of a Web site can be addressed by personalizing the requirement of their recommendations.

B. Web Mining

Web Mining is an application of new generation data mining techniques to collect bulk of data for mining, analyze them on the web and then extracting useful patterns (knowledge) from collected data in the form of web documents and services. This is a research area from several communities, such as Databases, Information Retrieval, Machine Learning, Artificial Intelligence and Natural Language Processing.

There are three general classes of information that can be discovered by web mining [2]:

- Web activity- This activity is basically the primary activity happen while browsing. The request of a particular query goes to the server which stores as server logs and client side browsing activity which refers to as web browser activity tracking.
- Web graph- This information is generated from the links between pages, people and the other data. This links are represented as a structure of graph.
- Web content-This data is the actual data user is searching for or browsing in the entire session. It consists of the Web pages and the content inside it.

Types of Web Mining [3]:

1) Web Content Mining: Web Content Mining is a process of extracting useful information from the contents of Web documents or services. Web Content Mining is related to both data mining and text mining. It is related to data mining as most of the data mining techniques can be applied to web content mining and it is related to text mining as most of the content of the web is text. It involves techniques for summarizing, classification and clustering of the web contents. These contents are the combination of structured data like list and tables, unstructured data in the form of text, semi-structured data like combination of text, images and collection of multimedia documents such as images, videos, audios, which are embedded in or linked to the Web pages. Research activities in this field involve information retrieval techniques, text mining techniques such as classification and clustering and finally for multimedia files audio or video processing techniques.

2) Web Structure Mining: Web Structure Mining is a tool used to identify the relationship between inter-linked web pages by information. This direct link connection forms a structure known as web structure schema which can be easily discoverable through database techniques for web pages. This web structure schema can also be organized in graph which consists of web pages as nodes and hyperlinks as the edges which connects the two related web pages also this web schema can also be organized in a tree structured format, based on the HTML and XML tags within the page.

   When a query is entered in the web browser, this connection allows a search engine to pull data related to inputted query directly to the linking web pages from the web site on which the connection depends on. This connection is carried out by means of spiders which scan the web sites thoroughly then retrieving the home page, linking the information through the reference links to bring forth the required specific page containing the information. Thus, Web structure Mining, in short can be regarded as the process of discovering structure of information from the web. Web structure Mining can be performed at both intra page level also called as

document structure level and inter page level also called as hyperlink level. An intra page component connects more than one webpage in the same location. An inter-link page component is a structural component which connects one web page to another at different location.

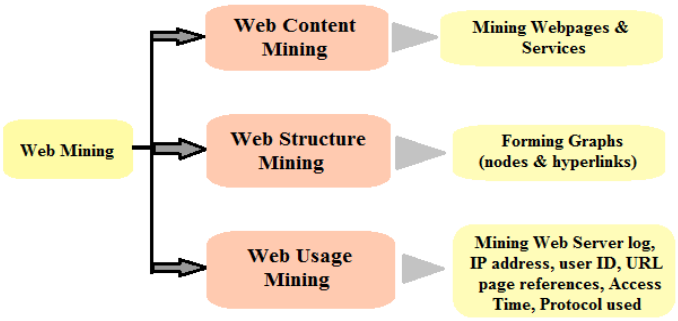

Fig 1. Web Mining

3) Web Usage Mining: In Web Usage Mining type, most of the information is generated automatically by Web servers and is collected in server log. Therefore, it is also termed as Web log Mining. It is the process of analyzing and discovering the interested usage patterns from the collected web access log repositories and then produces the results to serve the need of the Web-based applications. Web log mining uses secondary data from server access logs; user inputted queries or saved bookmarks, browser log, browser history, stored email-ids and passwords and also the registration data.

Web usage mining is a mining process of the recorded data when user goes through some interaction with the Web. It also records the browsing behavior of the user. The complete transaction of a user over Web is store as Web server log access of what file the user had saved or accessed. The usage data which is collected includes IP addresses, user id as identity, URL page references, protocol used and time required by the user to make complete access of the web. Thus, the Web-page recommendation system is based on the Web Access Sequence (WAS) from web usage data [9].

II. LITERATURE SURVEY

The Web Personalization has become an important tool both from web application point of view as well as business point of view. Due to the exponential growth of information resources and services on web, it has been very troublesome to investigate websites. Therefore, in order to minimize efforts and for relevancy over web, there is not only a need for new information services like Web Personalization but also high demand and distinguished personalization schemes have been suggested in the recent decade.

According to H. Lieberman (1995) in [4], Letizia is considered to be first system that records the user's navigational behavior and gives interesting recommendations to the user. The system is the client-site intelligent agent that monitors the browsing behavior of the user and searches for the potentially interesting pages for recommendation. The system further utilizes the best-first search algorithm supplemented by heuristic information of the user's inferred pages which is discovered by user's navigational behavior and offer suggestions for other neighboring pages.

According to T. Joachims, D. Freitag, and T. Mitchell (1997) in [5], there should be a system which uses personal profiles of users and recommends other items or pages based on their content similarity to the items or pages that are in the user's profile. They named this system as WebWatcher. WebWatcher is a content based system that provides navigation hints to the user on the basis of the knowledge acquired by the user's interests, the location and the relevance of the items in the site and the way in which other users interacted in the past. This system performs well from the perspective of the users who are searching the web for information but it gave less satisfaction for the E-commerce applications.

According to Masseglia et al (1999) in [6][7], data mining techniques such as association rules and sequential pattern forming on Web log files are used to modify server hypertext organization dynamically as per the developer's taste. The mining process in this paper is improved by means of Web Tool, prototype system and from visual query language. Dynamic links are generated, further they uses rules developed from sequential patterns or association rules. Navigation pattern is analyzed and each time the pattern of a user matches the association rule, the hypertext organization is dynamically modified. They also addressed the problem of incremental Web usage mining. ISEWUM method handles the problem of mining user patterns when new transaction is being added to the Server web log file just by considering user patterns obtained by earlier mining.

According to Mobasher, Cooley & Srivastava (2000) in [8], Web Personalization is defined as any action that tailors the Web experience by a particular user or a set of user. They defined Web usage mining as a three phase process: preprocessing, pattern discovery and pattern analysis. WebSIFT was their prototype system, which first performs cleaning and preprocessing of the raw data for identifying users, their server sessions and inferring cached page preprocessing. They use general statistic algorithm and data mining techniques such as association rules, sequential pattern analysis, clustering and classification. Finally the analysis is made through knowledge discovery algorithms in which the preprocessed content and the structure information are already being filtered and passed.

According to J. Lai et al (2005) in [10], it is difficult to determine user's search and browsing behavior on the basis of the keywords entered by the user in the web browser, the rate of clicking and the time they spent on each site. For this inconvenience, they proposed a method to derive user searching profiles and document profiles based on the similarity score of the documents. Further, they combined these both to present customized results to the users.

According to P. Palleti et al (2007) in [11], a user profile basically consists of probabilistic correlations among query and document terms which are utilized for providing personalized search results. Therefore, they implemented a probabilistic query expansion model for personalized web search. In this model, the system uses collaborative filtering for changing user interests by making user profile.

According to Mohammad Hasan et al (2006) in [12], link prediction is a supervised learning task. Their main aim is to predict the prospective links in co-authorship graph. Co-authorship is a social network where scientists work together to achieve a mutual goal. They work upon the datasets from the co-authorship graph taken from scientific publication data and identified set of features that are key to the superior performance under the supervised learning step which can easily and effectively solve the link prediction problem. They compared the performance of the various classifiers like decision tree, SVM, k-nn, multi-layer perceptron, rbf-network using various performance measures like accuracy, precision recall, F-values, squared error. As per the results obtained, SVM proved best of all the classifiers with narrow margin in all performance measures.

According to Doina Caragea, Vikas Bahirwani, Waleed Aljandal and William H. Hsu (2009) in [13], the extraction of the graph features from the graph associated with the network does not only always improve the predictions. This is because the features are constructed from the large number of users interests fail to capture the implicit semantics or views of the interests. For this reason, author used a clustering approach to build interest ontology and explore the ability of algorithm at predicting friendship links in both case when interest-based features are used alone or in combination with graph-based features. Machine learning classifier like random forest classifier is used for predicting the links in the Live Journal social network.

Accoding to Vincent Leroy, B. Barla Cambazoglu, and Francesco Bonchi (2010) in [14], in the traditional link prediction problem, prediction was done by considering the snapshot of a social network as a starting point to predict the links that are likely to appear in the future by using graph theoretic measures. In this research work, authors introduced cold start link prediction which is two-phase method based on the bootstrap probabilistic graph. This involves the generation of an implicit social network under the form of a probabilistic graph and then applied the probabilistic graph measures to produce the final prediction. The data was collected from Flickr using interest groups as the initial information.

According to Lars Backstorm and Jure Leskovec (2011) in [15], lot of research had been done on the problem of link prediction from a given snapshot of a network, but the challenge of combining the information from the network structure with rich node and edge attribute data remains open. Therefore, for combing this information, the author of this paper developed an algorithm based on Supervised Random Walks. The goal of this Walks is to assigns strength to edges in the network so that the random walker is more likely to visit the nodes to which new links will be created in the future. From that edge strength estimation function is evaluated using efficient training algorithm. The experiments on Facebook social and large collaboration network showed that these were far better than the unsupervised and feature extraction approaches.

Chuang Zhang, Bing Yu Zhai, Ming Wu (2013) in [16], considered the social attributes for predicting links in the social network. Authors developed a sociological model named exponential Random Graph Model (ERGM) for Microblog link prediction by utilizing the data of the user attributes and network topology, a link prediction model has been established. This model considers to the community as a global network where all the nodes and the edges contribute to the prediction of links. Every model has some pros and cons. This model is very effective over small and medium-sized network but it was computationally infeasible for relatively dense networks.

Keman Huang, Yushun Fan, Wei Tan, Xiang Li (2013) in [17], reveals that the number of web accessible services and their compositions were increased. According to the authors, the public available services are not enough used, and when they are used, they are used only at the cut offs. This phenomenon inspired authors to propose a methodology to help users understands the usage pattern of the service ecosystem, the interaction among services and the generation or the beginning of these interactions. Therefore a service recommendation method is derived that involves both the services and their composition with timely curve. First they constructed the network evolution model from the heuristic usage of the services and then a rank-aggregation-based link prediction method is applied to predict the evolution of the ecosystem.

According to Xiao Hun, Leye Wang, Son N. Han (2015) in [18], the previous studies by researchers uses the cross-platform approaches to predict a new-user's link on a certain Online Social Networks (OSNs) by porting his existing links from other OSNs. But this condition is not fruitful when OSNs are not willing to share their data or if users itself do not want to connect t different OSN accounts. Owing to the shortcomings, author used a single-platform approach to carry out link prediction. This model utilizes three types of social features: basic feature, derived feature and the latent relation feature which include all the available information from the new user, along with the attributes and links from existing users for link prediction using Support Vector Machines. The data sets is taken from the Facebook consisting of 479,000 users.

A.Vinupriya, S.Gomathi (2016) in [19], proposed a novel scheme of Web Page Personalization for effective web page recommendation. Personalization is being performed using parameters like page hit count, total time user is spending on a particular link, the number of downloads and link separation. FLAME Clustering and event monitoring and user's click behaviour is recorder to get better item relevance estimation. The system is developed on a generalized inverted index framework for quick personalization. The accuracy of the proposed system is determined by the performance measures like clustering delay and index process delay. The result accuracy of the proposed WPP is more than the previous NOM (Novel Optimization Method).

III. EXISTING APPROACHES

There have been various algorithms for clustering data to get optimized recommendation results. In previous studies, the recommended system is based on the interested domains by giving back a similar record in the knowledge base. So for achieving this, the first task was to cluster the users by k-means algorithm. And the second task was to find out the similarity when the new user was joined and to which cluster representative it is matching. This was achieved by using k-Nearest Neighbor (k-NN).

A. k-means Algorithm

From the term itself, k-means works on finding the mean of the nearest cluster, partitions n observation into k clusters where every new observation belong to the cluster of nearest mean. This leads to a partitioning of the data space into Voronoi cells. k-means clustering tends to search out clusters of comparable spatial extent, whereas the expectation-maximization mechanism permits clusters to possess completely different shapes. The k- means clustering algorithm starts by dividing the input points into k initial sets. It can be done either randomly or with the help of some historic information or manually done. It further calculates the mean point of each set obtained after partition called as centroid of that set. It further forms a new partition by associating each point with the closest centroid. These centroids are again recalculated for the new clusters. This procedure is repeated until convergence means the point is no longer switch to clusters or centroid no longer changed [20].

B. k-Nearest Neighbor Algorithm

On the basis of finding the similar users, their interested domains should be same. When a new user comes, his or her search history is viewed by the system and by means of k-NN algorithm, it could obtain k similar users' interested domain list from the knowledge base. The k-Nearest Neighbors' algorithm may be non parametric technique used fro classification and regression. The k-NN algorithm is among the only of all machine learning algorithms. Both for classification and regression, it will be helpful to assign weight to the contributions of the neighbors, in order that the nearer neighbors contribute additional to the typical than the additional distant ones. A disadvantage of the k-NN algorithm is that it's sensitive to the native structure of the data.

C. FLAME Clustering Algorithm

FLAME is a Fuzzy Clustering by Local Approximation of Memberships is a data clustering algorithm that defines in the dense parts of data sets and performs cluster assignment solely based on the neighborhood relationships among objects.

Flame algorithm is divided into three steps:

1) Extraction of the structure information from the dataset:
   In this step, construction of the neighborhood graph is achieved to connect each object to its K-Nearest Neighbor. Density of each object is estimated based on its proximities to its neighbor. Objects that are created are classified as Cluster Supporting Object (CSO) which have the highest density than all its neighbor, Cluster Outliers are the objects with density lower than all its neighbor as well lower than a predefined threshold, and the third category is all the rest that are not in these two category.
2) Local/ Neighborhood Approximation of fuzzy membership:
   In this step, initialization of fuzzy membership is done where all the CSO are assigned with fixed and full membership to itself to represent to one cluster, all outliers are assigned with fixed and full membership to the outliers group and the rest is assigned with equal membership to all cluster and outlier group.Then the fuzzy membership of all type 3 (rest) objects are updated by repetition in convergence procedure called Local/Neighborhood Approximation of Fuzzy Memberships which is a iterative converging procedure where fuzzy membership of each object is updated by a linear combination of the fuzzy memberships of its nearest neighbors.
3) Cluster construction from fuzzy membership:
   It can be achieved in two ways. One way to assign each object to the cluster in which it has highest membership called one-to-one object-cluster assignment. And the other is to assign each object to the cluster in which it has membership higher than a threshold called one-to-multiple object-cluster assignment.

IV. LIMITATIONS OF EXISTING SYSTEM

A key limitation of k-means is its cluster model. The concept over this algorithm is that it is predicated on the spherical clusters that are separable in a means so the mean value converges towards the cluster centre. The cluster is of comparable size so that the assignment to the nearest cluster centre is that correct assignment. The main disadvantage of the KNN algorithm is that it is a lazy learner, it uses the employment information for classification and do not learn anything from it. Again this algorithm should figure the gap and sort all the employment information at each prediction, which can be slow if there is an oversized kind of employment examples.

V. METHODOLOGY

The problem within the existing systems is that, there was no additional personalization supported ontology and domain or historical data available for personalization. So the personalized recommendation method is developed which uses domain ontology and history data to boost the personalization that should be there to use for website recommendation purpose. The system proposes a new scheme named as WPP (Web Page Personalization) for effective web usage mining.

A. Phases of the Web Page Personalization

1) Collection of Web Data:
   This phase will judge the user's behavior on the web. Data can be implicit or explicit. Implicit data includes the historical activity of the user, clicks that user made which is recorded in the web server log. Explicit data includes the data of the user at the time of registration. In this system, the user data includes the registration details of the user, the link prediction entries i.e. these are the entries of the link that user will provide. It include link name, link url, short description of a link and the date.
2) Pre-Processing of the Web data:
   Collected data is preprocessed before applying any algorithm to it. Preprocessing includes filtering out data; compress the data to wash out all the irrelevant data and cleaning inconsistencies.
3) Analysis of the Web data:
   This phase typically deal with the preprocessed data to apply machine learning or Data Mining techniques to discover interesting usage patterns, applying clustering algorithms and statistical correlation between web pages and user groups. This phase in Web Page Personalization is also known as Web Usage Mining.

4) Decision Making And Recommendation :

The last phase in personalization makes use of the results of the previous analysis step to deliver recommendations to the user. The recommendation process typically involves generating dynamic Web content on the fly, such as adding hyperlinks to the last web page requested by the user. This can be accomplished using a variety of Web technology options such as CGI programming.

To achieve better item connection estimation, the system uses the following parameters.
(1) Event observation and click behaviors from net search.
(2) FLAME clustering.

This paper introduces an online personalized recommendation learning framework for customized content recommendation further as the key parts of this framework with more accuracy than FLAME clustering algorithm. For this system, generalized data set is created. For example, the information about Mahatma Gandhi, Jawaharlal Nehru is stored in the SQL server.

From the user's click stream identified needs and interest. If a new user entering a new query in the browser, then Google open number of search results in the response of that query as correlated results. The proposed system checks the similarity in the links shown in the correlated results. If first link matches with the second link with more than 70% similarity then the system merges both the link and will show only one link. This is for minimizing the search result and to limit the irrelevant data. The details about the link is not stored in cache memory, in spite it is stored in the SQL server to see which link is revisited by the new user i.e. which new user type related query in the browser.

Search Engine like Google possesses a limitation that it shows multiple numbers of links. From this 20 % of the links are useful and the rest of the data links shows the irrelevant data which is not at all related to the query in the browser. In this procedure, the system will not show more links to that query and the search result will get minimized, so that if user is searching for a particular query, Google will show the perfect result to that query. Hybrid Clustering Algorithm is used to obtain more accurate result. Working of this algorithm is similar to the FLAME Clustering Algorithm.

VI. RESULTS AND DISCUSSION

The proposed system is assessed by the performance measures like clustering delay, index process delay, performance time comparison and the result accuracy. Clustering delay is the time taken to the formation of cluster. In the previous algorithm, the clustering delay is more as compared to the proposed hybrid algorithm with WPP. That means the time required to the formation of cluster is reduced in this proposed hybrid algorithm with WPP. It is calculated by the formula

Clustering Delay = Process Start Time – Process End Time.

Figure shows the comparison between the clustering delays. In this graph, proposed hybrid algorithm with WPP shows less time in clustering delay while comparing with the previous algorithm.

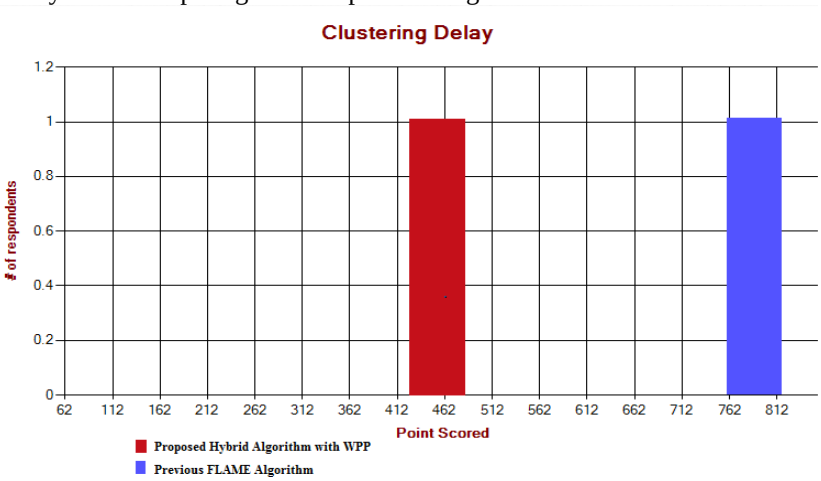

Fig. 2 Clustering Delay

Second parameter for assessment is index process delay. Index process delay is the time required to show the correlated results when a required query is typed in the browser. Indexing shows the sequence in which the data is shown the search engine. It is calculated by the formula

Index Process Delay = Process Start time –Process End Time

Figure shows the comparison between the index process delays of the proposed hybrid algorithm with WPP and the

previous algorithm. In this graph, proposed hybrid algorithm with WPP shows less number of iteration.

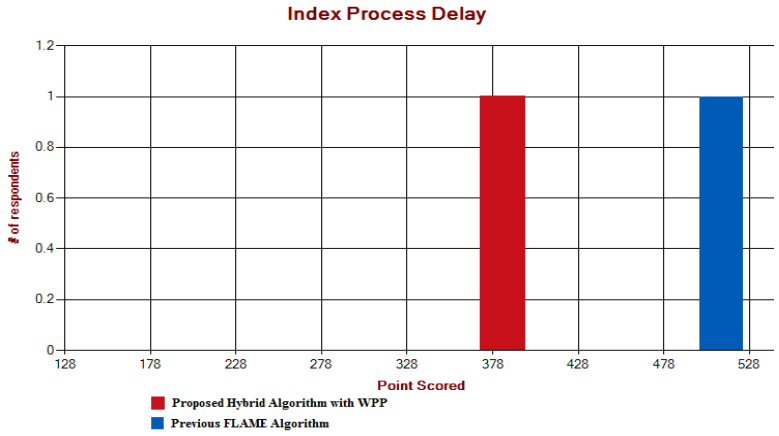

Fig. 3 Index Process Delay

Third parameter for assessment is time performance comparison. It is time required by the system to show the relevant data based on the query typed in the browser. It is the total time taken by the system to show the recommendations. It is calculated by the formula

Performance time comparison = Process Start Time –Process End Time

Figure shows the performance time comparison of the proposed hybrid algorithm with WPP and the previous algorithm. In this graph, proposed hybrid algorithm with WPP shows less number of iteration.

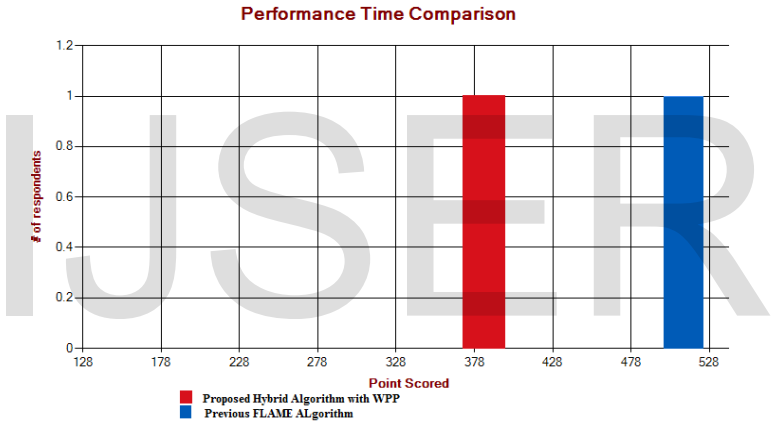

Fig. 4 Performance Time Comparison

The fourth parameter is the result accuracy. This is in the terms of how exact and minimized recommendation is shown by the proposed system. It is calculated by the formula

ACC=TP+TN/P+N

Figure shows the comparison of the result accuracy. In this graph, proposed system shows high accuracy than the existing system.

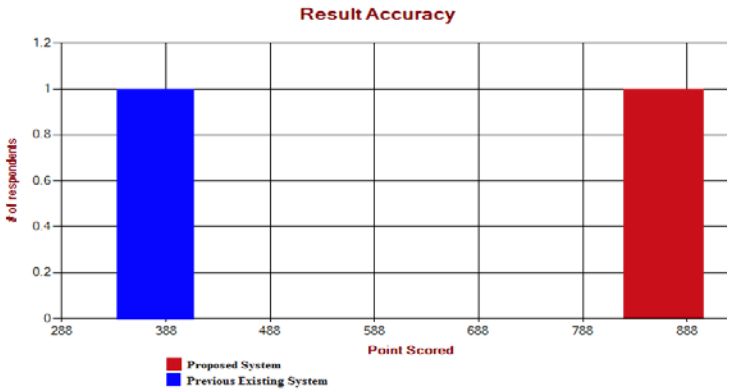

Fig. 5 Result Accuracy

IV. CONCLUSIONS

This paper discuses the importance of evaluating the performance of the employees in virtual organization. It is notices based on the given literature survey that there is a need to develop the proposed system because employee's performance evaluation in virtual organization is tedious work. Therefore, a hybrid approach is defined to increase performance monitoring of employee's using Domain Driven Data Mining(D3M) approach using 360 degree feedback and sentiment analysis .Utilizing the concept of using four multi-factorial evaluation model in the performance appraisal system could alleviate the changes need to be made in this system whenever it is necessary. Further, the rating will be provided based on the survey results. When it comes to making long term and short term goals, virtual organization must carefully calculate who will champion their initiatives. Therefore, it is important to select the appropriate employee for the tasks that will affect the position of the organization in business scenario. Placing the wrong person can result in devastating problem which are subject to strong public scrutiny. These problems can range from lack of employees morale to financial destruction.

REFERENCES

[1] Modraj Bhavsar, MRs. P. M. Chavan, "Web Page Recommendation using Web Mining", International Journal Of Engineering Research and Applications (IJERA), www.ijera.com, Vol. 4, No. 7, pp.201-206, July 2014.
[2] http://www.scaleunlimited.com/about/web-mining
[3]Beatric,Ryan Fernandes, Leo. J. Peo, Nikhila Kamat, Sergius Miranda, "New Approaches to Web Personalization Using Web Mining Techniques", International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 5, No. 2 , pp. 2195-2201,2014.
[4] H. Lieberman, "Letizia: An agent that assists web browsing", Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI), pp. 924– 929, 1995
[5]T. Joachims, D. Freitag, T. Mitchell," Webwatcher: A tour guide for the world wide web", Proceeding of the 15th International Conference on Artificial Intelligence (ICAI), Nagoya, Japan, 1997.
[6] Masseglia .F., Poncelet. P., Cicchetti. R., "WebTool: An integrated framework for data mining", Proceedings of the 9th International Conference on Database and Expert Systems Applications(DEXA), pp. 892-901, Florence, Italy, August 1999.
[7] Masseglia .F., Poncelet. P., Teisseire, M.,"Using Data Mining Techniques on Web Access Logs to Dynamically Improve Hypertext Structure"ACM SigWeb Lett, Vol.8, No. 3, pp. 13-19, October 1999.
[8] Jaideep Srivastava, Robert Cooley, Mukund Deshpande, Pang-Ning Tan, "Web usage mining:discovery and applications of usage patterns from web data," SIGKDD Explorations, Vol. 1 No. 12 , January 2000.
[9] B. Liu, B. Mobasher, O. Nasraoui, "Web usage mining", in Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data, B. Liu, Edition. Berlin, pp. 527–603, Germany: Springer-Verlag, 2011.
[10] J. Lai and B. Soh, "Personalized Web search results with Profile comparisons," 3rd International Conference on Information Technology and Applications (ICITA), Vol. 1, pp. 573 – 576,2005.
[11]P. Palleti, H. Karnick, P. Mitra, "Personalized Web Search Using Probabilistic Query Expansion," International Conferences on Web Intelligence and Intelligent Agent Technology Workshops (IEEE/WIC/ACM), pp. 83 – 86, 2007.
[12] M. Al Hasan, V. Chaoji, S. Salem, M. Zaki, "Link prediction using supervised learning," in SDM : Workshop on Link Analysis, Counter- terrorism and Security, 2006.
[13]D. Caragea, V. Bahirwani, W. Aljandal, W. H. Hsu, "Ontology-based link prediction in the live journal social network." in SARA, Vol. 9, 2009.
[14] Vincent. Leroy, B. Barla. Cambazoglu, and Francesco Bonchi, "Cold start link prediction," in KDD, pp. 393–402, 2010.
[15]L. Backstrom, J. Leskovec, "Supervised Random Walks: Predicting and Recommending Links in Social Networks", in WSDM (ACM), pp. 635-644, Feb 9 to 12, 2011.
[16]Chuang Zhang, Bing Yu Zhai, Ming Wu, "Link Prediction of Community in Microblog Based on Exponential Random Graph Model", Pattern Recognition and Intelligence System, Lab Beijing University of Posts and Telecommunications, Beijing, China 100876, 2013.
[17]Keman Huang, Yushun Fan, Wei Tan, Xiang Li, "Service Recommendation in an Evolving Ecosystem: A Link Prediction Approach", IEEE 20th International Conference on Web Services , pp.501-514, 2013.
[18]Xiao Hun, Leye Wang, Son N. Han," Link Prediction for New Users in Social Networks", IEEE International Conference on Communication (ICC), 10 September, 2015.
[19]A.Vinupriya., S.Gomathi, "Web Page Personalization and Link Prediction Using Generalized Inverted Index and Flame Clustering", International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, INDIA, Jan.07- 09, 2016.
[20] Minxiao Lei, Lisa Fan," A Web Personalization System Based on Users' Interested Domains", Proceedings of the 7th IEEE International Conference on Cognitive Informatics (ICCI), pp.153-159, 2008.