Global Journal of Computer Science and Technology
Volume 12 Issue 9 Version 1.0 April 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Web Page Recommendation Approach Using Weighted Sequential Patterns and Markov Model

By K. Suneetha & Dr.M.Usha Rani
Sree Vidyanekethan Engineering College/JNTU

Abstract - Web page recommendation aims to predict the user's navigation through the help of web usage mining techniques. Currently, researchers focus their attention to develop a web page recommendation algorithm using the well known pattern mining techniques. Here, we have presented a web page recommendation algorithm using weighted sequential patterns and markov model. To mine the weighted sequential pattern, we have modified the prefixspan algorithm incorporating the weightage constraints such as, spending time and recent visiting. Then, the weighted sequential patterns are utilized to construct the recommendation model using the Patricia trie-based tree structure. Finally, the recommendation of the current users is done with the help of markov model that is the probability theory enabling the reasoning and computation as intractable. For experimentation, the synthetic dataset is utilized to analyze the performance of W-Prefixspan algorithm as well as web page recommendation algorithm. From the results, the memory required for the W-prefixSpan algorithm is less than 50% of memory needed for PrefixSpan algorithm.

Keywords : Web page recommendation, Weighted sequential pattern, Prefixspan, Patricia-trie, Markov model.

GJCST Classification: I.5.m

WEB PAGE RECOMMENDATION APPROACH USING WEIGHTED SEQUENTIAL PATTERNS AND MARKOV MODEL

Strictly as per the compliance and regulations of:

Web Page Recommendation Approach Using Weighted Sequential Patterns and Markov Model

K. Suneetha [α] & Dr.M.Usha Rani [σ]

Abstract - Web page recommendation aims to predict the user's navigation through the help of web usage mining techniques. Currently, researchers focus their attention to develop a web page recommendation algorithm using the well known pattern mining techniques. Here, we have presented a web page recommendation algorithm using weighted sequential patterns and markov model. To mine the weighted sequential pattern, we have modified the prefixspan algorithm incorporating the weightage constraints such as, spending time and recent visiting. Then, the weighted sequential patterns are utilized to construct the recommendation model using the Patricia trie-based tree structure. Finally, the recommendation of the current users is done with the help of markov model that is the probability theory enabling the reasoning and computation as intractable. For experimentation, the synthetic dataset is utilized to analyze the performance of W-Prefixspan algorithm as well as web page recommendation algorithm. From the results, the memory required for the W-prefixSpan algorithm is less than 50% of memory needed for PrefixSpan algorithm.

Keywords : Web page recommendation, Weighted sequential pattern, Prefixspan, Patricia-trie, Markov model.

I. Introduction

Sequential pattern mining, an advance of association rule mining, is an imperative subject of data mining, often applied for extracting the useful information [9]. Sequential pattern mining algorithms deals with the problem of determining the frequent sequences in a given database [6]. Sequential pattern is a sequence of itemsets that often occur in a specific order and thus, all items in the same itemset are expected to encompass the same transaction time value or within a time gap. Each sequence have a temporally ordered list of events, wherein each event is a compilation of items (itemset) occurring simultaneously. The temporal ordering among the events is induced by the absolute timestamps associated with the events [10]. Generally, all the transactions of a customer are collectively viewed as a sequence, called customer - sequence, where each transaction is represented as an itemset in that sequence and all the transactions are scheduled in a particular order based on the transaction-time [8]. Recently, there has been a substantial interest in using sequential mining approaches to construct web page recommendation systems.

Here, we have presented a web page recommendation algorithm using weighted sequential patterns and markov model. The overall process of page recommendation based on Web usage mining consists of three phases: data preparation, mining of weighted sequential patterns and recommendation.

Data preparation: This step consists of, (i) identifying the different users' sessions from the usually very poor information available in web log files and (ii) reconstructing the users' navigation path within the identified sessions.

Mining of weighted sequential patterns: The traditional sequential pattern mining problem is extended by allowing a weight to be associated with each page in a user session to reflect interest of each page within the user session. In turn, this provides with an opportunity to associate a weight parameter with each page in a resulting sequential pattern, which called a weighted sequential pattern (WSP).

Recommendation: After mining the weighted sequential patterns, the patricia-based tree is constructed. From the Patricia tree, a recommendation model is developed based on markov model for predictions of users to find web pages they want to visit.

II. Related Work

Literature presents a lot of web page recommendation algorithms based on web used mining, collaborative filtering, and rule-based filtering. Here, we portray some recent works presented in the literature for web page recommendation.

Forsati, R *et al* [3] have developed an effective and scalable approach to deal with the web page recommendation problem. Here, a distributed learning machine has been employed to learn the performance of previous users' and to cluster the pages based on learned pattern. Dealing with unvisited or recently added pages was one of the difficult and challenging tasks in recommendation systems. As they would never be

Author α : Sr. Lecturer Dept. of MCA, SVEC, Tirupati Andhrapradesh, INDIA-517102. E-mail: umasuni.k@gmail.com
Author σ : Assoc.Professor & HEAD, Dept. of Computer Science, SPMVV, Tirupati, Andhrapradesh, INDIA- 517102.
E-mail: musha_rohan@yahoo.com

recommended, it is indispensable to provide a chance for these seldom visited or recently added pages to be incorporated in the recommendation set. By considering this problem, a Weighted Association Rule mining algorithm has been presented for the recommendation purposes. Also, a HITS algorithm has been exploited to extend the recommendation set. Furthermore, they have analyzed the proposed algorithm under various settings, and revealed the efficiency of this approach in enhancing the overall quality of web recommendations.

Yicen Liu *et al* [15] have introduced an automatic tag recommendation algorithm that has been employed in the large-scale and real-time data process successfully and efficiently. Most of the prior researches on tag suggestion have focused on initially, discovering the relationship between testing and training data and then, assigning the top ranked tags of the most related training data to the testing object. But, they not paid any attention in determining the internal relationship between the tags and weblogs. In their current research, more than 43% of tags, which have been labeled by weblog users, have really been employed in the body of the text. In the meantime, the term frequency distribution, the paragraph frequency distribution, and the first occurrence position of tags were dissimilar from the ones of non-tags in the text. As well, the tags of a weblog have been assigned in two steps. Initially, some probability distributions of the word attributes have been trained through the labeled training weblogs, and some keywords of a testing weblog have been extracted as one part of the tags based on the probability distributions. Subsequently, with the aid of Latent Semantic Indexing (LSI) model, the other parts of the tags have been obtained from the first part ones. Experiments conducted on an extensive tagging dataset of weblogs 12 have confirmed that the average tagging time for a new weblog was less than 0.02 sec, and more than 74% of testing weblogs have been properly labeled by means of the top 15 tags.

An approach for recommendations of unvisited pages has been presented by Forsati, R *et al* [11]. They have focused on the recommender systems based on the user's navigational patterns and provided proper recommendations to cater to the current needs of the user. The group of users with analogous browsing patterns has been identified by employing an offline data preprocessing and clustering technique. The experiments conducted on real usage data from a commercial web site have demonstrated a considerable enhancement in the recommendation efficiency of the proposed system.

Web Personalization is viewed as an application of data mining and machine learning approaches to create models of user behavior that can be applied to the task of forecasting the user needs and adapting future interactions with the eventual goal of enhanced user satisfaction. An extensive overview of intelligent methods for Web Personalization has been presented by Sarabjot Singh Anand and Bamshad Mobasher [12]. They have studied the state-of-the-art in Web personalization. Initially, a depiction of the personalization process and a classification of the current techniques to Web personalization have been presented. Also, they have discussed the different sources of data available to personalization systems, the modeling techniques utilized, and the current techniques to analyze these systems. Numerous challenges faced by the researchers in developing these systems and also the solutions to these challenges proposed in literature have been described. They have concluded with a discussion on the open challenges that must be addressed by the research community if this technology is to create a positive impact on user satisfaction with the Web.

Due to the increasing number of Web sites such as e-businesses contain a huge number of pages, users find it very hard to swiftly reach their own target pages. Thus, Hiroshi Ishikawa *et al* [5] have proposed two approaches to Web usage mining as a key solution to these problems. First of all, an efficient recommendation system called the L-R system has been depicted, which creates user models through classifying the Web access logs and by mining access patterns based on the transition probability of page accesses, and then, recommend the significant pages to the users based on both the user models and the Web contents. The prototype system has been analyzed and obtained the positive effects. Secondly, another approach has been employed for creating user models that clusters the Web access logs based on the access patterns. Moreover, the user models assist to find the unexpected access paths corresponding to ill-formed Web site design. In addition, Daniel Mican and Nicolae Tomai [2] have proposed WRS, architecture for robust web applications. Within the structure, usage data was being implicitly obtained by data collection sub-module. Here, the usage data has been extracted, online and in real time, via a proactive technique. They have efficiently exploited association rule mining among both frequent and infrequent items for successful pattern discovery. This was due to the fact that the pattern discovery module transactionally processes users' sessions and employs incremental storage of rules. Also, they have proved that the Wise Recommender System (WRS) has been straightforwardly implemented within any web application, because of the efficient integration of the three phases into an online transactional process.

III. Proposed Algorithm to Web Page Recommendation Based on Weighted Sequential Access Patterns

Web page recommendation is significant research over the past decade due to its real world application. With the intention of real world applicability, we have developed an approach for web page recommendation using weighted sequential pattern and markov model. Here, the traditional sequential pattern mining algorithm is modified significantly by incorporating the significant measure to mine more useful patterns. Then, the markov model described in [4] is used to recommend the web pages. The block diagram of the proposed approach for web page recommendation is given in Figure 1. The important steps for generating recommendations to the user is as follows,

- Data Preprocessing
- W-PrefixSpan for mining of weighted sequential web access patterns
- Building Pattern tree model
- Generation of recommendations using markov model

Block diagram

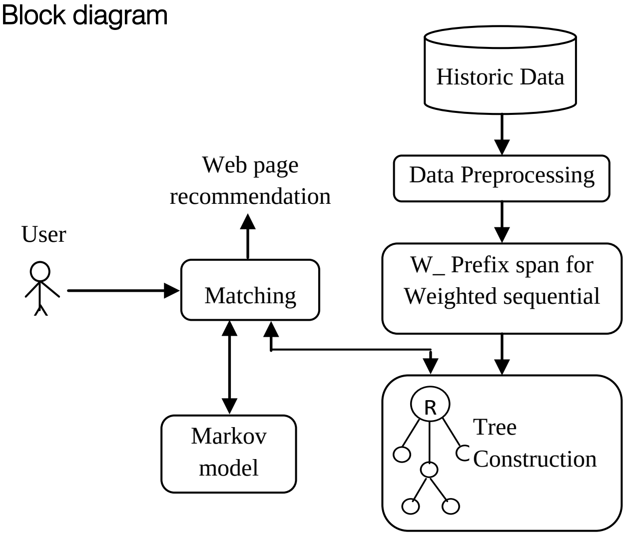

Fig. 1: Block diagram of the proposed approach to web page recommendation

a) Data Preprocessing

This section describes the preprocessing steps of web log file that is the input of the proposed web page recommendation approach. In general, the web log file consists of, IP address, access time, HTTP request method used, URL of the referring page and browser name (for an example, *Web server log file: 192.162.37.21 [23/Feb/2012:08:17:25]"GET / HTTP/1.1" "http://www.sigkdd.org/kddcup/index.php" IE/7.0 Windows 07*). The initial process of the proposed web page recommendation approach is to preprocess the web log file such a way that the mining process should be applied. Here, we make use of the sequential pattern mining process so there is a need to convert the web log file into the sequential database that should be in the proper format to mine the weighted sequential patterns.

User Identification: User identification is an important step for constructing the sequential database. IP address and user session are utilized here to track a unique user from the web log file. Unique IP address is a new user but at the same time, the user session should be fixed for particular time period. If the user session is reached to a particular duration for the same IP address, then the new session is acted like new user. Based on this, user transaction is formed from the web log file.

Weighted Sequential database generation: Once the user transactions are identified, the weighted sequential database is generated including the sequence of web pages visited by the user, time spent by the user on corresponding web page and its recent information.

Let assume the web log database D having IP address, access time, HTTP request method used, URL of the referring page and browser name. After applying the data preprocessing steps, we generate the weighted sequential database that is represented as follows, W_{ij} in which, 'i' belongs to the set of users and 'j' signifies the set of pages visited by the corresponding user. Here, every element of w_{ij} contains three tuples $w_{ij} = (p_{ij}, s_{ij}, r_{ij})$ in which the first tuple belongs to the web pages, the second one belongs to the time spent within that page and the third one belong to whether its recent one or not. For example, $w_{ij} = (p_1, 20, 0)$ is a tuple, which denotes that the user spent 20 seconds on page p_1 and '0' signifies that the page p_1 is not recently accessed by the user.

b) W-Prefixspan for Mining of Weighted Sequential Web Access Pattern

Once the weighted sequential database is constructed, the mining procedure is carried out to find the interesting sequential patterns. Here, PrefixSpan [7] is modified as W-PrefixSpan (Weighted-PrefixSpan) by incorporating the ***spending time*** and ***recent view*** into the mining procedure. The weightage measure assumed in the proposed W-PrefixSpan algorithm is spending time and recent view. The two aspects taken for providing the weightage of the sequential patterns are,

Spending time: One of the fields in the *web log data* is the duration of the web page which is viewed by the user. Generally, time spent by the user within a particular page is necessary to identify the importance of web pages. From the web page which having long duration, we can conclude that this particular web page has been referred by the user in a long occasion because of its worth. Thus, the *spending time* is an important measure for the researchers who are attempting to identify the interest of the users. So, if we incorporate the time duration into the mining of sequential patterns, the interesting relationships can be found out from the mined sequential patterns that can be effectively applied to web page recommendation process.

Recent view: Another significant measure taken for sequential pattern mining is recent view that describes whether the page is accessed recently or not. The reason behind taking the recent view for mining the sequential pattern is that the more importance should be given for the web pages which are accessed recently than the older one. The behavior of the user surely vary depend on the time so the recent behavior of the user is significant for finding the sequence analysis. With the intention of behavior variation over time, the recent view is also incorporated into the sequential pattern mining algorithm to achieve a subset of more interesting sequential patterns (SR-Patterns).

W-PrefixSpan algorithm: In this section, we describe an efficient algorithm, W-PrefixSpan, for mining all the SR- patterns from weighted sequence databases. The W-PrefixSpan algorithm is developed by modifying the eminent PrefixSpan algorithm, which uses the pattern growth methodology for mining the frequent sequential patterns recursively. Let $W_{ij} = \left\langle (p_{11}, s_{11}, r_{11}), (p_{12}, s_{12}, r_{12}), \cdots, (p_{1n}, s_{1n}, r_{1n}) \right\rangle$ be a data sequence of weighted database W_{ij}, where p_{ij} is web page, s_{ij} is a spending time and r_{ij} signifies the recent view. A sequence $W_s = \left\langle (p_{11}, s_{11}, r_{11}), (p_{12}, s_{12}, r_{12}), \cdots, (p_{1m}, s_{1m}, r_{1m}) \right\rangle$ is said to be a sub sequence of W_{ij} only if, (1) W_s is a subsequence of W_{ij}, $W_{ij} \in W_s$

(2) $t_1 < t_2 < \cdots < t_m$ where, t_1 is the time at which p_{ij} occurred in W_s, $1 \le r \le m$. A sequence is said to be SR sequence W_{ij} if and only if, (1) W_s is a subsequence of W_{ij}, (2) the W-support should be satisfied.

At first, the weighted sequential database W_{ij} is given to the proposed W-PrefixSpan algorithm that discovers the 1-length weighted sequential patterns from the weighted sequential database by scanning the database once. The 1-SR patterns (spending time with recent view) which satisfy the predefined support threshold are mined from the sequential database by simply scanning the database. The W-support for the 1-length pattern is computed as follows,

$$W\_\sup(p) = \frac{1}{N} \frac{\sum_{i=1}^{N_T \in p} I_s(i) * R(i)}{\sum_{i=1}^{N_T \in p} R(i)}$$

Where, $N \rightarrow$ Number of user transaction in the weighted sequential database $N_T \rightarrow$ Number of transaction that contains the web page p, $R(i) \rightarrow$ recent information

$$I_s(i) = \sum_{i=1}^{N_T} \left(\frac{s_i}{\sum_{i=1}^{M_T} s_i} \right)$$

Where, $M_T \rightarrow$ Total number of web pages in one transaction, $s_i \rightarrow$ spending time Then, the projection database is formed by projecting the collection of postfixes of mined 1-SR sequence. In projection database, 'n' disjoint subsets are generated if the mined 1-SR patterns contain 'n' number of sequence. Then, the 2-length SR-patterns are mined from the projected database by computing the weighted support on the projected database. Again, the projected database is formed with the help of mined 2-SR patterns and this process is repeated recursively until all SR sequential patterns are mined. The following provides a detailed explanation of the important steps involved in the proposed W-PrefixSpan algorithm.

Input : A weighted sequence database W_{ij}, and the minimum support threshold $\min\_W\_\sup$.
Output : The complete set of weighted-sequential patterns β.
Method : Call *W-PrefixSpan*$(\langle\ \rangle, 0, W_{ij})$
Subroutine : *W-PrefixSpan* $(\alpha, l, W_{ij}|_{\alpha})$
Parameters: α : SR-sequential pattern; l : the length of α; $W_{ij}|_{\alpha}$: the α-projected database, if $\alpha \ne \langle\ \rangle$; otherwise, the weighted sequence database W_{ij}.
Method:
1. Scan $W_{ij}|_{\alpha}$ once, find the set of SR- pattern s such that
   a) 's' can be assembled to the last element of α to form a SR-sequential pattern; or
   b) $\langle s \rangle$ can be appended to α to form a SR-sequential pattern.
2. For each SR- web page s, append it to α to form a SR-sequential pattern α'.
5. For each α', construct α'-projected database $W_{ij}|_{\alpha}$, and call *PrefixSpan* $(\alpha', l+1, W_{ij}|_{\alpha})$.

c) Building of Pattern Tree Model

Once we mine the weighted sequential patterns, the pattern tree is constructed using the procedure defined in [14, 13]. Initially, trie-based data structure given in [1] is used to construct the pattern tree for web page recommendation. Later, the modification was done by [14, 13], who utilized the patricia-based data structure for web page recommendation due to the advantages of particia structure over the trie structure. Here, the procedure defined in [14, 13] is applied to the proposed approach for constructing tree structure. The method for constructing the pattern tree in the proposed web page recommendation approach is as follows: 1) Generate an empty root node, 2) Add the most sub pattern in the SR-sequential pattern set into a node next to the root node, 3) Insert the postfixes of pattern into child node only if the current pattern to be inserted is a super pattern of inserted patterns, 4) Otherwise, current pattern is inserted into the node next to the root node, and 5) Step 3 and step 4 is repeated for every pattern in the mined SR-pattern set.

d) Generation of Recommendations Using Markov Model

This section describes the markov model utilized for web page recommendation. Here, we make use of the markov model described in [4] that is used in the identication of the next page to be accessed by the Web site user based on the sequence of previously accessed pages. Here, whenever a new user comes to get the recommendation, the sequence path of the new user is matched with the patricia-trie structure. Then, the subsequent web page whether it may be from same node or from its child node is retrieved. Now, the sequence path of the new user is used to find the accurate recommendation using the probability definition used in the previous work [4].

Let the input sequence visited by the user be $s_1, s_2, \ldots, s_n$. At first, the sequence, $s_1, s_2, \ldots, s_n$ is matched with the patricia-trie structure that can provide the matched result like, $s_1, s_2, s_n, s_{n+1}, \ldots, s_k$, $s_1, s_2, s_n, s_{n+1}, \ldots, s_l$ and $s_1, s_2, s_n, s_{n+1}, \ldots, s_m$. Every sequence has the weighted support value that is given in the node of constructed tree. Then, the probability of computation is carried out to find the most important sequence for the user. The following equation provides the probability of estimated value for the matched result.

$$s_{n+1}^{(1)} = \{P(s_{n+1} = (s \in s_1, s_2, s_n, s_{n+1}, \ldots, s_k) \mid s_1, s_2, \ldots, s_n)\}$$

$$s_{n+1}^{(2)} = \{P(s_{n+1} = (s \in s_1, s_2, s_n, s_{n+1}, \ldots, s_l) \mid s_1, s_2, \ldots, s_n)\}$$

$$s_{n+1}^{(3)} = \{P(s_{n+1} = (s \in s_1, s_2, s_n, s_{n+1}, \ldots, s_m) \mid s_1, s_2, \ldots, s_n)\}$$

This probability, $pro(s_{n+1} \mid s)$, is estimated by using all sequences of all users in tree structure constructed from the weighted sequential database W_{ij}.

$$P(s_{n+1} = (s \in s_1, s_2, s_n, s_{n+1}, \ldots, s_m) \mid s_1, s_2, \ldots, s_n) = \frac{W\_\sup(s_1, s_2, \ldots, s_n)}{W\_\sup(s_1, s_2, s_n, s_{n+1}, \ldots, s_m)}$$

Then, the final recommendation is based on the following equation:

$$s_{n+1} = \arg\max\{s_{n+1}^{(1)}, s_{n+1}^{(2)}, s_{n+1}^{(3)}\}$$

IV. Results and Discussions

This section presents the results obtained from the experimentation and its detailed discussion about the results. The proposed approach of web page recommendation is experimented with the synthetic dataset and the result is evaluated with the precision, applicability and hit ratio.

a) Experimental Set Up and Dataset Description

The proposed web page recommendation approach is implemented in Java (jdk 1.6). Here, the synthetic dataset is generated as like the same format of real datasets and the performance of the proposed approach is evaluated with the evaluation metrics. The generated synthetic dataset is divided into two parts such as, Training dataset (It is used for building the pattern tree model and test dataset (It is used for testing the web recommendation approach).

b) Performance of the W-Prefixspan Algorithm

Here, the performance of the W-prefixSpan algorithm is analyzed with the execution time, memory usage and patterns mined. At first, the training data is given to the prefixSpan and W-prefixSpan algorithm to mine the sequential patterns and then, the mining performance of the algorithms are analyzed. The values obtained from the experimentation are plotted as graphs, shown in figure 2, 3 and figure 4.

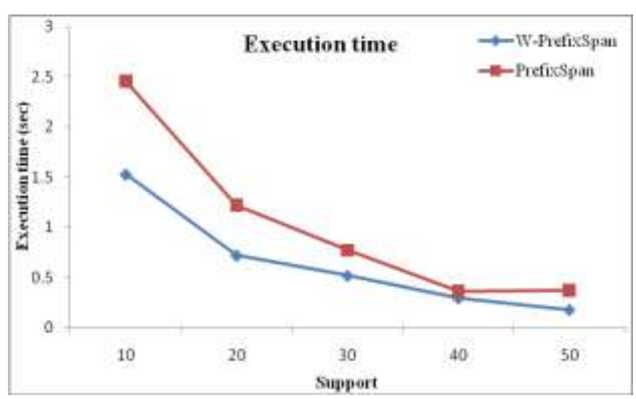

Fig. 2 : Computation time

Global Journal of Computer Science and Technology Volume XII Issue IX Version I April 2012
5

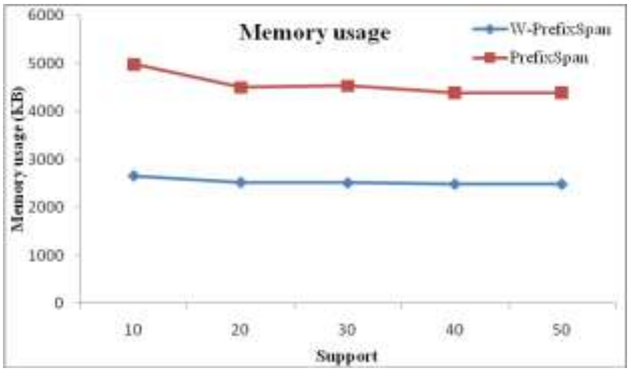

Fig.3: Memory usage

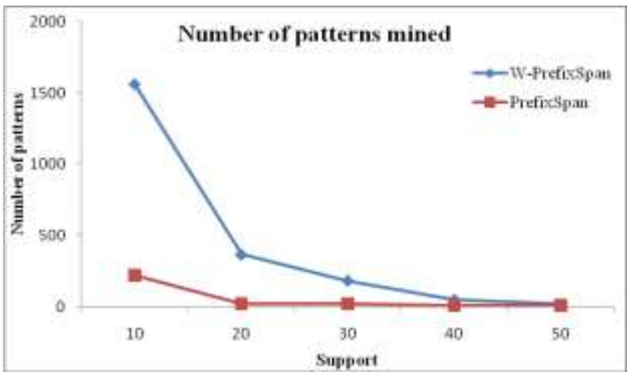

Fig.4 : Number of patterns mined

The figure 2 shows that the execution time of the W-prefixSpan algorithm is 1.5 Sec that is less compared with the time taken of the PrefixSpan algorithm. In memory usage, the W-prefixSpan algorithm needed only less than 50% memory compared with Prefixspan algorithm. Figure 4 shows that number of patterns mined using W-Prefixspan is high compared to Prefixspan algorithm.

V. Conclusion

We have proposed a web page recommendation algorithm using weighted sequential patterns and markov model. Here, we have presented W-PrefixSpan algorithm that is developed by incorporating the weightage constraints such as, spending time and recent visiting with the prefixspan algorithm. The mined weighted sequential patterns are then utilized to construct the recommendation model using the patricia trie-based tree structure. At last, markov model-based recommendation is carried out for the current users by matching the visiting path with the tree and markov model. The experimentation is done with the help of synthetic dataset and the performance of W-Prefixspan algorithm as well as web page recommendation algorithm is analyzed. From the results, the memory required for the W-prefixSpan algorithm is less than 50% of memory needed for PrefixSpan algorithm.

References Références Referencias

1. Baoyao Zhou, Siu Cheung Hui, Kuiyu Chang, "An Intelligent Recommender System using Sequential Web Access Patterns", in proceedings of IEEE Conference on Cybernetics and Intelligent Systems, 2004.
2. Daniel Mican, Nicolae Tomai, "Association-Rules-Based Recommender System for Personalization in Adaptive Web-Based Applications", Proceeding ICWE'10 Proceedings of the 10th international conference on Current trends in web engineering, pp.85-90, 2003.
3. Forsati, R., Meybodi, M.R., Rahbar, A. "An efficient algorithm for web recommendation systems", IEEE/ACS International Conference on Computer Systems and Applications, pp.579-586, 2009.
4. Faten Khalil, Jiuyong Li, Hua Wang, "Integrating Recommendation Models for Improved Web Page Prediction Accuracy", in Proceedings of the thirty-first Australasian conference on Computer science, Vol. 74, 2008.
5. Hiroshi Ishikawa, Manabu Ohta, Shohei Yokoyama, Junya Nakayama, Kaoru Katayama, "On the Effectiveness of Web Usage Mining for Page Recommendation and Restructuring", Proceeding Revised Papers from the NODe 2002 Web and Database-Related Workshops on Web, Web-Services, and Database Systems, pp.253-267, 2003.
6. Jatin D Parmar and Sanjay Garg, "Modified Web Access Pattern (mWAP) Approach for Sequential Pattern Mining", Journal of computer Science, Vol. 6, No.2, pp.46-54, June 2007.
7. Jian Pei, Jiawei Han, Mortazavi-Asl B., Pinto H., Qiming Chen, Dayal U., Mei-Chun Hsu, "PrefixSpan,: mining sequential patterns efficiently by prefix-projected pattern growth", in Proceedings of 17th International Conference on Data Engineering, 2001.
8. Qiankun Zhao and Sourav S. Bhowmick, "Sequential Pattern Mining: A Survey", Technical Report, CAIS, Nanyang Technological University, Singapore, No.118, 2003.
9. Sizu Hou, Xianfei Zhang, "Alarms Association Rules Based on Sequential Pattern Mining Algorithm", In proceedings of the Fifth International Conference on Fuzzy Systems and Knowledge Discovery, vol. 2, pp.556-560, Shandong, 2008.
10. Salvatore Orlando, Raffaele Perego and Claudio Silvestri, "A new algorithm for gap constrained sequence mining", In Proceedings of the ACM Symposium on Applied Computing, pp. 540 – 547, Nicosia, Cyprus, 2004.
11. C.P. Sumathi,R. Padmaja Valli,T. Santhanam, "Automatic Recommendation of Web Pages in Web

Usage Mining", (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 09,pp.3046-3052, 2010.
12. Sarabjot Singh Anand, Bamshad Mobasher, "Intelligent Techniques for Web Personalization", ITWP 2003, LNAI 3169, pp. 1-36, Springer-Verlag Berlin Heidelberg, 2005.
13. Utpala Niranjan, R. B. V. Subramanyam and V. Khanaa, "Developing a Web Recommendation System Based on Closed Sequential Patterns", Communications in Computer and Information Science, Vol. 101, no. 1, pp. 171-179, 2010.
14. Utpala Niranjan, R.B.V. Subramanyam, V.Khana, "An Efficient System Based On Closed Sequential Patterns for Web Recommendations", IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 3, No 4, May 2010.
15. Yicen Liu, Mingrong Liu, Xing Chen, Liang Xiang, Qing Yang "Automatic Tag Recommendation for Weblogs", International Conference on Information Technology and Computer Science, pp.546-549, 2009.

This page is intentionally left blank