An Efficient Web Prediction Model Using Modified Markov Model with ANN

M. Sivasankari M.C.A[1]

[1] *Mphil.Scholar, Department of Computer Science, KG College of Arts and Science,Coimbatore. Tamil nadu, India*

Abstract— **Web prediction is a classification problem in which we try to predict the preceding set of Web pages in which a user may visit supported on the knowledge of the previously visited pages. While serving the Internet user's behavior prediction can be applied effectively in various critical applications. Such application has usual tradeoffs between modeling complexity and prediction accuracy. In this paper we proposed artificial neural network (ANN) for predicting web by the user. In addition modified Markov model has been analysed and presented in prediction of web. A prediction framework uses ANN based on the training samples. By doing this the proposed framework shows the improved prediction time without compromising prediction accuracy.**

Keywords— Markovian model, Artifical neural network, data mining-gram

I. Introduction

Web prediction is a classification scheme in which the next set of Web pages is predicted that user may visit pages depends on the knowledge of the previously visited pages. This information of user's history of navigation within a period of time is said to be *session*. These sessions provides source of data for training which are extracted from the logs of the Web servers, and they hold series of pages that users have visited with the visit date and duration. The Web prediction problem (WPP) can be simplified and utilized in major industrial applications such as wireless applications, caching systems and search engines. Consequently, it is critical to look for practical solutions in which it improves both prediction and training processes. The prediction process can be improved to reduce the user's access times during browsing, and it can simplify the network traffic by neglecting visiting needless pages.

Once a prediction model for a definite Web site is available, the search engine can utilize it to cache the next set of pages that the users might visit. Those caching eases the latency problem of out looking Web documents certainly at the time of Internet traffic congestion periods. An additional extensive application of Web prediction is "personalization," wherein users are sorted based on their desires and interests. The interests and desires of users are confined along with the previously visited sorted Web pages. Moreover, in the wireless domain, Web prediction in mobile is utilized to minimize the number of clicks required in wireless devices such as PDAs and smart phones that assists to alleviate problems related to the display size restrictions. The challenges in web prediction of both preprocessing and prediction are an issue now-a-days. Preprocessing challenges comprise handling huge amount of data that cannot fit in the computer memory, selecting optimum sliding window size, recognizing sessions, and extracting domain information. Prediction challenges comprise long prediction/ training time, memory limitation and low prediction accuracy.

Recently, various location-based social networking sites such as Brightkite, Gowalla, and Foursquare and location based smart phone applications such as Loopt and Google Latitude have appeared. Additionally, a social- networking capability service offers users the skill to post short status updates recognizing their existing activity location by "checking in" at a location. This check-ins can then be shared within a person's social network which is frequently linked to a Twitter account; therefore users are allowed to (1) study about new venues for socializing and (2) assemble with friends at their existing location. It also explores to analyze the connections that exist between customer mobility models and customer behaviors. Formerly most human mobility data which has been learned arrived from cell-phone tower tracking. However, it requires both functional information about the venues visited by the entity and social information about connections. In addition, the majority of customers' behavior prediction depends on the historical preferences by never considering mobility characteristics and manipulation of social network. As a result, incorporating location- based information mutually with social network authority, could possibly improve customer behavior prediction..

.

II. RELATED WORK

Researchers have used a wide range of prediction models with *k*-nearest neighbor (*k*NN) [1], SVMs [2], [3], Bayesian model [4],Markov model [5], [2], and others.

Joachims *et al.* [1] presented the Web Watcher which is a path-based recommender technique based on *k*NN and reinforcement learning. This recommendation uses the combination of previous tours of similar users and reinforcement learning.

Hassan *et al.* [4] presented Bayesian model to spot on definite patterns such as short and long sessions rank of page categories ,page categories and range of page views.

Nasraoui *et al.* [6] presented a Web recommendation system which uses fuzzy inferences. The concept of Clustering is enhanced to group profiles by utilizing hierarchical unsupervised niche clustering. Context-sensitive URL associations are inferred and using a fuzzy approximate-reasoning-based engine.

Mobasher *et al*. [7] use the ARM technique in WPP and presents the frequent item set graph to match an active user session with frequent item sets and predict the next page that user is likely to visit. Still, ARM suffers from well-known limitations such as scalability and efficiency.

Anderson *et al*. [8] use dynamic links that gives shorter path to attain the final destination. Perkowitz and Etzioni [9] make use of adaptive Web sites the browsing activities. Su *et al.* [10] presented the *N*-gram prediction model and applied the all-*N*-gram prediction model in which a number of *N*-grams are built and used in prediction.

Fu *et al.* [10] presented the use of *N-gram* model in mobile Web navigation. The majority of this work has shown that, the lower the order of the *N-gram*, the better the prediction accuracy and performance.

.

III. PREVIOUS WORK

In previous work, web prediction has been done by using Example classifier with markovian model. A prediction model, namely, Example classifier (*EC)*, will be produced and consulted to assign examples to the most suitable classifier. In particular, each predictor in the generated group of prediction models confines strengths and weaknesses of the model based on various factors such as the set of training examples, the flexibility, the structure, and the noise resiliency of the prediction system.

A. MARKOV MODEL

The fundamental concept of Markov model is to predict the subsequent action based on the result of preceding actions. In Web prediction, the subsequent action corresponds to predicting the subsequent page to be visited. The previous actions relates to the previous pages that have been visited already. Particularly in web prediction, the Kth-order Markov model is defined as the probability in which user will visit the k^{th} page offered that the user has visited the ordered k – 1 page. As, in the second-order Markov model, prediction of the subsequent Web page is calculated based on the two Web pages which is visited previously. The major advantages of Markov model are its performance and efficiency in case of model prediction time and Model building. This can be simply exposed that building the k^{th} order of Markov model is linear with the size of the training set. Main idea is to use an well-organized data structure such as hash tables to construct and following each pattern with its probability. Prediction is achieved in constant time since the running time of accessing an admission in a hash table is constant. Particularly, a precise order of Markov model cannot predict for a session that was not viewed in the training set because such session will have zero probability.

B. EXAMPLE CLASSIFIER

Initially all classifiers are trained on the training set T. The output of the training process is N-trained classifiers. While at the mapping phase, each training example e in T is mapped to one or more classifiers that make it to predict its target. For instance, t_1 is mapped to classifier C2, while t_3 is mapped to the set of classifiers C1, C2. The mapped training set T undertakes a filtering process wherein each example is mapped to only one classifier in proportion to the confidence or strength of the classifiers. For illustration, after filtering stage, t_3 in T is mapped to C1 rather than C1, C2 because C2 predicts t_3 accurately with higher probability, for instance. If the models have equivalent prediction confidences, one model is selected randomly. At last, the filtered data set FT is used to train the Example classifier (EC) as the final output.

Algorithm :EC classifier

Step 1: Input set of prediction model of size n, and set of training samples T

Step 2: For each classifier model m in set of prediction model train m on T

Step 3:For each training example e in T and a classifer model m in M

Do

If m predicts the target example e correctly then

Map e to m and store the confidence of m in prediction

Step 4: For each example e in Training T, if e is mapped to more than one model then filter the labels so that only one label is kept

Step 5: Example classifier EC is trained on the training set T’, where T’ is the training set that has all examples in T and each example is mapped to the set of prediction model.

IV. PROPOSED WORK

In this section, we propose another variation of Markov model and prediction framework by reducing the number of paths in the model so that it can fit in the memory and predict faster.

A. *MODIFIED MARKOVAIN MODEL*

Assume the user sessions such as S1=<P1, P2> and S2 = <P2, P1> are two different sessions; therefore, each session can have different prediction probability. The essential idea in the modified Markov model is to assume a set of pages in constructing the prediction model to minimize its size.

For instance consider the session's such as < P1, P2, P3> < P1, P3, P2> < P2, P1, P3>,<P2, P3, P1>,<P3, P1, P2> as one set P1, P2, P3.

The aim of the proposed work is that job on the Web can be completed using distinct paths despite of the ordering that the users choose. Additionally, the size of prediction model is minimized by discarding sessions which have repeated pages.

These sessions results when the user unknowingly clicks on a link and hits the back button.

The K^{th} order of modified Markov model calculates the users probability of visiting the k^{th} page by providing that the user has visited the $k - 1$ pages in any order as in

$$\mathbf{Pr(P(k)|\{P(k\text{-}1),\ldots\ldots\ldots P(k\text{-}n)\})=Pr(P(k)|P(T))}$$

Note that the last page of the session is considered to be the final destination and it is divided from the sessions. For instance the prediction of P1, P2 is P3 and P4 with probabilities 5/6 and 1/6 correspondingly and thus prediction is resolved to P3.

B. PROPOSED PREDICTION FRAMEWORK

.

We present a novel framework for Web navigation prediction. The fundamental idea is to generate different prediction models either by using different classification techniques or by using different training samples. The proposed work uses Modified Markovian model with using artificial neural network for prediction of web user's behaviour.

1. *Artificial Neural Network*:

Artificial neural network is used for the users' behaviour prediction model. In general neural network have been used for predicting users web prediction model behaviour. By depending on the data preprocessing result, multilayer neural network perceptron is trained by back propagation model to learn the relationship between input and outputs. Due to the usage of high dimensionality feature 67 features are used as a input layer. One neuron is used for obtaining the status of time.13 neurons for interactive impact of user. 52 neurons for users' historical decisions. The output layer is obtained for prediction of web behaviour by the user by utilizing thirteen neurons.

Assume that edges between neurons are initialized randomly. In each and every iteration, the given node weighted inputs of the hidden node is summed to obtain the total input of that node..Then the sigmoid activation function is applied at the scale value of interval [0,1] which determines the output. The final output of node in the hidden layer is given by

$$\frac{\mathbf{1}}{\mathbf{1 + \exp}\left(-\sum_{i=1}^{n} \boldsymbol{W_{ij} x_i}\right)}$$

The output of the hidden layer is feds into the output layer and scaled by the other sigmoid activation function to obtain the output values. In training phase, for a given training instance the correct output neuron is given as a target value of one. While others are given the target value zero. The weight between all layers is adjusted based on square error from the target value by using back propagation algorithm. In the testing phase, the output neurons receives larger activation which is nearest to one has been chosen as predicted result.

Algorithm: Prediction framework using Artificial neural network

Step 1: Initialize network with random values of each network layers

Step 2: Initialize weights W_{ij} of each edges randomly

Step 3: Determine total input to the node

Sum the weighted input to a given node in the hidden layer

Step 4: Apply sigmoid activation function to the scale [0,1]

Step 5: Feed output of hidden layers to the output layer of neurons and scale by sigmoid activation function

Step 6: During training, the correct output value of the given training instance gives the target value one and other neuron give the target value zero.

Step 7: Adjust the weight of all layers by back propagation algorithm

Step 8: The prediction output is obtained in the test phase by choosing the largest activation which is chosen as one.

V. EXPERIMENTAL RESULTS

A. DATA SET INFORMATION

The proposed system can be evaluated using NASA data. Additionally reprocessing of dataset can be done. The preprocessing of a data set comprises the following: grouping of sessions, identifying the start and the end of each session, conveying a unique session ID for each session, and filtering irrelevant records. In our experiments, we pursue the cleaning steps and the session identification techniques.

1. ***Prediction Setup:***

Consider a testing session (t) of length L, Then the prediction can be conducted using the $(L - 1)$-gram Markov model and attain the prediction to assess the accuracy of the model. Note that the last page of t is the final outcome that the correctness of the mode is evaluated against; therefore, $(L - 1)$- gram has been used . In case t is larger than the highest N-gram used in the experiment, a sliding window of size L on t has been applied.

The following comparison table shows the changes of Prediction accuracy in existing and proposed system.

TABLE I
COMPARITIVE TABLE FOR EXISITNG AND PROPOSED SYSTEM

No of datas	Markov model with EC	Modified markov model with ANN
1 Gram	0.284	0.315
2 Gram	0.313	0.35
3 Gram	0.200	0.261
4 Gram	0.089	0.116

The following shows the prediction accuracy for proposed and Existing system with respect to the values shown in comparative table 1:

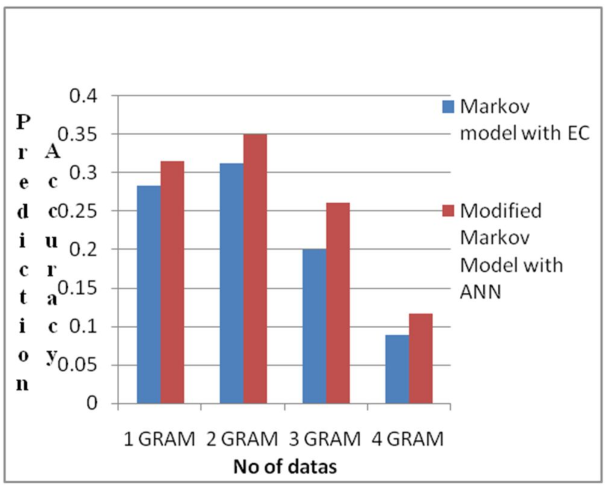

Figure 1: Prediction accuracy graph for proposed system

The above graph in Figure 1 shows the prediction accuracy of Modified Markov model with ANN. From the graph we can observe that the proposed work gives better prediction accuracy than the existing work of markov model with EC with respect to the increase in datas.

VI. CONCLUSION

The present work proposed Modified markov model with Artificial Neural network for web prediction by the user. The proposed work analysed K^{th} Markov model and evaluated prediction accuracy. Modified Markovian model is proposed to minimize the complexity of the original markov model which is used in the earlier method. Modified Markovian model reduces the size of the markov model and simultaneously achieves the higher prediction accuracy of web pages. The prediction framework has been proposed which uses ANN to reduce prediction a time and preserves accuracy. It has been proved that when the network data increases, ANN takes longer training phase and holds information about the datas. Experimental result provides better prediction accuracy result when compare with the Existing work.

REFERENCES

[1] T. Joachims, D. Freitag, and T. Mitchell, “WebWatcher: A tour guide for the World Wide Web,” in *Proc. IJCAI*, 1997, pp. 770–777.

[2] M. Awad, L. Khan, and B. Thuraisingham, “Predicting WWW surfing using multiple evidence combination,” *VLDB J.*, vol. 17, no. 3, pp. 401– 417, May 2008

[3] M. Awad and L. Khan, “Web navigation prediction using multiple evidence combination and domain knowledge,” *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 37, no. 6, pp. 1054–1062, Nov. 2007.

[4] M. T. Hassan, K. N. Junejo, and A. Karim, “Learning and predicting key Web navigation patterns using Bayesian models,” in *Proc. Int. Conf. Comput. Sci. Appl. II*, Seoul, Korea, 2009, pp. 877–887.

[5] J. Pitkow and P. Pirolli, “Mining longest repeating subsequences to predict World Wide Web surfing,” in *Proc. 2nd USITS*, Boulder, CO, Oct. 1999

[6] O. Nasraoui and C. Petenes, “Combining Web usage mining and fuzzy inference for Website personalization,” in *Proc. WebKDD*, 2003, pp. 37–46.

[7] B. Mobasher, H. Dai, T. Luo, and M. Nakagawa, “Effective personalization based on association rule discovery from Web usage data,” in *Proc. ACM Workshop WIDM*, Atlanta, GA, Nov. 2001.

[8] C. R. Anderson, P. Domingos, and D. S. Weld, “Adaptive Web navigation for wireless devices,” in *Proc. IJCAI Workshop*, Seattle, WA, 2001

[9] M. Perkowitz and O. Etzioni, “Adaptive Web sites: An AI challenge,” in *Proc. IJCAI Workshop*, Nagoya, Japan, 1997.

[10] Y. Fu, H. Paul, and N. Shetty, “Improving mobile Web navigation using *N*-Gram prediction model,” *Int. J. Intell. Inf. Technol.*, vol. 3, no. 2, pp. 51–64, 2007.