Analysis of user's Navigational Patterns using Linear Temporal Logic Model

[1]K.Nafeesa Begum [2]R.Pallavi Reddy

[1]PG Student (CSE), G.Narayanamma Institute of Technology & Science (for women), Shaikpet, Hyderabad, Telangana.

[2]Assistant Professor, G.Narayanamma Institute of Technology & Science (for women), Shaikpet, Hyderabad, Telangana.

Abstract:-*Online shopping is becoming more typical in our everyday lives. Understanding clients' interest in which they shop is important in order to adjuste-exchange sites to customers' prerequisites. The records about clients' interest are not considered in the previous server logs. The investigation of such insights has focused on applying mining procedures where static portrayal is utilized to display clients' behaviorand the sequel of clickstream data isn't considered in earlier techniques. Consequently, joining a perspective of the process followed by clients among a session can be of extraordinary scheme to recognize more unpredictable standards behavior. This paper proposes a straight forward Linear Temporal Model checking approach for the assessment of web logs. By characterizing a typical way of mapping log data as per the e-exchange sites, web logs can be effortlessly changed into event logs where the performance of clients is caught. This information is used to identify the clients' behavior based on personal interests.The outcomes analyzed propose a few upgrades to the web business with the aim of developing its proficiency.*

1. INTRODUCTION

In the presentassociated world, the manner in which individuals shop has changed.Individuals are purchasing progressively more over the Internet instead of going customary buying. Web based

business presents clients with the chance of perusing vast item indexes, assessing costs, being constantly proficient, building up a list of things to get and getting offers on higher supplies depending on their interests. This developing trend focused onhighlighting the likelihood of a benefactor to effectively go from one e-exchange to other when their prerequisites are not appeased. As a result, the web based business experts require to know and comprehend customers' conduct while those explore by means of the site, attempting to select the reasons that provoked them to purchase an object, or not. Getting this conduct comprehension will permit internet business sites to convey a more prominent customized service to customers, saving clients and expanding benefits. Pin-pointing buyer' conduct and the reasons that manual purchasing procedure requires is a totally difficult assignment.

Web based business sitesdeliver clients with an extensive variety of navigational options and moves, a client can search by means of item classifications, conform to two or three navigational ways to go to a particular item, or utilize bestcomponents to search for stock, for instance. More often, those individual

conducts are recorded in the web server logs. Web server logs keep, in an arranged way, the sequel of web events created by every individual. The extremely valuable clients' conduct is stored in these logs, which should be found and investigated. A right assessment can be used to enhance the webpage format, to develop and tweak substance, to favoritems, or to perceive the enthusiasm of clients in purchasing one of kind of items, for instance.

Web mining systems have demonstrated their practicality on above issues. Its principal objective is to discover procedure models endeavoring to give a clarification based on the clients' interests. Different techniques have been successfully used in the field of e-commerce, such as classification techniques, clustering, association rules or sequential patterns. Such techniques are a piece of the business insight zone and apply specific algorithms to find concealed examples and connections in expansive datasets.

In this paper the Linear Temporal Logic and model checking techniques has been practised as a substitute. These frameworks have proved their proficiency for open systems. It is a framework presented for structured e-exchange websites. The goal is to analyse the usage of e-exchange websites and explore clients' complex behavioural patterns by means of checking temporal logic formulas describing such behaviours against the log model.

2. RELETED WORK

In the case of e-exchange sites, most data mining procedures consider server logs to extract the groupings of buyer patterns as events. In any of the procedures, those groupings are not instantly mined; as a substitute, each arrangement is changed into a session portrayal. The substance of those frameworks can be different. In the portrayal, it conveys the internet browser utilized by the customer, the assortment of visited web pages, the time the buyer spent on each page, or the key expressions used. Consideration on the clients' enthusiasm for item classifications and their portrayal comprise of the number of visited classes and the recurrence of such visits.

In contrast to the past strategies, by utilization ofdata mining techniques to discover the greatest successive terminologies contained inside the Web pages a customer visits, and discovering the portrayal from these expressions. This framework attempts to recognize the client's interests from the number of the visited pages. Different frameworks develop the portrayal from surveysatisfied by methods for clients or utilize a blend of buyer's acquiring, statistic and individual records. Regardless, when clients' portrayals have been registered,estimatesare ordinarily used to find the arrangements of classes showing a comparable conduct or a couple of regular interests.

This information can at some point or another be utilized to enhance the website substance, redo substance, to prescribe objects, to comprehend clients' conduct related to the purchasing way or to catch the enthusiasm of clients in particular object, for example.Another analyst pursues mining techniques to foresee the individual's conduct. Focus is on the customers' navigational groupings to make probabilistic models prepared to expect the customers' next instant. These frameworks are detailed as Markov chains. In any case, these methods have a few flaws; the strategy of making these designs is computationally pricey. The combo of these clustering frameworks and Markov chains

enhances the expectations of these measurable frameworks. The thought is to initially assemble client sessions applying a couple of clustering frameworks and later create a specific Markov chain for every one of the groups identified.

3. LINEAR TEMPORAL LOGIC MODEL CHECKING

Temporal Logic and modelchecking has been proposed as an alternative to data mining procedures. Such procedures have demonstrated their appropriateness for open frameworks. A methodologyfor its utilization in organized e-exchange sites has been proposed. The intention is to look at the utilization of online business sites and to discover clients' social conduct by methods for checking temporal log formulas describing such behavior against the log model. At the beginning, webserver logs are preprocessed to extricate the detailed traces(groupings of eventsof a buyer session). Traces might be client or framework activities done while a client visits an item or item classification page, while she or he adds an item to the list of things to get, while the web index is utilized, and so on.

Linear temporal logic is a model checking technique which is used for analysis of dynamic scenarios. A trace is considered as a sequence of propositional interpretations. It has operators which are more complex. They are also known as the causality operators. The operators used are G (globally: true now and forever), F (eventually (finally): true now or sometime in the future), X (next: true in the next state), U (until: something true until something else happens).

In LTL model, a formula in a trace can be true or false.

$x \lor \neg y \rightarrow$ true: *is true in the first state of the trace*

$X \neg y \rightarrow$ true, *because y is false in the second state*

$XXy \rightarrow$ false, *because y is false in the third state*

$Gx \rightarrow$ true, *because x is true in all states*

$Gy \rightarrow$ false, *because y is not true in all states*

$GFy \rightarrow$ true: *for all states of the sequence, there is some future point where y is true*

$xU \neg y \rightarrow$ true: *x is true until $\neg y$ becomes true;*

$X(yU \neg x) \rightarrow$ false: *in the next state, y is false but x not (yet) true*

$yUXXy \rightarrow$ true: *in the first state y is true, in the second state XXy is true*

$G(yUXXy) \rightarrow$ not straightforward to check: all states have to be checked foryUXXy in turn, check XXy in all states

The business examiner can utilize a firm of (predefined) temporal logic examples to plan questions that could help him, to discover and perceive the way clients utilize the site. Considering the websitestructureand substance notwithstanding the different sorts of shopper's moves, those questions can investigate the presence of complex causality relationships between events contained in the customer time frames. From thetool point of view, the need of having control on the disappointing performance impacts acquired with the use of a couple of variant checking toolsavailable, uniquely while utilized against enormous models, drove us

towards the enthusiasm of building up a specific model checking gadget named LTL model checking.

For instance, the assessment of proposed method is depicted for an e-exchange site. The instance of study portrays the way rawlogs had been handled, how the traceshave been separated, how clients' personal conduct standards were planned and checked towards the log. Some possible interpretations of the results obtained for the queries as well as some possible actions which could help in the re-design of the website whose aim is to improve it have been provided.

Figure 1 demonstrates the ordinary structure used in web based business sites to arrange and sort items. From the home section different sections can be percolated to category section. Different sub-categories can also be percolated from category section. Some unique sorts of areas might be noticeable.
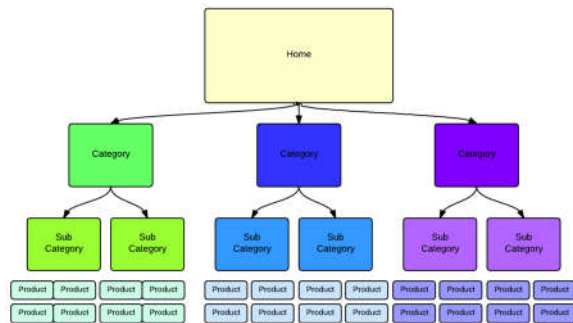

Fig. 1. Product categorization in an e-exchange site

Category segments, which relate to the primary item order. These segments enable access to all items. In general,the item arrangement is generally disjoint, yet this isn't required: in some e-commerce sites a similar item could belong to various areas.

Sub-category segments, give an auxiliary arrangement of the site items, whose goal is to identify a subset of items with some normal and explicit features.In contrast to the past case, not all items must be open from these optional areas. Besides, two unique kinds of optional segments can be recognized relying upon whether items in such areas are for all time or briefly added to the segment. A case of segments with impermanent items would be offers or areas with new items that are periodically renewed.An instance of sub-category sections with permanent links could be sections where products can be accessed by manufacturer, theme, etc.
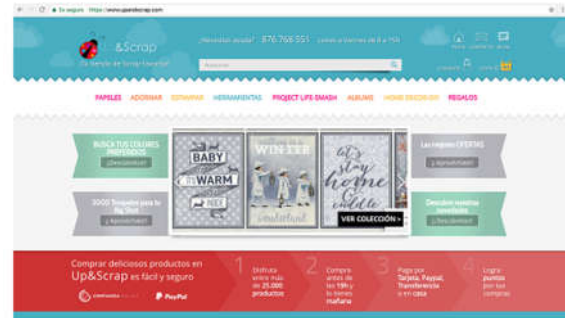

Fig. 2.Homepage of an e-exchange site

• Category sections. The website consists of a major menu with the primary sections. From this menu each of the classes and subcategories, can be accessed. There are eight sections (they are listed in the order as they seem in Figure):

1. Papers

2. Decorate

3. Stamp

4. Tools

5. Project life-smash

6. Albums

7. Home décor-diy

8. Gifts

• Sub-category sections. They are composed of sections to show new merchandise and unique offers, as well as secondary categorizations according to multiple criteria. We can identify 6 sections.

I. Sections with temporary products.

1. Offers

 2. New products

II. Sections with permanent merchandise.

1. Brand

2. Thematics

3. Collections.

4. Designers

Also, the site incorporates a look for search enginethat grants clients to immediately search for items without utilizing the proposed segments. The website, despite its organized format, lets in clients to achieve stock after numerous particular route ways or, even, immediately from the internet searcher. Dissecting clients' conduct could be of a greathelp to improve clients' interactionswith the web webpage.

4.   IDENTIFICATION   OF   USER'S NAVIGATIONAL PATTERNS

With the end goal to empower the utilization of LTL-based model checking with respect to web logs, initially the dataset of an e-exchanges site is uploaded. Once the dataset gets uploaded, it can be viewed in a table format by clicking on the view dataset button. Then the logs are pre-processed in order to obtain a clean log, by removal of wearisome

requests, client session identification, and preparation of log for analysis. After pre-processing, the next step is to identify user behavior in purchasing each product present in different sections with consideration of session time and total time spent by each client. In the next step, the identification of navigational patterns is done by entering each of the clients IP address. The LTL formulae are applied to propose new links to a user based on his personal interests. This procedure helps in converting a one time buyer to a permanent buyer.

The algorithm is as follows.

```
1   Initialize start_time and end_time
2   Initialize session_data and time_data
3   if(arr[1].equals(ip))
4           if(index == 0)
5                   Note start_time and increment index
6           else if(index > 0)
7                   Note end_time calculate total time and session_data
8   Input the time threshold and path threshold
9   Get session_data and time_data
10  Initialize new_link and improve_link to 0
11  Initialize path_length to 1
12  Initialize flag as false
13  if(spent_time.after(time_thr))
14          Append intermediate pages
15          if(path_length > path_th)
16                  identify improved links
17                  add all improved links to new links
18          else
19                  improve_link = 0;
20                  new_link = 0;
21  else
22          increment path_length
23          append intermediate pages
24  if(!flag)
25          Display no links to improve
```

5.EXPERIMENTAL RESULTS

First of all the run.bat file is clicked which opens the Login screen, on that page the username & password is entered and login button is clicked.
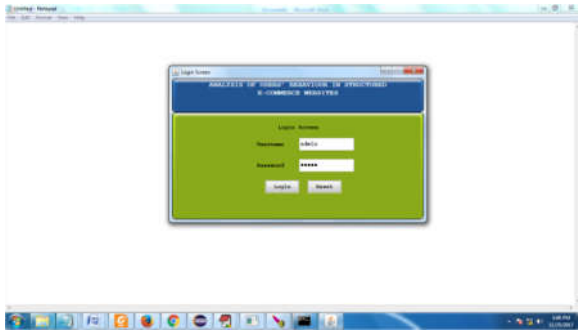
Fig. 3. Login page

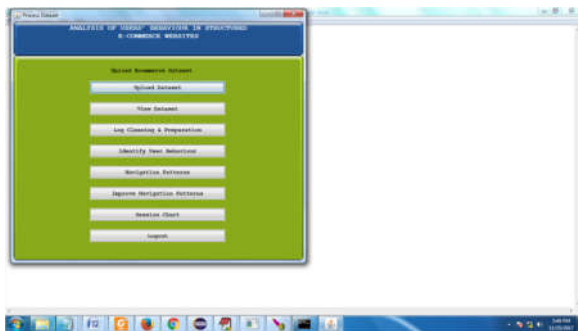It will display the below screen, from where the dataset gets uploaded.

Fig. 4. Home page

After the abovestep, the dataset uploaded can be viewed by clicking on view dataset button. It contains Record numbers, User ip addresses, request types, visiting pages,date and time. After that clicking on log cleaning and preparation button does cleaning of the unwanted data and unnecessary data. After cleaning the dataset, it contains record number, record id, user ip addresses, time stamp, event name, Relative url, operation, status code, L1 section,L2 section.

Fig. 5. Display after log cleaning and preperation

After cleaning once the Identify user behavior button is clicked, user behavior of different products gets displayed.The navigational patterns of an IP address are identified by clicking on Navigational Patterns button and entering the IP address of a user, it shows user navigation pages and urls.After that clicking on improve navigation patterns button, the time threshold is entered and ok button is selected. Then path threshold value is entered and ok button is selected. It displays the improvement results of navigations pages and urls.
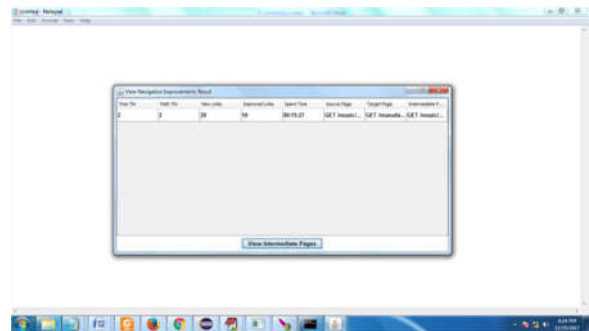

Fig. 6. Display of improved navigational details page

Then view intermediate pages button is clicked to view the pages viewed by a user.The session chart is viewed by clicking on session chart button. It shows record types and count of records.
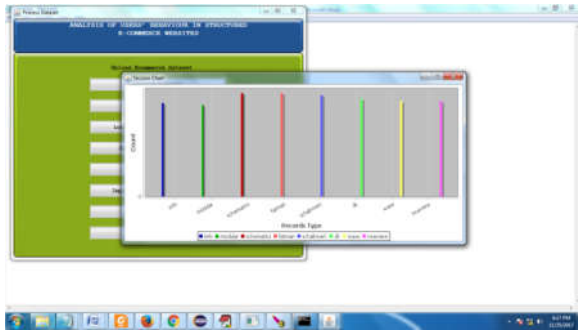
Fig. 7. Session chart

6. CONCLUSION

The main focus is to determine the clients' navigational patterns and also analyze sections that are hugely visited by the clients' and its impact of the purchasing procedure on the website. The improved links equipped to each individual client based on personal interests help the business analysts in remodeling the website. Primarily, the web log is contemplated in structured format and is arranged in a table format for easy access. The LTL model checking techniques are implemented on the pre-processed log to analyze the interrelationship among the sequences of events in the form of clients' navigational patterns that are related to the purchasing procedure. Based on the personal interest of each client in visiting web pages, the business analyst analyzes the log and presents the client with improved links. The reasoning proposes improved links relating to the categorization of products of different web site sections. These improved links not only help to retain a client as a permanent buyer but also help to improve the website in proficiency.

REFERENCES

[1] J. B. Schafer, J. A. Konstan, and J. Riedl, "E-commerce recommendation applications." Hingham, MA, USA: Kluwer Academic Publishers, Jan. 2001, vol. 5, no. 1-2, pp. 115–153.

[2] N. Poggi, D. Carrera, R. Gavalda, J. Torres, and E. Ayguade, "Char- ´ acterization of workload and resource consumption for an online travel and booking site," in Workload Characterization (IISWC), 2010 IEEE International Symposium on. IEEE, 2010, pp. 1–10.

[3] R. Kohavi, "Mining e-commerce data: the good, the bad, and the ugly," in Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2001, pp. 8–13.

[4] W. W. Moe and P. S. Fader, "Dynamic conversion behavior at ecommerce sites," Management Science, vol. 50, no. 3, pp. 326–335, 2004.

[5] G. Liu, T. T. Nguyen, G. Zhao, W. Zha, J. Yang, J. Cao, M. Wu, P. Zhao, and W. Chen, "Repeat buyer prediction for e-commerce," in Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 155–164.

[6] J. D. Xu, "Retaining customers by utilizing technology-facilitated chat: Mitigating website anxiety and task complexity," Information & Management, vol. 53, no. 5, pp. 554 – 569, 2016.

[7] Y. S. Kim and B.-J. Yum, "Recommender system based on click stream data using association rule mining," Expert Systems with Applications, vol. 38, no. 10, pp. 13 320–13 327, 2011.

[8] R. Kosala and H. Blockeel, "Web mining research: A survey," SIGKDD Explor. Newsl., vol. 2, no. 1, pp. 1–15, Jun. 2000.

[9] F. M. Facca and P. L. Lanzi, "Mining interesting knowledge from weblogs: a survey," Data & Knowledge Engineering, vol. 53, no. 3, pp. 225–241, 2005.

[10] C. J. Carmona, S. Ram´ırez-Gallego, F. Torres, E. Bernal, M. J. del Jesus, and S. Garc´ ´ıa, "Web usage mining to improve the design of an ecommerce website: Orolivesur. com," Expert Systems with Applications, vol. 39, no. 12, pp. 11 243–11 249, 2012.