

# Analysis of Users' Behaviour in Structured E-Commerce Websites

Kurabalakota Gurunadh<sup>1</sup>, Dr. M. Sreedevi<sup>2</sup>  
*M.C.A, M.Phil, Ph.D*

**Abstract-** Online shopping is becoming more and more common in our daily lives. Understanding users' interests and behavior is essential in order to adapt e-commerce websites to customers' requirements. The information about users' behavior is stored in the web server logs. The analysis of such information has focused on applying data mining techniques where a rather static characterization is used to model users' behavior and the sequence of the actions performed by them is not usually considered. Therefore, incorporating a view of the process followed by users during a session can be of great interest to identify more complex behavioral patterns. To address this issue, this paper proposes a linear-temporal logic model checking approach for the analysis of structured e-commerce web logs. By defining a common way of mapping log records according to the e-commerce structure, web logs can be easily converted into event logs where the behavior of users is captured. Then, different predefined queries can be performed to identify different behavioral patterns that consider the different actions performed by a user during a session. Finally, the usefulness of the proposed approach has been studied by applying it to a real case study of a Spanish e-commerce website. The results have identified interesting findings that have made possible to propose some improvements in the website design with the aim of increasing its efficiency.

## EXISTING SYSTEM

With the rapid development of Internet in China, the industry's business model has changed. At present, great process has been made in Web e-commerce platform for its convenience and transaction fast. Competition for users is the key factor for e-commerce business in the increasingly fierce competition. If you can grasp customer needs, develop targeted business activities, not only can provide convenient trading mode and a wide choice for customers, but also make the e-commerce business to retain customers better. One of the solutions is Web data mining technology. We can get the user behaviour from the browsing behaviour of

customers on Web and further analysis, then to find a solution. This will allow sellers know more about their customers' needs, and provide personalized according to customer preferences, then obtains the competitive advantage.

## PROPOSED SYSTEM

In today's ever connected world, the way people shop has changed. People are buying more and more over the Internet instead of going traditional shopping. E-commerce provides customers with the opportunity of browsing endless product catalogues, comparing prices, being continuously informed, creating wishlist and enjoying a better service based on their individual interests. This increasing electronic market is highly competitive, featuring the possibility for a customer to easily move from one e-commerce when their necessities are not satisfied. As a consequence, e-commerce business analysts require to know and understand consumers' behaviour when those navigate through the website, as well as trying to identify the reasons that motivated them to purchase, or not, a product. Getting this behavioural knowledge will allow e-commerce websites to deliver a more personalized service to customers, retaining customers and increasing benefits.

- In the characterization contains the web browser used by the customer, the number of visited webpages, the time the customer spent on each page, or the keywords used in search engine focus on the users' interest in the different product categories and their characterization consist of the list of visited categories and the frequency of such visits.
- Unlike the previous approaches, uses text mining techniques to discover the most frequent words contained in the Web pages a customer visits, generating the session characterization from these words. This solution tries to identify the

user's interests from the contents of the visited pages.

- clustering algorithms are generally used to discover the sets of sessions showing a similar behaviour or some common interests.
- This information can subsequently be used to improve the website contents and structure, to adapt and personalize contents to recommend products to understand customers' behaviour related to the buying process or to understand the interest of users in specific products.

Another researchers apply alternative mining techniques to predict the user's behaviour. extract the users' navigational sequences to create statistical and probabilistic models able to predict the user next click. These models are represented as Markov chains. Nevertheless, these approaches present some drawbacks: the process of creating these models is computationally very expensive, and, besides, this type of models responds to very short-term reasoning (the model does not have information to know how the current navigational state has been reached and how future states representing long term goals can be reached). The combination of clustering algorithms and Markov chains improves the predictions of these statistical models, as shown in. The idea is to first group user sessions applying some clustering algorithms and, after, to generate a specific Markov chain for each of the obtained clusters. Currently, there are powerful commercial tools for analyzing logs of e-commerce websites, being Google Analytics one of the main ones. Google Analytics controls the network traffic, collects information about user sessions (first and last web page visited, pages visited, time spent on each page, etc.), and displays reports synthesizing users' behaviour. These traffic-based data can also be combined with other users' personal and geographic information. Google Analytics is not able to import the web server logs of a website, but it works analyzing the information collected by means of page tagging techniques. Another interesting feature of the followed mining approach is the fact of being able to analyze sequences of detailed events. The fact of considering the causal relations of events inside a user session, allowing to look for intra-session patterns (and not only patterns repeated in different sessions) can provide the analysts with a much more detailed perspective of a user behaviour.

MODULE DESCRIPTION

1. Clustering Module
2. Behavioural Module
3. linear-temporal logic Model

Clustering Module:

- clustering algorithms are generally used to discover the sets of sessions showing a similar behaviour or some common interests.

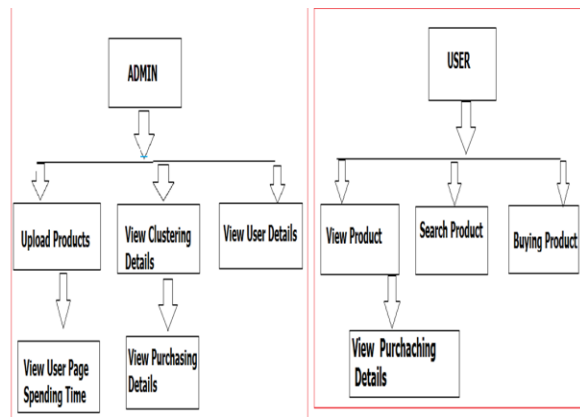
Behavioural Module:

- Behavioural knowledge will allow e-commerce websites to deliver a more personalized service to customers, retaining customers and increasing benefits .
- The goal is to analyse the usage of e-commerce websites and to discover customers' complex behavioural patterns by means of checking temporal logic formulas describing such behaviours against the log mode

Linear - Temporal logic Model:

Linear temporal logic or linear-time temporal logic is a modal temporal logic with modalities referring to time. In LTL, one can encode formulae about the future of paths, e.g., a condition will eventually be true, a condition will be true until another fact becomes true, etc. It is a fragment of the more complex, which additionally allows branching time and quantifiers. Subsequently LTL is sometimes called *propositional temporal logic*, abbreviated Linear temporal logic is a fragment of SIS monadic second-order logic of one successor

ARCHITECTURE DIAGRAM

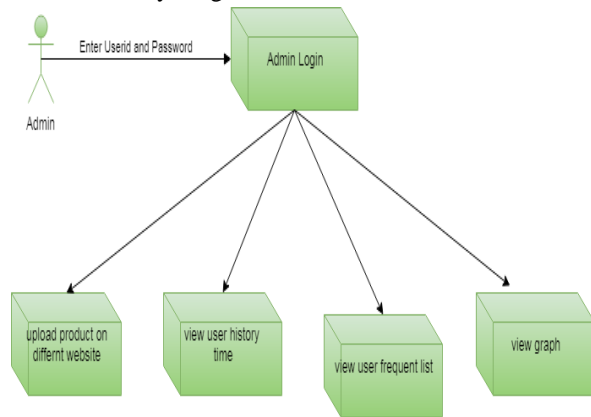


CLUSTERING ALGORITHM

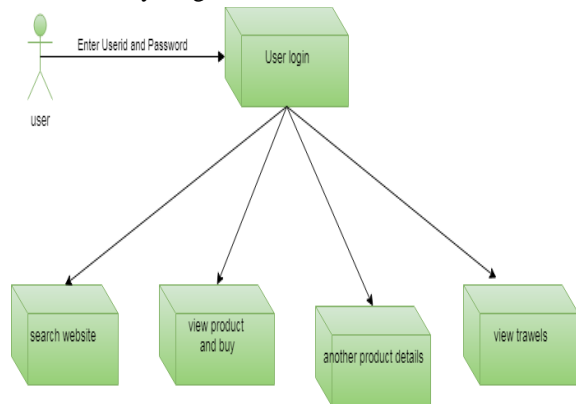
Cluster analysis or clustering is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). It is a main task of exploratory data mining, and a common technique for statistical data analysis, used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, bioinformatics, data compression, and computer graphics.

UML DIAGRAM

Admin activity diagram:



User activity diagram:



CONCLUSION

In the case of open systems, where the sequences of interactions (stored as system logs) are not constrained by a workflow, process mining techniques whose objective is to extract a process

model will usually provide with either over fitting spaghetti models or under fitting flower models, from which little interesting information can be extracted. A more flexible approach is required. In the paper we apply LTL-based model checking techniques to analyse e-commerce web logs. To enable this analysis, we have proposed a common way of representing event types and attributes considering the e-commerce web structure, the product categorization and the possibilities of users to navigate through the website according to such organization. From this structural point of view, The analysis carried out has allowed us to identify several issues and to propose improvements regarding the product categorization and the organization of some of the website sections, which have been transferred to the enterprise managers. Although the paper is strongly related to that website, the proposed approach is general and the methodology is applicable to structured e-commerce websites. The first phase of the methodology, the preprocessing phase, is the one which is specific for each e-commerce website, since it depends on the specific system log and, meanwhile the analysis technique and the queries can be completely reused. it can be executed in parallel, deploying different parallel servers with different parts of the log and executing the queries in parallel. We also plan to extend the set of studied patterns in order to analyze more behavioural patterns and to facilitate their automatic discovery. For that, a side-by-side work with specialists of the problem domain is required in order to define a set of interesting queries as wide as possible. Additionally, extending the web server logs with information about users or online customer reviews is going to be studied. User's information would allow us to study multi session patterns and correlate results with demographic information; while, online reviews would allow us to analyze customer's feedbacks in order to recommend products .