

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

An Implementation for User Behavioral Model of a Website

Mohammed Mahmood Delzy¹, Rıdvan Saraçoğlu², Hakem Beitollahi³

M.Sc. Student, Department of Electrical and Electronics Engineering, Institute of Natural and Applied Science,

Yüzüncü Yıl University, Van, Turkey¹

Associate Professor, Department of Electrical and Electronics Engineering, Institute of Natural and Applied Science,

Yüzüncü Yıl University, Van, Turkey²

Ph.D. in computer engineering, Department of Computer Science, Lecturer at Soran University ,Erbil, Kurdistan

Region, Iraq³

ABSTRACT: This manuscript studies and analyzes the user behavioral model of a website. Our analysis show that users have different views when encounter with a problem and structure of a website. Issues such as sociology, psychology, economics and interest to computer science products make these different views. By studying the behaviors of users, a website can modify and reconstruct its structure to present its information with high quality to its users. Moreover, analyzing the behavior of users makes a great opportunity for a website to predict the behavior of its users in future sessions. Also, it assists a website to recognize the robot machines from human users without using the CAPTCHA puzzles. Normally, any website logs the action of its users in a file. For instance, the time and the date that an IP address creates a session; or the requests that an IP address has issues during a sessions. These helpful information assist us to analyze the behavior of users. We have implemented a software tool that extract various and valuable data from log files of any website. For instance the tool extracts most visiting pages, the favourite pages of users, download rate and session rate of users, the predicted times that a particular user visits a site, uptime and downtime of a particular user, and several other valuable information. We provide comprehensive analysis based on extracted data from the software tool.

KEYWORDS: Web User Behavior, Web Log, Web Mining.

I. INTRODUCTION

We start the problem statement with a useful question. What should be the structure and content of a website in order to attract the most benefit of visitors? By having the accurate users' behavioral model of a website, the website can predict the behavior of its users and update its contents and structure. Predicting the future demand of a user as he visits of pages of the web has got importance with the increasing effectiveness on the web-based [1]. The analysis of a human behavior shows that users have different views when encounter with a problem. Consequently, Studying and analyzing the behaviour of visitors of a website can be benefiting topic for both website owners and computer scientists. The main source of data for a website is the web log files (WLF), which store the action of any user that has contact the website. Such file contains a million records depending on the website traffic. These records are the main source to study the behaviour of users of the website. When a user accesses the website, log files (LF) are generated. LF recorded information about each user [2]. On the web server LFs consist of the purpose visited the user, the order, and the timestamps of these visits [3]. Uses of web server logs improve the performance of the search on the website [4]. The

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

LF can be located in three various places: client browsers, Web proxy servers, and Web servers [5,6]. The main aim of this study is to model the behavior of users of a specific website. The model should be able to predict the steps of a user when contact the website such as user prediction for next page. The model should be able to tell us what the favorite pages for a website are, for which duration they hold a session, what the time spend on a webpage is, how many sessions normally a user open, and many other questions. A software tool is implemented to model the behavior of its visitors. We applied data mining (DM) to analyse access LF records collected on a website [7]. Data mining (DM) is the automated process of finding interesting patterns or information, as well as understandable, descriptive, and predictive models from (huge-scale) data. DM tools predict future trends and behaviors [8, 9]. Web Mining (WM) is the application of DM methodologies to extract information from the data on the web [10]. WM consists of three different types: Web usage mining (WUM), Web content mining (WCM), and Web structure mining (WSM) [11, 12]. WUM also involves three various stages: Preprocessing, Pattern analysis, and Pattern discovery [6]. Our software tool includes several attributes of web users and user traffics such as: Dates, Requests per Day, Requests per Week, Most Page, Requests for Most Pages, Today Visitors, IP (Internet Protocol) address, Download Rate, Time Spent (TS), Location of IP, Web Pages, and Predict for next page. Using our model, a website can modify its structure, design and contents depend on the interests of its visitors. In this case, the website will be more attractive for its users and welcome more users.

II. RELATED WORK

Catledge and Pitkow (1995), analyzed browser user-event WLs to better know user surfing conduct. In specific, they distinguished among directed browser, general goal searching, and random (undirected) searching. They were able to describe behavior of the user in anticipation of improving design of browser [20].

Radinsky et al. (2012), studied how to model and predict behavior of the user along time. They provided various temporal representations, procedures of learning, and represented how they can build models that predict next user behaviors from data of historical [21].

Sarukkai(2000), described a technique for modeling log data and web server to afford foretell of visitor path next link and next page. The model is use of probabilistic making of Markov chains [22].

Zhu et al.(2002), Proposed Markov Chain and web mining to predict the browsing behaviors of users. The Markov chain is proposed to predict the largest probability of future surfing page by calculating every eventuality of every link according the structure of the web and the behaviors of historical browsing [23].

Nagy & Gaspar(2009), investigated the time spent on web pages (TSP) as an indicator ignore of quality of contents of online. They introduced the idea of the sequential surfing and re visitation to more exactly restore the navigation pattern of users depend on TSP and the restored pile of explorer. They provided a clustering path to create groups of analogous pages of web by classifications of spent time. They introduced three activities of the user and connected spent time classifications: browsing links, backward stepping and reading [24].

Verma et al.(2011), a condition for predicting needs of the user to better the usability and the retention of user of a website can be addressed by processing WLF effectively. While browsing the sites, interactions of the users' on the web site are registered in WLF. These WLF are plentiful source of knowledge [25].

Beitollahi &Deconinck(2014), proposed a defense technique against the denial of the application layer of service attacks based on user behavior model. Number of sessions, request rate per session, session duration, number of clicks per each session, mean time between two connections, favorite pages, sequence order between pages, etc. were the main factors to model the user behavior. They used a heuristic technique to model the behavior [26].

Román & Velásquez(2009), have studied behavior of the user with a psychologically-based publishing model. They showed that the series of pages obtained by synthetic agents are found to be nearer to ~5% of the average of the visits number to every page [27].

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

III. MATERIALS AND METHODS

A. Data Mining (DM)

DM techniques is used to extract or discover the patterns and characteristics of users of a website. DM techniques are applied on structures and contents on WLF to achieve personalization and performance of the web [5, 10].

B. Web Mining (WM)

WM is the application of DM methodologies to extract information from the data on the web [13]. WM is one of the main and important fields of DM. WM consists of three various types: Web Usage Mining (WUM), Web Content Mining (WCM), and Web Structure Mining (WSM) [14-16], [10] (As shown in Figure 1).

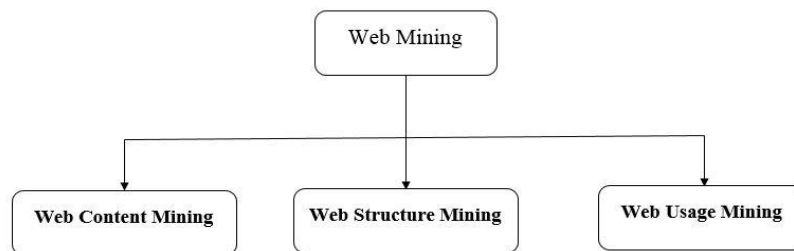


Figure 1: Types of web mining

1. Web content mining (WCM)

WCM is the task of extracting or discovering valuable knowledge from the contents of the website documents. WCM may include of images, video, audio, lists, text and tables. The goal of WCM is to provide an effective mechanism to help the users to find the knowledge they seek [17]; [18].

2. Web Structure mining (WSM)

WSM is the type of WM of discovering the structure of hyperlinks within the page of the web [17]. This includes the use of graph theory to analyze the node structure and connections of the site. According to the nature and kind of the web structure data, it is consist again into two types, the first patterns extraction from the hyperlink on the network. The second mining of the document structure [6].

3. Web Usage Mining (WUM)

WUM focuses on the finding of hidden information from the web user surfing patterns. The initial data source used in WUM is the Web Logs (WL) or log file. In WUM or web log mining, behavior of the user's or benefits are revealed by applying DM techniques on WLF [10,13].

C. Using LF data in WUM

The contents of the WLF are used in WUM. WUM also involves three various stages: Preprocessing, Pattern analysis, and Pattern discovery [6].

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

D. Web Logs and Web User in WM

1. Web User Behavior (WUB)

User behavior is the normal actions performed by web users when surfing and searching, like selecting text, bookmarking, clicking links, printing, etc. [19]. When user login to the website, he is browsing web pages and spent many times on the website for get his purpose which he wanted previously. Web User BehaviorAnalyzer (WUBA) focuses on the evolution of approaches and tools to study behavior of the user's web navigation. When users connect with a website, data recording their conduct which is saved in the logs of the web server. These records are the provenance source to study user's behavior on the website [2].

2. Web Log (WL)

The major source of website data is the WLs, which store the action of any user that has contact the website. Such file contains million records depending on the website traffic. These records are the main source to study the behavior of users of the website. LFs holds more knowledge about the behavior of the users like IP (Internet Protocol) Address, Time Spent, Number of Bytes Transferred, Access Request, and so on. [6]. By analysing these LFs gives a well concept about the web user [2]. When user connect with a website, LF are generated [2]. LF recorded information about every web users [2]. WLF is saved as text file [10]. For every request from browser of a user's to a web server, a response is created in automatically, called a WLF, LF, or WL [10].

Time	Computer	Group	User	Caption	URL
2012-01-06 14:24:57	Kevin	Admin	Kevin	eTinySoft Total Video Converter, Total V...	www.effectmatrix.com/index.htm
2012-01-06 14:22:25	Kevin	Admin	Kevin	PowerPoint convert software - E.M. Pow...	www.effectmatrix.com/PowerPoint-Video-Converter/in...
2012-01-06 14:22:03	Kevin	Admin	Kevin	eTinySoft Total Video Converter, Total V...	www.effectmatrix.com/index.htm
2012-01-06 14:21:41	Kevin	Admin	Kevin	Free PowerPoint to AVI - Free PowerPol...	www.effectmatrix.com/PowerPoint-Video-Converter/Fr...
2012-01-06 14:20:02	Kevin	Admin	Kevin	ppt to video converter - Google Search - ...	https://www.google.com/#clients=psy-ab&hl=en&new...
2012-01-06 14:19:08	Kevin	Admin	Kevin	PowerPoint to Video Converter, Free do...	www.xilisoft.com/ppt-software.html
2012-01-06 14:18:46	Kevin	Admin	Kevin	Xilisoft Video Converter, DVD Ripper, iPa...	www.xilisoft.com
2012-01-06 14:17:28	Kevin	Admin	Kevin	PowerPoint to Video, convert PowerPoi...	www.dvd-ppt-slideshow.com/ppt-to-video/
2012-01-06 14:17:17	Kevin	Admin	Kevin	ppt to video converter - Google Search - ...	https://www.google.com/#clients=psy-ab&hl=en&new...
2012-01-06 14:15:39	Kevin	Admin	Kevin	ppt to video converter - Google Search - ...	https://www.google.com/#clients=psy-ab&hl=en&new...
2012-01-06 14:14:44	Kevin	Admin	Kevin	PPT to DVD Converter: Convert PowerP...	www.wondershare.com/pro/ppt2dvd-pro.html
2012-01-06 14:03:27	Kevin	Admin	Kevin	4Videosoft Video Converter Tools, Best ...	www.4videosoft.com
2012-01-06 11:54:32	Kevin	Admin	Kevin	Download Video Tools, DVD Tools, Pow...	www.leanvsoft.com
2012-01-06 11:52:44	Kevin	Admin	Kevin	Freemith Freeware - Free video convert...	www.freemith.com
2012-01-06 11:19:32	Kevin	Admin	Kevin	Cucusoft Video Converter, DVD to ipod...	www.google.com
2012-01-06 11:18:27	Kevin	Admin	Kevin	Cucusoft Video Converter, DVD to ipod...	www.cucusoft.com
2012-01-06 11:08:11	Kevin	Admin	Kevin	Cucusoft Video Converter, DVD to ipod...	www.cucusoft.com
2012-01-06 10:44:40	Kevin	Admin	Kevin	VSO Software: Video Converter and Bur...	www.vso-software.fr
2012-01-06 10:41:12	Kevin	Admin	Kevin	Google Trends for Websites: www.any-v...	trends.google.com/websites?q=www.any-video-conver...
2012-01-06 10:23:12	Kevin	Admin	Kevin	Google Trends for Websites: www.any-v...	trends.google.com/websites?q=www.any-video-conver...

Web Logs

Figure 2: Sample of web logs

E. Location of a LF

The LF can be situated in three various places: client browsers LF, web proxy servers LF, and web servers LF [6].

IV. IMPLEMENTATION

This is a data mining program designed with the software tool, a software tool is implemented to model the web users behavior and to analyse user behaviors on the website according to web log data file.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

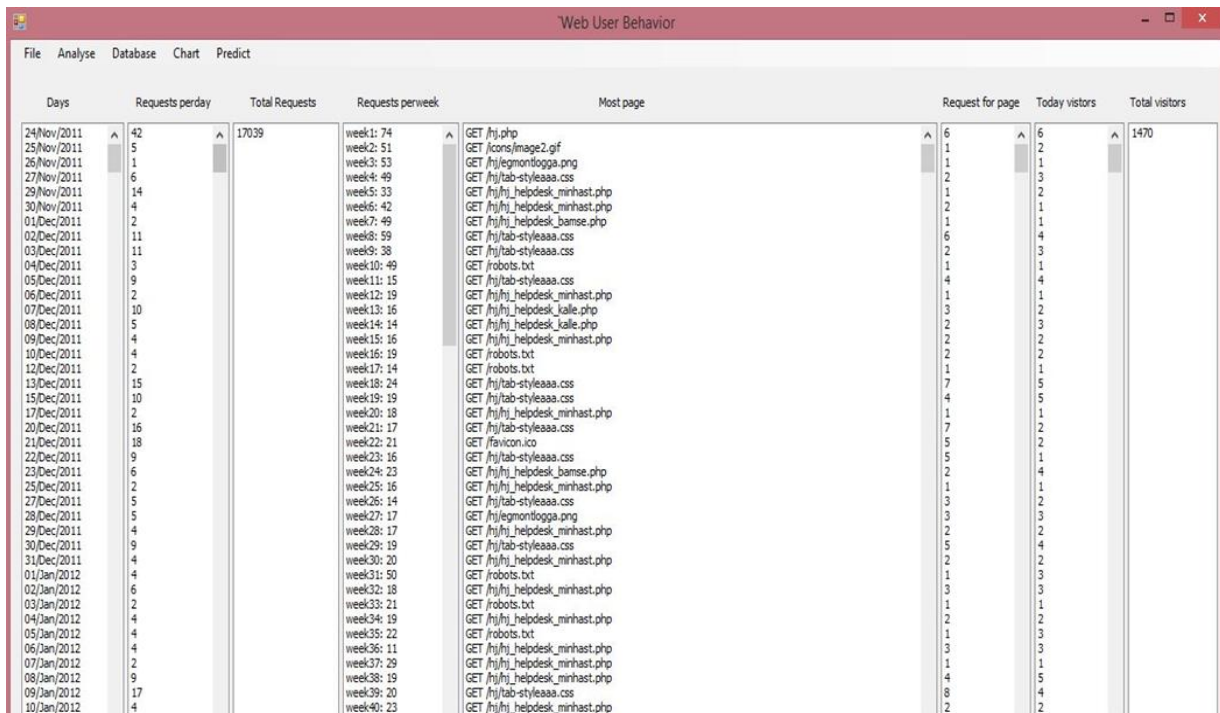


Figure 3: Implementation

Attributes

A server can include different attributes of web users and user traffics.

A. Analyse

To analyse the user's behavior on the website. It includes two sections: General and Specific.

1. General

General is to analyse the web user's behavior in general and it includes several attributes: (As shown in Figure 3)

Days or Dates

The dates or the days show every days which users visit the website. It includes days, months and years. This section holds dates from 24/Nov/2011 to 2/May/2016 (As shown in Figure 3).

Request Rate

The request rate shows the requests number that sends the user to the server in a given period time. The server evaluates request rate for various irregular users through various random times a day, various days, various weeks, etc.

Requests per Day

The request per day shows the requests number that sends the user to the server in a particular day. Requests per day include all the requests which users ask for from web pages in a day. In our program we have 17039 requests as total requests (As shown in Figure 3).

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

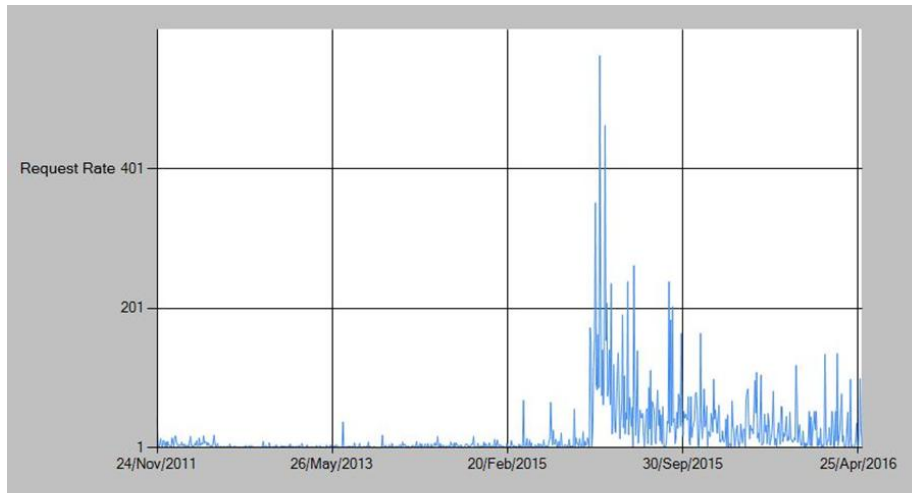


Figure 4: Request rate for request per day

Figure 4 shows request rate for request per day while Y-axis demonstrates request rate and X-axis demonstrates request per day. As it's shown over 65% of users have the rate of the request per day. It implies that over 65% of users have sent a request to the site.

Requests per Week

The requests per week shows the requests number that sends the user to the server in a particular week, requests per week include all those requests which users send to web pages in a week (As shown in Figure 3).

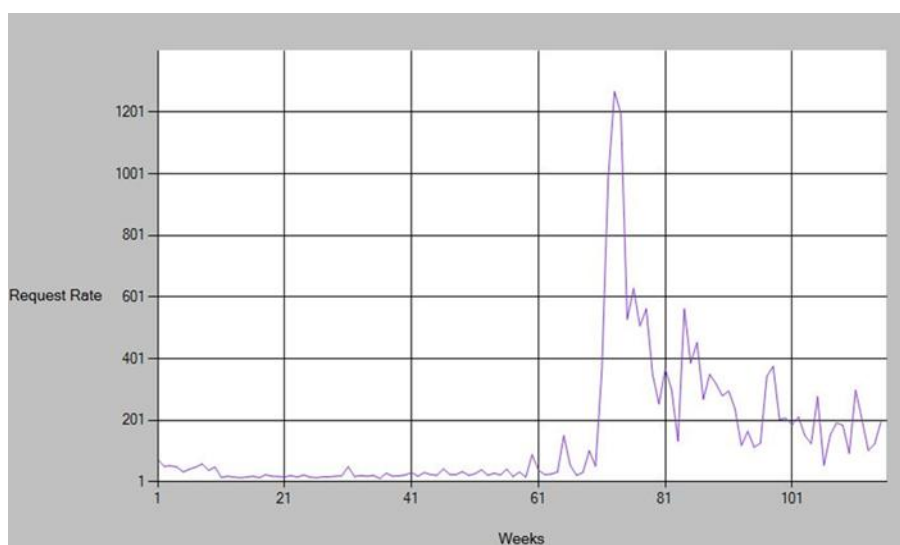


Figure 5: Request rate for request per week

Figure 5 shows request rate for request per week while Y-axis demonstrates request rate and X-axis demonstrates request per week. As it's shown 1268 of users have request in week 73.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

Most Page and Requests for Most Pages

The most page represents the most visited page that a user sends requests to the server for each day. Most page includes all the pages that users open most of the times according to a specific date. The requests for most pages represents the number of requests on the most pages in a specific day (As shown in Figure 3).

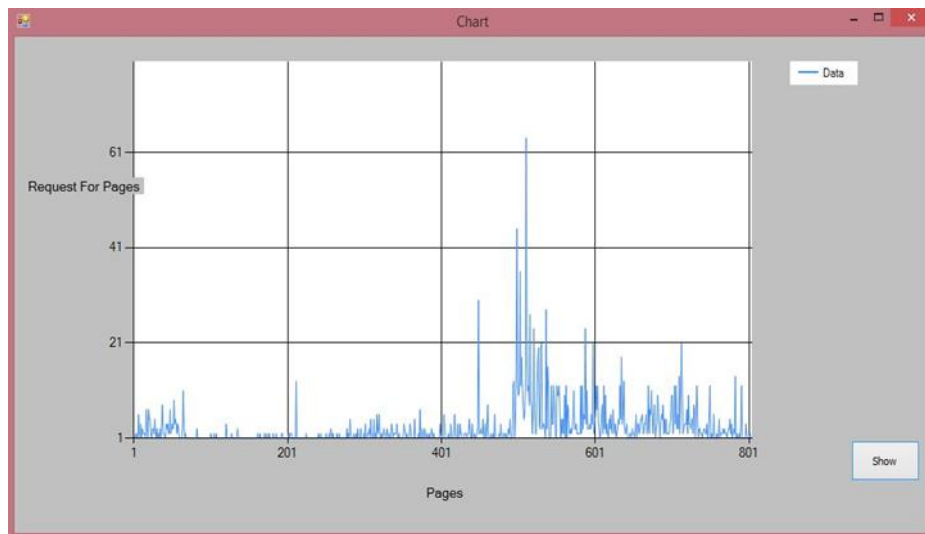


Figure 6: Request rate for most pages

Figure 6 shows request for most pages for pages where Y-axis demonstrates request for most pages and X-axis demonstrates pages. From 01/Jul/2015 users have 64 requests for most page for the page POST/index.php?option=com_jdownloads&Itemid=0&view=upload.

Today Visitors and Total Visitors

Today visitors represents the number of visitors who have visited the website today, namely, it only includes those visitors that have already visited the website in a specific date. total visitors hold all visitors that visit the website during this web log data from 24/Nov/2011 to 02/May/2016 (As shown in Figure 3).

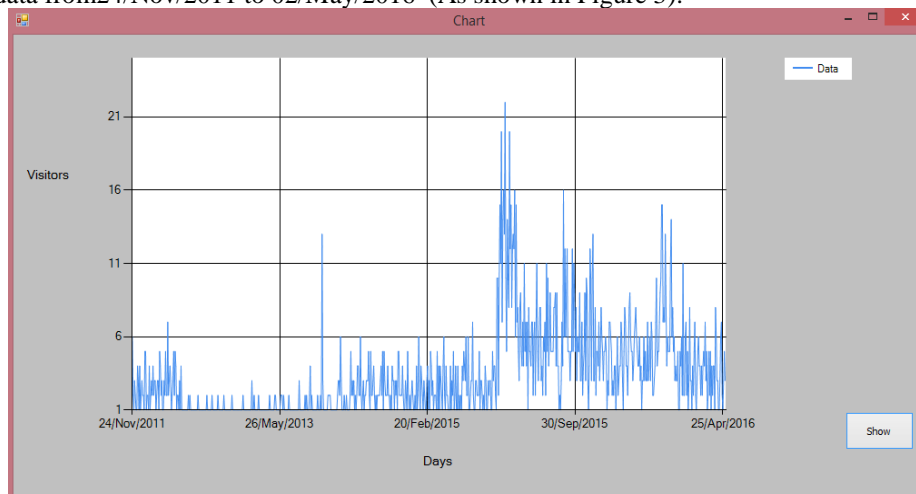


Figure 7: Visitors for days

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

Figure 7 shows visitors for days while Y-axis demonstrates visitors and X-axis demonstrates days. From 22/Jun/2015 As it's shown more of users have visit to the website which is 22 visitors.

2. Specific

The specific is to analyse the behaviors of web users in specific and includes two attributes: IP and Web Pages.

IP (Internet Protocol) address

An IP address is a numerical identity given to every networked PC. IP addresses in LFs are a crude source of knowledge for identifying users. It is supposed for simplicity and for goals in here that a single IP shows as a single user, with some behaviors.

In IP attribute, there are two sections: IP and Day. IP contains all IPs that have visited a website. Day includes those dates which IP have visited the website in this day or date which shows in the section of Day. IP has three parts: File, Analyse and Location of IP (As shown in Figure 8).

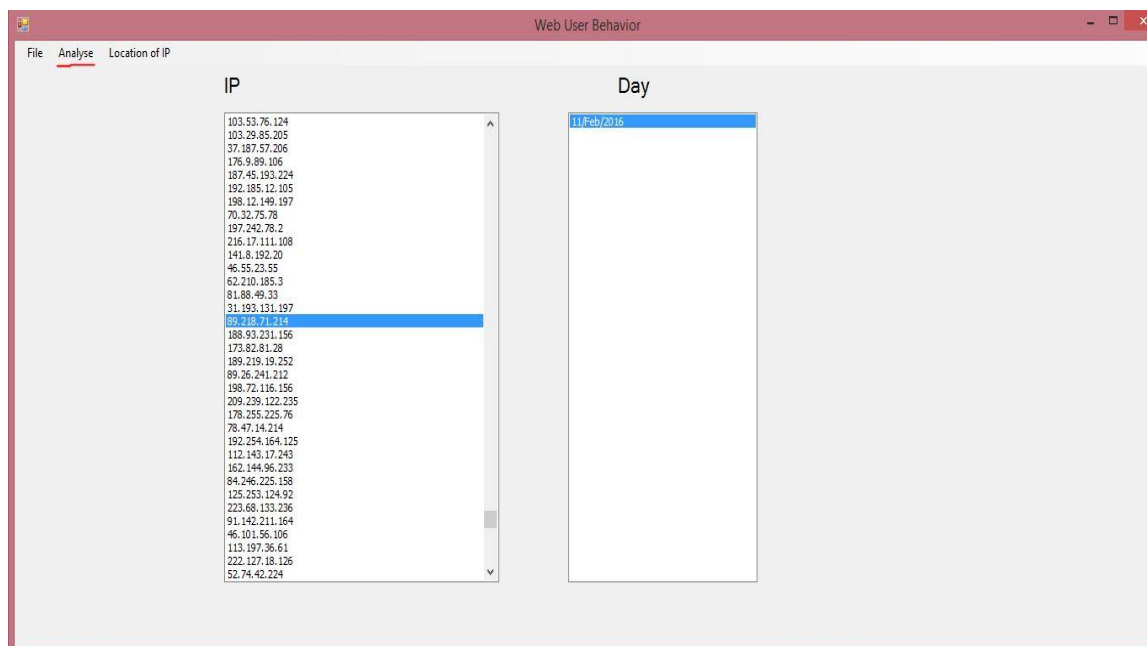


Figure 8: Sample of Analyses IP

File

The file consists of two sections: Open and Exit. Open is for loading the WL data file and Exit is for exiting the section of specific.

Analyse

The Analyse is to analyse user's behavior in the website according to IP and IP date visitings.

For example, here if we want to analyse user's behavior, we choose one IP in section IP meanwhile choosing date in section Day together. Then we choose analyse. We will see many users' behaviors who have visited the web site (As shown in Figure 8).

Analysing web user behavior includes many attributes such as: Requested Pages, Most Page, Request (Request for Most Page), Sequence, Download Rate and Time Spent (Up Time (Online Time) and Down Time (Offline Time)) (As shown in Figure 9).

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

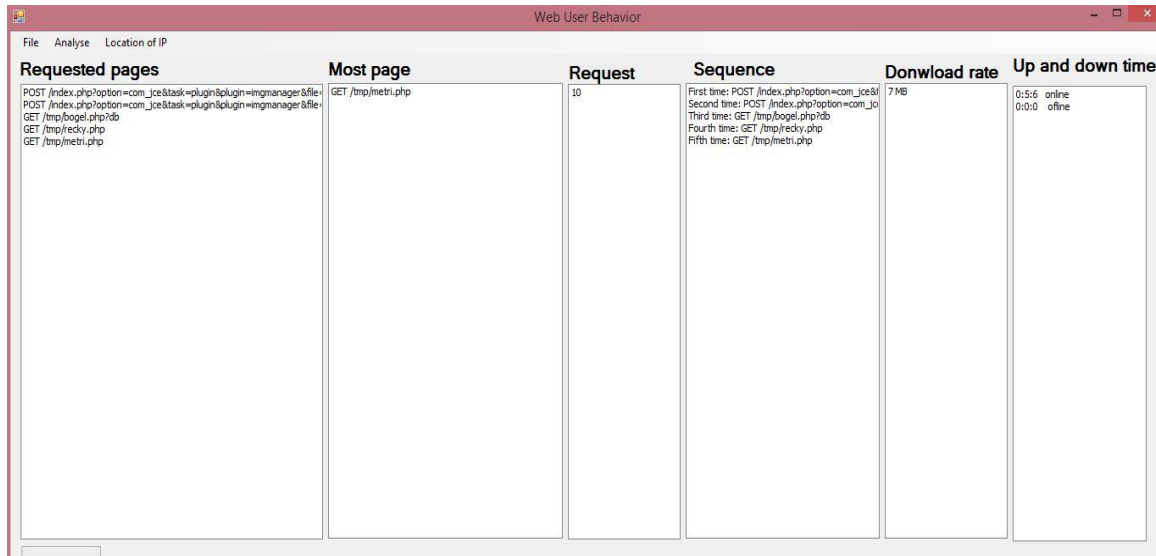


Figure 9: Analysing web user behavior

Requested Pages

The Requested Pages are the ones which represent the users' behaviors. They are the pages which users have requested in a web site (As shown in Figure 9).

Most Page

The Most Page is one of the user's behavior which represents the page which users have requested or opened in the web site the most (As shown in Figure 9).

Request (Requests for Most Page)

The Request is one of the user's behavior representing the requests number for the most page of a web site (As shown in Figure 9).

Sequence

The Sequence is one of the user's behavior which represents the pages that user has requested from the first time to the last time that forms sequence in a web site (As shown in Figure 9).

Download Rate

The download rate is the bytes number that downloads by the user from the server side through a particular interval of time. The server can evaluate download rate by selecting various random users through various times a day, various days, various weeks, etc. The Download Rate is one of the user behavior represents the rate of requests or downloads etc. that a user spends or stays on the web site (As shown in Figure 9).

Time Spent (TS)

The Time Spent is one of the user's behavior that includes two sections: Up Time (Online Time) and Down Time (Offline Time) (As shown in Figure 9).

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

Uptime (Online Time)

The uptime shows a time which a user connects with the site even she cuts off the connection with the server. The server chooses users randomly through various times, various days, various weeks, etc. and computes their online (As shown in Figure 9).

Downtime (Offline Time)

The downtime shows the break time from the time which a particular user disconnects from the server till she joins again to the server. Though it cannot be evaluated for every users, for example, certain users may either come back to the server at no time or connect again to the site too late, certainly particular users re-join every few days or various times a day (As shown in Figure 9).

Location of IP

The Location of IP represents the location of users that connect to the website. Location of IP includes IP, Country and City (As shown in Figure 10).



Figure 10: Location for one IP

Web Pages

Web pages are reference data for web log files. A page of the web is what you see on the screen when you write in the web address, put a query or click on a link. A web page can include color, text, sound, graphics, animation and it can include any kind of information.

Example: we have a lot of web pages, for analysing all items in web pages. When we select any web page, it shows us requests, namely, we will see how many requests each web page has received in the opposite column (As shown in Figure 11).

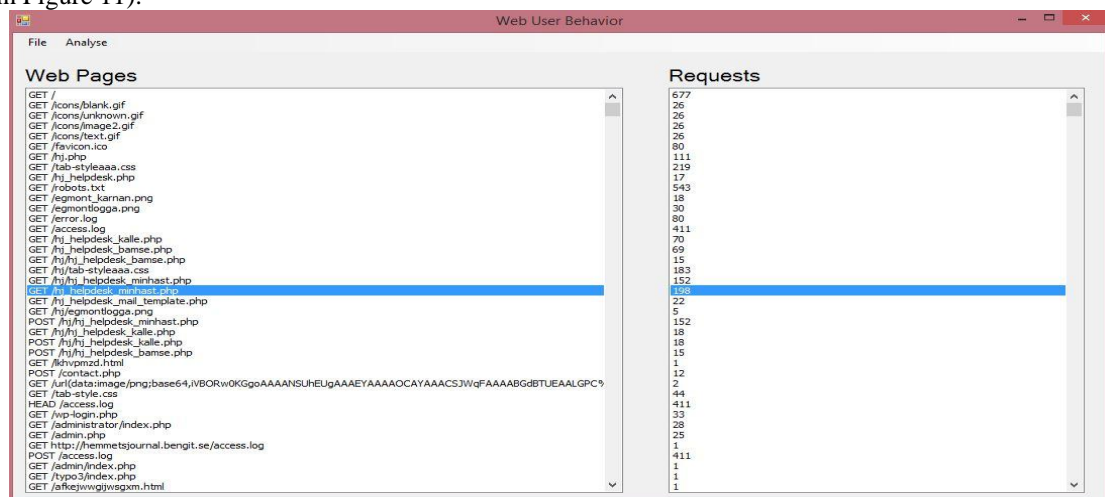


Figure 11: Web Pages with Requests

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

B. Predict

The predict represents user prediction for next page depend on the history of users we can predict for users for future page or next page. Predicting the future demand of a user as he visits of pages of the web has got importance with the increasing effectiveness on the web-based. For each activity in a session of user, the history is defined like the sequence of the observed ordered of conducts. That predicts the future request or next action given the history

Example: we have many IPs, when we choose one IP in section IP. Then we click on the button of User prediction for next page, in section user prediction for next page column, we will see the page which we predicted for this IP according to the history of users (As shown in Figure 12).

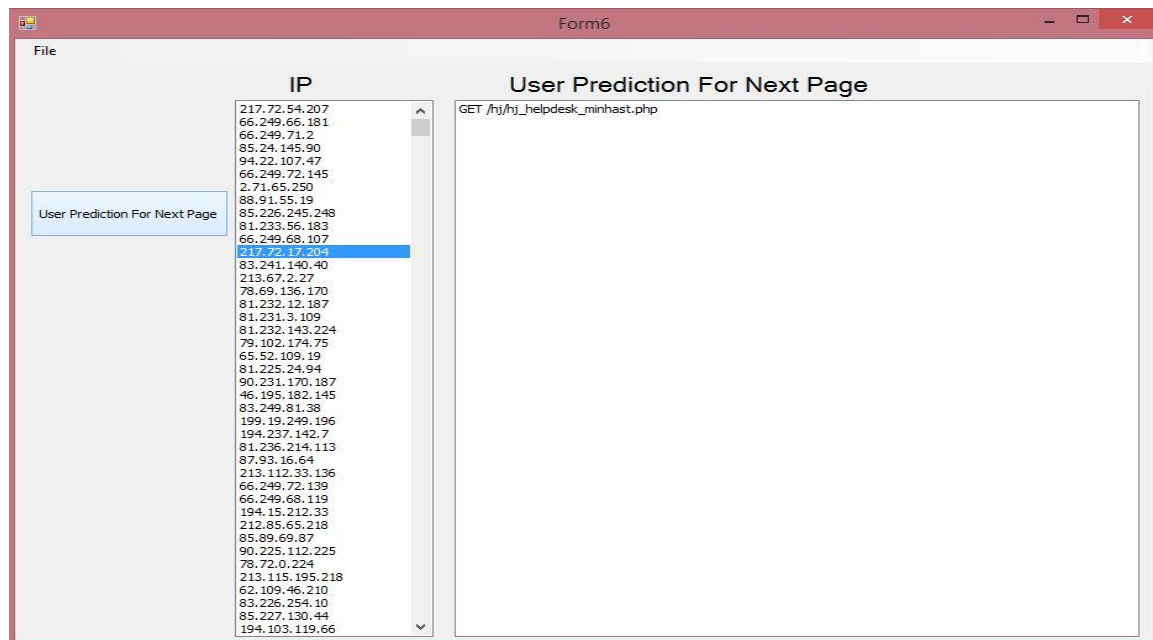


Figure 12:User prediction for next page

V. CONCLUSION

This paper proposes the user behavioral model of a website. The web log of websites is the main source data for our analysis. A software tool is implemented to model the behavior of users of a website. The software tool extracts various helpful information from the web logs of a website. Our data assists the web site administrator to improve the structure of the web and to get the most valuable information about user's behaviors.

Several analysis about users are provided with this paper such as access times of users, the duration of a session, number of requests per session and number of sessions per users during a day, week or month, the favorite pages of users, categorizing the pages based on the willing of users, the location of users and several other main interest information's for a website's administrator. Finally, using these data, we could predict the behavior of users. For instance, we can predict when our well-known users appear in future or what a user requests in his future behavior.

Our work assists a web site to recognize his daily or weekly users and distinguish them from new users; thereby provide them a more and high quality services. Moreover, our work helps a website to recognize human and normal users from robot machines and attackers. For instance in a DDoS attack, our work can be the basis of a real countermeasure.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 3, March 2017

ACKNOWLEDGMENTS

The authors are grateful for the support provided by YüzüncüYıl University.

REFERENCES

- [1] Deshpande, Mukund, and George Karypis. "Selective Markov models for predicting Web page accesses." *ACM Transactions on Internet Technology (TOIT)*, 4.2, 163-184, 2004.
- [2] Satish, Belsare, and Patil Sunil. "Study and Evaluation of user's behavior in e-commerce Using Data Mining." *Research Journal of Recent Sciences ISSN 2277* 2012.
- [3] Herder, E., Characterizations of user Web revisit behavior. In: Workshop on Adaptivity and User Modeling in Interactive Systems ABIS05, pp. pp. 32-37, 2005.
- [4] Patil, SavitaDevidas. "Use of Web Log File For Web Usage Mining." *International Journal of Engineering Research and Technology IJERT*. Vol. 2. No. 4, 2013.
- [5] Suneetha, K. R., and Raghuraman Krishnamoorthi. "Identifying user behavior by analyzing web server access log file." *IJCSNS International Journal of Computer Science and Network Security* 9.4, 327-332, 2009.
- [6] Grace, L. K., V. Maheswari, and Dhinaharan Nagamalai. "Analysis of web logs and web user in web mining." *arXiv preprint arXiv:1101.5668*, 2011.
- [7] Batista, Paulo, and Mário J. Silva. "Mining on-line newspaper web access logs." *12th International Meeting of the Euro Working Group on Decision Support Systems (EWG-DSS 2001)*.
- [8] Aburrous, Maher, et al. "Predicting phishing websites using classification mining techniques with experimental case studies." *Information Technology: New Generations (ITNG), 2010 Seventh International Conference on*. IEEE, 2010.
- [9] Fayyad, Usama M. "Mining databases: Towards algorithms for knowledge discovery." *IEEE Data Eng. Bull.* 21.1, 39-48, 1998.
- [10] El-Yazeed, NM Abo. "An overview of preprocessing of web log files for web usage mining." [Online] Available: etmseg.org/cms/upload/1354650330.Pdf, 1-16, 2010.
- [11] Gupta, Richa. "Journey from data mining to Web Mining to Big Data." *arXiv preprint arXiv:1404.4140*, 2014.
- [12] Alam, Shafiq, Gillian Dobbie, and Patricia Riddle. "Particle swarm optimization based clustering of web usage data." *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology-Volume 03*. IEEE Computer Society, 2008.
- [13] Zubi, Z. S., & Riani, M. S. Applying web mining application for user behavior understanding. *Recent Advances in Image, Audio and Signal Processing*, 2013.
- [14] Borges, Jose, and Mark Levene. "Data mining of user navigation patterns." *Web usage analysis and user profiling*. Springer Berlin Heidelberg, 92-112, 2000.
- [15] Madria, S. K., Bhowmick, S. S., Ng, W. K., Lim, E. P. Research issues in web data mining. In *International Conference on Data Warehousing and Knowledge Discovery* (pp. 303-312), 1999.
- [16] Siddiqui, Ahmad Tasnim, and Sultan Aljahdali. "Web mining techniques in e-commerce applications." *arXiv preprint arXiv:1311.7388*, 2013.
- [17] Iváncsy, Renáta, and István Vajk. "Frequent pattern mining in web log data." *Acta Polytechnica Hungarica* 3.1, 77-90, 2006.
- [18] Srivastava, T., Prasanna Desikan, and Vipin Kumar. "Web mining—concepts, applications and research directions." *Foundations and advances in data mining*. Springer Berlin Heidelberg, 275-307, 2005.
- [19] Velayathan, Ganesan, and Seiji Yamada. "Behavior-based web page evaluation." *Proceedings of the 2006 IEEE/WIC/ACM international conference on Web Intelligence and Intelligent Agent Technology*. IEEE Computer Society, 2006.
- [20] Catledge, Lara D., and James E. Pitkow. "Characterizing browsing strategies in the World-Wide Web." *Computer Networks and ISDN systems* 27.6, 1065-1073, 1995.
- [21] Radinsky, K., Svore, K., Dumais, S., Teevan, J., Bocharov, A., Horvitz, E. Modeling and predicting behavioral dynamics on the web. In *Proceedings of the 21st international conference on World Wide Web*. ACM, (pp. 599-608), 2012.
- [22] Sarukkai, Ramesh R. "Link prediction and path analysis using Markov chains." *Computer Networks* 33.1, 377-386, 2000.
- [23] Zhu, J., Hong, J., Hughes, J. G. Using markov models for web site link prediction. In *Proceedings of the thirteenth ACM conference on Hypertext and hypermedia* (pp. 169-170), 2002.
- [24] Nagy, Istvan K., and Csaba Gaspar-Papanek. "User behaviour analysis based on time spent on web pages." *Web Mining Applications in E-commerce and E-services*. Springer Berlin Heidelberg, 117-136, 2009.
- [25] Verma, Vikas, A. K. Verma, and S. S. Bhatia. "Comprehensive analysis of web log files for mining." *International Journal of Computer Science Issues (IJCSI)* 8.6, 2011.
- [26] Beitollahi, Hakem, and Geert Deconinck. "ConnectionScore: a statistical technique to resist application-layer DDoS attacks." *Journal of Ambient Intelligence and Humanized Computing* 5.3, 425-442, 2014.
- [27] Román, Pablo Enrique, and Juan Domingo Velásquez. "Analysis of the Web User Behavior with a Psychologically-Based Diffusion Model." *AAAI Fall Symposium: Biologically Inspired Cognitive Architectures*. 2009.