

# An Extensible Web Mining Framework for Real Knowledge

N. Pushpalatha

Marri Laxman Reddy Institute of Technology & Management, Hyderabad, India

N. Subhash Chandra

Holy Mary Institute of Technology & Science Hyderabad, India

**Abstract**—With the emergence of Web 2.0 applications that bestow rich user experience and convenience without time and geographical restrictions, web usage logs became a goldmine to researchers across the globe. User behavior analysis in different domains based on web logs has its utility for enterprises to have strategic decision making. Business growth of enterprises depends on customer-centric approaches that need to know the knowledge of customer behavior to succeed. The rationale behind this is that customers have alternatives and there is intense competition. Therefore business community needs business intelligence to have expert decisions besides focusing customer relationship management. Many researchers contributed towards this end. However, the need for a comprehensive framework that caters to the needs of businesses to ascertain real needs of web users. This paper presents a framework named eXtensible Web Usage Mining Framework (XWUMF) for discovering actionable knowledge from web log data. The framework employs a hybrid approach that exploits fuzzy clustering methods and methods for user behavior analysis. Moreover the framework is extensible as it can accommodate new algorithms for fuzzy clustering and user behavior analysis. We proposed an algorithm known as Sequential Web Usage Miner (SWUM) for efficient mining of web usage patterns from different datasets. We built a prototype application to validate our framework. Our empirical results revealed that the framework helps in discovering actionable knowledge.

**Keywords**- Knowledge discovery; Web usage mining, Fuzzy clustering, Business intelligence.

## I. INTRODUCTION

As enterprises in the real world need to know the web usage patterns of their customers, this research is useful to ascertain customer behavior and make strategies to improve customer satisfaction. The web usage mining for user behaviour analysis has many real world utilities. Building a framework that can cater to the needs of enterprises for user behaviour analysis is a challenging task. However, it is very useful to business community to make expert decisions. The user behaviour analysis needs different kinds of algorithms. In the proposed framework placeholders are provided for accommodating any kind of usage mining algorithms and fuzzy logic with combined processing. The fuzzy clustering can provide soft clusters that can be subjected to web usage mining for finding

useful patterns. The patterns when interpreted by domain experts can result in business intelligence. Thus the proposed research has impact on the business community and consumer base as well. We used Wycombe District Council (WDC) dataset collected from Internet sources [1]. An excerpt of dataset is shown in Figure 1.

Publisher Label	Publisher URL	Start date	End date	1 page per visit	2 pages per visit	3 pages per visit	4 pages per visit	5 pages per visit	6 pages per visit	7 pages per visit					
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/01/2008	00:00:00	31/01/2008	00:00:00	19512	4236	4165	5907	2775	1794	1663	1070	930	4534
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/02/2008	00:00:00	29/02/2008	00:00:00	10757	4412	5085	4219	2902	2053	1657	1230	1061	4374
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/03/2008	00:00:00	31/03/2008	00:00:00	28635	4279	5408	4536	3170	2033	1504	1201	1025	5065
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/04/2008	00:00:00	30/04/2008	00:00:00	13048	3554	5307	4209	3008	2007	1532	1356	956	4799
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/05/2008	00:00:00	31/05/2008	00:00:00	21731	4216	4851	3791	2644	1763	1319	956	800	3090
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/06/2008	00:00:00	30/06/2008	00:00:00	10397	4183	4783	3679	2644	1684	1341	972	820	3171
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/07/2008	00:00:00	31/07/2008	00:00:00	20356	4320	5007	3455	2661	1793	1438	1021	867	3963
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/08/2008	00:00:00	31/08/2008	00:00:00	19485	4521	4482	3453	2593	1632	1413	963	835	3570
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/09/2008	00:00:00	30/09/2008	00:00:00	13077	4544	4435	3743	2714	1712	1413	991	840	3096
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/10/2008	00:00:00	31/10/2008	00:00:00	17920	4425	4546	3742	2532	1532	1179	860	724	2545
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/11/2008	00:00:00	30/11/2008	00:00:00	17466	4639	4540	3347	2320	1532	1179	860	724	2545
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/12/2008	00:00:00	31/12/2008	00:00:00	15208	3644	3859	2986	1865	1352	1022	724	589	2751
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/01/2009	00:00:00	31/01/2009	00:00:00	20738	4932	5114	3731	2748	1878	1445	1045	840	3023
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/02/2009	00:00:00	29/02/2009	00:00:00	23085	4392	5079	3751	2707	1833	1479	1062	815	3971
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/03/2009	00:00:00	31/03/2009	00:00:00	21099	4156	5020	3893	2852	1868	1556	1114	867	3916
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/04/2009	00:00:00	30/04/2009	00:00:00	23039	4593	4394	3675	2700	1868	1442	1066	861	3818
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/05/2009	00:00:00	31/05/2009	00:00:00	23575	4891	5164	3643	2735	1820	1488	1089	814	3820
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/06/2009	00:00:00	30/06/2009	00:00:00	25639	5876	5389	4320	2976	1990	1489	1179	956	4298
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/07/2009	00:00:00	31/07/2009	00:00:00	25615	5716	5083	4181	2941	1975	1508	1166	942	4218
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/08/2009	00:00:00	31/08/2009	00:00:00	24629	5588	5087	3896	2936	1885	1433	1077	851	4079
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/09/2009	00:00:00	30/09/2009	00:00:00	24356	5586	5680	3940	2721	1888	1513	1022	819	3653
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/10/2009	00:00:00	31/10/2009	00:00:00	24356	5449	5434	3804	2643	1788	1430	977	861	3435
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/11/2009	00:00:00	30/11/2009	00:00:00	16786	3242	4280	3285	2115	1436	1121	796	596	2597
"Wycombe District Council"	"http://opendatacommunities.org/oc/district-council/wycombe"	01/12/2009	00:00:00	31/12/2009	00:00:00	16446	4619	4182	3756	2529	1490	1081	833	647	1785

Figure 1. An excerpt from WDC dataset.

Different techniques came into existence in order to have user behavior analysis. However, we felt that a framework that can provide extensible features to have user behavior analysis is needed. In this paper we proposed and implemented a framework which is generic but provides placeholders for various future technologies. The framework is extensible and even supports personalized settings for user behavior analysis. The remainder of the paper is structured as follows. Section 2 provides review of literature. Section 3 presents the proposed framework. The section 4 presents experimental results while section 5 concludes the paper besides giving directions for future work.

## II. RELATED WORK

Vythoulkas and Koutsopolous [16] employed neural networks, approximate reasoning, and fuzzy set theory for modelling choice behavior. They assumed that simple rules are used by decision makers. Rules and rule weights are used in the process of behavior analysis. Botha and Solms [17] used trend analysis and fuzzy logic for modeling behavior of intruders. Their approach is proactive based on the combination of trend

analysis and fuzzy logic. Xia, Ho and Capretz [18] Neuro-Fuzzy logic for analyzing software usage trends in the industry. Anderson et al. [19] focused on the analysis of user behavior with respect to fall in human life. Fuzzy logic and voxel person are used to detect fall behavior of humans using pre-recorded videos. Wang et al. [20] employed neural networks and fuzzy logic for behavior of attackers with respect to intrusions. They used KDD CUP 1999 dataset for analyzing behavior of intruders. Mitrovic et al. [21] explored behavior of bloggers with quantitative analysis. They focused on the behavior that is reflected by emotions. They combined machine learning and statistical physics to analyze emotional behavior of web users.

Adadeh, Mohamadi, and Habibi [22] used genetic fuzzy systems for analyzing malicious users. They employed iterative rule learning to have knowhow on the user behavior. A good survey of fuzzy web mining can be found in [23] where techniques pertaining to fuzzy web structure mining, fuzzy web content mining and fuzzy web usage mining. Velesquez [24] combined both web usage mining and eye-tracking technologies for classifying web site key objects. This has provided more effective means of mining pertaining to web usage. He [25] focused on case based reasoning (CBR) and text mining for understanding user experience and improves it. They observed that text mining and Web 2.0 usage can bring about more useful information towards user behavior. Cruz-Benito et al. [26] explored educational virtual world for user behavior. They discovered usage behavior of users in the education domain in the learning process. Conti et al. [27] studied user behavior pertaining Android application usage. They focused on user actions and the trends in the user behavior in using Android applications. Vu et al. [28] focused on travel behavior of tourists. They used the notion of geotagging photos for user behavior analysis. Abello et al. [29] made a survey of semantic web technologies used for Online Analytical Processing (OLAP) which can be used for user behavior analysis.

The different approaches employed in the literature are good for specific purposes. However, we found that there is a need for a comprehensive framework with flexible and extensible technologies that can cope with future technologies as well. In this paper we proposed and implemented a framework which is generic but provides placeholders for various future technologies. The framework is extensible and even supports personalized settings for user behavior analysis.

### III. PROPOSED FRAMEWORK

We proposed a framework which is generic in nature and accommodates future technologies in order to have better performance in user behavior analysis which helps in finding knowhow on web usage. The framework is named as eXtensible Web Usage Mining Framework (XWUMF). The framework provides reusable components or building blocks that can be used along with customized logic. The framework supports a hybrid approach which can have fuzzy clustering techniques and web mining techniques working together to have effective user behavior analysis mechanism. Memory usage and time taken are the two performance evaluation parameters it supports for every operation in the framework.

The framework accommodates new pre-processing techniques, fuzzy clustering techniques and web mining techniques so as to make the proposed framework flexible and extensible. Before presenting our framework, the overview of general web usage mining is shown in Figure 2.

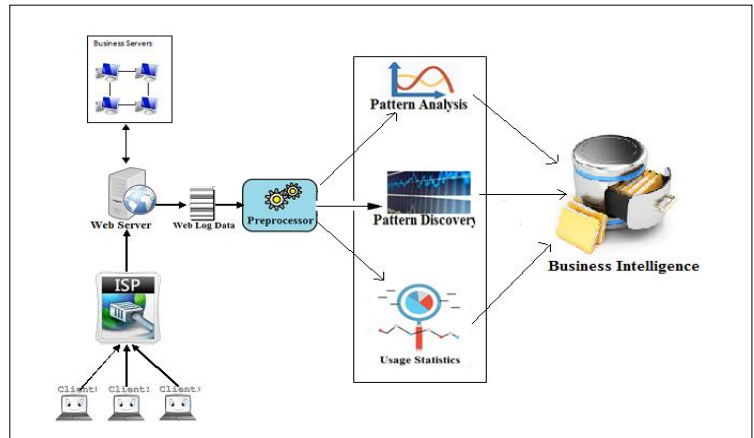


Figure 2. General Web Usage Mining Overview.

Web log data collected from web servers is subjected to pre-processing and then user behavior analysis in order to obtain business intelligence. In this approach the web log data comes from different sources. The data is pre-processed to improve the data by handling missing values. Then the data is subjected to usage mining, pattern discovery and pattern analysis. Finally the method results in business intelligence. The BI is in the form of well-structured patterns that have been interpreted by domain expert.

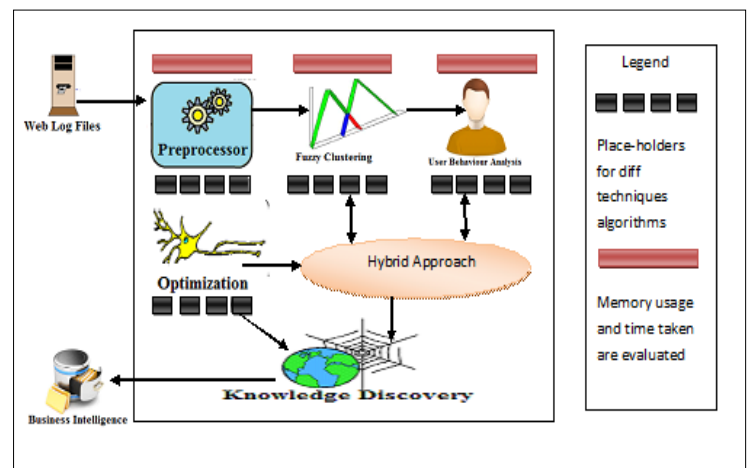


Figure 3. Proposed framework named XWUMF.

As shown in Figure 2, the web log files collected from various sources are subjected to pre-processing. It is similar to that of first method. After pre-processing, the framework supports a hybrid approach which combines fuzzy clustering and user behavior analysis. The framework has placeholders for pre-processing techniques, fuzzy clustering techniques, user behavior analysis and optimization. The framework is flexible and helps developers to build new techniques by referring to existing ones. The framework reflects that fact that future enhancements to the proposed framework are accommodated.

The flexible and generic framework can be realized with web mining as solution for extracting business intelligence from various sources of data. The measures considered for evaluating the proposed work are memory usage and time taken to perform various operations. Our empirical study in this paper is limited to web mining technique only. The full implementation details are presented in our next research paper. The algorithm we implemented as part of the framework is shown here.

---

**Algorithm:** Sequential Web Usage Miner

Inputs : Sequence Web Log Data, *mintime*, *minconf*

Output : User Behavior in terms of Usage Patterns

---

```

01 Initialize usage time vector UT
02 Initialize web log data vector WLD
03 Initialize rules vector R
04 Initialize final rules vector R'
05 Load sequence web log data into WLD
Compute Usage Time
06 For each row in WLD
07   Compute usage time ut
08   Add ut to UT
09   Associate UT with WLD
10 End For
Prune Search Space
11 For each row in WLD
12   IF ut >= mintime THEN
13     Remove row from WLD
14   END IF
15 End For
Compute Rules from WLD
16 Compute rules into R
Validate Rules
17 For each r in R
18   Compute  $conf_{WLD}(r)$ 
19   IF  $conf_{WLD}(r) \geq minconf$  THEN
20     Add r to R'
21   END IF
22 End For
23 Output R'
    
```

---

Algorithm 1. Sequential web usage miner

The algorithm is named Sequential Web Usage Miner (SWUM) which takes sequence web log data, minimum time, and minimum confidence as inputs. It generates patterns that reflect user behavior. First of all usage time of the web pages is computed based on the data provided in the dataset. For instance usage time of users for different web pages is shown in Fig. 1. The usage time is considered to filter out processing by using MinTime parameter. The minimum confidence provides further statistical measure to have quality patterns. The search space is pruned using MinTime and minimum confidence parameters. This process can improve the performance of algorithm for user behavior analysis. Then rules are computed from the web log data. The rules reflect the trends in web usage. Then the rules are validated to have final results in the form of web usage trends.

#### IV. EXPERIMENTAL RESULTS

We made experiments in a PC with i7 processor running Windows 10 operating system. We built a prototype application using Java platform. Java IO package is used to work with dynamics of datasets. Java Swing Application Programming Interface (API) is used for managing Graphical User Interface (GUI). We used Java Collections API for storing data and performing web usage mining. We used the proposed algorithm for mining user behavior in terms of usage patterns. Four datasets are used for experiments. The first dataset is known as Wycombe District Council (WDC) which is collected from [1]. The other three datasets are synthesized datasets to have better evaluation of the proposed algorithm. The important observations are the web usage patterns obtained from the dataset besides the performance measures like execution time and memory usage. Execution time is the measure used to know how fast the proposed algorithm is working while the memory usage is another performance measure used to know how much main memory is needed to process the data for user behavior analysis.

##### A. Results with MinTime between 5000 and 20000

We made experiments with all the four datasets using minimum time usages between 5000 and 20000 seconds. In fact minimum time is the support kind of statistical measure in data mining used to obtain quality results. Table 1 shows the execution times for different datasets observed while performing user behavior analysis.

Min Time	Execution Time (Sec)			
	WDC	Dataset 2	Dataset 3	Dataset 4
5000	0.220	2.828	6.409	14.084
10000	0.193	2.351	5.043	12.295
15000	0.196	2.508	5.354	13.685
20000	0.223	3.520	6.241	13.098

Table 1. Execution time (s) for different datasets

The execution time is more when MinTime parameter is increased. The rationale behind this is the time taken for filtering out tuples in the dataset. There is gradual increase in time taken when MinTime parameter value is increased. In case of the fourth dataset the execution time is 14.08 seconds and 12.29 seconds for MinTime 5000, 10000 similarly the execution time is 13.68 seconds and 13.09 seconds for Dataset four when MinTime is 15000 and 20000. The visualization of the results is shown in Figure 3.

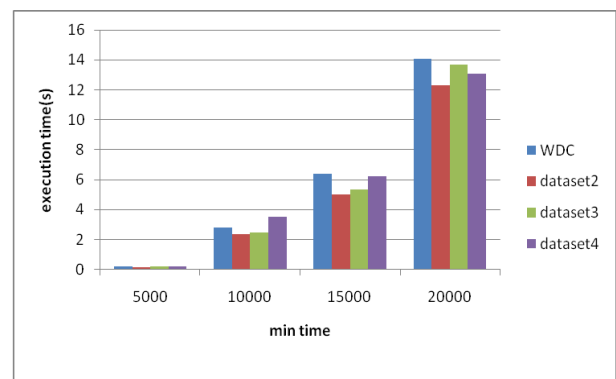


Figure 4. Shows execution time for different datasets when MinTime parameter range is between 5000 and 20000.

The MinTime parameter has its influence on the time taken for user behavior analysis. At the same time there is difference in time taken for different datasets due to the size of datasets.

**B. Results with MinTime between 25000 and 40000**

Min Time	Execution Time (Sec)			
	WDC	Dataset 2	Dataset 3	Dataset 4
25000	0.218	2.724	5.562	13.609
30000	0.312	3.126	5.425	13.882
35000	0.199	3.113	7.428	12.191
40000	0.143	2.559	5.781	14.513

Table 2. Execution time for different datasets when MinTime parameter range is between 25000 and 40000

This experiment is made with different MinTime range that is between 25000 and 40000. WDC shows gradual increase in time taken when MinTime increases. In case of the fourth dataset the execution time is 13.6 and 13.8 seconds for MinTime 25000, 30000 similarly the execution time is 12.2 and 14.5 seconds for Dataset four when MinTime is 35000 and 40000.

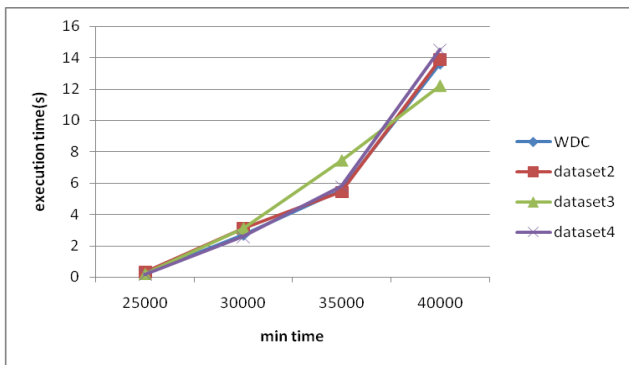


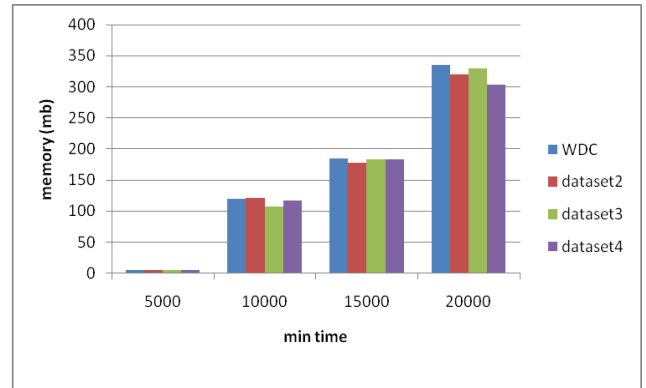
Figure 5. Execution times taken for user behavior analysis for different datasets with MinTime parameter range between 25000 and 40000.

As shown in Fig. 4, it is evident that there is increasing trend with all datasets in execution time with some exceptions. With MinTime 20000, the datasets showed increasing execution time from WDC to Dataset 4. Here the WDC also shows increasing execution time as MinTime increases with exception for Mintime 14000.

MinTime	Memory (MB)			
	WDC	Dataset 2	Dataset 3	Dataset 4
5000	5.947838	119.8514	184.6231	335.3677
10000	5.98597	121.2761	177.7986	320.0214
15000	5.928062	107.6817	183.1708	330.2894
20000	5.928093	116.6732	182.8204	303.5607

Table 3. Memory Usage performance of user behavior analysis for different datasets with MinTime parameter range between 5000 and 20000.

As shown in Table 3, the proposed algorithm needed 5.94 MB of main memory for user behavior analysis using WDC dataset. The three synthetic datasets show almost similar memory usage with slight difference. The maximum memory is consumed by the algorithm to process Dataset 4 while minimum memory is consumed by the algorithm for WDC



dataset.

Figure 6. Memory Usage taken for user behavior analysis for different datasets with MinTime parameter range between 5000 and 20000.

As shown in Fig. 5, the memory usage dynamics are presented. WDC needed less main memory when compared with other datasets. The proposed algorithm makes use of main memory for user behavior analysis.

MinTime	Memory (MB)			
	WDC	Dataset 2	Dataset 3	Dataset 4
25000	5.920708	109.8285	190.4586	329.9598
30000	5.90757	116.37	186.1025	320.612
35000	5.923828	120.343	181.3618	306.5138
40000	5.9076	117.943	182.3883	334.3728

Table 4. Memory Usage performance of user behavior analysis for different datasets with MinTime parameter range between 25000 and 30000

As shown in Table 4, the proposed algorithm needed 5.92 MB of main memory for user behavior analysis using WDC dataset. The three synthetic datasets show almost similar memory usage with slight difference. The maximum memory is consumed by the algorithm to process Dataset 4 while minimum memory is consumed by the algorithm for WDC dataset.



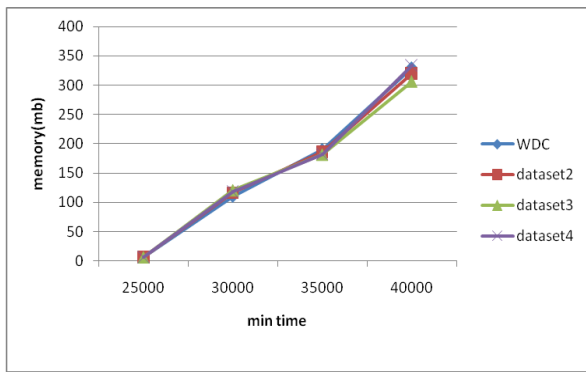


Figure 7. Memory Usage taken of user behavior analysis for different datasets with MinTime parameter range between 25000 and 30000.

As shown in Fig. 6, the memory usage dynamics are presented. WDC needed less main memory when compared with other datasets. The proposed algorithm makes use of main memory for user behavior analysis.

### V. CONCLUSIONS AND FUTURE WORK

In this paper we studied the problem of web usage mining. We proposed a framework named eXtensible Web Usage Mining Framework (XWUMF). The framework supports hybrid approach for processing web log data. Web log data provides usage behavior of customers. User behavior analysis is made using the combination of web mining and fuzzy logic. The proposed framework is flexible and extensible so as to support different combination of techniques in future. User behavior analysis can be made using the proposed algorithm in different domains. The framework is not tied with any kind of domain. Customer-centric approach can be leveraged by enterprises by using the framework for user behavior analysis. The business intelligence in this regard is essential as there is intense competition among business in the real world. In this paper we implemented web usage mining algorithm named Sequential Web Usage Miner (SWUM) for efficient mining of web usage patterns from different datasets. We used four datasets for validating efficiency of the proposed algorithm. We built a prototype application to validate our framework. Our empirical results revealed that the framework helps in discovering actionable knowledge. We implement the rest of the framework in our future work for having accurate results in web usage mining.

### REFERENCES

[1] K. R. Suneetha and Dr. R. Krishnamoorthi. (2009). Identifying User Behavior by Analyzing Web Server Access Log File. *IJCSNS*. 9 (4), p.1-6.

[2] Eugene Agichtein, Eric Brill and Susan Dumais. (2006). Improving Web Search Ranking by Incorporating User Behavior Information. *ACM*. p.1-8.

[3] Gerd Stummea, Andreas Hotho a, Bettina Berendt b. (2006). Semantic Web Mining State of the art and future directions. *ELSEVIER*. p1-20.

[4] Osmar R. Zaiane and Jun Luo. (2001). Towards Evaluating Learners' Behaviour in a Web-Based Distance Learning Environment. *IEEE*. p1-5.

[5] MAGDALINI EIRINAKI and MICHALIS VAZIRGIANNIS. (2003). Web Mining for Web Personalization. *ACM*. 3 (1), p1-27.

[6] Doru Tanasa and Brigitte Trousse. (2004). Advanced Data Preprocessing for Intersites Web Usage Mining. *IEEE*. p1-7.

[7] Cristóbal Romero. (2010). Educational Data Mining: A Review of the State of the Art. *IEEE*. 40 (6), p1-20.

[8] Raymond Kosala and Hendrik Blockeel. (2000). Mining Research: A Survey. *SIGKDD*. 2 (1), p1-15.

[9] Ricardo Baeza-Yates, Carlos Hurtado, Marcelo Mendoza and Georges Dupret . (2005). Modeling User Search Behavior. *IEEE*. p1-10.

[10] Ed H. Chi, Peter Pirolli, Kim Chen, James Pitkow. (2001). Using Information Scent to Model User Information Needs and Actions on the Web. *ACM*. p1-8.

[11] Vincent S. Tseng, Kawuu W. Lin. (2006). Efficient mining and prediction of user behavior patterns in mobile web systems. *ELSEVIER*. p1-13.

[12] Nathan Eagle & Alex (Sandy) Pentland. (2005). Reality mining: sensing complex social systems. *Springer*. p1-14.

[13] Olfa Nasraoui, Esin Saka and Antonio Badia. (2008). A Web Usage Mining Framework for Mining Evolving User Profiles in Dynamic Web Sites. *IEEE*. 20 (2), p1-14.

[14] Raghuram Krishnapuram, Anupam Joshi, Olfa Nasraoui and Liyu Yi. (2001). Low-Complexity Fuzzy Relational Clustering Algorithms for Web Mining. *IEEE*. 9 (4), p1-13.

[15] Yunjuan Xie and Vir V. Phoha. (2001). Web User Clustering from Access Log Using Belief Function. *ACM*. p1-7.

[16] Petros C. Vythoulkas and Haris N. Koutsoulos . (2003). Modeling discrete choice behavior using concepts from fuzzy set theory, approximate reasoning and neural networks. *elsevier*. (n.d), p51-53.

[17] Martin Botha and Rossouw von Solms. (2003). Utilising fuzzy logic and trend analysis for effective intrusion detection. *elsevier*. (n.d), p423-434.

[18] Wei Xia, Danny Ho and Luiz Fernando Capretz. (2008). A Neuro-Fuzzy Model for Function Point Calibration. *5* (n.d), p22-30.

[19] Derek Anderson , Robert H. Luke , James M. Keller , Marjorie Skubic , Marilyn Rantz and Myra Aud . (2009). Linguistic summarization of video for fall detection using voxel person and fuzzy logic. *elsevier*. (n.d), p80-89.

[20] Gang Wang ,Jinxing Hao , Jian Mab, Lihua Huang . (2010). A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering. *elsevier*. (n.d), 6225-6232.

[21] Marija Mitrovi, Georgios Paltoglou and Bosiljka Tadi. (2011). Quantitative analysis of bloggers' collective behavior powered by emotions. *josmte*. (n.d), p1-17.

[22] Mohammad Saniee Abadeh , Hamid Mohamadi and Jafar Habibi. (2011). Design and analysis of genetic fuzzy systems for intrusion detection in computer networks. *Elsevier*, p7067-7075.

[23] Chun-Wei Lin and Tzung-Pei Hong. (2013). A survey of fuzzy web mining. *WIREs Data Mining and Knowledge Discovery*. 3 (n.d), p190-199.

[24] Juan D. and Velasquez . (2013). Combining eye-tracking technologies with web usage mining for identifying Website Keyobjects. *elsevier*. (n.d), p1-10.

[25] Wu He, Improving user experience with case-based reasoning systems using text mining and Web 2.0. *elsevier*. 40 (n.d), 500-507, 2013.

[26] Juan Cruz-Benito , Roberto Therón , Francisco J. García-Peñalvo and Emiliana Pizarro Lucas. (2014). Discovering usage behaviors and engagement in an Educational Virtual World. *elsevier*. (n.d), p1-8.

[27] Mauro Conti, Luigi V. Mancini and Riccardo Spolaor. (2015). Can't You Hear Me Knocking: Identification of User Actions on Android Apps via Traffic Analysis. *acm*. (n.d), p297-304.

[28] Huy Quan Vu , Gang Li , Rob Law and Ben Haobin Ye . (2015). Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. *elsevier*. 46 (n.d), p222-232.

[29] Alberto Abell , Oscar Romero, Torben Bach Pedersen, Victoria Nebot, Maria Jose Aramburu and Alkis Simitis. (2015). Using Semantic Web Technologies for Exploratory OLAP: A Survey. *IEEE*, p1-21.

[30] Wycombe District Council (2016). Website user behaviour. <http://data.wycombe.gov.uk/api/esdInventory>

#### AUTHOR PROFILE

**Mrs. N. Pushpalatha** working as an Associate Professor in Marri Laxman Reddy Institute of Technology and Management, Hyderabad. She has 10 years teaching experience and good knowledge in computer science courses. She completed B.Tech. & M.Tech. Computer Science & Engineering from University College of Engineering, JNTU Campus, Kakinada. She is presently pursuing Ph.D. from JNTU, Hyderabad, India



© 2017 by the author(s); licensee Empirical Research Press Ltd. United Kingdom. This is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license. (<http://creativecommons.org/licenses/by/4.0/>).