



INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATION AND MANAGEMENT

CONTENTS

Sr. No.	TITLE & NAME OF THE AUTHOR (S)	Page No.
1.	FEASIBILITY STUDY OF E-SERVICING ON IRANIAN MUNICIPALITIES (G2C): A CASE STUDY OF AHWAZ MUNICIPALITY <i>DR. MEHRDAD ALIPOUR & SHAHIN KOLIVAND AVARZAMANI</i>	1
2.	ANALYSIS OF MOBILE AGENT BASED E-SUPPLY CHAIN MANAGEMENT SYSTEM USING QUEUING THEORY: A COMPARATIVE STUDY BETWEEN M/M/1 AND M/D/1 MODELS <i>DR. RIKTESH SRIVASTAVA</i>	7
3.	PREPARING PRE-SERVICE TEACHERS TO INTEGRATE EDUCATIONAL TECHNOLOGY IN THE COLLEGES OF EDUCATION CURRICULUM IN THE CENTRAL REGION OF GHANA <i>ABREH MIGHT KOJO</i>	18
4.	THE RELATIONSHIP BETWEEN THE INFORMAL AND FORMAL FINANCIAL SECTOR IN NIGERIA: A CASE STUDY OF SELECTED GROUPS IN LAGOS METROPOLIS <i>ABIOLA BABAJIDE</i>	24
5.	AN APPRAISAL OF SERVICE QUALITY MANAGEMENT IN MANAGEMENT EDUCATION INSTITUTIONS: A FACTOR ANALYSIS <i>DR. BHANWAR SINGH RAJPUROHIT, DR. RAJ KUMAR SHARMA & GOPAL SINGH LATWAL</i>	33
6.	AN EFFECTIVE TOOL FOR BETTER SOFTWARE PRODUCT <i>DR. V.S.P. SRIVASTAV & PIYUSH PRAKASH</i>	44
7.	HUMAN RESOURCE MANAGEMENT ISSUES FOR IMPROVING THE QUALITY OF CARE IN HEALTH SECTOR: AN EMPIRICAL STUDY <i>SAJI MON M.R, N.MUTHUKRISHNAN & DR. D.S. CHAUBEY</i>	49
8.	THE EFFECT OF E-MARKETING AND ITS ENVIRONMENT ON THE MARKETING PERFORMANCE OF MEDIUM AND LARGE FINANCIAL SERVICE ENTERPRISES IN ETHIOPIA <i>TEMESGEN BELAYNEH ZERIHUN & DR. V. SHEKHAR</i>	57
9.	ERGONOMICS RELATED CHANGES ON TRADITIONAL BANKS IN KERALA CONSEQUENT ON CHANGES IN TECHNOLOGY AND ITS IMPACT ON EMPLOYEES <i>DR. P. M. FEROSE</i>	66
10.	MODERN FACES OF FINANCIAL CRIMES IN ELECTRONIC BANKING SYSTEM <i>VIKAS SHARMA</i>	70
11.	QUALITY OF SERVICE (QOS) BASED SCHEDULING ENVIRONMENT MODEL IN WIMAX NETWORK WITH OPNET MODELER <i>ARUN KUMAR, DR. A K GARG & ASHISH CHOPRA</i>	73
12.	A DECENTRALIZED INDEXING AND PROBING SPATIAL DATA IN P2P SYSTEM <i>T. MAHESHWARI & M. RAVINDER</i>	78
13.	CONVERGENCE TO IFRS - AN INDIAN PERSPECTIVE <i>CA SHOBANA SWAMYNATHAN & DR. SINDHU</i>	81
14.	COMPARING EFFICIENCY AND PRODUCTIVITY OF THE INDIAN AUTOMOBILE FIRMS – A MALMQUIST –META FRONTIER APPROACH <i>DR. A. VIJAYAKUMAR</i>	86
15.	EMERGING TRENDS IN KNOWLEDGE MANAGEMENT IN BANKING SECTOR <i>DR. DEEPIKA JINDAL & VIVEK BHAMBRI</i>	93
16.	A STUDY ON CONSUMER ACCEPTANCE OF M-BANKING IN TIRUCHIRAPPALLI CITY <i>S. MOHAMED ILIYAS</i>	97
17.	TECHNICAL ANALYSIS AS SHORT TERM TRADING STRATEGY IN THE INDIAN STOCK MARKET- AN EMPIRICAL EVIDENCE IN THE PUBLIC SECTOR BANKS <i>S. VASANTHA</i>	102
18.	SOFTWARE DEFECTS IDENTIFICATION, PREVENTIONS AND AMPLIFICATION IN SDLC PHASES <i>BHOJRAJ HANUMANT BARHATE</i>	114
19.	A STUDY ON TIME MANAGEMENT IN EMERGENCY DEPARTMENT THROUGH NETWORK ANALYSIS IN A CORPORATE HOSPITAL <i>DR. L. KALYAN VISWANATH REDDY & HENA CHOWKSI</i>	118
20.	MAINTAINING CENTRALIZED BANK INFORMATION FOR GETTING QUICK ACCESS OF INFORMATION OF ALL OTHER ACCOUNTS USING DENORMALIZATION OF DATABASE CONCEPT OF COMPUTER <i>AMIT NIVARGIKAR & PRIYANKA JOSHI</i>	124
21.	DIGITAL OPPORTUNITIES IN NORTH INDIA: A STUDY ON DIGITAL OPPORTUNITY PARAMETERS AMONG NORTH INDIAN STATES <i>DEEP MALA SIHINT</i>	126
22.	BUSINESS ETHICS & GOVERNANCE <i>ARIF SULTAN, FATI SHAFAT & NEETU SINGH</i>	131
23.	EMPLOYEES' PERCEPTION ON TRAINING AND DEVELOPMENT (A STUDY WITH REFERENCE TO EASTERN POWER DISTRIBUTION OF AP LIMITED) <i>DR. M. RAMESH</i>	134
24.	AN OPTIMAL BROKER-BASED ARCHITECTURE FOR TRANSACTIONAL AND QUALITY DRIVEN WEB SERVICES COMPOSITION <i>KAVYA JOHNY</i>	140
25.	WEB USAGE MINING: A BOON FOR WEB DESIGNERS <i>RITIKA ARORA</i>	148
	REQUEST FOR FEEDBACK	151

A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories

Indexed & Listed at: [Ulrich's Periodicals Directory](#) ©, [ProQuest, U.S.A.](#), [Index Copernicus Publishers Panel, Poland](#), [Open J-Gate, India](#),

[EBSCO Publishing, U.S.A.](#), as well as in [Cabell's Directories of Publishing Opportunities, U.S.A.](#)

Circulated all over the world & Google has verified that scholars of more than Hundred & Eighteen countries/territories are visiting our journal on regular basis.

Ground Floor, Building No. 1041-C-1, Devi Bhawan Bazar, JAGADHRI – 135 003, Yamunanagar, Haryana, INDIA

www.ijrcm.org.in

CHIEF PATRON

PROF. K. K. AGGARWAL

Chancellor, Lingaya's University, Delhi
Founder Vice-Chancellor, Guru Gobind Singh Indraprastha University, Delhi
Ex. Pro Vice-Chancellor, Guru Jambheshwar University, Hisar

PATRON

SH. RAM BHAJAN AGGARWAL

Ex. State Minister for Home & Tourism, Government of Haryana
Vice-President, Dadri Education Society, Charkhi Dadri
President, Chinar Syntex Ltd. (Textile Mills), Bhiwani

CO-ORDINATOR

MOHITA

Faculty, Yamuna Institute of Engineering & Technology, Village Gadholi, P. O. Gadhola, Yamunanagar

ADVISORS

DR. PRIYA RANJAN TRIVEDI

Chancellor, The Global Open University, Nagaland

PROF. M. S. SENAM RAJU

Director A. C. D., School of Management Studies, I.G.N.O.U., New Delhi

PROF. S. L. MAHANDRU

Principal (Retd.), Maharaja Agrasen College, Jagadhri

EDITOR

PROF. R. K. SHARMA

Professor, Bharti Vidyapeeth University Institute of Management & Research, New Delhi

CO-EDITOR

MOHITA

Faculty, Yamuna Institute of Engineering & Technology, Village Gadholi, P. O. Gadhola, Yamunanagar

EDITORIAL ADVISORY BOARD

DR. RAJESH MODI

Faculty, Yanbu Industrial College, Kingdom of Saudi Arabia

PROF. PARVEEN KUMAR

Director, M.C.A., Meerut Institute of Engineering & Technology, Meerut, U. P.

PROF. H. R. SHARMA

Director, Chhatarpati Shivaji Institute of Technology, Durg, C.G.

PROF. MANOHAR LAL

Director & Chairman, School of Information & Computer Sciences, I.G.N.O.U., New Delhi

PROF. ANIL K. SAINI

Chairperson (CRC), Guru Gobind Singh I. P. University, Delhi

PROF. R. K. CHOUDHARY

Director, Asia Pacific Institute of Information Technology, Panipat

DR. ASHWANI KUSH

Head, Computer Science, University College, Kurukshetra University, Kurukshetra

DR. BHARAT BHUSHAN

Head, Department of Computer Science & Applications, Guru Nanak Khalsa College, Yamunanagar

DR. VIJAYPAL SINGH DHAKA

Head, Department of Computer Applications, Institute of Management Studies, Noida, U.P.

DR. SAMBHAVNA

Faculty, I.I.T.M., Delhi

DR. MOHINDER CHAND

Associate Professor, Kurukshetra University, Kurukshetra

DR. MOHENDER KUMAR GUPTA

Associate Professor, P. J. L. N. Government College, Faridabad

DR. SAMBHAV GARG

Faculty, M. M. Institute of Management, Maharishi Markandeshwar University, Mullana

DR. SHIVAKUMAR DEENE

Asst. Professor, Government F. G. College Chitguppa, Bidar, Karnataka

DR. BHAVET

Faculty, M. M. Institute of Management, Maharishi Markandeshwar University, Mullana

ASSOCIATE EDITORS**PROF. ABHAY BANSAL**

Head, Department of Information Technology, Amity School of Engineering & Technology, Amity University, Noida

PROF. NAWAB ALI KHAN

Department of Commerce, Aligarh Muslim University, Aligarh, U.P.

DR. ASHOK KUMAR

Head, Department of Electronics, D. A. V. College (Lahore), Ambala City

ASHISH CHOPRA

Sr. Lecturer, Doon Valley Institute of Engineering & Technology, Karnal

SAKET BHARDWAJ

Lecturer, Haryana Engineering College, Jagadhri

TECHNICAL ADVISORS**AMITA**

Faculty, Government M. S., Mohali

MOHITA

Faculty, Yamuna Institute of Engineering & Technology, Village Gadholi, P. O. Gadholi, Yamunanagar

FINANCIAL ADVISORS**DICKIN GOYAL**

Advocate & Tax Adviser, Panchkula

NEENA

Investment Consultant, Chambaghat, Solan, Himachal Pradesh

LEGAL ADVISORS**JITENDER S. CHAHAL**

Advocate, Punjab & Haryana High Court, Chandigarh U.T.

CHANDER BHUSHAN SHARMA

Advocate & Consultant, District Courts, Yamunanagar at Jagadhri

SUPERINTENDENT**SURENDER KUMAR POONIA**

CALL FOR MANUSCRIPTS

We invite unpublished novel, original, empirical and high quality research work pertaining to recent developments & practices in the area of Computer, Business, Finance, Marketing, Human Resource Management, General Management, Banking, Insurance, Corporate Governance and emerging paradigms in allied subjects like Accounting Education; Accounting Information Systems; Accounting Theory & Practice; Auditing; Behavioral Accounting; Behavioral Economics; Corporate Finance; Cost Accounting; Econometrics; Economic Development; Economic History; Financial Institutions & Markets; Financial Services; Fiscal Policy; Government & Non Profit Accounting; Industrial Organization; International Economics & Trade; International Finance; Macro Economics; Micro Economics; Monetary Policy; Portfolio & Security Analysis; Public Policy Economics; Real Estate; Regional Economics; Tax Accounting; Advertising & Promotion Management; Business Education; Business Information Systems (MIS); Business Law, Public Responsibility & Ethics; Communication; Direct Marketing; E-Commerce; Global Business; Health Care Administration; Labor Relations & Human Resource Management; Marketing Research; Marketing Theory & Applications; Non-Profit Organizations; Office Administration/Management; Operations Research/Statistics; Organizational Behavior & Theory; Organizational Development; Production/Operations; Public Administration; Purchasing/Materials Management; Retailing; Sales/Selling; Services; Small Business Entrepreneurship; Strategic Management Policy; Technology/Innovation; Tourism, Hospitality & Leisure; Transportation/Physical Distribution; Algorithms; Artificial Intelligence; Compilers & Translation; Computer Aided Design (CAD); Computer Aided Manufacturing; Computer Graphics; Computer Organization & Architecture; Database Structures & Systems; Digital Logic; Discrete Structures; Internet; Management Information Systems; Modeling & Simulation; Multimedia; Neural Systems/Neural Networks; Numerical Analysis/Scientific Computing; Object Oriented Programming; Operating Systems; Programming Languages; Robotics; Symbolic & Formal Logic and Web Design. The above mentioned tracks are only indicative, and not exhaustive.

Anybody can submit the soft copy of his/her manuscript **anytime** in M.S. Word format after preparing the same as per our submission guidelines duly available on our website under the heading guidelines for submission, at the email addresses: **1** or info@ijrcm.org.in.

GUIDELINES FOR SUBMISSION OF MANUSCRIPT

1. COVERING LETTER FOR SUBMISSION:

DATED: _____

THE EDITOR

IJRCM

Subject: SUBMISSION OF MANUSCRIPT IN THE AREA OF _____.

(e.g. Computer/IT/Engineering/Finance/Marketing/HRM/General Management/other, **please specify**).

DEAR SIR/MADAM

Please find my submission of manuscript titled ' _____ ' for possible publication in your journals.

I hereby affirm that the contents of this manuscript are original. Furthermore, it has neither been published elsewhere in any language fully or partly, nor is it under review for publication anywhere.

I affirm that all author (s) have seen and agreed to the submitted version of the manuscript and their inclusion of name (s) as co-author (s).

Also, if my/our manuscript is accepted, I/We agree to comply with the formalities as given on the website of journal & you are free to publish our contribution in any of your journals.

NAME OF CORRESPONDING AUTHOR:

Designation:

Affiliation with full address, contact numbers & Pin Code:

Residential address with Pin Code:

Mobile Number (s):

Landline Number (s):

E-mail Address:

Alternate E-mail Address:

2. **MANUSCRIPT TITLE:** The title of the paper should be in a 12 point Calibri Font. It should be bold typed, centered and fully capitalised.
3. **AUTHOR NAME (S) & AFFILIATIONS:** The author (s) **full name, designation, affiliation (s), address, mobile/landline numbers, and email/alternate email address** should be in italic & 11-point Calibri Font. It must be centered underneath the title.
4. **ABSTRACT:** Abstract should be in fully italicized text, not exceeding 250 words. The abstract must be informative and explain the background, aims, methods, results & conclusion in a single para. Abbreviations must be mentioned in full.
5. **KEYWORDS:** Abstract must be followed by list of keywords, subject to the maximum of five. These should be arranged in alphabetic order separated by commas and full stops at the end.

6. **MANUSCRIPT:** Manuscript must be in **BRITISH ENGLISH** prepared on a standard A4 size **PORTRAIT SETTING PAPER**. It must be prepared on a single space and single column with 1" margin set for top, bottom, left and right. It should be typed in 8 point Calibri Font with page numbers at the bottom and centre of the every page. It should be free from grammatical, spelling and punctuation errors and must be thoroughly edited.
7. **HEADINGS:** All the headings should be in a 10 point Calibri Font. These must be bold-faced, aligned left and fully capitalised. Leave a blank line before each heading.
8. **SUB-HEADINGS:** All the sub-headings should be in a 8 point Calibri Font. These must be bold-faced, aligned left and fully capitalised.
9. **MAIN TEXT:** The main text should follow the following sequence:

INTRODUCTION**REVIEW OF LITERATURE****NEED/IMPORTANCE OF THE STUDY****STATEMENT OF THE PROBLEM****OBJECTIVES****HYPOTHESES****RESEARCH METHODOLOGY****RESULTS & DISCUSSION****FINDINGS****RECOMMENDATIONS/SUGGESTIONS****CONCLUSIONS****SCOPE FOR FURTHER RESEARCH****ACKNOWLEDGMENTS****REFERENCES****APPENDIX/ANNEXURE**

It should be in a 8 point Calibri Font, single spaced and justified. The manuscript should preferably not exceed 5000 words.

10. **FIGURES & TABLES:** These should be simple, centered, separately numbered & self explained, and **titles must be above the table/figure. Sources of data should be mentioned below the table/figure.** It should be ensured that the tables/figures are referred to from the main text.
11. **EQUATIONS:** These should be consecutively numbered in parentheses, horizontally centered with equation number placed at the right.
12. **REFERENCES:** The list of all references should be alphabetically arranged. The author (s) should mention only the actually utilised references in the preparation of manuscript and they are supposed to follow **Harvard Style of Referencing**. The author (s) are supposed to follow the references as per following:
- All works cited in the text (including sources for tables and figures) should be listed alphabetically.
 - Use **(ed.)** for one editor, and **(ed.s)** for multiple editors.
 - When listing two or more works by one author, use --- (20xx), such as after Kohl (1997), use --- (2001), etc, in chronologically ascending order.
 - Indicate (opening and closing) page numbers for articles in journals and for chapters in books.
 - The title of books and journals should be in italics. Double quotation marks are used for titles of journal articles, book chapters, dissertations, reports, working papers, unpublished material, etc.
 - For titles in a language other than English, provide an English translation in parentheses.
 - The location of endnotes within the text should be indicated by superscript numbers.

PLEASE USE THE FOLLOWING FOR STYLE AND PUNCTUATION IN REFERENCES:**BOOKS**

- Bowersox, Donald J., Closs, David J., (1996), "Logistical Management." Tata McGraw, Hill, New Delhi.
- Hunker, H.L. and A.J. Wright (1963), "Factors of Industrial Location in Ohio," Ohio State University.

CONTRIBUTIONS TO BOOKS

- Sharma T., Kwatra, G. (2008) Effectiveness of Social Advertising: A Study of Selected Campaigns, Corporate Social Responsibility, Edited by David Crowther & Nicholas Capaldi, Ashgate Research Companion to Corporate Social Responsibility, Chapter 15, pp 287-303.

JOURNAL AND OTHER ARTICLES

- Schemenner, R.W., Huber, J.C. and Cook, R.L. (1987), "Geographic Differences and the Location of New Manufacturing Facilities," Journal of Urban Economics, Vol. 21, No. 1, pp. 83-104.

CONFERENCE PAPERS

- Garg Sambhav (2011): "Business Ethics" Paper presented at the Annual International Conference for the All India Management Association, New Delhi, India, 19-22 June.

UNPUBLISHED DISSERTATIONS AND THESES

- Kumar S. (2011): "Customer Value: A Comparative Study of Rural and Urban Customers," Thesis, Kurukshetra University, Kurukshetra.

ONLINE RESOURCES

- Always indicate the date that the source was accessed, as online resources are frequently updated or removed.

WEBSITE

- Garg, Bhavet (2011): Towards a New Natural Gas Policy, Political Weekly, Viewed on December 17, 2011 <http://epw.in/user/viewabstract.jsp>

WEB USAGE MINING: A BOON FOR WEB DESIGNERS

RITIKA ARORA

M. TECH. STUDENT

SHAHEED BHAGAT SINGH COLLEGE OF ENGINEERING & TECHNOLOGY

FEROZEPUR

ABSTRACT

Web usage mining is one of the categories of web mining which extract useful usage patterns from web data particularly web servers. This paper provides a brief overview on Web Usage mining, its data sources, Processing techniques and Applications.

KEYWORDS

Web mining, Web Usage Mining, Pre-processing, Pattern Discovery, Pattern analysis.

INTRODUCTION

In today's world we are overwhelmed with data and information from various sources but starved of knowledge. This has led to the need of mining useful data especially from the huge source of data that is World Wide Web. Hence Web mining, has been the focus of several recent research projects and papers. Web mining is that area of data mining which deals with the extraction of interesting knowledge from World Wide Web [1]. Depending upon which type of data is to be mined on web the Web mining is categorized into three different classes. These categories are web content mining, web structure mining, web usage mining. In web content mining, we discover useful information from the contents of web site which may include text, hyperlinks, metadata, images, videos, and audios. Search engines and web spiders are used to gather data for content mining. In web structure mining, we mine the structure of website on the basis of hyperlinks and intra-links inside and outside the web pages. In web usage mining (WUM) or web log mining, users' behavior or interests are revealed by applying data mining techniques on web log file [2].

NEED FOR WEB USAGE MINING [2]

Web usage mining is the type of Web mining activity that involves the automatic discovery of user access patterns from one or more Web servers. Today organizations rely more on internet and World Wide Web to conduct business. Most of the data of these organizations is generated automatically by web servers and collected in server access logs. Analyzing such data can help these organizations in the following:-

- a) Design cross marketing strategies
- b) Determining Life time value of clients
- c) Evaluate the effectiveness of the promotional campaigns
- d) Optimizing functionality of web-based application
- e) Website Personalization
- f) Targeting advertisements to specific group of users.
- g) Better Website Structure.

WEB DATA

For knowledge discovery in databases one of the primary steps is to create a suitable data set for the data mining tasks. The different type of data collected for web mining can be client side data, server side data, proxy data or organization database. These data collected not only differ in terms of location of the sources from which they are collected but also the kinds of data available, segment of population from which the data was collected and method of implementation used. The different kinds of data that can be used for web mining is classified as following:

Usage data

Data that describes the usage patterns of Web pages, such as IP addresses, page references, and the date and time of accesses and various other information depending on the log format [3].

Content data

Data which is imparted for the user and is actually the visible data on the web page comprising of mainly text and graphic.

Structure data

Structure data gives the organization of the content. This includes the intra-page and inter-page structure. Intra-page gives the arrangement of html and xml tags within a web page and inter-page information refer to the hyper links connecting the web pages.

User Profile

Data which provide demographic information about the user of the web site such as customer profile information and registration data [3].

Data sources

The data sources used in Web Usage Mining may include web data repositories like:

Web Server Logs

Web servers are surely the richest and the most common source of data for performing web mining as it record the browsing behavior of site visitors. The data in them reflects the concurrent access of web site by multi users. These logs usually contain basic information e.g. name and IP of the remote host, date and time of the request line exactly as it came from the client, etc. This information is usually represented in standard format e.g. Common Log Format, Extended Log Format and LogML [2]. But one cannot completely rely on the server logs due to the various level of caching in web environment as cached web pages are not recorded in the server logs. Packet sniffing technology can also be used to collect usage data through server logs. Other type of information like cookies and query data is also stored by the web servers in separate logs. Server side also provide content data and structure information and web page meta information [3].

Web Proxy Logs

Web proxy acts as an intermediate level of caching between the client browser and web servers. Proxy caching reduces the loading time of web page as well as network traffic load. Collecting navigation data at the proxy level is basically the same as collecting data at the server level. The main difference is in the case that proxy servers collect data of groups of users accessing huge groups of web server. Proxy reveals the actual HTTP requests from the multiple clients to multiple Web servers [3].

Web browser Logs

Usage data can be tracked also on the client side by using Java Script, java applet, or even modified browsers. Though performance of java applets is same as that of server logs but they incur some additional over head especially when they are loaded. Java scripts on the other hand consume little interpretation time but cannot capture all user clicks. Java Scripts and java applet collect only the single-user and single client browsing behavior whereas the modified browser are

more versatile as they collect data about a single user over multiple websites. These techniques avoid the problems of users' session identification and the problems caused by caching. In addition, they provide detailed information about actual user behaviors'. But these approaches rely heavily on the users' cooperation and rise many issues concerning the privacy law [2].

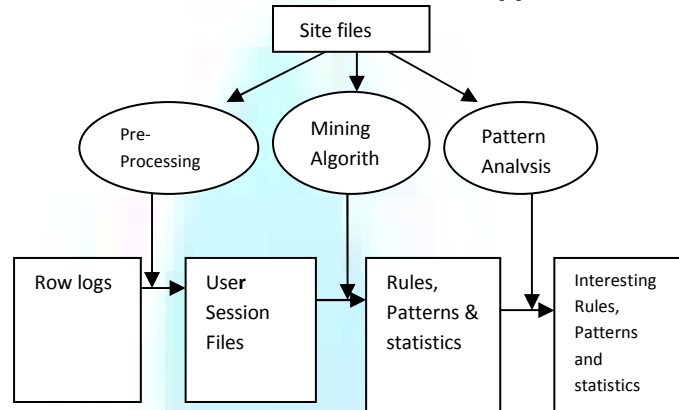
Data obtained from the above data sources can be categorized into several data abstractions such as users, page view, client session, server session, click streams and many more.

WEB USAGE MINING PROCESS

Web usage mining consists of three main processes as shown in Fig. 1.

- A. Pre-processing(Data preparation)
- B. Pattern discovery
- C. Pattern analysis and visualization

FIG. 1: WEB USAGE MINING PROCESS [5]



Pre-Processing

Data pre-processing transforms the data into a format that will be more easily and effectively processed for the purpose of the user. In context to web usage mining the click stream data is cleaned and partitioned into a set of user transaction representing the activities of each user during different visits to site. Other sources of knowledge such as the site content or structure as well as the semantic domain knowledge from site ontologies may also be used in pre-processing or to enhance user transaction data [6]. The different types of pre-processing in Web Usage Mining are:

Usage Pre-Processing: Pre-Processing relating to Usage patterns of users.

Content Pre-Processing: Pre-Processing of content accessed.

Structure Pre-Processing: Pre-Processing related to structure of the website [3].

Pattern Discovery

In this stage the statistical, database and machine learning operations are performed to obtain hidden patterns reflecting the typical behavior of the user as well as the summary statistics on web resources, session and user [7]. Web Usage mining can be used to uncover patterns in server logs but is often carried out only on samples of data. The mining process will be ineffective if the samples are not a good representation of the larger body of data. The following are the pattern discovery methods [6].

- a) Statistical Analysis – Statistical technique are the most common method to extract knowledge about visitors to a web site. By analyzing the session file, one can perform different kinds of descriptive statistical analyses (frequency, mean, median) on variables such as page view, viewing time and length of a navigational path [3].
- b) Association Rules – Association rule learning is popular and well researched method for discovery relations between variables in large data bases. They show attributes value conditions that occur frequently together in a given dataset [13].
- c) Clustering - In unsupervised classification, called clustering or exploratory data analysis, no labeled data are available. The goal of clustering is to separate a finite unlabeled data set into a finite and discrete set of "natural," hidden data structures [12].
- d) Classification – Classification is a data mining (machine learning) technique used to predict group membership for data instances. For example, you may wish to use classification to predict whether the weather on a particular day will be "sunny", "rainy" or "cloudy". Popular classification techniques include decision trees and neural networks [11]. Classification is useful in such areas as topic aggregation and Web community identification [4].
- e) Sequential Patterns - This Technique discovery inter-session pattern such that the presence of the set of items is followed by another item in time-ordered set of sessions or episodes. With the help of this technique the Web marketers can predict future visit patterns which will be helpful in placing advertisement aiming at particular group of users [3].
- f) Dependency Modeling – This is another useful pattern discovery technique in Web Mining. The primary main is to develop a model, representing significant dependencies among various variables in Web domain. Such information may help develop strategies to increase sales of a product offered by a website [6].

Pattern Analysis

This is the final step in the Web Usage Mining process. After the pre-processing and pattern discovery, the obtained usage patterns and statistics are further processed, filtered to remove uninteresting information and result is aggregated into user model that can be used for market predictions and decision making [6]. The methods like SQL (Structured Query Language) processing and OLAP (Online Analytical Processing) can be used [3].

WEB USAGE MINING APPLICATION

Letizia

Letizia is an application that assists a user browsing the Internet. As the user operates a conventional Web browser such as Mozilla, the application tracks usage patterns and attempts to predict items of interest by performing concurrent and autonomous exploration of links from the user's current position. The application uses a best-first search augmented by heuristics inferring user interest from browsing behavior [6].

WebSift (Web Site Information Filter)

Website is an application which performs Web Usage Mining from server logs recorded in the extended NSCA format (includes referrer and agent fields), which is quite similar to the combined log format. The pre-processing algorithms include identifying users, server sessions, and identifying cached page references through the use of the referrer field. It identifies interesting information and frequent item sets from mining usage data [3].

Adaptive Websites

An adaptive website adjusts the structure, content, or presentation of information in response to measured user interaction with the site, with the objective of optimizing future user interactions. Adaptive websites are web sites that automatically improve their organization and presentation by learning from their user access patterns. User interaction patterns may be collected directly on the website or may be mined from the web server logs.

Web Personalization

Is personalizing the browsing experience of a user by dynamically tailoring the look, feel, and content of a Web site to the user's needs and interests. Personalizing the website broaden and deepen customer relationships, provide continuous relationship marketing to build customer loyalty, help automate the process of proactively market products to customers by lights-out marketing and cross-sell/up-sell products ,provide the ability to measure customer behavior and track how well customers are responding to marketing efforts [8].

CONCLUSION

Web Usage Mining is current field of research in the area of web mining. As the world is shrinking due to globalization and more and more organizations are using internet as a platform for their business, the need for mining the web data is increasing. Web Usage Mining provides an excellent way to deal and process web data particularly the server data to extract useful web usage patterns. This paper proves to be a starter kit to the beginners on Web Usage Mining.

REFERENCES

- [1] Federico Michele Facca and Pier Luca Lanzi, "Recent Developments in Web Usage Mining Research", 5th international conference on Data Warehousing and Knowledge Discover, 2003.
- [2] S.P Nina, Rahman, M Bhuiyan, K.I Ahmed, "Pattern Discovery of Web Usage Mining", International conference on Computer Technology and Development, 2009.
- [3] Jaideep Srivastava, Robert Cooley, Mukund Deshpande, Pang-Ning Tan, Web Usage Mining: "Discovery and Applications of Usage Patterns from Web Data", SIGKDD Explorations, vol. 1, Issue 2, pp. 12-23, 2000.
- [4] Pranam Kolari and Anupum Joshi, "Data Mining: Research and Practise", IEEE journal on computing in science and engineering, vol. 6, Issue 4, pp. 49-53, 2004.
- [5] Brijendra Singh and Hemant Kumar Singh, "Web Data Mining Research: Survey", IEEE international conference on Computational Intelligence and Computing Rearsch, 2010.
- [6] <http://web.media.mit.edu/~lieber/Lieberary/Letizia/Letizia-Intro.html>.
- [7] Bing Liu, Web Data Mining: Exploring Hyperlink, Contents and Usage Data, Springer, December, 2006.
- [8] <http://www.web-datamining.net/internet-marketing/>
- [9] http://dataminingwarehousing.blogspot.com/2008_10_01_archive.html.
- [10] <http://ecmlpkdd.cs.helsinki.fi/pdf/berendt-2.pdf>
- [11] <http://databases.about.com/od/datamining/g/classification.htm>
- [12] R. Xu and D.Wunsch, "Survey of Clustering Algorithm", IEEE Transactions on Neural Network, vol. 16, pp. 645-678, 2005.
- [13] Liang Zhong Shen, "A Novel Efficient Mining Association Rules Algorithm for Distributed Databases", Advance Materials Research, vol. 108-111, pp 50-56, 2010.
- [14] Doru TANASA, "Web Usage Mining: Contributions to Intersites Logs Preprocessing and Sequential Pattern Extraction with Low Support" Phd.thesis, June2005.

REQUEST FOR FEEDBACK

Dear Readers

At the very outset, International Journal of Research in Computer Application and Management (IJRCM) acknowledges & appreciates your efforts in showing interest in our present issue under your kind perusal.

I would like to request you to supply your critical comments and suggestions about the material published in this issue as well as on the journal as a whole, on our E-mails i.e. **infoijrcm@gmail.com** or **info@ijrcm.org.in** for further improvements in the interest of research.

If you have any queries please feel free to contact us on our E-mail **infoijrcm@gmail.com**.

I am sure that your feedback and deliberations would make future issues better – a result of our joint effort.

Looking forward an appropriate consideration.

With sincere regards

Thanking you profoundly

Academically yours

Sd/-

Co-ordinator