

## A Mining Approach for Web Engineering In Respect Of Business Intelligence Application

Tapan Nayak,<sup>1</sup> Prof. B Lakshma Reddy<sup>2</sup>

<sup>1</sup>Department of Computer Science, CMJ University, Shillong, Meghalaya, INDIA

<sup>2</sup>Department of Computer Science, Garden City College, Bangalore, INDIA

**Abstract:** Using mining approach in Engineering process of Web Applications is a complex problem, due to the variety of languages and technologies that are contemporary used to realize them. Indeed, the benefits that can be obtained are remarkable: the presence of documentation at different abstraction levels will help the execution of maintenance interventions, migration and reengineering processes, reducing their costs and risks and improving their effectiveness. The main objective of this paper is to identify the application of web engineering process from mining techniques for a intelligence application. In this paper, all aspects are studied that are basically required for web mining and web engineering process. Since the last few decades, there has been an immense increase in use of World Wide Web (WWW) for a wide and variety of web based business applications. The web based intelligent application plays a leading role in e-commerce business applications. As a consequence, there is a need to improve Intelligence of Web Engineering Applications in the context of Business and IT industry. To achieve this objective web engineering process must be able to identify some useful insights for business intelligence. The proposed research work attempts to initiate Business Intelligence from Web Based Applications. The issues of research work are uniformly accommodated in five steps which are provided in this paper.

**Keywords:** Web Engineering, Web Mining, Web site Errors, Web Development, Web Crisis

### I. INTRODUCTION

Although the development of Web based applications may seem easy, it is often more complex and challenging than many of us think. . Two key attributes distinguish web-based systems development from traditional software development: rapid growth of the requirements of Web-based systems and the continual change of their information content. Web Applications are evolutionary. For many web applications, it's not possible to specify fully what they should or will contain at the start of their development, because their structure and functionality will evolve over time. Hence, Web based system development is not a one-time event as currently perceived and practiced by many web developers; it is, instead, an iterative process with a long life cycle.

#### 1.1 Motivation

Reasons for mining approach for web engineering on business application

- **Interfacing.** An overall web architecture describing how the network and the various servers such as web servers, applications servers, and database servers interact and application can be used when a system is required to interface to another system and how both systems would negotiate is to be established.
- **Commercial espionage.** Learning about an enemy's or competitor's latest research by stealing or capturing a prototype and dismantling it. It may result in development of similar product.
- **Usability and User-Centered Designs.** Effective Web site design requires attention to usability. Web-based systems need to be designed for easy navigation and customer attractive, User-centered design methods for Web sites is presented in , while presents a User-Centric Approach to Modeling Web Information Systems.
- **Improve documentation shortcomings** Using, the web log file contains information about the user IP address, the requested page, time of request, the volume of the requested page, its referrer, and other useful information. Web log file is the main source of data analysis in web mining. Web log file contains data about requested URL, time and date of request, method used.
- **Obsolescence.** Application Integrated circuits often seem to have been designed on obsolete, proprietary systems, which means that the only way to incorporate the functionality into new technology is to web-engineer the existing chip and then re-design it.
- **Software Modernization.** Web mining is the usage of data mining techniques to extract interesting information from web data. Patterns extracted from applying web mining techniques on web data can be used to maintain websites by improving their usability.
- **Product Security Analysis.** To examine how a product works, what are specifications of its components, estimate costs and identify potential patent infringement. Acquiring sensitive data by disassembling and analysing the design of a system component.
- **Academic/learning purposes.** The essence of Web engineering is to successfully manage the diversity and complexity of Web application development and hence, avoid potential failures, which can have serious implications. It is a proactive approach to building Web applications.

## II. RELATED WORK

The rapid growth of the Internet, Intranets, Extranets, and the World Wide Web has already had a significant impact on business, industry and commerce, sports, finance, education, government and entertainment sectors, and our personal and working life. Many legacy information and database systems are being migrated to the Internet and the Web environments. A Web based business application software model of the system describes various components of the system and how they are linked for mining techniques used for extracting relevant information.. The application architecture model shows that, a map of various information and functional modules. An information module may provide the same information to all the users or provide customized or personalized information to each user. Functional modules such as login page, registration pages, web forms for data collections and the shopping carts used in e-commerce systems – collect and process the users input. Specific requirements for look and feel consistency of information, scalability, maintainability and quality control mechanisms determine the software architecture [4].

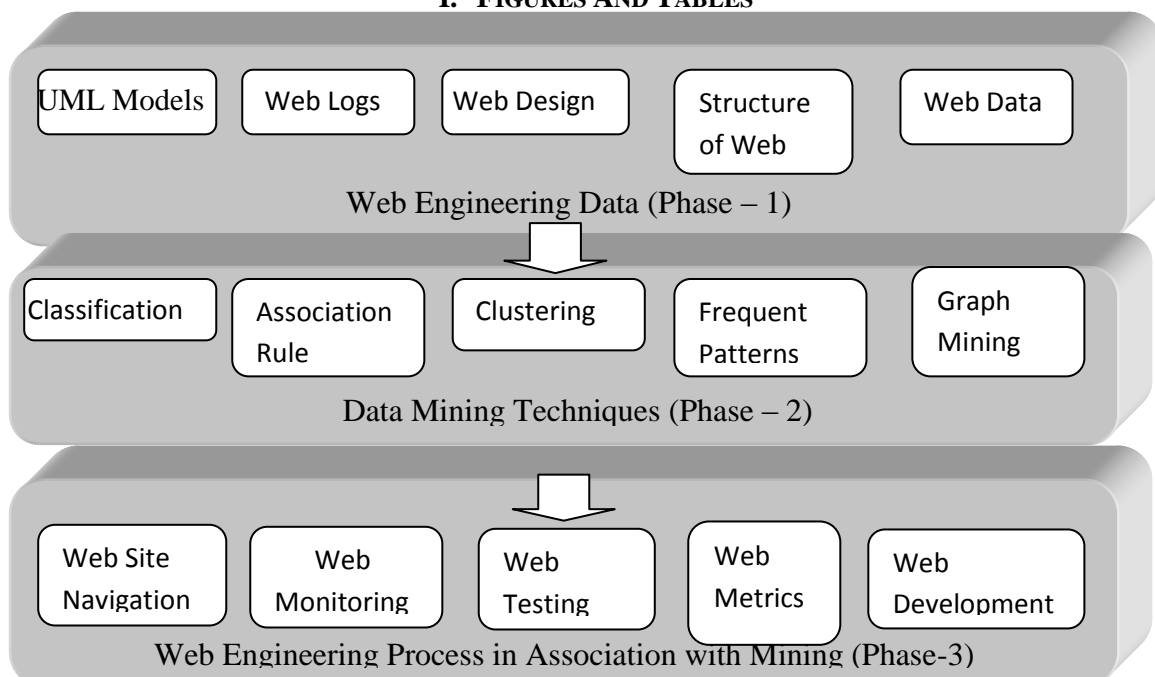
The primary cause of Web system failures are the flawed design and development process and poor management of their development. The emerging Web engineering discipline deals with the process of developing Web-based systems and applications.. To tackle the above mentioned problems, Web engineers are increasingly applying Data Mining algorithms to various Webs engineering process. For instance, mining algorithms can be applied to trace the patterns in which a user browses the Web pages and group together those Web pages making the retrieval of the information fast, thus increasing the responsiveness. Web development is a process, not simply a one-time event. Thus, Web Engineering deals with all aspects of Web-based systems development, starting from conception and development to implementation, performance evaluation, and continual maintenance. Web Engineering, therefore, covers a range of areas: requirements elicitation and analysis; Web system modeling; Web architecture; Web system design; Web page design; scripting/coding; interface with databases, ERP systems, and other Web-based systems; Web quality; Web usability; Web security; Web system performance evaluation; Web testing; Web development methodologies; Web development process; Web metrics; and Web project management [3].

## III. Mining Approach for Web Engineering

A successful Web Engineering Applications is one which can improves insights of web based business intelligence by sharing results among various modules in the process of effective decision making by using basic mining techniques as a tool. Many attributes of quality Web-based Systems such as ease of navigation, accessibility, scalability, maintainability, usability, compatibility and interoperability, security, readability, and reliability are not given due consideration during development. Many developers seem to be unaware of the real issues and challenges facing major Web based application development and its continual maintenance in respect of mining approach. Web Engineering must be a systematic discipline and quantifiable approach for development, operation and maintenance of web based business applications for a successful business. To achieve the goal of business intelligence in web based applications, one has to concentrate on various aspects of web content. A mining approach for web engineering application is necessary to study the role of various components of web application in business intelligence.

An algorithm was developed as a web mining approach which can investigate the components of web that are used in business intelligence [4]. The different phases in mining for web engineering are shown in figure 1.

### I. FIGURES AND TABLES



**Figure 1: Phases of Mining for Web Engineering**

**Phases in Mining Process:**

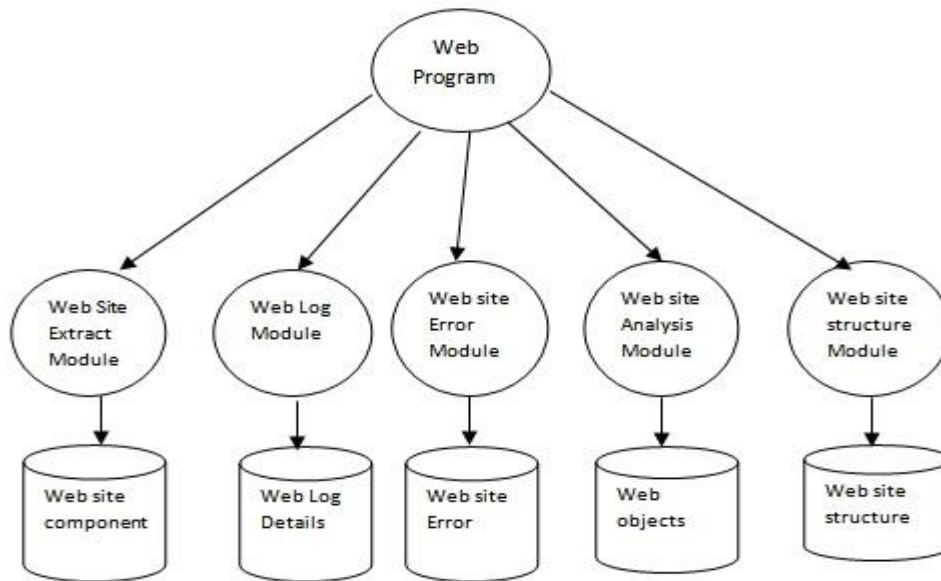
An algorithm consists of 3 phases. They are

Phase 1: Extracting Web Data from web application

Phase 2: Web Mining Approaches to extract features of web application

Phase 3: Web engineering process associated with Web Mining

**Phase 1: Extracting Web Data from web application:** An algorithm was developed using the web program in extracting the components of web such as Web site Structure, Web site Error Reports, Web site objects used in web design, and contents of Web log. The Web Program retrieves all necessary web data elements using a standard set of web tools. These tools include Web site Extractor, Web page Analyzer, W3C HTML Validator, Power Mapper, and Web log Expert. The overall structure of the algorithm is shown in figure 2.



**Figure 2: Architecture of Algorithm**

**Phase 2: Web mining techniques to extract features of web application:**

Web mining is the usage of data mining techniques to extract interesting information from web data. Patterns extracted from applying web mining techniques on web data can be used to maintain websites by improving their usability. The patterns can also be used to study user behaviour and interests, facilitate support and services introduced to the website navigator, improve the structure of the website, and facilitate personalization and adaptive websites. After extracting web engineering data, various web mining techniques are applied on data base which consists of web information.

**1. Association Rule Mining:**

In ARM a couple set of association rules are applied on elements of web site structure to identify relationships among various modules of web engineering application. Association rule mining is finding all association rules with support and confidence values that are greater than or equal a user-specified minsup and minconf respectively. In general, the process of extracting interesting association rules consists of two major steps.

Step 1: The first step is finding all item sets that satisfy minsup (known as Frequent-Item set generation).

Step 2: The second step, is generating all association rules that satisfy min conf using item sets generated in the first step. After generating frequent item sets, association rules that are greater than or equal to min conf are generated. Those rules are called interesting association rules. Those rules can be invested in many different applications. One of those applications is improving the structure of the company's website that the mined database belongs to. This is done during the website's design phase by creating links between items that seem to be sold together, or highlight those links if they are already exist, and/or create index pages which are pages that have direct links to some products that may be of interest for some group of customers. All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

**2) Classification:**

Classification is dividing an existing set of events or transactions into other predefined sets or classes based on some characteristics. In web usage mining, classification is used to group users into predefined groups with respect to their navigation patterns in order to develop profiles of users belonging to a particular class or category. A web mining strategy for web personalization based on a novel pattern recognition strategy which analysis and classifies users taking into account both user provided data and navigational behaviour of the users.

### 3) Clustering:

Clustering is the process of partitioning a given population of events or items into sets of similar elements, so that items within a cluster have high similarity in comparison to one another, but are very dissimilar to items in other clusters. In web usage mining there are two main interesting clusters to be discovered: usage clusters and pages clusters [4]. The authors in [5] present an approach to cluster web pages to obtain high quality clusters of web pages and use those clusters to produce index pages, where index pages are web pages that have direct links to pages that may be of interest of some group of website navigators. In [5] clustering techniques are applied to web log file to discover those subsets of web pages that need to be connected and to improve the already connected pages.

### Phase 3: Web Engineering Process associated with web mining

The basic knowledge of Web Mining techniques that are applied on Web data produce various process related to Web Engineering frame work. These process associated with web services, web architectures, web configuration management data, web application classification, web testing, requirements elicitation and analysis; Web system modeling; Web architecture; Web system design; web communities, Website Navigation, Web page design; Web quality; Web usability; Web security; Web system performance evaluation; Web testing; Web development methodologies; Web development process; Web metrics; and Web project management [6]. These services are very much essential in any type of web application. Patterns extracted from applying web mining techniques on web data can be used to maintain websites by improving their usability through simplifying user navigation and information accessibility and improving the content and the structure of the website in a way that meets the requirements of both website owner and user which will consequently increase the overall profit of the business or the industry that the maintained website belongs to [8].

## IV. CONCLUSION

This paper presents a development structure of a mining module for web engineering and also describes tool architecture to extract phases from existing web engineering structure, but also with the more challenging mining content. Here an attempt is made to find all areas involved in mining process. Further the work can be extended to find services of web mining process required for web engineering.

## REFERENCES

- [1] Ajith Abraham, "Business Intelligence from Web Usage Mining", Department of Computer Science, Oklahoma State University, USA.
- [2] Anupam Joshi and Pranam Kolari, "Web Engineering Column: Web Mining: Research and Practice", IEEE Computing and Science and Engineering, July/August 2004 pp. 49-53.
- [3] Athula Ginge and San Murugesan, "Web Engineering: A Methodology for Developing Scalable, Maintainable Web Applications", Cutter IT Journal, Vol.14, No.7, July 2001, Page. 24-35.
- [4] G sreedhar, Dr. A A Chari, Dr. V V Venkata Ramana, " A Qualitative and Quantitative Frame work for effective Web Site Design", IJCA, Vol 2, May 2010, PP 48-56.
- [5] Jiawei Han, Kevin Chen-Chuan Chang, "Data Mining for Web Intelligence", IEEE Computer November, 2002 pp. 64-70
- [6] Martha Koutri and Sophia Daskalaki. "Improving Web Site Usability through a Clustering Approach", In Proceedings of the 10th International Conference on Human-Computer Interaction HCI, Crete, Greece, 2003, PP. 11-19.
- [7] Tao Xie, Jian Pei, Ahmed E. Hassan "Mining Software Engineering Data", 29<sup>th</sup> International Conference on Software Engineering, 2007, IEEE.
- [8] Wingyan Chung "Designing Web-based Business Intelligence Systems: A Framework and Case Studies: In DESRIST", February 24-25, California CA USA 2006, pp. 147 – 171.
- [9] M V Kamal, "Mining for web engineering-business intelligence application case study", IJESS, Vol2, Issue 6 (June 2012).
- [10] Nayak, "Reverse Engineering: Methodologies for Web Applications" IJSWS 12-209, Vol. 2, Issue, 1, pp.39-44, August-November, 2012",
- [11] San Murugesan, Yogesh Deshpande "Web Engineering: A New Discipline for Development of Web-based Systems".
- [12] Scharl, A., A Conceptual, User-Centric Approach to Modeling Web Information Systems , Proc Australian Web Conference (Ausweb99), Ballina, Australia, April 1999.
- [13] Detroyer, O.M.F., and C.J. Leune, WSDM: A User-Centered Design Methods for Web Sites, Proc. WWW7Conference, Brisbane, 1998.