Web Application Regression Testing: A Comparative Evaluation of User Session and Slicing Based Approaches

Thesis Report

A THESIS PRESENTED TO

FACULTY OF BASIC & APPLICED SCIENCES

DEPARTEMENT OF COMPUTER SCIENCE AND SOFTWARE ENGINEERING

IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE

OF

MS IN SOFTWARE ENGINEERING

BY

Sabeen Masood
287-FBAS/MSSE/F09

Department of Computer Science and Software Engineering Faculty of Basic and Applied Sciences International Islamic University Islamabad

(June 2012)

1-Application programs

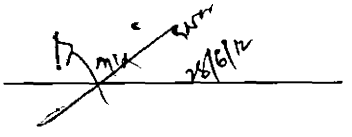
2-Computer software

APPROVAL

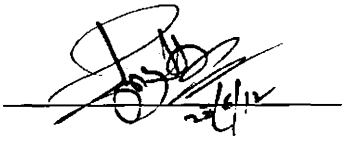
SUBJECT: APPROVAL OF THE RESEARCH THESIS "Web Application Regression Testing: A Comparative Evaluation of User Session and Slicing Based Approaches"

It is certified that we have read this research thesis report and have fully evaluated the research undertaken by **Saheen Masood** Registration No. **287-FBAS/MSSE/F09**. This research thesis fully meets the requirements of Department of Computer Science and Software Engineering and hence, the International Islamic University, Islamabad.

External Examiner:
Dr.Sajid Anwar
Assistant Professor,
Department of Computer Science,
Institute of Management Sciences, Peshawar

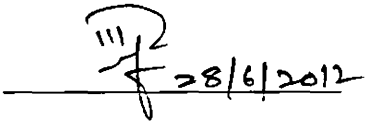


Internal Examiner:
Dr.Ayyaz Hussian
Assistant Professor,
Department of Computer Science and Software Engineering,
Faculty of Basic and Applied Sciences,
International Islamic University, Islamabad

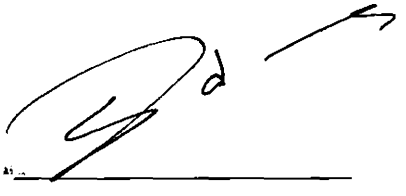


Supervisor:

Dr.Abdul Rauf
Assistant Professor,
Department of Computer Science and Software Engineering,
Faculty of Basic and Applied Sciences,
International Islamic University, Islamabad



Co Supervisor:
Ms.Salma Imtiaz
Assistant Professor,
Department of Computer Science and Software Engineering,
Faculty of Basic and Applied Sciences,
International Islamic University, Islamabad



DEDICATION

I would like to dedicate my research work to the HOLIEST man Ever Born on Earth, ***PROPHET MUHAMMAD (Peace Be Upon Him)*** *and I also dedicate my work to my*

PARENTS

Whose sincere love and prayers are a source of strength for me and made me to do this research work successfully.

A dissertation submitted to the

Department of Computer Science and Software Engineering, Faculty of Basic and Applied Sciences, International Islamic University, Islamabad,

as a partial fulfillment of the requirements for the award of the degree of

MS in Software Engineering (MSSE)

DECLARATION

I hereby declare that this Thesis **"Web Application Regression Testing: A Comparative Evaluation of User Session and Slicing Based Approaches"**, neither as a whole nor as a part therefore, has been copied out from any source. It is further declared that I have written this thesis entirely on the basis of my personal efforts, made under the proficient guidance of my thesis supervisor, **Dr. Abdul Rauf** and my thesis co supervisor, **Ms. Salma Imtiaz.**

If any part of this research thesis proved to be copied or found to be a research of some other individual, I shall standby the consequences.

No portion of the research work presented in this thesis report has been submitted in support of any other degree or qualification of this or any other University or institute of learning.

Sabeen Masood

287-FBAS/MSSE/F09

TABLE OF CONTENTS

List OF Tables

List OF FIGURES

CHAPTER # 1

Introduction

1.1. Web Application Regression Testing

Web applications are broadly used and are considered to be the backbone of many organizations. Almost all large organizations and commercial companies have very demanding e-commerce web sites for attracting the clients. The huge monetary loss is faced by the organizations, if there is failure in web application domain. The one way that can reduce this loss is to make websites meeting to the non-functional requirements like usability, reliability, security and availability.

Web applications use heterogeneous technologies and languages, as a result the requirements of a web based application acquire more complexity than other applications. This heterogeneity can be effectively managed during the testing phase by thoroughly testing the environment of these applications.

Frequent changes are unavoidable in software now days; these frequent changes are due to dynamic nature of usage as well as because of advancement in technologies. These rapid changes in software demand re-testing the software very frequently. Re-testing software completely requires numerous resources, so the usage of regression testing is to overcome this issue. If we want to include a new component in web application or eliminate an out-dated one, then regression testing should be carried out to ensure the modifications which have not affected the other portions of the web application.

Testing is performed on the basis of test cases, which specify the functionality to be tested, what the inputs are and what the expected result is. Developing test cases is very hectic and time consuming activity. So, testers mostly reuse test cases for validating a modified system along

with some new test cases. Regression testing is significant part of maintenance activity. Seventy to ninety percent of the total software lifecycle cost is attributed to maintenance phase.

Regression testing is basically retesting of the software and ensures that a new modification has not introduced any bug in the software. Regression testing has two common approaches that handle same type of issues mentioned in [8] namely user session based testing and slicing based testing.

1.1.1. User Session Based Testing

User session is mostly used in situations where program specifications are not mentioned, but it has no affect on the user session test case generation process if web technology changes. In user session testing, only those parts of the application are tested, which are being used by the user. Execution of large number of captured user session is impractical due to their huge number. Testing all user sessions would require more test development effort and time for test execution.

User session based testing suggests that large number of sessions is difficult to test, thus we need to tradeoff between quantity of session data and effort required for software testing. Efficiency of user session will be affected by the data collection process.

User Session automatically creates test cases on user profiles. Tester easily access logs of user real usage data in web application domain. User session based testing is basically the convergence of these real usage user data into test cases. This technique is good for detecting bugs but it is unscaleable with large number of sessions. So testers should tailor the number of sessions using different techniques such as clustering and reduction etc.

User sessions can easily be recorded in web applications. It is also used for regression testing. User session based testing also provides an effective partitioning or coverage along with identifying the defects that were not highlighted in any other approach in web application domain. It can be conducted even with poor requirements where the requirements are unclear or incomplete.

In software testing, utilizing user sessions is an automatic process to reduce the overall cost of testing. Other information such as web site links is also easily extracted from user sessions. Two tools (Rational Robot and Web King) are also available for automated testing of web application through user session

1.1.2. Slicing Based Testing

Slicing is the second sub technique of regression testing. It was introduced by Mark Weiser [41] to assist students understand and debug their programs. The program slicing reduces the size of certain program though maintaining the original behavior of the certain program.

It helps to understand the internal structure of the system as well as revealing the significant information [32]. It is also very important for selective regression testing, but it is not suitable for verifying and validating an initial copy of software.

According to the definition [41], a slice is an executable division of program statements. It saves the actual program behavior with respect of the division of variables of interest and at a given point of program. In slicing based testing, testers only focus specific part of application for testing.

Several software engineering jobs need to decompose a larger program into smaller programs e.g., program comprehension, debugging, downsizing, maintenance and testing. Slicing is very useful for this concern.

In slicing based testing, a remarkable amount of papers have been published to present different kinds of program slicing [20] but very little work reported on slicing in web application regression testing domain.

1.2. Motivation

Practically the regression testing is not much beneficial in large software systems due to high cost and time demands. Therefore many organizations reuse test cases for regression testing [4]. Retesting the entire test suite will consume large amount of time and resources if new test cases are written each time a modification is made. Thus test cases are reused for validating the modified system.

Web application testing is an emerging and challenging field of current era because many test cases will be generated even for small changes due to web dynamic and complex nature. Large number of researchers has worked in web application testing domain.

Each researcher has focused on different areas of web applications like static and dynamic pages, links, frames, architecture, model and scenarios [21] but limited work reported in web application regression testing [16].Several kinds of web application regression testing methodologies have been proposed in the literature but unfortunately no detailed review has been conducted. It is very important to explore types of methodologies are being used in this domain.

The research work presented in this thesis discusses different methodologies for web application regression testing and then compares two methodologies (user session and slicing) in terms of their effectiveness and performance. User session is termed as the finest technique for regression testing [8], where as slicing is stated as similar to user session testing in terms of problem coverage. Therefore we have chosen to compare these two techniques in domain of regression testing for web based application.

The previous research work [8] highlights that user session and slicing have same capability to detect the errors in web testing and both are not appropriate for testing initial copy of web applications. This research finds out whether both methodologies can be applied together or not and if yes whether they can enhance the results in web application regression testing domain.

1.3. Aims and Objectives

The aims and objectives of this research is

- Comparative evaluation of the two approaches (user session and slicing).
  - ✓ Compare performance level of user session and slicing.
  - ✓ Check whether user session or slicing is better for web application regression testing w.r.t performance
  - ✓ Check whether they provide optimized performance:
    - o When both are applied in conjunction or separately.

1.4. Research Questions

The purpose of this study is to address the following research questions through experiment.

RQ1. Which of the following techniques, user session and slicing, is better in regression testing of web application with respect to performance?

The purpose of this question is to check that which technique is better for web application regression testing and also compares the effectiveness and performance of two techniques (user session and slicing) when they are used alternatively.

RQ2. What is the impact of user session and slicing on performance when they are used in conjunction?

The purpose of this question is to check that whether the performance increases when these two techniques are used at the same time. The question is based on the fact that both techniques have been used for web application regression testing and both are used to handle same type of problems.

1.5. Research Process

The steps involved in the research process are as follows:

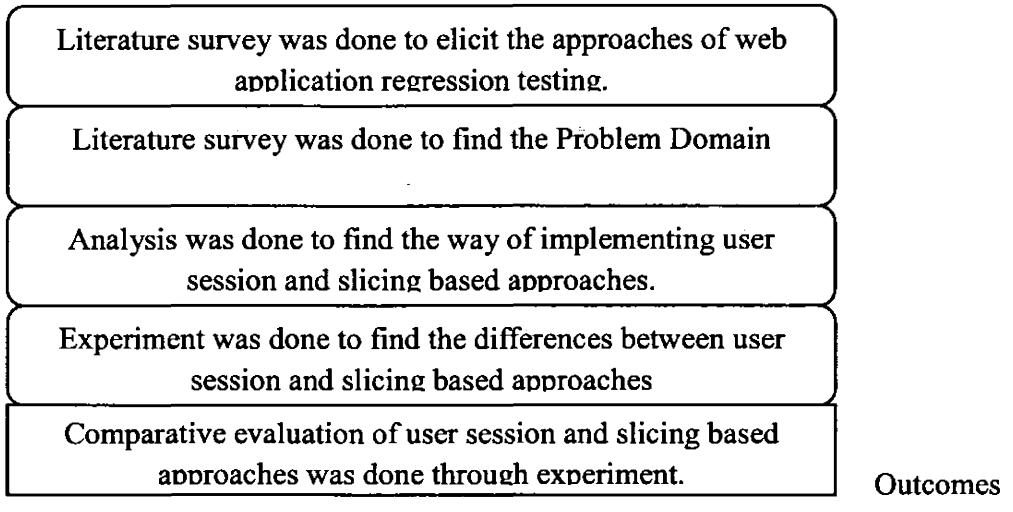


Figure 1. 1:Research Process

1.6. Significance of Research

The research will helpful not only in summarizing the data but will also be helpful in knowing the following at least

- What types of methodologies are being used in the domain of regression testing for web application?
- Which technique is better (user session or slicing) in regression testing of web application with respect to performance?
- What is the effect of using both techniques in conjunction on performance?
- What is the effect of using both techniques separately on performance?

1.7. Thesis Outline

Remaining of the thesis is organized as follows:

Chapter 2: Chapter 2 discusses web application regression testing approaches that have been reported in literature.

Chapter 3: The proposed methodology is discussed in Chapter 3. In this chapter; implementations of user session based approach, slicing based approach and hybrid approach is also discussed.

Chapter 4: This chapter presents the findings of the experiment and discusses them thoroughly. Results have been validated by applying different test cases.

Chapter 5: Conclusion and Future work is given in Chapter 5. The contributions of this research work have also been discussed in the same chapter.

CHAPTER # 2

Background and Related Work

2.1. Introduction

A web application development is different from development of traditional applications due to its dynamic and complex nature. Web applications design must take into account some quality characteristics like scalability, security, availability, reliability to efficiently convey the information and to effectively distribute functionalities among server side and client side to optimize performance. Furthermore, a web application testing differs from traditional application testing especially with the utilization of web services that often support recurrently changing business environments.

High user demands and quick web technological changes have subject web applications rapid change and maintenance, which requires the development and execution of efficient regression testing techniques [23].

Regression testing refers to testing the tailored version of a system v', using test set T used earlier to test the actual system v. The appropriate selection of test cases T can be made through number of ways and various regression testing techniques has been proposed. These techniques have different objectives [32]. To re-test a program after adaptation, pick the test suite subset which will increase confidence of changes to be covered. Regression testing techniques are essential for sufficient collection of these subsets of test cases [16].

2.2 Background: Web Application Regression Testing

In this section, the prior work in regard to the web application regression testing is presented. Large number of researchers has worked in web applications testing area but we still lack a mature and well developed model to test and examine web applications up till now. All the

previous work deals with different characteristics of the web applications while ignoring the rest. Even poorer, very little regression testing techniques have been used in web application domains [23] which are as following:-

Table 2. 1: Literature Survey of Web Application Regression Testing

Paper title	Focus	Published in	Proposed Approach	Limitation/ Disadvantages
"Improving Web Application Testing with User Session Data".	Review the techniques that were used in web application regression testing domain.	Int. Conf. on Software Engineering, 2003	Technique that collected user session data from the web and performed web testing [1]	- User navigation and filtering session's techniques were not considered.
"Automated Regression Testing of Web Applications".		Project Report, 2005	Technique that would perform automated web application regression testing using user sessions[14]	Tool was created that could not be checked most complex fields. Expected results examined manually.
"Automated Session Data Repair for Web Application Regression Testing".		IEEE Computer Society, 2008	Web application regression testing approach that repair user session data [3]	-Unaltered web application was used. -Behavior of the repair algorithm was checked on single web application.
"Harnessing Web-based Application Similarities to Aid in Regression Testing".		International Software Reliability Engineering Symposium, 2009	Automated oracle comparator technique based on HTML / XML output semantics and underlying inherent similarities [6]	-Model was not able to detect unusual scenarios e.g. small alterations in natural language text
"Automating regression testing using web based application		International Software Tools for Technology Transfer (STTT) Journal	Automated regression testing approach that automatically	- Their Annotations were conventional -Annotators were accountable for

similarities".		, 2011	compares the output of two versions [7]	telling the features of the model.
"Lessons learned from a survey of web applications testing".		Int. Conf. on IT, 2009	Conducted a survey of different testing techniques used for web applications [8]	Approaches were reviewed from literature and their details, benefits and deficiencies are based on literature review. Not conducted any experiment.
"Regression Testing for Web Applications Based on Slicing".		Int. Conf. on Software and Applications, 2003	Technique for web application regression testing that was based on slicing [12]	-Only presented basic framework and related methods. -Not considered method optimization.
"Regression Testing Web Applications"		Int. Conf. on Advanced Computer Engineering, 2008	Technique which selected test cases based on event dependency graph for web application regression testing [16]	-Only considered the web application in event driven environment.
"A Meta-Model to Support Regression Testing of Web Applications".		International Software & Knowledge Engineering Conference, 2008	UML 2.0 profiles meta-model for web application regression testing[20]	-Manually detected model constraint violations. -Current prototype provided coverage for only a subset of web widgets and controls.
"Automatic Generation of Regression Test Cases for Web Components using Domain Analysis and		International Computer Applications Journal, 2010	Domain analysis and modeling technique for web application testing [21]	- Graphical Web Model was manually constructed.

Modeling".				
"Modeling and Automated Black box Regression Testing of Web Application".		Journal of Theoretical and Applied IT, 2005	Analysis model for testing and modeling web application [23]	-Not considered external content sources and testing of server side logic in a model.
"Testing Web Applications Focusing on Their Specialties".		ACM Notes of Software Engineering, 2005	Web application regression testing was proposed[30]	-Some realization techniques for web testing were considered.
"Regression Testing Web Services based Applications".		IEEE Int. Conf. on Computer Systems and Applications, 2006	Modeled a web application and its components as an abstract model [31]	-Not implemented the technique on web application.
"A Regression Testing Method for Composite Web Services".		Int. Conf. on Biomedical Engineering and Computer Science, 2010	Composite web services regression testing method [33]	-Not monitored the services in the test model.

Detailed literature survey of web application regression testing is presented below:

Sebatien at el [1] proposed a new technique that collected user session data from the web and performed testing. New and existing techniques of test generation were compared. Results showed that user session formed effective test suites than white-box techniques.

Alshahwan [14] provided fully automated web application regression testing approach. The approach uses previously recorded user requests of last release and also maintains new session data for the new release. Alshahwan and Harman has [3] proposed data repair approach of user

session for web regression testing and also introduced an algorithm based on the session repair concept.

Weimer and Dobolyi [6] proposed a fully automated technique that uses automated oracle comparator based on the semantics of XML/ HTML output and underlying inherent similarities among web application.

Dobolyi at el [7] has extended previous work mentioned in [6] by providing a fully automated regression testing approach that automatically compares the output of two versions. The introduction of "Smart" a highly precise oracle comparator was given. It is an efficient comparator for locating actual faults and also reducing the cost of regression testing.

Kam and Dean [8] conducted a survey of different testing techniques used for web applications and have concluded that a single testing technique is not able to test all the bugs of a web application. Some methods can only be used to test controlled flow of the transactions while others can only handle uncontrolled flows. Thus we need multiple techniques integrated together, to validate and verify a web application.

Xu et al [12] modeled web applications through System Dependent Graph (SDG) and proposed a technique for web application regression testing that was based on slicing. The System Dependent Graph will raise the cost and workload of the testing process.

Tarhini et al [16] proposed a technique which selected test cases based on event dependency graph for web application regression testing. Test cases for regression testing were selected based on identifying changed components.

Hernandez et al. [20] proposed UML 2.0 profiles meta-model and developed automated testing scripts for web application testing by applying a model-driven approach. A test implementation prototype for an e-commerce application was presented and also explained the model transformations to port regression tests to various platforms.

Gagandeep and Sengupta[21] proposed domain analysis and modeling technique for web application testing that was based on model driven architecture of the system. This approach drastically reduced the cost and effort of rigorous cycle of software development and testing process. Analysis of web application using a graphical web model was also accomplished along with the optimization and automation of the test generation process.

Shaar and Haraty [23] presented an analysis model for testing and modeling web application. It includes a divisional analysis model consisting of three sub models; the architectural environment model, the client side model and the server side programs model. The automated black box regression testing technique was also proposed but external content sources and server side logic testing were not considered in their model.

Xu et al[30] proposed web application testing in five parts i.e. web modeling for regression testing, testing techniques and methods, test case generation, testing execution and measurement process and then proposed four methods of testing which were applicable to test web application. Some specialties of web application i.e. distributed structures, numerous users, interactive and dynamic functions of web applications etc were also considered in their research. All their research based on previous research work.

Tarhini et al[31] proposed a technique of regression testing for re-testing customized web application. This technique modeled a web application and its components behavior using abstract level model. Furthermore, the safe algorithm of regression testing was suggested to test the web applications.

Yang et al[33] proposed a method of regression testing for composite web services. Prototype was developed for testing. The testing method was also presented which analyzed the process of implementation. This method is very much efficient for Composite Web service testing.

2.3. Related Work

B Kam and T R. Dean listed six potential web application problems, categories of research groups and the core features of web applications with the conclusion that user based approach and slicing based approach have same capability to detect same type of web application problems in web testing [8]. These two techniques are shown below in the table:

Table 2. 2:Comparison between User Session and Slicing [8]

Method	Testing Problems						Scores
	Static Link	Dynamic Link	Form link	Dynamic Page Creation	Syntax Error	Uncontrolled flow transaction	
User Session		X	X	X			3
Slicing		X	X	X			3

Table 2.2 shows that both methods have the capability to handle same type of problems and both are not suitable for testing the initial copy of the web application [8].

Prior work on regression testing has focused on the problems that are related to managing the intrinsic complexity of re execution of a large test data suite on a modified version of application [10, 11, 26]. While considering this technique in web applications, testers mainly focus on design issues and white box techniques [12]. Regression test suites are executed by the tester to ensure that the modified version of the software is functioning as predictable.

User session and slicing based approaches have same type of testing problems. Many researchers worked on user session based approaches in web application regression testing but limited work reported on slicing based approaches in web application regression testing. On the other hand, no one has ever compared these two same capable approaches that handle same type of issues.

2.4. Summary

A large number of researchers have worked on more general features of web application testing, apart from regression testing issues. Each researcher has focused different aspects of web applications like static and dynamic pages, links, architecture, scenarios and model. A large number of researchers have worked in the area of user session based testing but a few researchers have worked on slicing in web application regression testing domain. Most of the work reported on slicing concerns extension and improvement to slice construction algorithms and different forms of slicing.

CHAPTER # 3

Methodology

3.1. Introduction

Testing is the process that is used to assist testers to recognize the completeness, quality and accuracy of developed software. Web applications need to be available 24 hours a day thus tester need to fix bugs within a short time frame.

User session based approach, slicing based approach and hybrid approach of regression testing in web application domain is implemented in this chapter. A comparison of these techniques is performed to identify better one among them with respect to performance.

3.2. Research Methodology

Research methodology is the detailed steps where as steps of research method giving here. In this research, experimental research methodology has been used which has drawn the results of comparative evaluation of two techniques. The following steps are involved in this research:

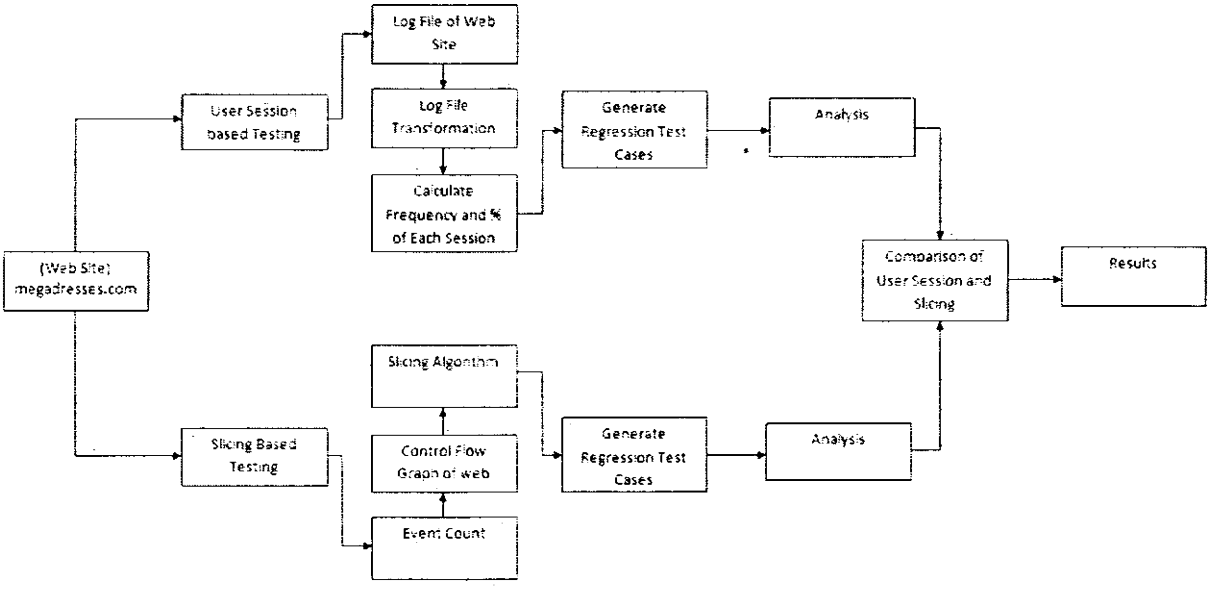


Figure 3. 1: Steps of Experiment for Research Question 1

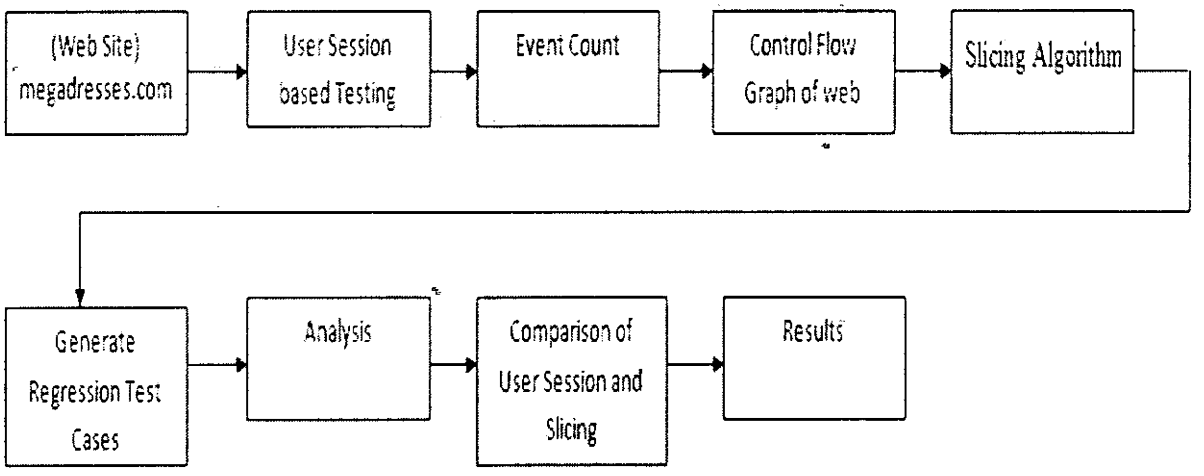


Figure 3. 2: Steps of Experiment for Research Question 2

3.3. Implementation of Approaches

User session, slicing and hybrid approach has been implemented on website named as "www.megadresses.com".

3.3.1. Implementation:User Session Based Testing

Transformation of user sessions into test cases is called user session based testing. Test cases contain number of HTTP requests which are linked with each user session. Every session is a group of user requests in the form of name-value pairs and URLs in user session based testing. A user session is initialized and ended whenever a new user or an IP address accesses the website or leaves respectively. User sessions are identified by the IP addresses of user but it is considered as a new session after 45 minutes.

User sessions are important for many reasons. First, it can help to customize users according to their preferences. Secondly, it gives information of the traffic on the website and thirdly, it easily

examines defects through recorded website logs. The user behavior is the best representation of several common requirements.

When a user visits a Web application, it records actions in a log file and uses these actions afterward for testing. These results ensure quality of web and provide very effective way for testing. Test case reduction removes unnecessary test cases from the test suite, which do not affect the efficiency.

Log file Extraction

The data of user session technique is taken from server. A data set of server website is known as www.megadresses.com, the IP address indicates different domains that make a request for a page from a web server. The session is the time in web logs of a server when the user makes the requests for that web page; which is the time during which user accesses that web page. The data sets acquired from web servers is in notepad (.txt) format but this data is always going to be in raw format making it impossible to analyze. In order to alter the data sets into logical format the data sets are converted into the Excel tables. The process of converting Notepad data file into MS Excel table is given below:

First open the MS Excel

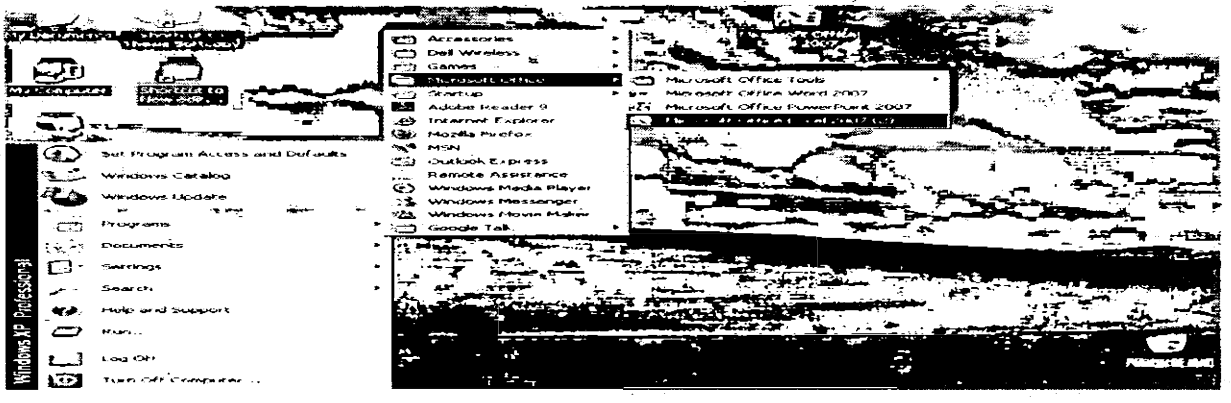


Figure 3. 3:Open web log data set File Saving in MS Excel

Browse the path of web server data file in MS Excel.

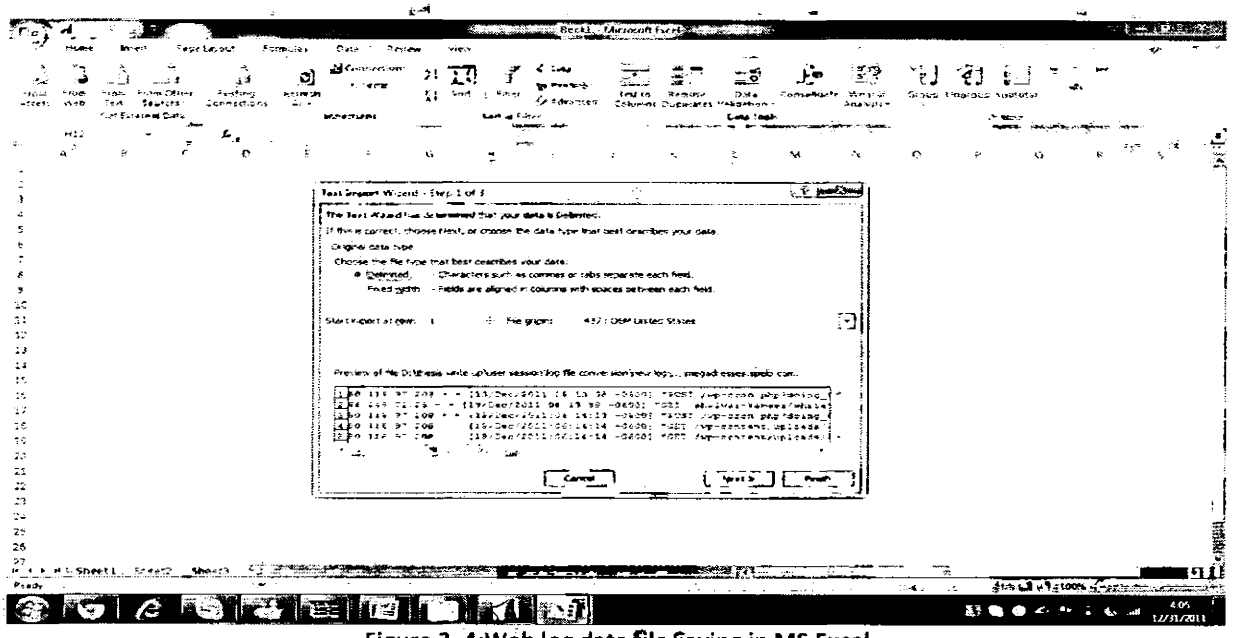


Figure 3. 4:Web log data file Saving in MS Excel

The space delimited is selected then click next and finish. It has to save the web server data file.

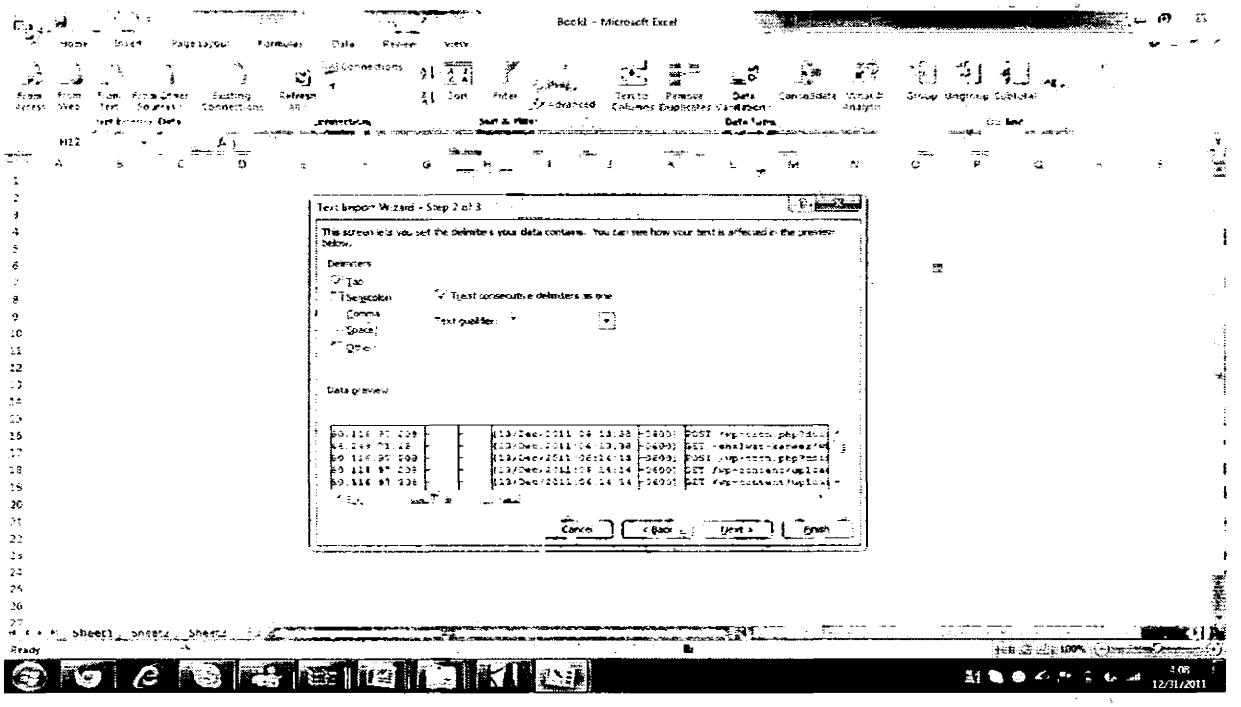


Figure 3. 5:Web log data File Saving in MS Excel

After finish, the data looks in a logical way. Then it is saved.

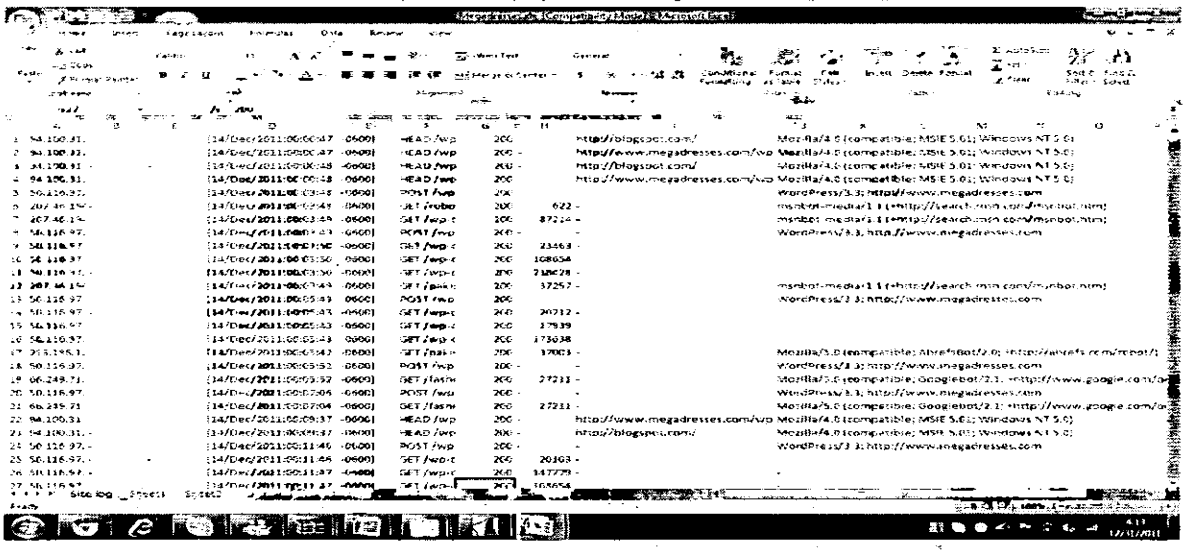


Figure 3. 6: Log data File Saving in MS Excel

The first column is denoted as the internet protocol (IP) address, the second and the third column are utilized as a separator of the data.

Calculate Frequency

The user sessions are sorted out in descending order according to the link frequency. If a page is accessed by the user most frequently, its probability to access the web page will be high. This makes the data understandable, as well as easy to analyze.

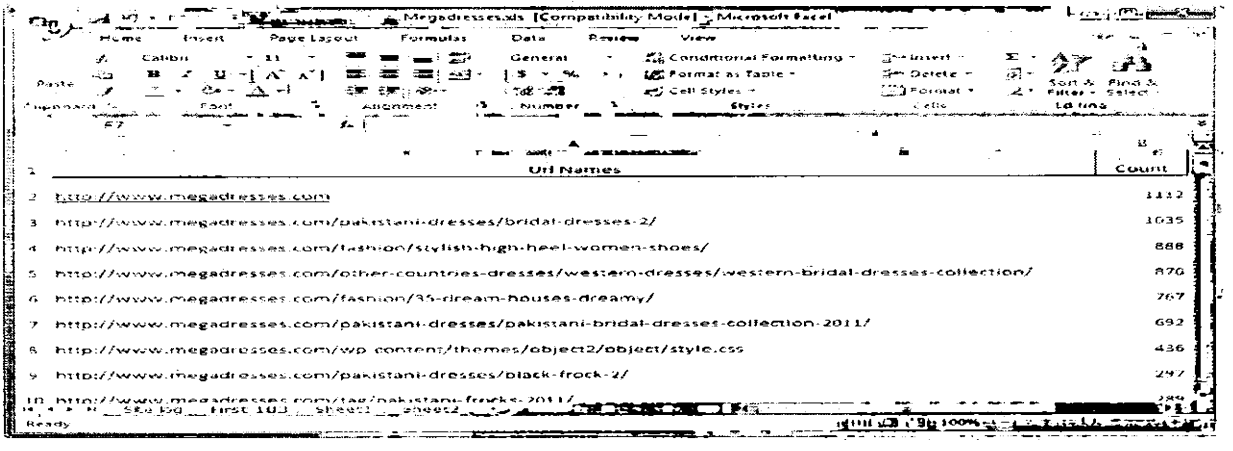


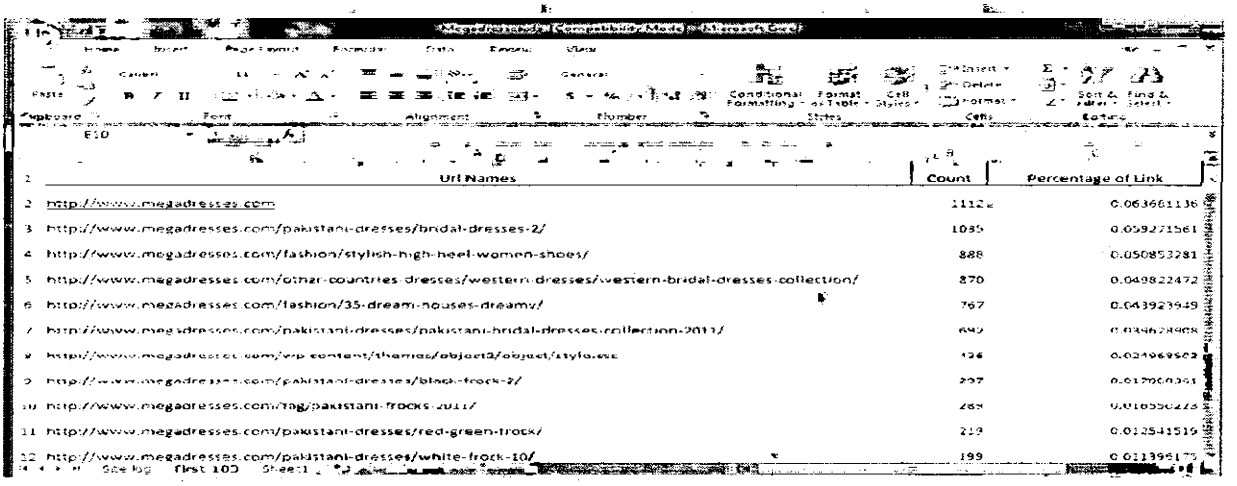
Figure 3. 7:Frequency Count of a log file

After that, the percentage of each link is computed relative to the total amount of frequency links is shown in the below mentioned formula:

Total number of links frequencies

Link Percentage = Frequency of the link

Total No of Link Frequencies

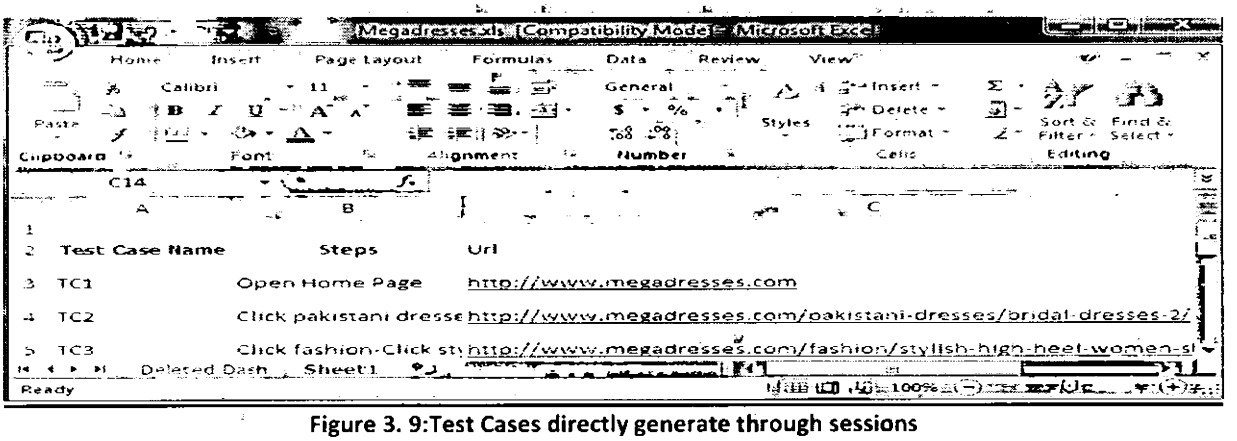


	Url Names	Count	Percentage of Link
2	http://www.megadresses.com	1112	0.063661136
3	http://www.megadresses.com/pakistani-dresses/bridal-dresses-2/	1035	0.059271561
4	http://www.megadresses.com/fashion/stylish-high-heel-women-shoes/	888	0.050853281
5	http://www.megadresses.com/other-countries-dresses/western-dresses/western-bridal-dresses-collection/	870	0.049822472
6	http://www.megadresses.com/fashion/35-dream-houses-dreamy/	767	0.043923949
7	http://www.megadresses.com/pakistani-dresses/pakistani-bridal-dresses-collection-2011/	692	0.039628908
8	http://www.megadresses.com/wp-content/themes/object2/object/style.css	426	0.024396952
9	http://www.megadresses.com/pakistani-dresses/black-frock-2/	297	0.017006361
10	http://www.megadresses.com/tag/pakistani-frocks-2011/	288	0.016490228
11	http://www.megadresses.com/pakistani-dresses/red-green-frock/	219	0.012541519
12	http://www.megadresses.com/pakistani-dresses/white-frock-10/	199	0.011396179

Figure 3. 8:Percentage of each link

Test Case Generation

Test cases are directly generated through sessions based on frequency and link percentage then covered and newly covered path found through test cases. Newly covered are those paths that are previously uncovered.



	A	B	C
1			
2	Test Case Name	Steps	Url
3	TC1	Open Home Page	http://www.megadresses.com
4	TC2	Click pakistani dresse	http://www.megadresses.com/pakistani-dresses/bridal-dresses-2/
5	TC3	Click fashion-Click sty	http://www.megadresses.com/fashion/stylish-high-heel-women-sl

Figure 3. 9:Test Cases directly generate through sessions

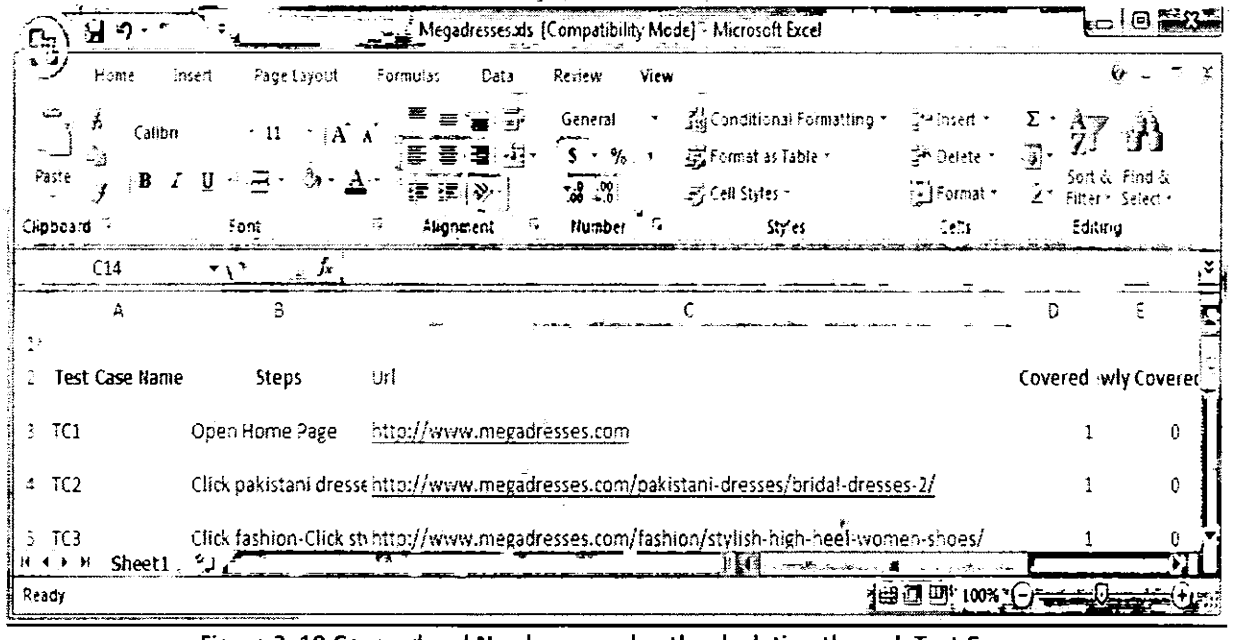


Figure 3. 10:Covered and Newly covered path calculation through Test Cases.

Calculate Test Efficiency and Coverage

Test efficiency and coverage are calculated through formula using covered and newly covered paths.

Formula of Test Efficiency and Coverage

- Coverage $= \frac{t``}{N}$
- Test Efficiency $= \frac{1 - t``}{n} = \frac{t`}{n}$

t` = Total no of newly covered path

t`` = Total no of covered path

n = Total no of test cases

3.3.2. Implementation: Slicing Based Testing

The Program Slicing technique is used for understanding, testing and maintenance of web applications [40]. Slicing method was used in order to perform regression testing effectively and efficiently. Here, slicing based testing is implemented in web applications.

Point of Interest

The first step is to identify point of interest, which is collected through event count. Event count is based on the paths of possible interactions with the web site (when users access the website they click on different links). Event count is basically a track of the user clicks on the web site.

URL	Clicks
http://www.megadresses.com/top-stores/	22
http://www.megadresses.com/pakistani-dresses/black-embroidered-frock/	16
http://www.megadresses.com/pakistani-dresses/black-maxi/	7
http://www.megadresses.com/	7
http://www.megadresses.com/category/pakistani-dresses/a-line-frock/	6
http://www.megadresses.com/pakistani-dresses/green-embroidered-dress/	6
http://www.megadresses.com/pakistani-dresses/lakhany-silk-gorgeous-wintery-collection-20112012/	6
http://www.megadresses.com/pakistani-dresses/red-dress-2/	6
http://www.megadresses.com/build-store/	4
http://www.megadresses.com/contact-us/	4
http://www.megadresses.com/category/pakistani-dresses/abaya/	3
http://www.megadresses.com/pakistani-dresses/lemon-yellow-party-dress/	3
http://www.megadresses.com/pakistani-dresses/orange-frock-2/	3
http://www.megadresses.com/pakistani-dresses/red-winter-beautiful-dress/	3
http://www.megadresses.com/pakistani-dresses/mehndi-maroon-bridal-dress/	2
http://www.megadresses.com/pakistani-dresses/multi-shaded-frock/	2
http://www.megadresses.com/pakistani-dresses/pakistani-bridal-dresses-collection-2011/	2
http://www.megadresses.com/pakistani-dresses/4349/	2

Figure 3. 11:An example of event counts

Event counts of web site are considered as a point of interest. After getting point of interest, control flow graph and slicing algorithm of website are made.

Construction of Control Flow Graph (CFG) and Slicing Algorithm

Slices are computed using a Control Flow Graph (CFG) and slicing criteria.

Figure 3.13 shows CFG.

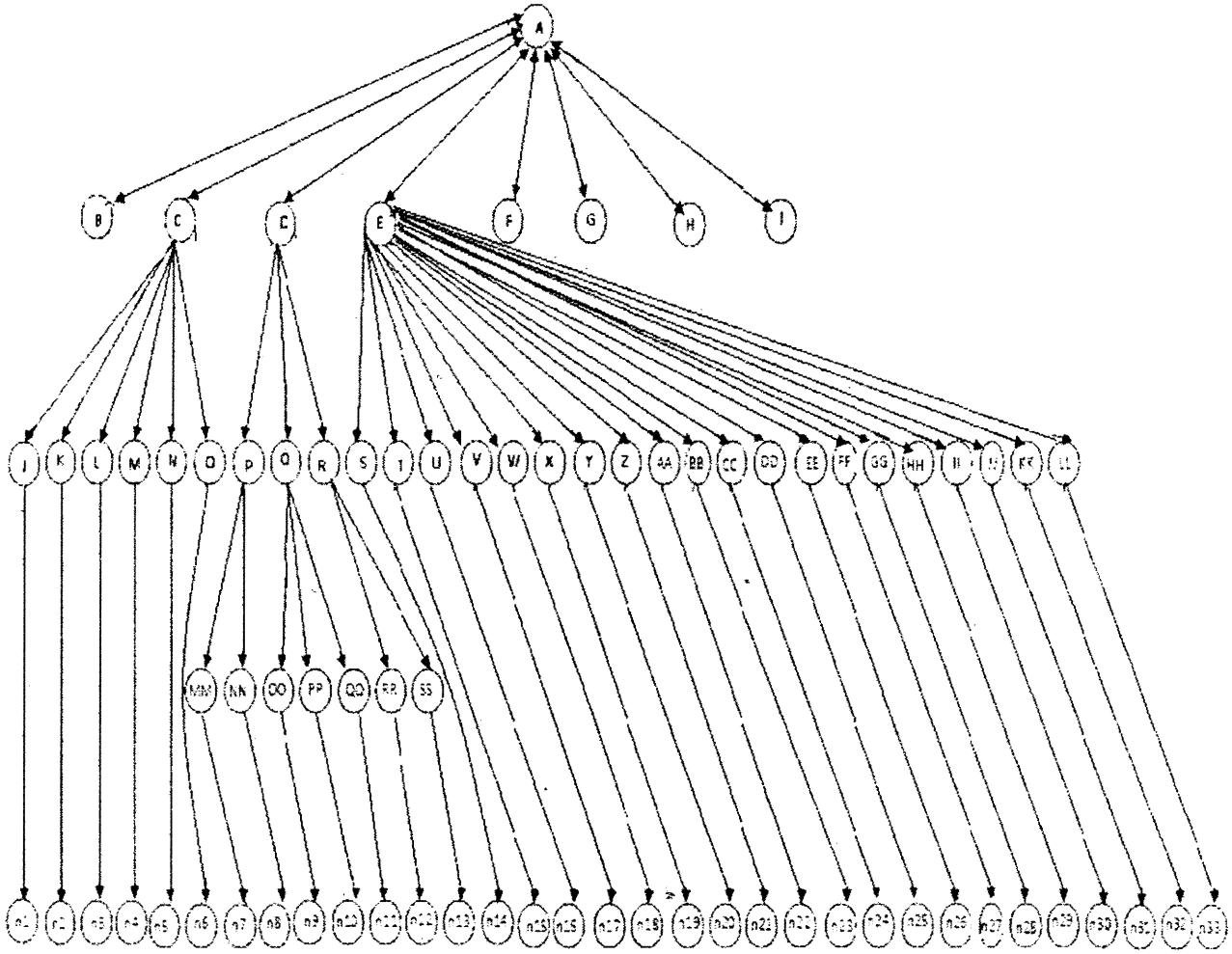


Figure 3. 12:Control Flow Graph(CFG)

This algorithm defines the steps of slicing.

Slicing Algorithm:

1. Generate the CFG.

2. For a particular slicing criteria <s, v>, mark s.

3. Starts back traversing from s and highlights the nodes and edges.

4. Map the highlighted node to the related links in the program.

The algorithm is used to find the slices for the event count link. The required slice could be acquired by backward traversing in the resultant graph and marking the edges and nodes. The Figure 3.14 is shown the resultant CFG.

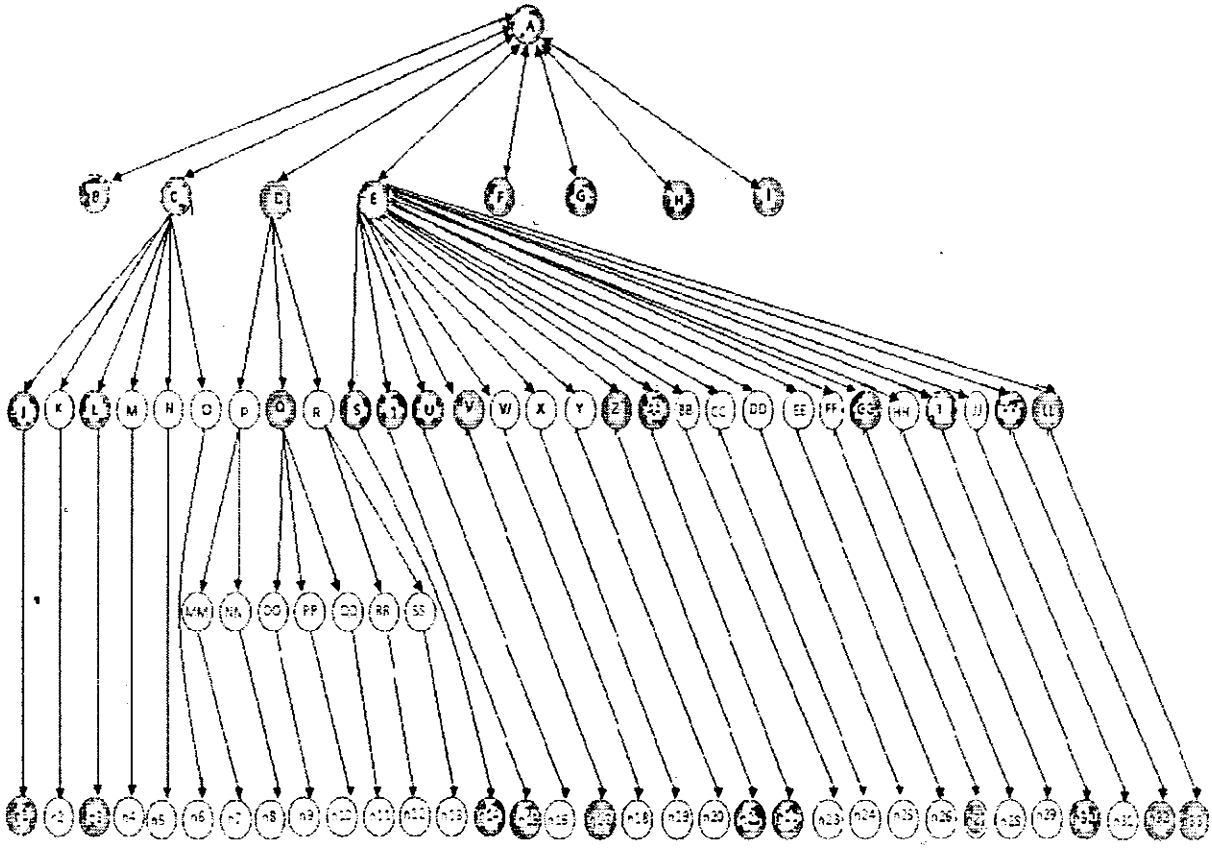


Figure 3. 13:The resultant CFG with marked node and edges

The slice is computed through backward traversing in the graph and marking the nodes and edges.

Test Case Generation

Test cases are generated through the slices of control flow graph and then covered and newly covered path are found through these test cases.

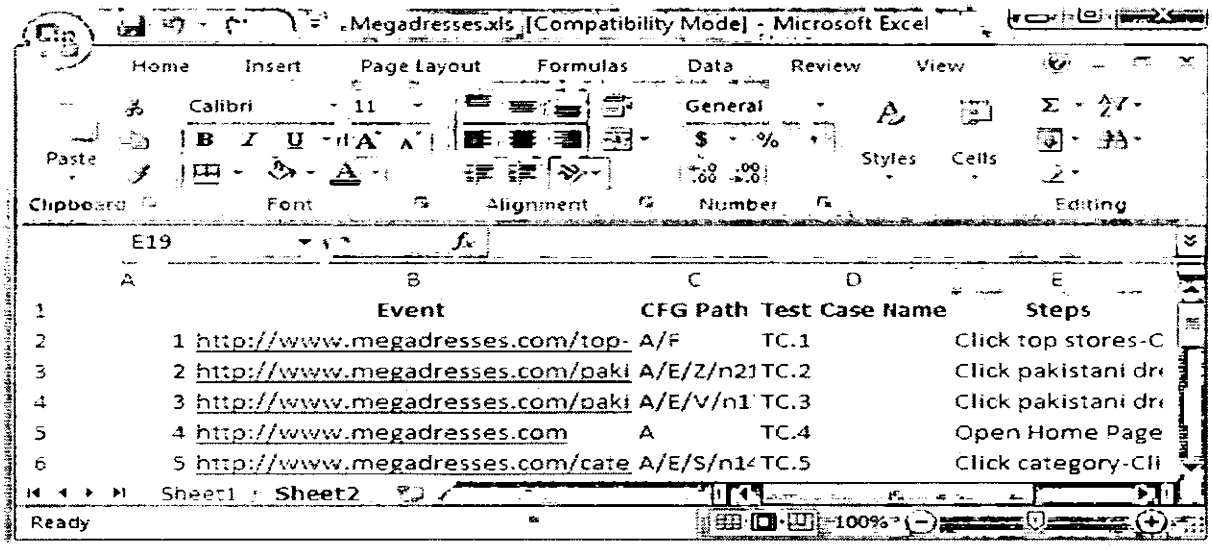


Figure 3. 14:Test Cases Through CFG

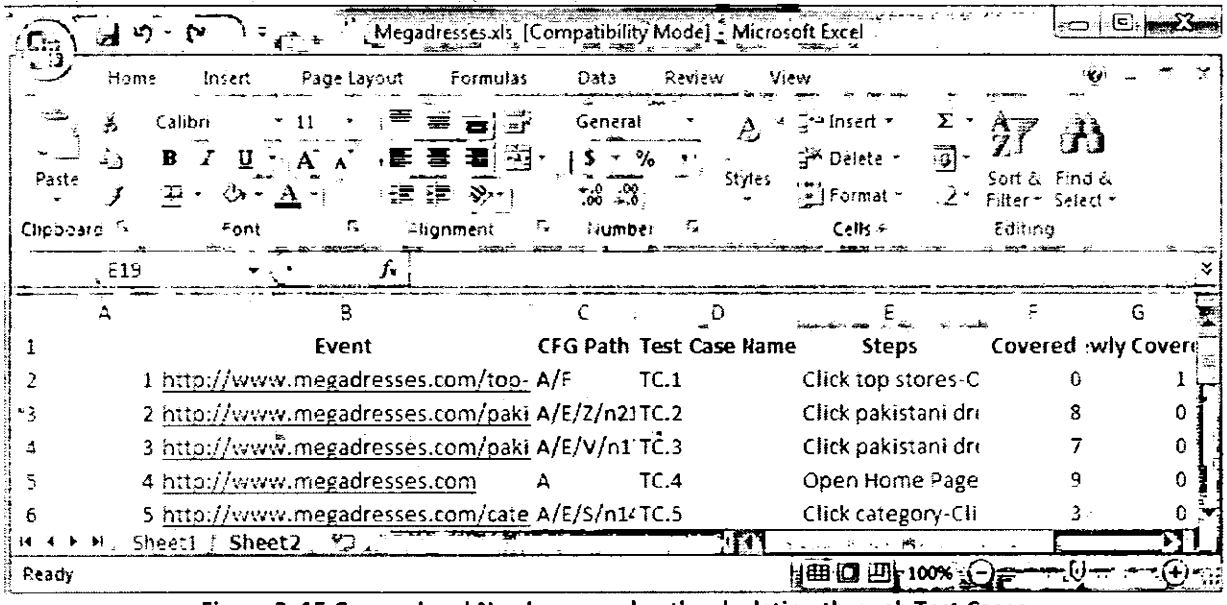


Figure 3. 15:Covered and Newly covered path calculation through Test Cases

Calculate Test Efficiency and Coverage

Test efficiency and coverage are calculated through formula using covered and newly covered path.

Formula of Test Efficiency and Coverage

- Coverage $= \frac{t``}{N}$
- Test Efficiency $= \frac{1 - t``}{n} = \frac{t`}{n}$

t` = Total no of newly covered path

t`` = Total no of covered path

n = Total no of test cases

3.3.3. Implementation:Hybrid Approach

In hybrid approach, the user session and slicing are used in conjunction. Half user session and half slicing URL paths are used in hybrid approach.

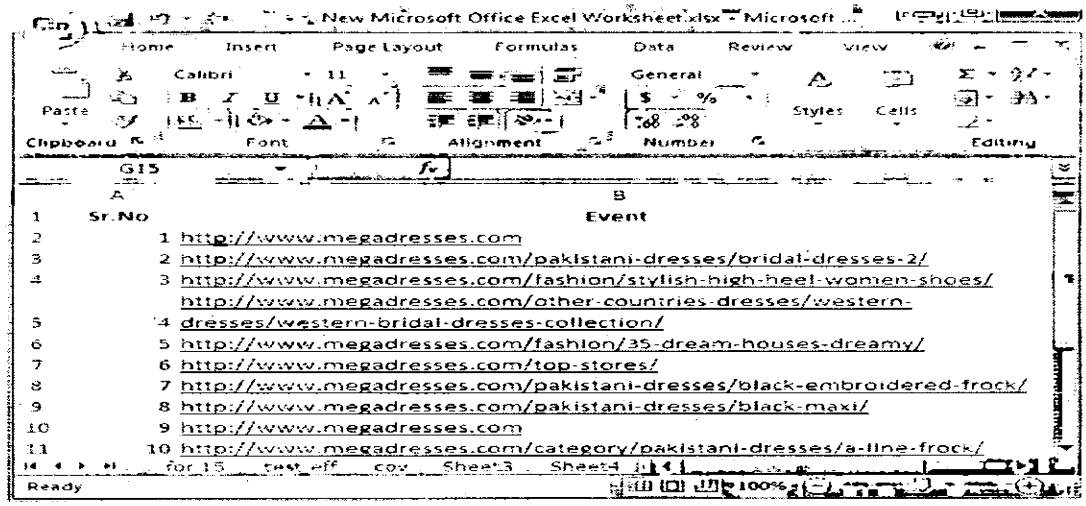


Figure 3. 16:Hybrid Approach Paths

Then control flow graph is made and slicing algorithm is applied. Figure 3.18 shows CFG after slicing.

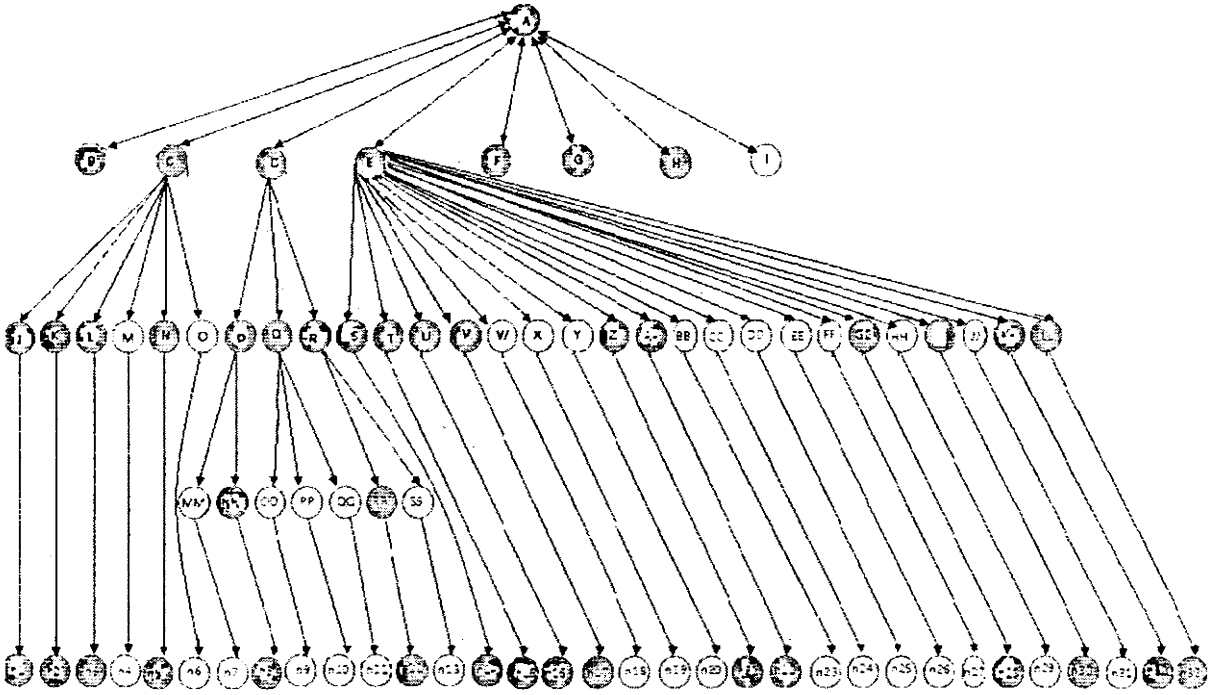
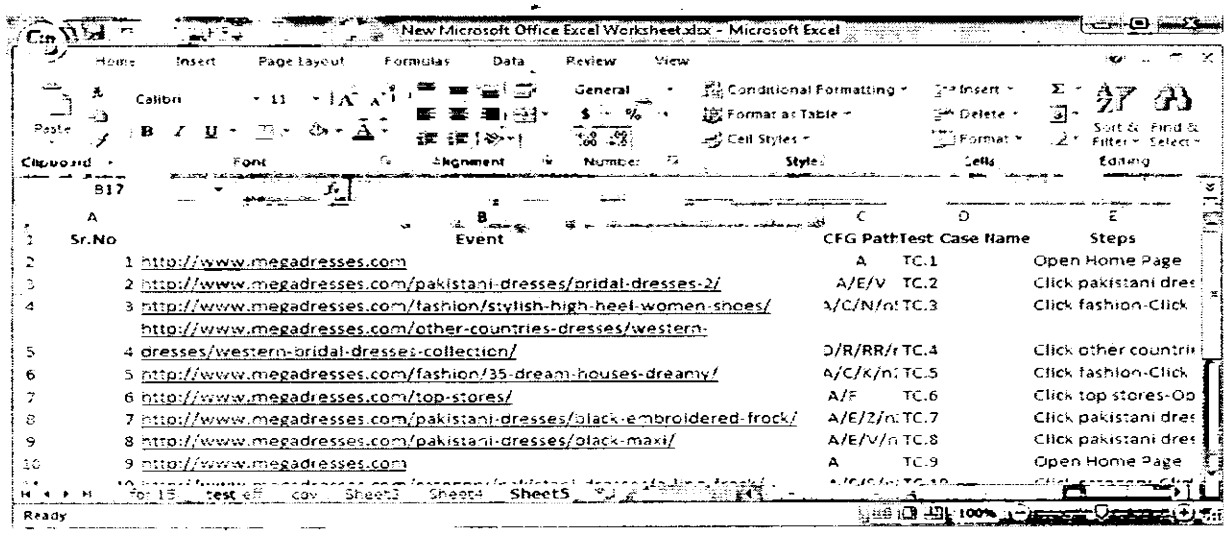


Figure 3. 17:Hybrid approach resultant CFG with marked node and edges

Then test cases are generated after the slices of website.



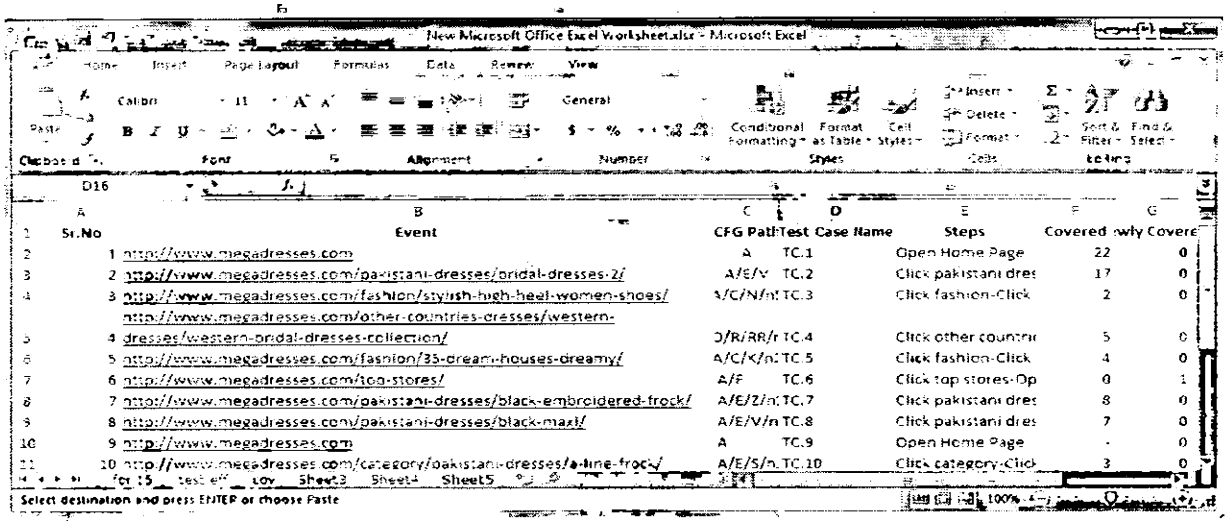
New Microsoft Office Excel Worksheet.xlsx - Microsoft Excel

Sr.No	Event	CFG Path	Test Case Name	Steps
1	http://www.megadresses.com	A	TC.1	Open Home Page
2	http://www.megadresses.com/pakistani-dresses/bridal-dresses-2/	A/E/V	TC.2	Click pakistani dres
3	http://www.megadresses.com/fashion/stylish-high-heel-women-shoes/	A/C/N/n	TC.3	Click fashion-Click
4	http://www.megadresses.com/other-countries-dresses/western-dresses/western-bridal-dresses-collection/	A/E/R/RR/r	TC.4	Click other countri
5	http://www.megadresses.com/fashion/35-dream-houses-dreamy/	A/C/K/n	TC.5	Click fashion-Click
6	http://www.megadresses.com/top-stores/	A/F	TC.6	Click top stores-Op
7	http://www.megadresses.com/pakistani-dresses/black-embroidered-frock/	A/E/Z/n	TC.7	Click pakistani dres
8	http://www.megadresses.com/pakistani-dresses/black-maxi/	A/E/V/n	TC.8	Click pakistani dres
9	http://www.megadresses.com	A	TC.9	Open Home Page

Figure 3. 18:Test Cases Through CFG.

Calculate Test Efficiency and Coverage

Test efficiency and coverage is calculated through formula using covered and newly covered path.



Sr.No	Event	CFG Path	Test Case Name	Steps	Covered	Newly Covere
1	http://www.megadresses.com	A	TC.1	Open Home Page	22	0
2	http://www.megadresses.com/pakistani-dresses/bridal-dresses-2/	A/E/V	TC.2	Click pakistani dres	17	0
3	http://www.megadresses.com/fashion/stylish-high-heel-women-shoes/	A/C/N/n	TC.3	Click fashion-Click	2	0
4	http://www.megadresses.com/other-countries-dresses/western-dresses/western-bridal-dresses-collection/	D/R/RR/r	TC.4	Click other countrie	5	0
5	http://www.megadresses.com/fashion/35-dream-houses-dreamy/	A/C/K/n	TC.5	Click fashion-Click	4	0
6	http://www.megadresses.com/top-stores/	A/F	TC.6	Click top stores-Op	0	1
7	http://www.megadresses.com/pakistani-dresses/black-embroidered-frock/	A/E/Z/n	TC.7	Click pakistani dres	8	0
8	http://www.megadresses.com/pakistani-dresses/black-maxi/	A/E/V/n	TC.8	Click pakistani dres	7	0
9	http://www.megadresses.com	A	TC.9	Open Home Page	-	0
10	http://www.megadresses.com/category/pakistani-dresses/a-line-frock/	A/E/S/n	TC.10	Click category-Click	3	0

Figure 3. 19:Covered and Newly covered path calculation through Test Cases

Whenever change comes these test cases are able to run for revalidating a customized system, but new test cases are also created. This fulfills the requirements of regression testing.

3.4. Summary

In this chapter, experimental research methodology has been explained. User session and slicing have implemented separately and then both approaches have implemented in conjunction. The implementation steps of each approach have discussed thoroughly.

CHAPTER # 4

Results

4.1. Introduction

All the approaches are compared using three sets of test cases (15, 30, 60) in this chapter. Initially, both approaches i.e., user session and slicing are applied on website for sixty test cases ,then applied on thirty test cases then on fifteen test cases . These approaches are compared recursively on these three sets.

4.2. Experimental Design

The experiment is conducted to find out the answer to following questions and evaluate the below given hypothesis.

- **Hypothesis Formulation:**

1. User session and slicing, both techniques when applied in conjunction have a positive effect on performance of web application regression testing.
2. User session and slicing, both of these techniques have varying performance values during web application regression testing when applied individually.

- **Null Hypothesis:**

1. User session and slicing, both techniques when applied in conjunction have no effect on the performance of web application regression testing.
2. User session and slicing, both of these techniques have same performance during web application regression testing when applied individually.

- **Treatment:**

User Session and Slicing approaches are applied on web application.

- **Experimental Type:**

Crossover design has been adapted to measure test effectiveness and coverage metrics.

- **Independent Variable:**

The approaches (user session, slicing and hybrid) are the independent variables during web application regression testing.

- **Dependent Variable:**

Performance is the dependent variable.

- **Metrics:**

The test effectiveness and coverage metrics act as the measurement tool to measure the performance.

- **Internal Validity:**

The experiment is applied on three different sets of test cases not applied on the whole test cases, which is supposed to be a threat of internal validity.

- **External Validity:**

Approaches are run recursively on web application through different test cases, thus the results are generalizable.

4.3. Data Set

A website "www.megadresses.com" has been developed and then uploaded on world wide web for the sake of experimentation. This website is currently running on the internet as well.

Coverage and test efficiency matrices are used to measure the performance of these approaches on website.

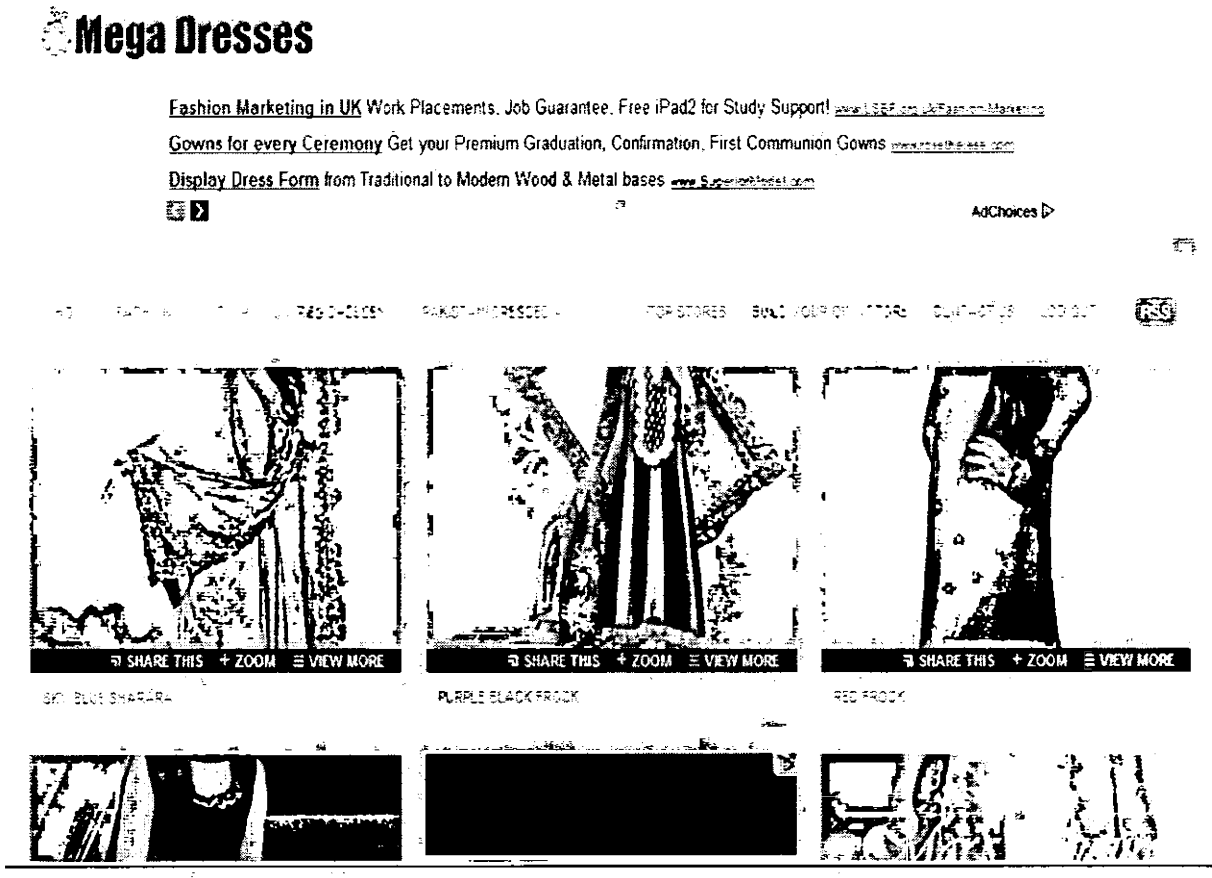


Figure 4. 1:Website Home Page

4.3. Experimental Results

The results of user session, slicing and hybrid approach are as follow:

4.3.1. For 60 Test Cases:

Initially, user session and slicing based approaches are applied on website for sixty test cases. The comparative results of user session and slicing obtained are depicted in the following graphs:

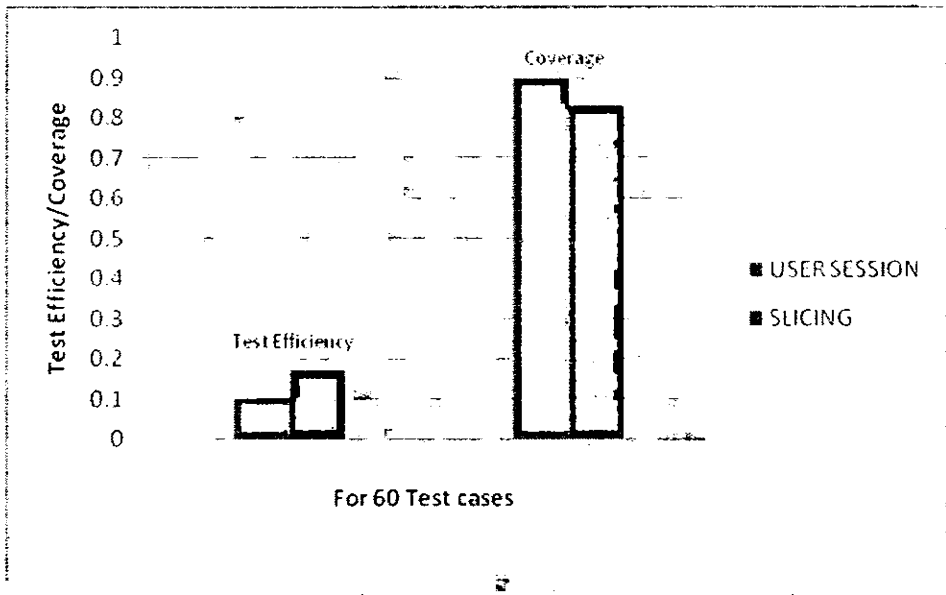


Figure 4. 2:Comparative Analysis of user session and slicing for 60 test cases

4.3.2. For 30 Test Cases:

Then user session and slicing based approaches are applied on website for thirty test cases. The comparative results of user session and slicing are:

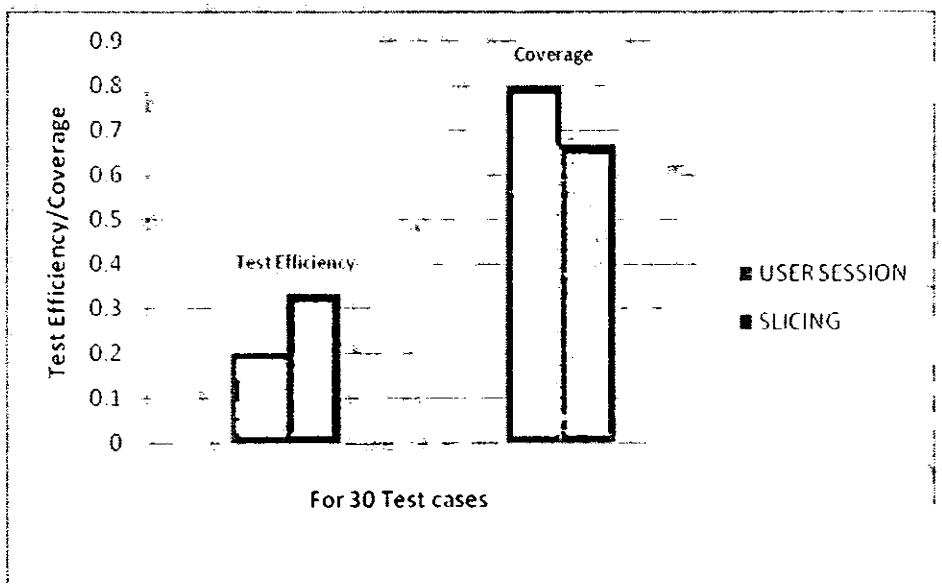


Figure 4. 3:Comparative Analysis of user session and slicing for 30 test cases

4.3.3. For 15 Test Cases

At the end, user session and slicing based approaches are applied on website for fifteen test cases. The comparative results of user session and slicing obtained are:

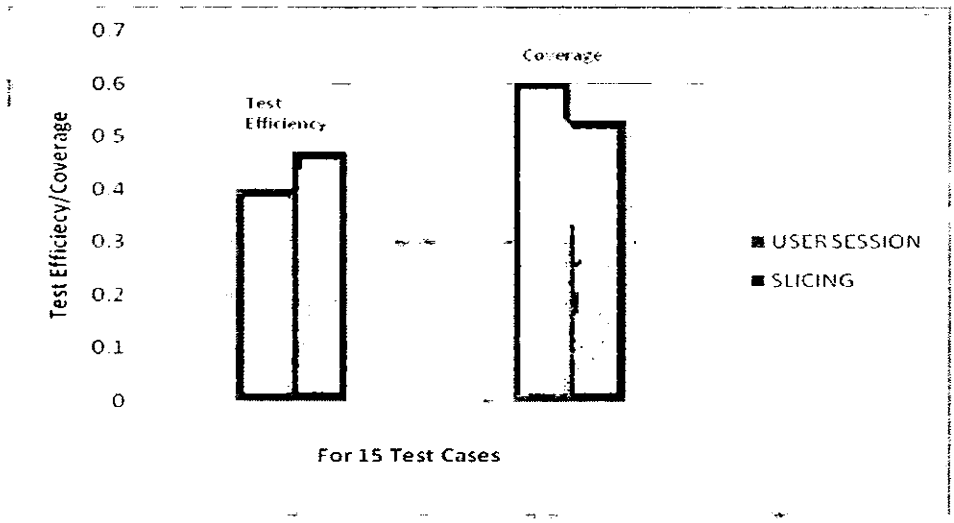


Figure 4. 4:Comparative analysis of user session and slicing for 15 test cases

Comparative Analysis of (15,30,60) test cases in terms of Coverage and Test Efficiency

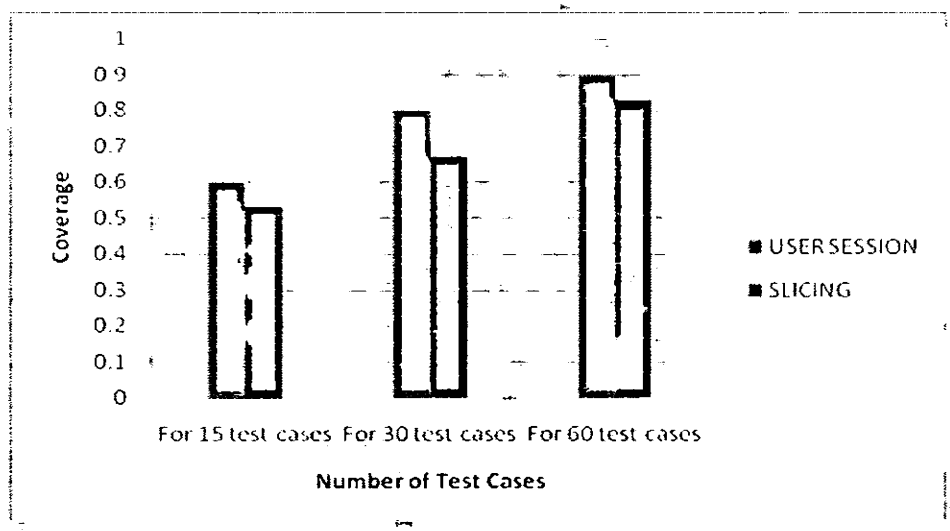


Figure 4. 5:Comparative Analysis of all three in term of coverage

When the results are compared in term of coverage, it is found that user session provides more coverage than slicing in all three sets of test cases because number of session's increases in the test suite.

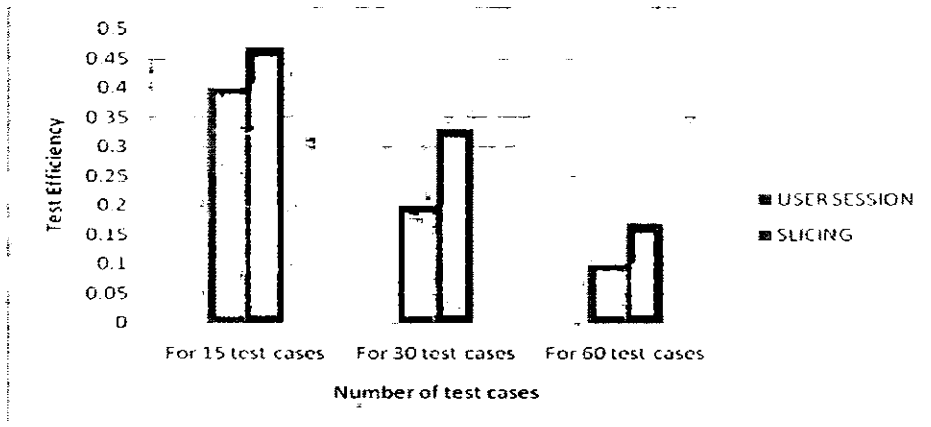


Figure 4. 6:Comparative Analysis of all three in term of test efficiency

When the results are compared in term of test efficiency, slicing provides more test efficiency than user session because less number of paths covered in slicing due to slices of web as compared to user session.

Comparative Analysis of all three (15,30,60) test cases in tabular form

Table 4. 1:Comparative Analysis of all three sets of test cases

	For 60		For 30		For 15	
	TEST EFFICIENCY	COVERAGE	TEST EFFICIENCY	COVERAGE	TEST EFFICIENCY	COVERAGE
USER SESSION	0.1	0.9	0.2	0.8	0.4	0.6
SLICING	0.17	0.83	0.33	0.67	0.47	0.53

Table 4.1 shows the comparative analysis of all three sets of test cases.

The efficiency of test generation process of user session approach is influenced by the data collection process. User session based approaches handles a number of alterations without the recollection of updated session data.

User session and slicing have different strengths and weaknesses so it depends on the requirement of testing project which is important in terms of coverage and efficiency, if, coverage is important then user session should be used and if test efficiency is important then slicing should be used.

In the end hybrid approach is implemented i.e. the user session and slicing based approaches are applied together. Hybrid approach is applied for three sets of test cases (15,30,60).This approach are compared recursively on these sets. The results obtained are showed in below given figures.

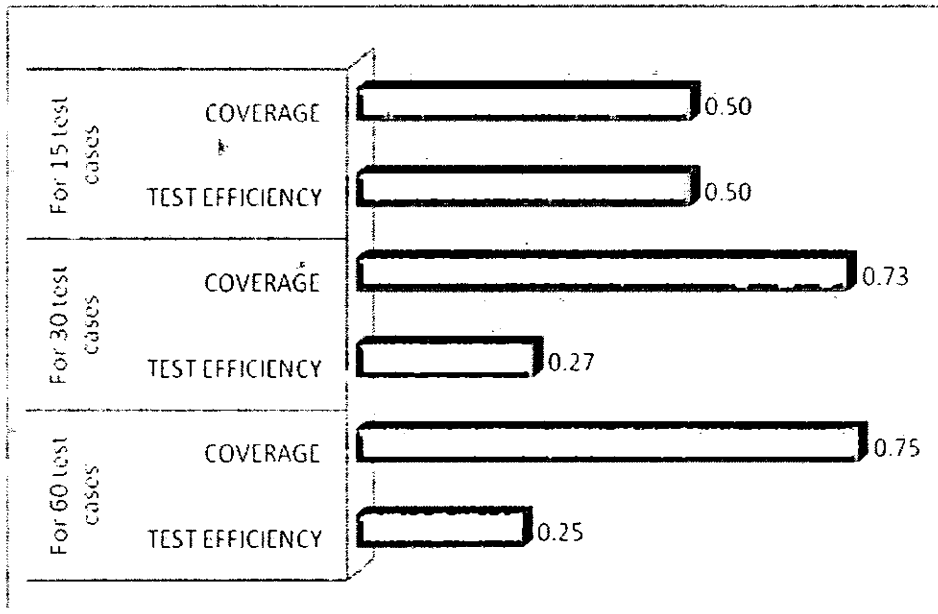


Figure 4. 7:Hybrid Implementation Result

The results of hybrid approach show that more test efficiency is achieved when both techniques are used in conjunction; however, coverage is minimally decreased in large number of test cases due to in conjunction usage of two techniques.

4.4. Summary

In this chapter, comparative results of the approaches i.e. user session, slicing and hybrid approach have been shown. The results show that user session increases the coverage and slicing increases the test efficiency. The hybrid approach is minimally decreases the coverage in large number of test cases but increases the test efficiency. User session and slicing have different strengths and weaknesses so it depends on the requirement of the testing project which one is used according to their project needs.

CHAPTER # 5

Conclusion and Future Work

5.1. Conclusion

Web applications consist of dynamic and static web pages which create dynamic responses based on user requests. Large web application software contains significant user interactions and in between objects interactions, thus testing of web is more complex than testing of traditional software.

Web applications have a number of properties such as rapid change in user demand and developing speed. Many bugs may hide in application due to limited time, assorted nature and cost of testing. Particularly these bugs/faults are incorporated in the application via changes and modifications of applications, so regression testing is very difficult to be performed in the Web applications.

Basically the aim of regression testing is to make sure that modifications do not negatively affect the original behavior of the program. Test cases are reused for validating a modified system during regression testing; thus new test cases are also generated.

User session and slicing based testing approaches are important approaches in web application regression testing domain. Both are not appropriate for testing the initial copy of the web application.

User session based testing has an advantage that real sessions can be recorded in web applications and these sessions are easily used for regression testing in web application domain. It is also useful in the absence of program requirements and specifications.

Program slicing is a suitable technology used in program understanding, analyzing, testing and maintenance. Program slicing is very important tool for incremental regression testing problem. A program slice consists of those portions of the program which affect the value calculated at some point of interest, referring to the slicing criteria.

User session and slicing have different strengths and weaknesses so they depend on the requirement of testing project which is important in terms of coverage and efficiency, if coverage is important then it is best to use the user session and if test efficiency is important then slicing should be used. When they are used in conjunction it is seen that test efficiency increases but they don't provide appropriate results in terms of performance.

5.2. Future Work

As manual slicing has been used, future work includes automate test case generation of slicing in web application regression testing domain. Another future direction can be extending the number of test cases, as both approaches have applied on 60 highest frequency test cases. Both approaches can be applied to whole test cases in future and then accordingly their effectiveness can be checked. Thirdly the results of this experiment can be used to compare with other techniques as well.

References

[1] S Elbaum, S Karre, G Rothermel, "Improving Web Application Testing with User Session Data". 25th Int. Conf. on Software Engineering, pp. 49 – 59. 2003.

[2] S Elbaum, G Rothermel, S Karre, M Fisher II, "Leveraging User-Session Data to Support Web Application Testing". IEEE Journal on Software Engineering, pp. 187 – 202. March 2005.

[3] M Harman, N Alshahwan, "Automated Session Data Repair for Web Application Regression Testing". Int. Conf. on Software Testing, Verification and Validation, pp. 298-307. 2008.

[4] F I. Vokolos, P G. Frankl, "Pythia: A regression test selection tool based on textual differencing". Int. Conf. on Reliability, Quality and Safety of Software Intensive Systems, pp. 3 – 21.1997.

[5] M Benedikt, J Freire ,P Godefroid,"VeriWeb: Automatically Testing Dynamic Web Sites". 11th Int. Conf. on WWW.2002.

[6] K Dobolyi, W Weimer, "Harnessing Web-based Application Similarities to Aid in Regression Testing". 20th International Software Reliability Engineering Symposium, pp. 71-80. 2009.

[7] K Dobolyi, E Soechting, W Weimer, "Automating regression testing using web-based application similarities". STTT–WSE 2009, Vol. 13.Issue 2,pp. 111-129 .2011.

[8] B Kam, T R. Dean, "Lessons learned from a survey of web applications testing". 6th Int. Conf on IT: New Generations. pp. 125-130.2009.

[9] R Gupta, M Jean , H Mary , L Soffa, "Program Slicing-Based Regression Testing Techniques". 1996.

[10] D Binkley. "The Application of Program Slicing to Regression Testing". Information and Software Technology Special Issue on Program Slicing, Vol. 40, Issue: 11-12, pp. 583-594.1999.

[11] R Gupta , M Jean , M J Harrold , M L Soffa. "An Approach to Regression Testing using Slicing". In Conference on Software Maintenance. 1992.

[12] L Xu , Z Chen , B Xu , H Chen , J Jiang . "Regression Testing for Web Applications Based on Slicing". 27th Int. Conf. on Computer Software and Applications. 2003.

[13] E Soechting , K Dobolyi , W Weimer. "Syntactic Regression Testing for Tree-Structured Output". 2009.

[14] N Alshahwan. "Automated Regression Testing of Web Applications". Project Report.2005.

[15] Q Zhongsheng. "Test Case Generation and Optimization for User Session-based Web Application Testing". Journal of Computer. 2010.

[16] A Tarhini , Z Ismail , N Mansour. "Regression Testing Web Applications". 2008.

[17] A Memon , I Banerjee , N Hashmi , A Nagarajan. "DART:A Framework for Regression Testing"Nightly/daily Build"of GUI Application".September 2003.

[18] C Mao ."Slicing Web Service-based Software".IEEE Int. Conf. on SOCA .2009.

[19] S Sampath, R C. Bryce,G Viswanath, A. G Koru, V Kandimalla. "Prioritizing User-session-based Test Cases for Web Applications Testing". 2008.

[20] Y Hernandez , T. M. King , J Pava . P. J. Clarke . "A Meta-Model to Support Regression Testing of Web Applications". 20th Int. Conf. on Software Engineering & Knowledge Engineering. 2008.

[21] Gagandeep , J Sengupta. "Automatic Generation of Regression Test Cases for Web Components using Domain Analysis and Modeling". International Journal of Computer Applications (0975 – 8887). Vol .11– 12. 2010.

[22] A. A. Andrews , J Offutt , R. T. Alexander . "Testing Web Applications by Modeling with FSMs". July 2005.

[23] H. Al Shaar , R Haraty ."Modeling and Automated Black box Regression Testing of Web Application". Theoretical and Applied IT Journal. 2005.

[24] D Jeffrey, N Gupta. "Test Case Prioritization Using Relevant Slices". 30th Annual Int. Conf. of Computer Software and Applications,Vol. 01. pp. 411–420. 2006.

[25] S Sprenkle, S Sampath, L Pollock, E Gibson, A Souter. "An Empirical Comparison of Test Suite Reduction Techniques for User-session-based Testing of Web Applications".2005.

[26] G Rothermel, R Untch, C Chu, M. J. Harrold. "Prioritizing Test Cases for Regression Testing".ACM SIGSOFT International symposium on Software testing and analysis, pp. 102 - 112. 2000.

[27] P Samuel , R Mall . "A Novel Test Case Design Technique Using Dynamic Slicing of UML Sequence Diagrams ".2008.

[28] S Sampath , S Sprenkle , E Gibson , L Pollock , A.S Greenwald . "Applying Concept Analysis to User-Session-Based Testing of Web Applications". IEEE Transactions on Software Engineering, vol. 33, pp. 643 – 658. 2007.

[29] F Ricca, P Tonnella."Analysis and testing of web applications", Int. Conf. on Software Engineering, pp. 25-34. 2001.

[30] L Xu, B Xu, J Jixiang. "Testing Web Applications Focusing on Their Specialties". ACM Notes on Software Engineering, pp. 10-16. 2005.

[31] A Tarhini, H Fouchal, N Mansour . "Regression Testing Web Services-based Applications", IEEE/ ACS Int. Conf. on Computer Systems and Applications. pp. 163-170. 2006.

[32] F Ricca , P Tonella . "Web Application Slicing". International Conference on Software Maintenance. 2001.

[33] B Yang, J Wu, C Liu , L Xo . "A Regression Testing Method for Composite Web Services". ICBECS, 23-25, pp. 1-4. April 2010.

[34] S Alsmadi, I Alsmadi, M Al-Kabi ."Generation of Test Cases From Websites User Sessions". 5th International Conference on IT. 2011.

[35] A Mansoor, S Zulfikar, A Bhutto."Analytical Survey on Automated Software Test Data Evaluation". 2010.

[36] P Tonella , F Ricca. "Statistical testing of web applications". Journal of Software Maintenance. 16(1-2):103–127. 2004.

[37] S Sampath. "Cost effective Techniques for user session based testing of web application" Doctoral Dissertation, University of Delaware Newark, DE, USA .2006.

[38] M. J. Harrold, R Gupta, and M. L. Soffa." A methodology for controlling the size of a test suite". 1993.

[39] S Sprenkle, E Gibson, L Pollock, S Sampath. "Automated replay and failure detection for web applications". Int. Conf. on Automated Software Engineering, pp. 253–262. 2005.

[40] Sahu, M.; Mohapatra, D.P. "Slicing Java Server Pages Application". Int. Conf. on Information Technology, pp. 252 – 253. 2008.

[41] De Lucia A ." Program Slicing: Methods and Applications". 2001.

[42] L Xu, B Xu, Z Chen, J Jiang, H Chen. "Regression Testing for Web Applications Based on Slicing". 27th Int. Conf. on Computer Software and Applications, pp. 652. 2003.

[43] T Reps, S Horwitz, M Sagiv, G Rosay." Speeding up Slicing".2nd ACM SIGSOFT symposium on software engineering Notes. Vol. 19. pp. 11-20. 1994.

[44] J Wu,B Xu, J Jiang. "Slicing web application based on Hyper Graph". International Conference on Cyberworlds. pp. 177 – 181. 2004.

[45] K.B.Galagher, J.R.Lyle ."Using Program Slicing in Software Maintence". August 1991.

[46] Z Qian ."User Session-Based Test Case Generation and Optimization Using Genetic Algorithm ". Journal of Software Engineering and Applications. Vol.3. pp 541-547. 2010.