Web Application Reliability Assessment using Error and Workload Data Obtained from Server Error and Access Logs

Laxmi Shanker Maurya
PhD (CSE) Research Scholar
Mewar University
Chittorgarh, Rajasthan, India

Anil Kumar Malviya
Associate Professor,
Department of Computer Science and Engineering
KNIT, Sultanpur, U.P, India

ABSTRACT

Software reliability assessment had been a major issue among the researchers in the recent past. Two important characteristics to assess software reliability are MTBF (Mean time between failure) and number of errors. In this paper, we capture and extract workload and failure data from server logs of a web application named College Excel. Classify the workload data based on four parameters. Calculating the cumulative values of four workload parameters and error for 30 days we plot graphs and verify the relationship between failure and different workload parameters. We also calculate the reliability of the College Excel web application using Nelson model (one of the most widely used input domain reliability model) and MTBF (mean time between failure) using failure and different workload units.

Keywords

Reliability, Web Application, Error, Workload

1. INTRODUCTION

Majority of the software systems developed these days are Web-based. As our reliance on Web-based applications continues to increase and the web systems supporting these applications become more complex, there is growing concern about the manner in which the Web-based systems/applications are created and their quality, integrity and reliability is assessed. The development of Web-based systems has generally been ad-hoc, resulting in poor quality and reliability. In the recent times, there have been many failures of Web applications due to a variety of problems and causes [3].

Web applications possess unique characteristics that make web testing and quality assurance different from the corresponding traditional techniques. Web applications can be characterized by the following aspects [1]. (i) Massive access of users. (ii) The difficulty of establishing the causes of the errors. (iii) The integration of diverse software elements for an application on the Internet. (iv) Diversity of frameworks to develop, to operate and to maintain a Web site.

Web-based systems are very sensitive to errors. Most work on web applications has been on making them more powerful, but relatively little has been done to ensure its quality. The most important quality attributes demanded by web-based systems are reliability, usability and security [2].

There are a number of views as to what software reliability is and how it should be quantified. Some people believe that this measure should be binary in nature so that an imperfect program would have zero reliability while a perfect one would have a reliability value of one. This view parallels that of program proving whereby the program is either correct or incorrect. Others, however, feel that software reliability should be defined as the relative frequency of the times that the program works as intended by the user [5].

The reliability of a software system is defined in [IEEE90] as the ability to perform the required functionality under stated conditions for specified period of time. Quality assurance and reliability assessment for web applications should focus on the prevention of web failures or the reduction of chances for such failures. Consequently, we contend that early reliability assessment is necessary for the reduction of testing efforts and for ensuring a level of operational reliability [4].

The rest of the paper is organized as follows: Section 2 analyzes some of the related works carried out by researchers. Section 3 describes the research methodology used in the proposed work. Section 4 presents our results and discussion. Section 5 deals with the conclusions and future work. Section 6 incorporates acknowledgements. Section 7 includes the references

2. RELATED WORKS

Marinho et al. [9] explore the feasibility to apply an architecture-based reliability analysis method in a real-life web application. Hangjung et al. [10] provide a basis for measuring reliability of an application system that is assembled using web services. Destercke et al. [11] propose a generic method to assess data reliability from a set of criteria using the theory of belief functions. Cherrier et al. [12] propose a solution to re-introduce coherence in the application, in order to keep the advantages of choreography. An overlay of logical check-points at the application layer defines links between the coherent states of a set of objects and triggers re-synchronization messages. Correcting statements are thus spread through the network, which enables fault recovery in Choreographies. Their work ends with a comparison between the checking cost and the reliability improvement. . Two fundamental types of software reliability models are: input domain reliability models (IDRMs) and time domain software reliability growth models (TDSRGMs) [3]. IDRMs can provide a snapshot of the web application's current reliability. For example, if total number of f failures is observed for n workload units, the estimated reliability R according to the Nelson Model [3], [6], one of the most widely used IDRMs, can be obtained as follows:
R= (n-f)/n=1-f/n=1-r.

Where r is the failure rate. When usage time t_i is available for each workload I, the summary reliability measure, mean-time-between-failures (MTBF), can be estimated as follows:
MTBF=1/f.

When the usage time t_i is not available, we can use the number of workload units as the rough time measure. In this case, MTBF=n/f.

3. RESEARCH METHODOLOGY

A web application named College Excel is used as an input in the research work. College Excel is comprehensive e-governance software that covers the entire gamut of campus activities in a college. The software is aimed at bringing in process refinement, efficiency, accountability and control in a college. College Excel is the product of BROADLLYNE. It is built using the Dot Net Technology from Microsoft. The software uses MS SQL Server as the Database. Apart from managing the administrative, financial, and academic functions of the institutions, College Excel provides a platform for sharing information between all the stakeholders, viz. Students, Parents, Teachers, Administrators, Management, Corporate and the Alumni.

MATLAB 6.1 is used for plotting various graphs shown in section 4 in the form of Fig. 1 to Fig.7. Nelson model [3] (an input domain reliability model) is used to assess the reliability and MTBF (mean time between failure) of the College Excel web application. Data were collected from the server logs using the content of error log and access log files.

4. RESULTS AND DISCUSSION

The results of the study are summarized in the form of figures ranging from Fig.1 to Fig.7 and in Table1.

Fig. 1 plots the number of errors over time under normal conditions. These daily errors vary considerably. However, the number of problems encountered per day is closely related to the actual usage intensity. The workload measurement results are plotted in Fig. 2 for daily megabytes transferred, in Fig. 3 for daily hits, and in Fig. 4 for daily sessions (top curve) and daily users (bottom curve). The workload distribution is highly uneven and varies from day-to-day, as represented by peaks and valleys in below mentioned workload plots, which conforms to previously observed traffic patterns [3], [7] and [8]. Among the four workload measures, daily megabytes and daily hits show larger variability in relative magnitudes than daily users or daily sessions. This result indicates that although the number of users or user sessions may remain relatively stable, some users may use the web application much more intensively than others, resulting in larger variations in detailed web application workload measurements over time.
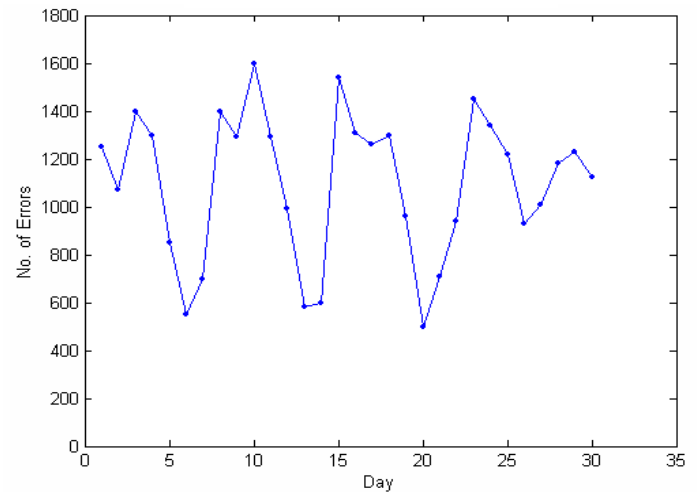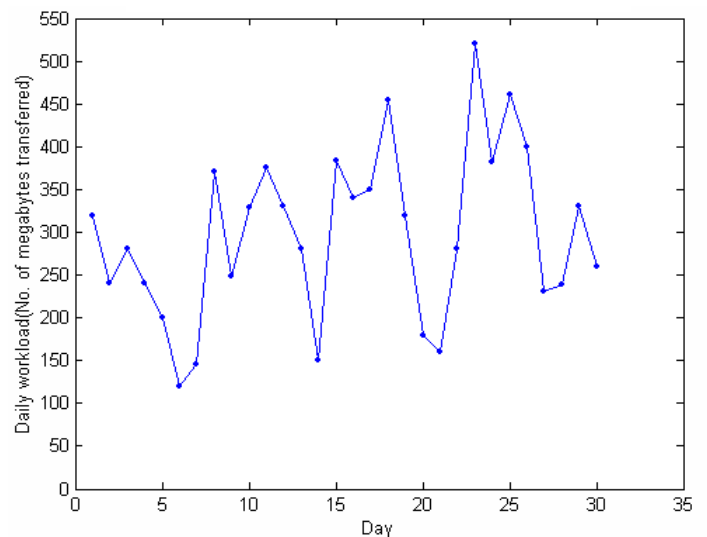

Fig.1.Error profile over time for College Excel

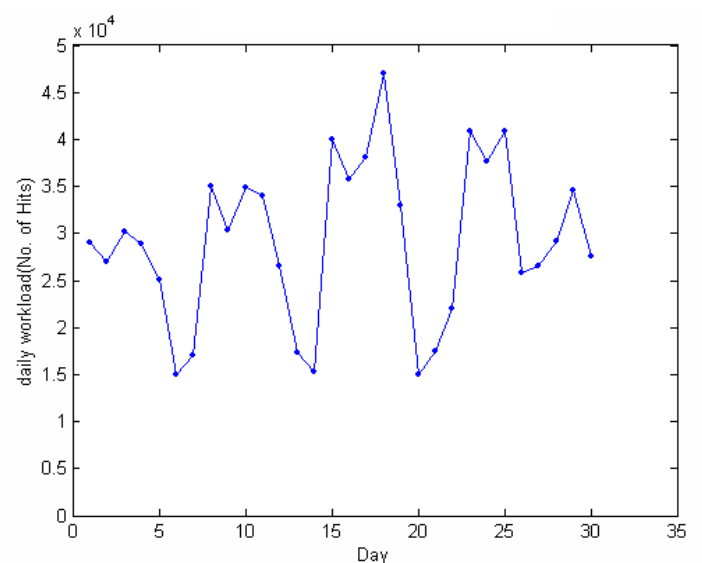Fig.2.Daily megabytes transferred over time for College Excel
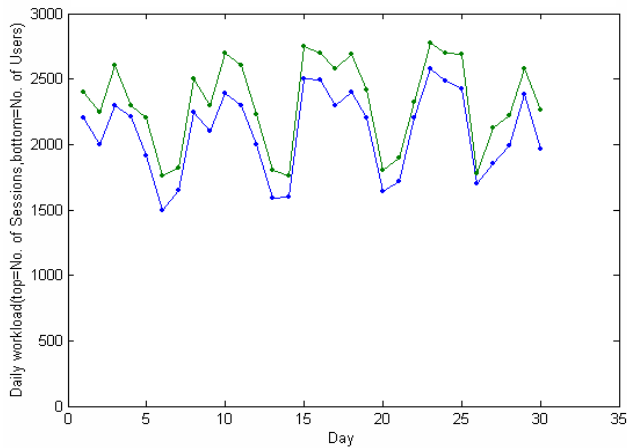

Fig.3.Daily hits over time for College Excel

Fig.4.Daily Seeions(top curve) or Users(bottom curve) over time for College Excel

Fig.7.Cumulative errors versus cum. users & sessions

The peaks and valleys in errors represented in Fig. 1 almost coincide with the peaks and valleys in workload, such as in Figs. 2, 3 and 4. This close relationship between workload measure and failure count are graphically examined as in Fig. 5, plotting cumulative errors versus cumulative megabytes transferred over the observation period. An essentially linear relation is detected between the two parameters. Similar results are obtained as shown in Fig. 6 by plotting cumulative errors versus cumulative hits and in Fig. 7 by plotting cumulative errors versus cumulative users (top) and cumulative sessions (bottom) respectively.
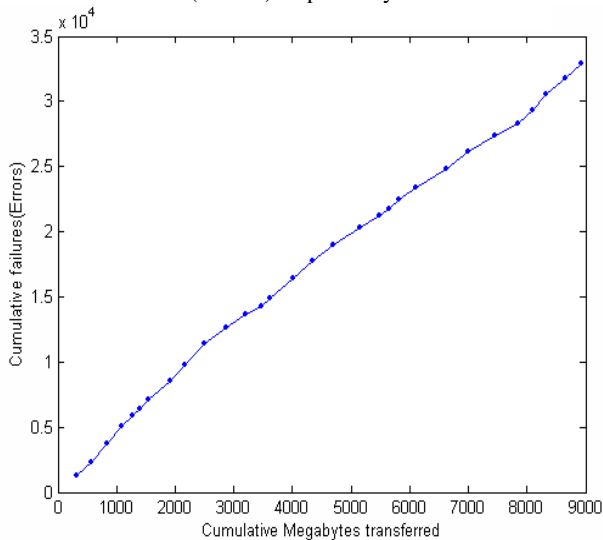
As per observation, since individual web failures are directly related with individual hits, megabytes transferred, no. of users and no. of sessions, we can use the Nelson Model described in section 2 to evaluate the overall web application reliability and MTBF using failures and workload units for the complete 30 days as shown in the Table1.

Table1: Calculated value of Reliability and MTBF for College Excel

Fig.5.Cumulative errors versus cumulative megebytes transferred
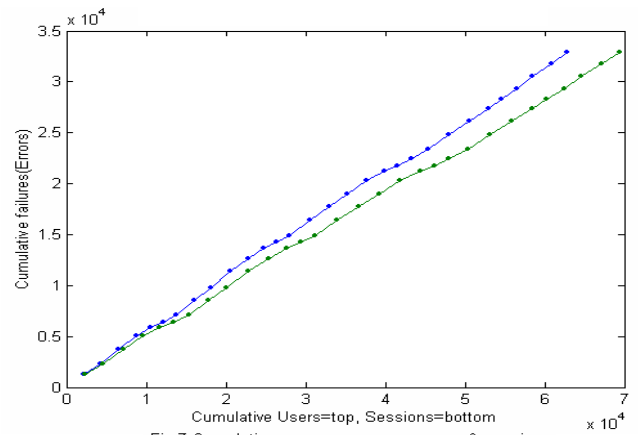

Fig.6.Cumulative errors versus cumulative hits

Sl. No.	Workload unit/Measure used	Reliability w.r.t. the measure selected(with explanation)	MTBF w.r.t. the measure selected(with explanation)
1.	No. of hits	0.963(or 96.3% of the individual web accesses/hits are successful)	26.68 hits(one error for every 26.68 hits)
2.	No. of kilobytes transferred	0.9964(or 99.64% of the kilobytes transferred are successful)	227.87 kilobytes(one error for every 227.87 kilobytes transferred)
3.	No. of users	0.477(or 47.7% of the users are successful)	1.912 users(one error for every 1.912 users)
4.	No. of sessions	0.527(or 52.7% of the sessions are successful)	2.115 sessions(one error for every 2.115 sessions)

5. CONCLUSIONS AND FUTURE WORKS

By analyzing the results of our study we conclude the following findings:-

- Reliability of a web application directly depends on the number of errors or failure count.
- Reliability of a web application not only depends on the errors or failure count but also depends on various workload measures such as number of hits, number of megabytes transferred, number of users and number of sessions. The linear relationships between errors and various workload measures verify the fact.
- More number of errors leads to less reliable web application.
- Increase in workload units increases number of errors which decreases reliability of the web application.

Some of the future dimensions in which our research may be extended are as follows:-

- We can collaborate with some industry to obtain latest server logs data.
- Automated tool may be developed to count error and workload data from the server logs.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] Leticia Davila-Nicanor, Pedro Mejia-Alvarez. "Reliability Evaluation of Web-based Software Applications". IEEE, 2005.

[2] Leticia Davila-Nicanor, Pedro Mejia-Alvarez. "Reliability Improvement of Web-Based Software Applications". IEEE, 2004.

[3] Jeff Tian, Sunita Rudraraju, Zhao Li. "Evaluating Web Software Reliability Based on Workload and Failure Data Extracted from Server Logs". IEEE, 2004.

[4] V. S. Alagar, O. Ormandjieva. "Reliability Assessment of Web Applications". IEEE, 2002.

[5] Amrit L. Goel. "Software Reliability Models: Assumptions, Limitations, and Applicability". IEEE, 1985.

[6] E. Nelson. "Estimating Software Reliability from Test Data". Microelectronics and Reliability, Vol. 17, No. 1, PP 67-73, 1978.

[7] M. F. Arlitt, C. L. Williamson. "Internet Web Servers: Workload Characterization and Performance Implications". IEEE/ACM Trans. Networking, Vol. 5. No. 5, PP 631-645, Oct. 1997.

[8] J. E. Pitkow. "Summary of WWW Characterizations". World Wide Web, Vol. 2, Nos. 1-2, PP 3-13, 1999.

[9] E. H. Marinho, A. A. Mendonca, G. N. Rodrigues, V. Alves, R. Bonifacio. "Exploring Architecture-Based Reliability Analysis of Current Multi-layered Web Applications". IEEE Conference Publications, PP 51-60, 2011.

[10] Zo Hangjung, D. L. Nazareth, H. K. Jain. "Measuring Reliability of Applications Composed of Web Services". IEEE Conference Publications, PP 278-282, 2007.

[11] S. Destercke, P. Buche, B. Charnomordic. "Evaluating Data Reliability: An Evidential Answer with Application to a Web-Enabled Data Warehouse". IEEE Journals & Magazines, Volume 25, Issue 1, PP 92- 105, 2013.

[12] S. Cherrier, Y. M. Ghamri-Doudane, S. Lohier, G. Roussel. "Fault-recovery and coherence in Internet of Things choreographies". IEEE Conference Publications, PP 532-537, 2014.