International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 4 | ISSN : 2456-3307

Testing for Sofware Quality in Search Engines

Ganga Bhavani K , Bhuvaneswari K
Department of Computer Applications, Rayalaseema Institute of Information and Management Science, Tirupathi, India

ABSTRACT

Metamorphic testing could be a testing technique which will be wont to verify the practical correctness of software within the absence of an ideal oracle. This paper extends metamorphic testing into a user-oriented approach to software verification, validation, and quality assessment, and conducts giant scale empirical studies with four major net search engines: Google, Bing, Chinese Bing, and Baidu. These search engines are terribly tough to check and assess victimisation typical approaches because of the dearth of an objective and usually recognized oracle. The results area unit helpful for each program developers and users, and demonstrate that our approach will effectively alleviate the oracle downside and challenges close a scarcity of specifications once verificatory, validating, and evaluating giant and complex software package systems.

Keywords: Software quality, verification, validation, quality assessment, oracle problem, lack of system specification, metamorphic testing, user-orientedtestingsearchengine

I. INTRODUCTION

The goal of software engineering practices is to develop top quality software. it's so crucial to develop analysis strategies for varied kinds of software qualities. Testing may be a wide used approach for evaluating software package qualities and serving to developers to search out and take away software faults. the bulk of software package testing techniques assume the supply of associate oracle, a mechanism against that testers will verify the correctness of the outcomes of action at law executions. In some things, however, associate oracle isn't accessible or is obtainable however is simply too pricy to be used – a situationknown because the oracle drawback, a elementary challenge for software testing.

A metamorphic testing (MT) technique has been developed to alleviate the oracle drawback. in contrast to standard testing strategies, MT doesn't concentrate on the verification of every individual output, however instead checks the relationships among the inputs and outputs of multiple executions of the program below take a look at. Such relationships square measure referred to as metamorphic relations (MRs), and square measure necessary properties of the meant program's functionality: If an mr violation is detected, then a fault is alleged to be unconcealed. MT has been accustomed check the useful correctness of varied applications and has additionally been applied to program proving and debugging. Its effectiveness has additionally been rigorously studied. The present analysis extends metamorphic testing into a quantitative approach for software package quality assessment, which has, however isn't restricted to, the verification and validation of software package correctness. we tend to applied our approach to alleviate the oracle drawback for the testing and quality assessment of (Web) search engines. Search engines are software systems designed to look for info on the globe Wide net, and are the most

interface through which individuals discover info on the net; net looking out is one amongst the foremost common functionalities of the net, second solely to email. As additional and additional services and knowledge square measure being created accessible on the net, search engines have become increasingly vital. In today's extremely competitive search market, it's imperative that search engines offer the specified result in line with the queries entered. It is, however, very troublesome to assess some key qualities of those search engines. as an example, as a result of the sheer volume of information on the net, it's terribly troublesome to verify or validate the correctness of the software package systems or to judge the accuracy and completeness of the search results. Also, given the apparent subjectiveness of various judges, objective assessment of search result relevancy and ranking quality is extremely troublesome. This paper addresses the on top of issues exploitation MT, and consequentially demonstrates new dimensions of the utility of MT. A series of empirical studies are conducted to check the software qualities of four major search engines, namely, Google (www.google.com), Bing (www.bing.com), Chinese Bing (Bing for Chinese users, www.bing.com.cn), and Baidu (www.baidu.com). in line with NetMarketshare , Google, Baidu, and Bing are the three most popular search engines in the world.

II. ALGORITHM

To apply MT to the automatic quality assessment of search engines, without the need for an oracle or human assessor, two groups of MRs were used: The "No Missing Web Page" group assesses the search engines' capability in retrieving appropriate Web pages to meet the users' needs; and the "Consistent Ranking" group assesses the ranking quality of the search results.

Metamorphic Relation: MPSite

MPSite belongs to the "No Missing Web Page" group of MRs, which assess the search engine's Web page retrieval capability. MPSite is focused on the search engine's reliability when retrieving Web pages that contain an exact word or phrase. It therefore assesses the keywordbased search feature. MPSite is described as follows:

Let A be a source query for which the search engine returns a non-empty list of results (called the source response), namely, (p1; p2; : : : ; pn), where 0 < n and pi is a Web page from domain di, $1 \leq i \leq n$. To enhance accuracy and validity of our approach, in MPSite we only consider situations where $0 < n \leq 20$ so that we can avoid the inaccuracy associated with large result sets (such as a large list being truncated by the search engine to improve response time). For the source response (p1; p2; : : : ; pn), n follow-up queries are constructed as follows: The ith follow-up query Bi ($1 \leq i \leq n$) is constructed in such a way that Bi is identical to A except that Bi includes an additional criterion which requires that all results be retrieved from domain di. Let FRi (a follow-up response) be the list of Web pages returned by the search engine for query Bi. The metamorphic relation MPSite requires that pi 2 Fri (note that there is no requirement on the ranking of pi in FRi). For example, let us test Google by issuing the following source query: "side effect of antibiotics in babies" where the quotation marks are part of the query.

Metamorphic Relation: MPTitle

For many search engines including those investigated in the present paper, if the words are not enclosed by double quotation marks, synonyms will be employed automatically. For instance, Google specifies that "Google employs synonyms automatically, so that it finds pages that mention, for example, childcare for the query [child care] (with a space), or California history for the query [ca history]." Synonyms are employed because the search engines attempt to

return Web pages that best meet users' information needs. In other words, the search engines attempt to imitate the behavior of a human operator, to which end, correct understanding of the Web pages and of the user intent are key. To test a search engine's information retrieval capability in situations where synonyms may be used for semantic search, a good strategy is to construct a test query q that best characterizes a target Web page p (the words in q may or may not directly appear in p). Furthermore, p must have been indexed in the search engine's database. The search engine can be tested on q. If p is not retrieved, then the user's perception of the search quality will be poor. A research question is: How can such q's and p's be identified in a fully automatic fashion for search engine assessment. The metamorphic relation MPTitle is designed to meet this challenge.

III. CONCLUSION

Metamorphic testing (MT) was initially proposed as a verification technique, where metamorphic relations (MRs) were identified by referring to the target algorithm to be implemented. In this paper, we have demonstrated the feasibility of MT being a unified framework for software verification, validation, and quality assessment. We conducted a study on search engines, where we identified MRs from the users' perspective without referring to the target algorithms or system specifications. More generally, this approach allows users to recognize whether or not a system is appropriate for their specific needs in the absence of complete software documentation, which is often the case with Web services, poorly evolved software, and open source software. We have applied our approach to assess several key software qualities of search engines under different operational profiles in the absence of an objective and generally recognized oracle. All ANOVA analyses returned statistically significant results with large effect size (_2 p) values. Most multiple-comparison results also had a statistical and practical significance (with large Cohen's d values in most situations), indicating that our approach is effective. The empirical results demonstrate that our approach is useful for both developers and users. Firstly, our approach can effectively detect various kinds of failures. Secondly, we found that the operational profiles have a significant impact on the quality of search. For a given search engine, its quality of search can be significantly different for different query languages, different types of query words, and different domains being searched. This finding provides a hint for the developers to identify the strength and weakness of their systems, and is also useful for the users to choose a suitable search engine or to better construct their queries.

IV. REFERENCES

[1]. A. Klein, F. Ishikawa, and S. Honiden. Efficient heuristic approach with improved time complexity for qos-aware service composition. In ICWS, pages 436–443. IEEE, 2011.

[2]. Tripathy, M.Khan, M.R.Patra, H.Fatima, P.Swain, " Dynamic web service composition with QoS clustering" IEEE , International Conference on Web services, 2014.

[3]. J. Hamilton, "The cost of latency," Perspectives Blog, 2009.

[4]. D. Bhagwat, K. Eshghi, D. D. Long, and M. Lillibridge, "Extreme binning: Scalable, parallel deduplication for chunk-based file backup," in Modeling, Analysis & Simulation of Computer and Telecommunication Systems, 2009. MASCOTS'09. IEEE International Symposium on. IEEE, 2009, pp. 1–9.

[5]. W. Xia, H. Jiang, D. Feng, and Y. Hua, "Silo: a similarity-locality based near-exact deduplication scheme with low ram overhead and high throughput," in Proceedings of the 2011 USENIX conference on USENIX annual technical conference. USENIX Association, 2011, pp. 26–28.

[6]. Y. Fu, H. Jiang, and N. Xiao, "A scalable inline cluster deduplication framework for big data protection," in Middleware 2012. Springer, 2012, pp. 354–373.

[7]. D. Teodosiu, N. Bjorner, Y. Gurevich, M. Manasse, and J. Porkka