

Survey of Automated Recommender System for Web Applications

Vinutha K.N¹, K.S. Sampada²

^{1,2} Department of Computer Science, VTU, Karnataka, India

Available online at: www.ijcseonline.org

Abstract— Online shopping new way of business in present days based on the previous surfing and purchasing products are recommended to the users. The existing method of recommending the product has to undergo several processes or functionalities and these processes or functionalities are manually tested for the accuracy. The manual testing method requires lot of time and money and other resources. To overcome the problem this paper proposes a Automation Testing for the recommender system, with Feature Vector Algorithm and perform a automation on each modules of the Feature Vector algorithm and also checks the Cross-Browser compatibility across the browser and also collecting the online reviews from by using Web Crawling Technique.

Keywords—Feature Vector, Recommender System, Cross-Browser Compatibility, Web Crawling Technique

I. INTRODUCTION

In online shopping consumers can directly buy the goods or the consumer can take the services from the seller in the real-time without any intermediate service over the internet. An e-shop, e-store is the Physical analogy of buying a product over the internet. In this Scenario recommender system plays an important role. Recommender System is the system in which provide recommendation to the user, and also Recommender systems are the information filtering that seeks to predict the rating or preference that user would give to the product or item. Now a day's recommender systems are most common and popular, and can be applied in various applications such as movies, news, books, twitter etc. Manual Testing in Recommender System, if new functionality is added the tester has to test from the beginning and it's highly impossible if the application is too large and it requires lot of resources like man power, time and money and system resources. And its limitations are as follows:

- There are certain actions that are difficult to do manually. For example Low level interface regression testing. This kind of testing is extremely difficult to perform manually, and, as a result, is prone to mistakes and oversight when done by hand.
- Manual testing can be repetitive and boring – no one wants to keep filling out the same forms time after time. As a result, many testers have a hard time staying engaged in this process, and errors are more likely to occur.
- If there is any change to the software, rerun the tests again by hand. This is valuable time lost.

But With automated tests, if we add new functionality to the program, you can rerun all of the required tests instantly

the tests are already set up. To addresses the issues in previous methods, In this Paper we introduces a method called the automation testing for the recommender system to provide the best product to the user in an effective way by using a series of algorithms(Feature Vector) like web crawler, data cleaning, tokenization, frequency computation, feature vector computation and ranking of products and also perform automation on each of these modules, and automatically generates the test cases for the web application 2.0 to provide best product to the users, based on the user queries.

II. RELATED WORK

Greg Linden[1]: Recommender algorithms are best for the e-commerce, e-shops, Internet shops etc. it uses the customers interests as input to the algorithm to generate the recommended items. In this survey addresses the recommender system problem and personalizes these problems by using common approaches are collaborative filtering, cluster models, and search-based methods. And compare these methods with the proposed algorithm, like item-to-item collaborative filtering. Rather than matching the user to similar customers, the algorithm finds items similar to each of the user's purchases and ratings, aggregates those items, and then recommends the most popular or correlated items. This computation is very quick, depending only on the number of items the user purchased or rated. Here remark is customer data is volatile and new customers have only limited information, based on only a few purchases or product ratings.

Emmanouil Vozalis[2]: This study gives the brief review of different recommender systems algorithms, and it presents the basic recommender systems challenges and problems. It will give the overview of the association rules, memory based, model-based and hybrid recommendation

algorithms. The recommender system faces the challenges like Quality of Recommendations, Sparsity, Scalability, Loss of Neighbour Transitivity, Synonymy, First Rater Problem: , Unusual User Problem .

Ali Mesbah[3]: This survey proposes a method for testing AJAX application automatically by using crawler to infer the flow graph for the user interface states and also identify the AJAX specific faults that occurs in user interface states and detects the faults in DOM tree invariants that serves as oracle to detect faults after this crawling is called crawljax. And also it provides the Fault revealing capabilities, scalability of the approach, generate test suite covering path obtained during the crawling. And the hardest path is deciding which application specific invariant to adopt and Manual effort required to use the approach. Introduces a tool ATUSA, it is an invariant checking components, a plugin mechanism to add application specific state validators, and generation of a test suite covering the paths obtained during crawling. This technique does not talk about cross-browser compatibility and black-box testing.

Matthias Schur[4]: This study presents a “procrawl” is a fully automatic tool, used to mine the behaviour models from the enterprise web applications for the sake of system testing and maintained, it observes the behaviour of the application through the user interface, it executes the tests and generates to explore the unobserved behaviour, “procrawl” produces the model it can used for effective model based regression testing. And the remark is that it will not infer the process variance based on the input data and it not investigates the further use cases, like program comprehension.

Valentin Dallmeier[5]: Quality assurance is the one of the challenge in the web application due to the large number and variance of involved components.web application 2.0 based on JavaScript , having a new challenge for testing, as a simple crawling through links covers only a small part of the functionality if any one of the element path is changed the test is broken because of lack of proper refactoring tools, small changes in the user interface often break the large part of the test suite which causes considerable maintains costs. So that most of the companies go for the manual testing or do not test at all. To address this problem web mate came into picture, it will automatically explore and navigate through web 2.0 application. It analyses the web application under the test it identifies all functionally different states, and then able to navigate each of these states at the user request’s all services delivered through in the combination of CSS , HTML and JavaScript. One remark in this approach is that it is not compatible with block-box testing.

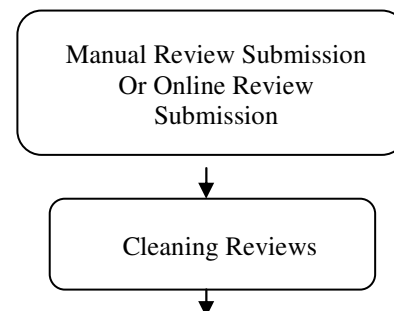
Corresponding Author: Prof. *K.S.Sampada, k.s.sampada@gmail.com*
Department of Computer Science, University of VTU, India

Shauvik Roy Choudhary[6]: The richness and diversity of the web platform’s cross-browser incompatibilities (XBIs) are the prominent issue for the web application developers to address this issue introduces a tool X-PERT, it can identify accurately and effectively and it can be applied to different technique to different aspects of the web application. By using different types of XBIs, the X-PERT integrates existing technique such as: Render given web application in two different browsers interest and extract the features. And Compare the features collected across the two browsers and if they differ, decide whether the difference is attributable to XBIs. The remark is that identify the technique that automatically eliminate the XBIs by using X-PERT through browser specific automated web page repairs. And also it does not identify cross-platform compatibility also.

A. Kumar [7]: Content based recommendation tries to recommend web sites similar to those web sites the user has liked, whereas collaborative recommendation tries to find some users who share similar tastes with the given user and recommends web sites they like to that user. Based on web usage data in adoptive association rule based web mining the association rules were applied to personalization. The technique utilize apriori algorithm to generate association rules. Even this method has some disadvantages. To overcome those disadvantages, This study gives the new algorithm for web recommendation system known as an effective Fuzzy Association Rule Mining Algorithm (FARM). This proposed Fuzzy ARM algorithm for association rule mining in web recommendation system results in better quality and performance of the recommender system.

III. AN EFFECTIVE APPROACH

In This Paper proposes an Automation Testing for the recommender system by using a automation tool Selenium Framework and Introduces a new algorithm called Feature Vector algorithm and automate the each modules of this algorithm by using automation tool Selenium Framework. The each modules of algorithm is as shown in the Fig.1. And discusses the each module in this section.



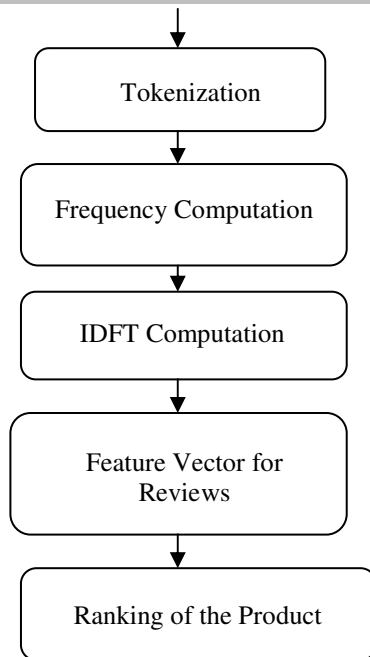


Fig.1. Algorithms steps involved in Manual Methodology.

A. Manual Review Submission

Submit the URL of flip cart select a Product type(Mobile, Camera etc) and the Product (Samsung Galaxy, Nokia Lumina etc) a Data Mining Technique known as web crawling will be applied in order to extract the reviews from HTML rendered DOM tree, the reviews are stored in the format of (Review id, Review Description, Product id, Product Type). Online Review Submission can be executed only if internet is available, in Manual Review Submission provide a Provision like submit users own reviews by selecting a product type and a product name.

B. Cleaning of Reviews

The data which is available in the form of reviews will be cleaned by using a series of standard stop words as well as customize stop words, stop words are words which do not have any meaning, this module also referred to as pre-processing.

C. Tokenization

Each of the cleaned reviews are converted into multiple tokens i.e., if a review is represented by 'R' then 'R' can be defined as a set $\{ T_1, T_2, \dots, T_n \}$ where 'n' is the total number of tokens for a review 'R', 'T_i' is the ith token.

D. Frequency Computation

The redundancy of the tokens in a review is removed by removing duplicates with the introduction of Frequency Computation. Where Frequency is number of times a tokens appears in a review.

E. IDFT (Inverse Document Frequency of Tokens) Computation

For each of the tokens find out the associated reviews and count the number of words present in the review and then compute IDFT given by $\log(N/f)$.

Where, N- Number of words in a review.
f- Frequency.

F. Feature Vector for Reviews

For each of the tokens Feature Vector (fv) is computed which is defined by
$$fv = IDFT * f$$

G. Ranking the Product:

The user will enter a search query, the search query is again divided into tokens for each of user entered tokens 'fv' is obtained across reviews and for all the products, if the $fv=0$ then replaces by frequency, like this process is repeated for all the tokens entered by the user and products are ranked based on descending order of 'fv'.

In order to automate the above discussed modules shown in Fig.1, introduces an automation methodology with the selenium framework to provide the best and rank the product to the users in an effective and automated way.

And the steps is as follows.

- Sequences of Test Cases are read for multiple screens. Each screen test cases are maintained in separate spread sheets.
- The List of Test Cases is executed and the complete algorithm is tested.
- If the test case is successful the results will have PASS for specific test case.
- If the test case fails the reason for failure is also captured.
- Web Driver Technology is used to automate the Web Application by sending the commands from the JVM to the browser.
- Apache POI is used to read and write the Test Results.

IV. CONCLUSION

In Manual Testing of recommender system requires lot of resources, and it needs retesting when new functionality is added to the testing. In order to overcome this problem introduced an effective approach, automation testing for the

recommender system to provide the best product to the users in an effective way and also it checks the cross browser compatibility, and automatically generates the test cases.

ACKNOWLEDGMENT

I would like to express my appreciation to Prof .K.S. Sampada for her valuable guidance during the term of my candidature. And I would like to thank to my college, Head of the Department, teaching and non-teaching staffs and great thanks to my mother and my grandparents for their great support.

REFERENCES

- [1] Greg Linden, Brent Smith, and Jeremy York Amazon.com Recommendations Item-to-Item Collaborative Filtering • Amazon.com JANUARY • FEBRUARY 2003 Published by the IEEE Computer Society.
- [2] Emmanouil Vozalis, Konstantinos G. Margaritis, Analysis of Recommender Systems' Algorithms, 2001.
- [3] A. Mesban and A van Deursen, "Invariant -based automatic testing of AJAX user interface." In ICSF'09: proceedings of the 2009 IEEE,31st International Conference on software Engineering. Washinton. DC, USA: IEEE Computer Society.2009,pp. 221-220.
- [4] M. Schur, A. Roth, and A. Zeller, "Mining behavior models from enterprise web applications," Sep. 2013, pp. 442-432.
- [5] V. Dallmeier, M. Burger, T. Orth, and A. Zeller, "Webmate: Generating test cases for web 2.0," in Software Quality. Increasing Value in Software and Systems Development. Springer, Jan. 2013, pp. 55-69.
- [6] S. R. Choudhary, M. R. Prasad, and A. Orso, "X-pert: accurate identification of cross-browser issues in web applications," in ICSE, 2013, pp. 702-711. [6] BMWI, "Exist," www.exist.de.
- [7] A. Kumar," Collaborative Web Recommendation Systems Based on an Effective Fuzzy Association Rule Mining Algorithm (FARM)" Indian Journal of Computer Science and Engineering Vol 1 No 3 184-191.

AUTHOR'S PROFILE

Prof.K.S.Sampada
Department of Computer Science, RNS Institute of Technology,
Bangalore, India

Vinutha K.N, M.tech
Department of Computer Science, RNS Institute of Technology,
Bangalore, India