

# A Comprehensive Review on Regression Test Case Prioritization Techniques for Web Services

Muhammad Hasnain<sup>1</sup>, Imran Ghani<sup>1</sup>, Muhammad Fermi Pasha<sup>1</sup>, Chern Hong Lim<sup>1</sup>, and Seung Ryul Jeong<sup>1\*</sup>

<sup>1</sup>School of Information Technology, Monash University Malaysia, Malaysia

[e-mail: muhammad.malik1; muhammad.fermipasha; Lim.ChernHong@monash.edu]

<sup>1</sup>Department of Mathematics and Computer Sciences, Indiana University of Pennsylvania, U.S.A.

[e-mail: imransaieen@gmail.com]

<sup>1</sup>Graduate School of Business Information Technology, Kookmin University, South Korea

[e-mail: srjeong@kookmin.ac.kr]

\*Corresponding author: Seung Ryul Jeong

*Received December 2, 2019; revised January 6, 2020; accepted January 30, 2020;  
published May 31, 2020*

---

## Abstract

Test Case Prioritization (TCP) involves the rearrangement of test cases on a prioritized basis for various services. This research work focuses on TCP in web services, as it has been a growing challenge for researchers. Web services continuously evolve and hence require reforming and re-execution of test cases to ensure the accurate working of web services. This study aims to investigate gaps, issues, and existing solutions related to test case prioritization. This study examines research publications within popular selected databases. We perform a meticulous screening of research publications and selected 65 papers through which to answer the proposed research questions. The results show that criteria-based test case prioritization techniques are reported mainly in 41 primary studies. Test case prioritization models, frameworks, and related algorithms are also reported in primary studies. In addition, there are eight issues related to TCP techniques. Among these eight issues, optimization and high effectiveness are most discussed within primary studies. This systematic review has identified that a significant proportion of primary studies are not involved in the use of statistical methods in measuring or comparing the effectiveness of TCP techniques. However, a large number of primary studies use 'Average Percentage of Faults Detected' (APFD) or extended APFD metrics to compute the performance of techniques for web services.

---

**Keywords:** Systematic Literature Review, Web Services, Regression Testing, Test Case Prioritization, Validation Metrics

## 1. Introduction

**R**egression Testing helps to determine the correctness of a web service, following functional and non-functional changes. A system integrated with the web services undergoes evolution and maintenance. It is due to the addition of a new function or development in the existing ones. Any error in a web service could impact the entire system. Regression testing ensures that changes do not have a significant adverse effect on the correctness of web services [1].

Regression testing is costly in terms of time and money because half of the cost accounts for software maintenance. Rerunning previously executed test case induces a high cost [2]. Regression testing can be performed in a systematic way to substitute the rerunning of all test cases by selecting test cases that cover modifications made in a system. In other words, a set of test cases are chosen, which helps to detect maximum faults within a web service. In addition to cost, the time factor also needs to be considered when executing test cases. These factors are significantly considered for the selection and reordering of test cases, as the size of the regression test suit is substantial usually [3]. Traditional industrial applications are written in the orchestration languages like 'Web Services-Business Process Execution Language' (WS-BPEL) and link with the web services via 'Web Services Descriptive Language' (WSDL) and XPath. Any error in these artifacts results in the extraction of wrong information from messages. Retesting of all services to reveal the incorrect information is constrained due to a shortage of resources [4]. To increase the efficiency of regression testing, 'regression test selection' (RTS) or 'test case prioritization' (TCP) is performed [5].

In the literature, there are several definitions of TCP. We have found the following two definitions to be more comprehensive because they both consider the importance of the scheduling (time) of test cases for effectiveness, cost, and performance goals.

*"Test case prioritization, i.e. scheduling test case executions in an order that attempts to increase their effectiveness at meeting some desirable properties, is a powerful approach to maximise the value of a test suite and to minimise the testing cost"* [6].

*"One approach to selecting test cases is to schedule the test cases according to some criterion to satisfy a performance goal; scheduling test cases in this manner is known as test case prioritization"* [7].

These two definitions of TCP discuss test cases regarding TCP execution and performance goals. The commonality between these two is the execution of test cases. The first definition focuses more on the sequence of executing test cases and their advantages. The second one focuses on the criteria for running test cases for the satisfaction of performance goals. Time and cost factors determine the effectiveness of a TCP technique, as revealed in the literature [8][9]. Walcott et al. mentioned that an industrial product with approximately 20k lines of code required seven weeks to test in its entirety. Seven weeks for regression testing is relatively expensive in terms of time [10]. To overcome this long processing time, Garg and Datta claimed that the parallel execution of the entire test suit required less than three hours [11]. A similar claim was verified by Liu et al. [12]. They found that it took 39.2 minutes to finish the parallel execution of all test cases.

Although TCP is time-costly in some cases, it is still advantageous because it reorders test cases within a limited time to ensure the earlier detection of faults [13]. TCP is used to automate the selection of a minimum number of test cases from the total numbers of test cases. The intention behind this selection is to uncover the maximum number of faults within the components of web services. Prioritized test cases are executed to prove that the selected test cases provide sufficient satisfaction for software functionalities in reduced time [14].

Catal and Mishra [15] examined regression testing techniques, which were widely used in primary studies published between 2001 and 2011. However, they did not include a comprehensive comparison of TCP techniques. Moreover, they did not include a discussion on TCP approaches regarding web services. Until recently, there is no systematic literature review (SLR) that identifies current issues regarding TCP techniques for web services. There is a growing call for a comprehensive SLR in order to provide an analysis of TCP techniques, models, frameworks, and algorithms as a solution to the emerged issues and challenges related to web services TCP techniques. Our efforts to provide a comprehensive SRL will fill this void, including all papers published up until 2017.

## 2. Background

To present this SLR's context, an overview of web services and TCP is provided in the subsequent subsections.

### 2.1 Web Services

A large number of web applications have procedures for WSDL [16], specifying messages and functional interface parameters. WSDL can be applied to XML documents for the representation of messages. A formal XML notation describes a web service. A WSDL document augments the interface information. Furthermore, external services observe XML documents. Farrag et al. have stated that WSDL has been mostly used in combination with XML schema and 'Simple Object Access Protocol' (SOAP) [17]. The combined services and schema have ensured communication over the internet.

In addition to the above-discussed procedures, web services have some standards. In literature, several definitions for the web standards of web services have been presented. Two of these definitions are given as follows.

*"The WS-standards define how middleware aspects (security, reliability, transactions, etc.) can be realized through web services"* [18].

*"The most commonly used standards and protocols include, but are not necessarily limited to, the Extensible Markup Language (XML), Simple Object Access Protocol (SOAP), the Web Services Definition Language (WSDL) and Universal Discovery Description and Integration (UDDI)"* [19].

The first definition of web standards, as presented by Simon et al. [18], focuses on various aspects of the middleware of web services, while the second definition of web standards only enumerates on many web standards that are applied to communicate transactions over the internet [19].

A service description contains all information required to interact with the web services, including location, transport protocol, and message format [20]. A web service is modular, self-describing, and self-contained. Above, the authors have described web services and web standards, mainly used for communication over the internet. Communication over the internet is increasingly being used, resulting in the evolution of web services. Alone, the WSDL specification of web services cannot verify the retesting of web services because web service providers have source codes regarding dynamic web services [21].

### 2.2 Test Case Prioritization (TCP)

TCP problem was defined in the earlier research works by Elbaum et al. [22][23], where they stated  $T$  as a test suit;  $PT$  as permutations of  $T$  and  $f$  represented a function from  $PT$  to real

numbers. In this definition,  $PT$  is taken as a set of possible orders, and  $f$  is applied to any such order to achieve an award value for the order.

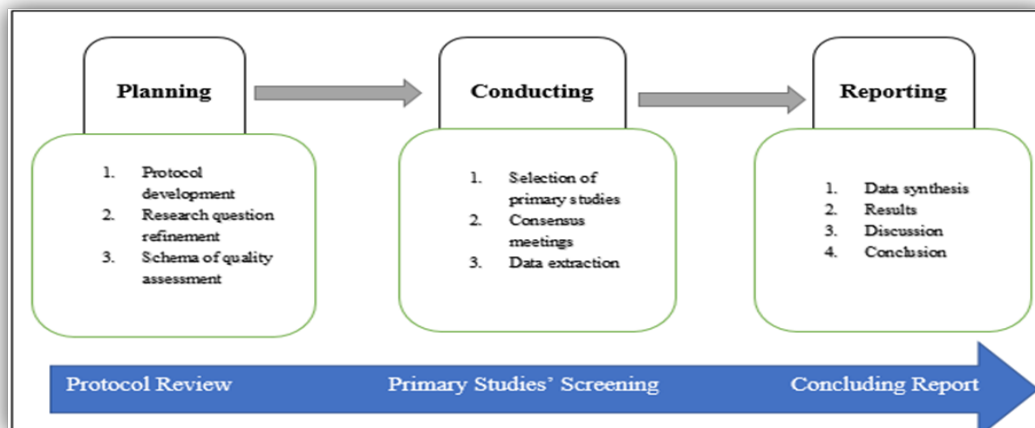
Problem definition stated in the works, as mentioned earlier, is similar to other works [24][25]. TCP problem in the proposed approaches is the same to date because researchers have attempted to resolve this problem in terms of time, and the cost of ordering test cases.

A hybrid TCP technique is not necessarily composed of two techniques of the same criteria-based group of techniques. For example, a hybrid TCP technique may involve a combination of users' requirements priority scores and users' session information. On many occasions, researchers have incorporated one TCP technique into another TCP technique to create a hybrid TCP technique. Marchetto et al. [26] proposed a hybrid TCP technique by including low code coverage and high requirement coverage information. The proposed hybrid TCP technique has been generalized to all size projects [27].

### 3. The Systematic Literature Review Method

Researchers use SLRs as an instrument for identifying, evaluating, and interpreting existing research on specific research questions and topic areas [28]. Contributing studies in SLR are considered primary studies. A well-defined and explicit process is required to conduct an SLR. For the proposal and execution of this SLR, this paper's authors have leveraged the recommended guidelines by Kitchenham and Charters [28], and the following adjustment recommendation made by Kitchenham et al. [29].

**Fig. 1** depicts three phases, and the activities performed in each phase. This three-phase model of SLRs is used by Brereton et al. [30] and Kazmi et al. [31] in their studies.



**Fig. 1.** The Phases of a Systematic Literature Review

#### 3.1 Research Questions

It is essential, in all types of systematic reviews, that research questions should be specified [32]. Research questions are used to build the search strings and determine the accurate extraction of data from selected primary studies. Research questions become part of the protocol and do not be changed after the protocol is accepted [30]. Questions to be addressed in this research include the following:

**RQ 1.** What approaches are proposed as a means to prioritize web services test cases?

**RQ 2:** How have the proposed web services TCP techniques been validated?

**RQ 3:** Are there statistical methods used in web services TCP techniques?

**RQ 4.** What issues related to TCP exist in regression testing, with respect to web services?

Next, a set of keywords are established to find relevant primary studies on the research topic. Finding suitable primary studies has supported the authors in seeking out answers to the above-given research questions.

### 3.2 Identification of Appropriate Search Keywords

Searching the string with the use of keywords is an essential part of SLR because there is a need to find relevant studies. Thus, the author constructed simple search strings from the topic of interest and research questions [33]. Alternative and synonymous keywords were also used. Both Boolean 'AND' and 'OR' operators were used for joining and using alternative keywords, respectively. Simple search strings need less effort on refinement. The search strings have been used in different repositories, as detailed in the following **Table 1**.

**Table 1.** Search strings/keywords

Sr. No.	Search Strings/Keywords
1.	Test Case Prioritization OR regression testing
2.	Web services test case prioritization
3.	Regression testing OR test case prioritization AND web services
4.	Web services incremental testing
5.	Web services test case prioritization AND Evaluation
6.	Web services AND test case prioritization AND Statistical methods

### 3.3 Primary Studies' Inclusion and Exclusion Criteria

Primary studies within regression testing, test case prioritization, and web validation have been examined, depending on the proposed criteria and empirical evidence. Primary studies with regression testing, test case prioritization techniques, and web services validation contained in their titles, keywords, abstracts, and full texts, have been examined.

#### 3.3.1 Inclusion Criteria

The inclusion criteria of research papers have been defined based on abstracts and titles:

- Studies which discuss regression testing for web services
- Studies which discuss test case prioritization for web services
- Studies which discuss web services validation
- Studies which discuss any type of regression testing
- Studies which discuss algorithms for the test case prioritization technique

If the titles and abstracts of a study contained any of the above points, serving as inclusion criteria, the study was subsequently included in this SLR.

#### 3.3.2 Exclusion Criteria

In addition to the inclusion criteria detailed above, this paper's authors also developed a series of exclusion criteria for the current SLR.

- Studies which did not focus on regression testing for web services were excluded
- Studies which did not focus on test case prioritization for web services were excluded

- Technical reports, Ph.D. dissertations, and studies, of less than five pages, were also excluded. Sufficient evidence was required from empirical studies regarding regression testing for web services.
- Duplicate studies were excluded. It was noticed that some of the researchers had published journal papers, which are versions of their previous conference or workshop papers. A journal paper has more detailed experiments, results, and discussion details as compared to conference or workshop papers.
- Studies that do not provide technical details on this paper's research topic were also excluded.

The authors have followed the grey literature exclusion criteria given by Kitchenham et al. [33] as they assumed that grey literature was not published due to publishing bias [34]. The quality of primary studies was assessed through quality assessment criteria, as detailed in the following section.

### 3.4 Quality Assessment Criteria (QAC)

Kitchenham et al. [29] have reported that quality assessment is mandatory for ensuring aggregation results, which are taken from the best available databases. To verify this research finding, Zhou et al. have suggested that a large number of SLRs, such as 110 out of 127, have used the explicit quality assessment criteria [35]. It means that the findings of Zhou et al. are in line with those of Kitchenham et al. regarding QAC's importance. Therefore, we have used the quality assessment criteria for evaluating chosen primary studies, as proposed in Kitchenham et al. [32], through the use of quality assessment questions.

This paper's authors have used four quality assessment questions to evaluate primary studies, which are the following:

- QA 1.** Is the study focused on the relevant research problem domain?
- QA 2.** Does the study explicitly involve the web services test case prioritization technique?
- QA 3.** Is the study about regression testing for web services?
- QA 4.** Is the study about the validation of web services?

The quality assessment questions have been scored, as detailed in the following table: Each question has been scored on a scale from Yes (1), Partly (0.5), to No (0 "Scores were summed up accordingly. In order to increase the reliability of this SLR, the authors involved other independent researchers in assessing the primary studies. Any confusion or disagreement was settled before reaching a final agreement.

### 3.5 Search Process

A traditional review is restricted to literature, which is already known by authors. The same studies are cited more often than cursory searches [36]. Like other SLRs, this research study adopts a broad research strategy that is based on predefined search strings and keywords. The following digital libraries (DLs) were used to find relevant literature on the research topic:

- ACM
- IEEE Xplore
- ScienceDirect (The Journal of System and Software, Information and software Technology
- Springer Link
- Wiley Online Library

The above-listed repositories are the most popular because many SLRs in the area of

software engineering use them for journal and conference papers. Other repositories were excluded because the search was restricted to journal papers and conference papers within the field of regression testing, and particularly regarding TCP for web services. **Table 2** presents the number of studies found in each of these databases after each search.

**Table 2.** Selection Procedure

Name of Repository	ACM	IEEE	Science Direct	Springer	Wiley Online	Total
First Search	779	739	69	94	375	2056
Second Search (Title, Abstract, AND Keywords)	140	182	16	19	87	444
Third Search (Applying AND Operators)	16	25	14	10	12	77
Removing duplicated research papers	12	22	13	8	10	65

The researchers applied a broad search to find these research papers. The researchers used keywords given in **Table 1** and identified many research papers that focus on regression testing and web services TCP. They applied this search process by choosing an advanced research option for the above-given repositories. They selected the titles, abstracts, and keywords for the first search, and then refined the search process by only selecting abstracts with searched keywords. However, the final search was based on studying articles in full text to ensure that the articles discussed regression testing and TCP for web services. Research articles written in languages other than English were not included [32].

### 3.6 Data Extraction

Data extraction is a critical part of an SLR. Data extraction processes are employed in an SLR, to extract findings from the chosen primary studies, through a consistent means. We used the data extraction procedure proposed by Kitchenham [34]. They extracted information regarding the focused aspects of regression testing and mainly web services TCP from each primary study.

**Table 3.** Summary of the Research Focuses of Primary Studies

ID	Ref.	Research Focus	Paper Type
1	[24]	Early efforts on 'test cases prioritization in regression testing.'	Journal
2	[37]	Proposing Coyote framework for web services TCP	Conf.
3	[23]	Comparing fine granularity and coarse granularity-based prioritization techniques	Journal
4	[38]	Proposing a new algorithm for test-suite prioritization	Journal
5	[39]	Mutational faults, as a means of assessing test case prioritization techniques	Journal
6	[10]	The time constraint prioritization technique for different environments	Conf.
7	[40]	Proposing a "multi-agent based framework" MAST framework	Conf.



8	[41]	A research study focused on search, greedy and metaheuristic algorithms, for the prioritization of test cases and the comparison of algorithm performance	Journal
9	[42]	TCP technique, using requirement priority and cost of test cases	Conf.
10	[43]	The user's session-based TCP technique	Conf.
11	[44]	Hidden Markov Models	Journal
12	[45]	A requirement-based approach	Journal
13	[46]	Presentation of a comparison of the FEP and 'Random Prioritization' RND TCP techniques	Journal
14	[47]	A UML2.0 sequence diagram-based prioritization technique, and the development of a ruled-based matrix	Journal
15	[48]	A clustering technique	Conf.
16	[49]	Proposing a risk-based web services TCP technique	Journal
17	[50]	The effectiveness of the test case reduction technique	Journal
18	[51]	'Adaptive Random Testing' (ART) and greedy techniques	Journal
19	[52]	The 'Modified Condition/Decision Coverage' MC/DC-based technique	Journal
10	[53]	A collaborative framework	Journal
21	[54]	The test case selection, minimization and prioritization technique	Journal
22	[55]	Test case prioritization through using a string distances to elaborate prioritization algorithm	Journal
23	[56]	Development of a requirement-based test cases prioritization technique for web services	Journal
24	[57]	eXtensible BPEL approach	Journal
25	[58]	A robustness testing approach	Journal
26	[59]	A test case refactoring approach	Journal
27	[60]	Focused on functional testing for web services	Journal
28	[61]	Hidden Markov Models	Journal
29	[62]	A survey on testing and verification from services-oriented architecture	Journal
30	[15]	A mapping study on test case prioritization	Journal
31	[63]	The approximation approach to optimize TCP	Conf.
32	[64]	The dependency graph-based TCP technique	Journal
33	[7]	Formulated hybrid criteria for regression testing	Journal
34	[65]	Fuzzy clustering approach	Conf.
35	[66]	Focused on regression testing for PHP applications	Journal
36	[67]	Similarity-based TCP technique	Journal
37	[68]	Developed a suite of metrics of the prioritization technique, for	Journal

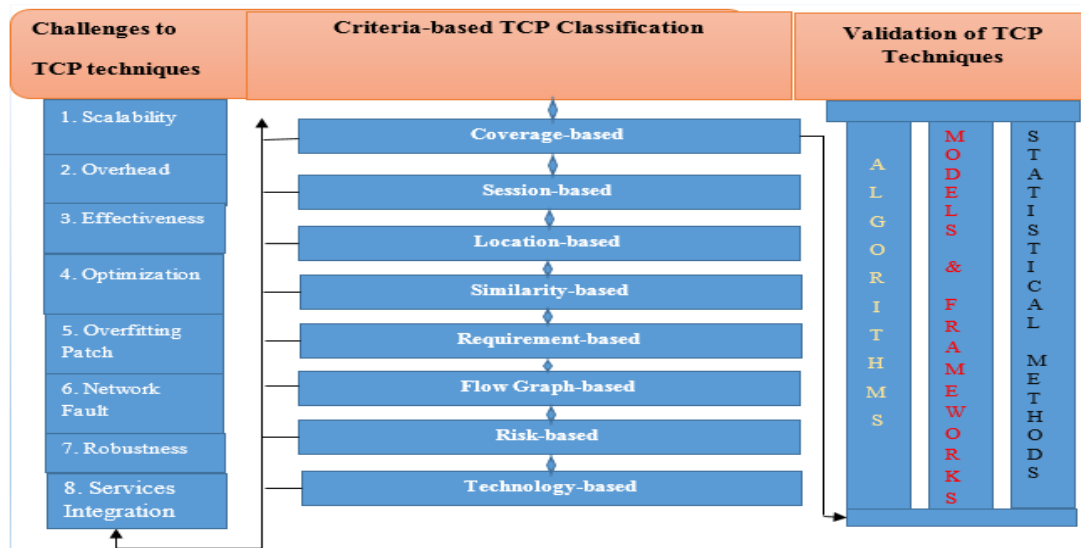


		location-based services	
38	[69]	Topic-based TCP technique	Journal
39	[2]	A mapping study on a regression study for web services	Journal
40	[70]	A cost-cognizant based TCP technique	Journal
41	[71]	Development of a pre-emptive approach of regression testing for web services	Journal
42	[72]	Development of a function call sequence-based profile, for cluster analysis in regression testing	Conf.
43	[73]	A graph-based TCP technique	Journal
44	[74]	'Reinforcement Learning' based Hidden Markov Models (HMM) for TCP	Journal
45	[75]	'Ant Colony Optimization' ACO Algorithm	Conf.
46	[76]	Using three web service metrics, namely APFD, 'Harmonic Mean of the rate of Fault Detection' HMFd, and 'Relative Position' RP.	Journal
47	[77]	Cuckoo Search algorithm	Conf.
48	[78]	Coverage-based TCP technique and evaluation of four common TCP techniques	Journal
49	[79]	A history-based TCP technique	Journal
50	[80]	A TCP technique for 'Location-Based Services' (LBS)	Journal
51	[81]	Using a fuzzy expert system for TCP	Journal
52	[82]	TCP based on coverage and optimally	Journal
53	[26]	TCP for multi-objective performed on 21 Java applications	Journal
54	[83]	A modified ant colony optimization based algorithm, used for finding the shortest path for test selection cases	Conf.
55	[84]	Requirements correlation used for test case prioritization	Conf.
56	[85]	Historical data (requirement prioritization) used for test case prioritization	Conf.
57	[86]	A slice-based TCP technique	Journal
58	[87]	The Hadoop MapReduce Framework	Journal
59	[88]	Refactoring-Based (RB) TCP	Journal
60	[89]	Cloud testing	Conf.
61	[9]	A fuzzy logic-based expert system	Journal
62	[90]	The automatic regression testing	Journal
63	[91]	Model-based TCP technique	Journal
64	[31]	An SLR on test case selection for regression testing	Journal
65	[92]	The PerfRanker prioritization technique of test cases, for collection intensive software	Journal

**Table 3** presents a summary of each primary study selected in this SLR. The summary includes the central idea of each proposed TCP approach, along with the venue. As mentioned earlier, the authors have focused more on journals and less on conference papers. As given in **Table 3**, 50 articles (or 77%) are journal papers, and 15 papers (or 23%) are conference papers. The publishing trend on the web services TCP techniques since 2011 has evidence that primary studies were either published in conferences or Journal papers. In summary, the topic web services TCP is in the middle phase as many journal and conferences studies have been already addressed web services TCP issues.

### 3.7 A Conceptual Framework of SLR on TCP for Web Services

The following **Fig. 2** presents a framework for the currently-presented SLR. The proposed conceptual framework is an abstraction of the relevant TCP approaches, the issues of TCP techniques, and the validation methods used in regards to the existing web services TCP approaches in primary studies. The proposed framework is built on the data we have gathered during the search process. Moreover, the proposed framework is validated by using the information from retrieved papers.



**Fig. 2.** A Conceptual Framework of TCP Techniques

The proposed framework in **Fig. 2**, in its composition, contains three main components. The first is the challenges to the TCP techniques, the second is the criteria-based classification of TCP techniques, and the third is the evaluation of web services TCP techniques. This framework has been proposed to group the TCP techniques along the lines of similar features and to find answers to the SLR's research questions.

## 4. Results & Discussion

### 4.1 (RQ1) What Approaches are proposed, as a Means to Prioritize Test Cases?

RQ1 revisits web services TCP approaches, including proposed TCP models, frameworks, and algorithms.

Rothermel et al. [24] and Tsai et al. [37] were among the earliest academic researchers who presented techniques for web service testing. In the latter study, researchers proposed an XML-based Coyote framework that converted the WSDL specifications into scenarios of testing in its first part and information tracing in the second part. That time distributed services required integration testing, and the Coyote framework was used for integration testing instead of module testing.

Four SLRs were identified, along with several non-SLRs in literature. Both SLRs and numerous non-SLRs attempted to address the TCP problem. A comprehensive comparison of SLRs, from the perspective of TCP techniques, has been contained in the following content.

In the first SLR, Qiu et al. [2] sought to explore regression testing techniques used within the context of five diverse stakeholders. In the same study, the roles of all five stakeholders within regression testing were studied, as well. In addition to the position of web services stakeholders, a large proportion of the SLR results have highlighted the TCP techniques, and have provided a limited classification of TCP techniques. In the second SLR, Kazmi et al. adopted a different approach to TCP techniques and stated that 'Regression Testing Selection' (RTS) involved fault, coverage, and cost-based adequacy criteria [31]. Kazmi et al.'s content did not include any other than the three factors listed above. However, they also investigated a comprehensive set of standard metrics applied in regression testing. While the first SLR by Qiu et al. identified evaluation as a validation method for TCP techniques, Kazmi et al. used 11 metrics as a means of evaluating techniques in their SLR. The third SLR by Yoo and Harman detailed state of the art regarding session-based and coverage-based techniques, from available research papers, published up to 2010 [54]. In both of these criteria-based techniques, Yoo and Harman more emphasized on an increasing fault detection rate [54]. In the same SLR, various types of TCP techniques were studied under the coverage-based criteria of the TCP techniques, such as branch total, statement total, and the fault-exposing potential information included in other criteria-based TCP techniques. The second SLR by Kazmi et al. [31] is in line with the third SLR, in detailing the types of regression testing. However, Yoo and Harman overemphasized the oracle problem of regression testing and did not focus on other TCP challenges [54]. In the fourth SLR, Catal and Mishra followed the pattern of a mapping study and included research papers published between 2001 and 2011 [15]. They did not describe the TCP techniques for researchers and practitioners in detail.

From the comparison of four SLRs, the authors have found that these SLRs have missed, including the newly-proposed TCP techniques for web services. Therefore, we have provided a sufficient literature review of the criteria-based existing and new-emerged web services TCP techniques. The TCP techniques are revisited, discussed in four SLRs conducted by Yoo and Harman [54], Qiu et al. [2], Kazmi et al. [31], and Catal and Mishra [15]. Based on data extraction, as shown in **Table 3**, we provide the classification of TCP approaches by following criteria, algorithms, frameworks, and models proposed in the selected studies.

**Fig. 3** shows that 63% of the total 65 research studies discussed the proposal of criteria-based web services techniques. Algorithm-based web services TCP techniques were found in 17% of the total 65 primary studies. Model-based web services techniques were examined in 11% of total 65 primary studies, and framework based TCP approaches were reported in 9% of total 65 primary studies. It showed that web services TCP is no more examined in the context of classification algorithms and clustering approaches.

Before this study, no SLR or mapping study investigated the models, frameworks, and algorithms separately in order to find a solution to the TCP problem. In summary, researchers have focused on proposing TCP techniques, models, frameworks, and algorithms as a means

of increasing the effectiveness and optimization of TCP techniques for web services.

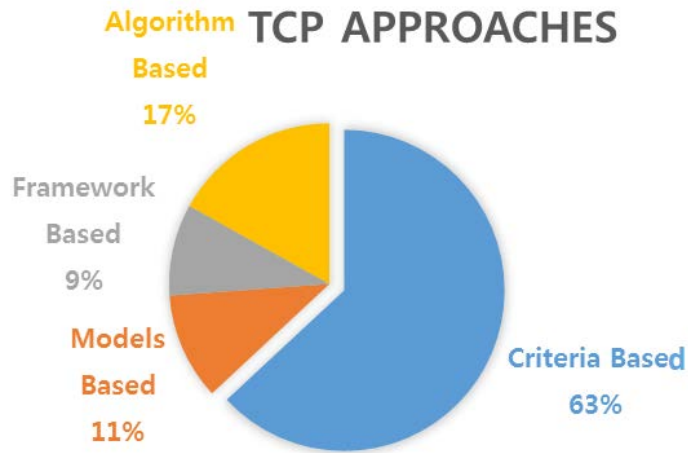


Fig. 3. TCP Approaches

Fig. 4 shows a map of selected primary studies over the TCP techniques and statistical methods. Note that the information includes not only data extracted from chosen primary studies [96]. Detailed information about TCP techniques has been provided within the above section of this SLR. Additionally, abbreviations have been used for each identified TCP technique. For example, the abbreviations coverage based (CB), session-based (SB), similarity-based (SiB), requirement based (RB), flow graph-based (FgB), risk-based (RiB) and location-based (LB) techniques, are all presented in Fig. 4.

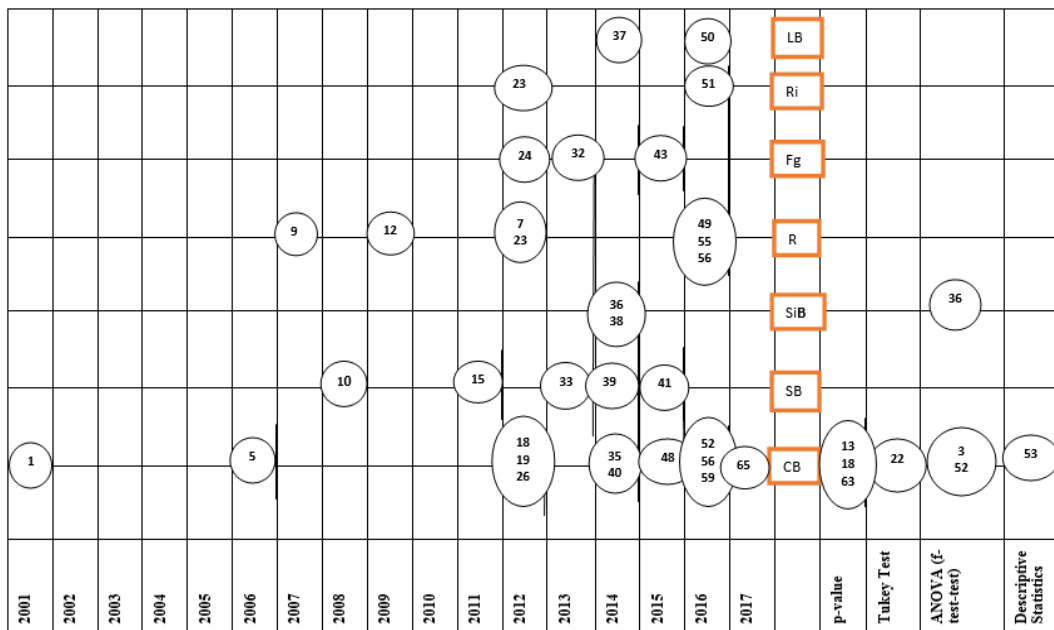


Fig. 4. A Bubble Chart of the Distribution of TCP Technique and Statistical Methods

Fig. 4 shows that CB approaches have been widely proposed to address the TCP issue. Moreover, the bubble chart, as Fig. 4 indicates that several papers introduced CB techniques, and evaluation of them was done by statistical methods including p-value, Tukey test, ANOVA, and descriptive statistics. Besides, CB techniques, only SiB techniques, have been validated by Anova Analysis statistics. Due to the availability of open-source and industrial interfaces of web services, researchers evaluated their proposed approaches [1][5][18][19][26][35]. Hence, they found it convenient to propose CB approaches of TCP. Other than CB techniques, RB approaches are popular techniques because many researchers proposed TCP approaches using functional specifications of web services. SiB and LB, and RiB based TCP approaches were least examined in the literature because the latter mentioned approaches remained less effective to detect the faults. To overcome the frequent changes in web services, code coverage based approaches performed better for patched web services. Cost-saving is another achievement from coverage based TCP approaches because multiple test cases can be executed at once [88]. Since regression testing and particularly TCP approaches' proposal and evolution have spawned in the last many years, coverage based approaches such as code based and requirement based TCP approaches have become popular due to their long history in comparison to other TCP approaches.

#### 4.2 (RQ2) How have the proposed web services TCP Techniques been validated?

To quantify the results of primary studies in regression testing for web services TCP approaches (combining TCP techniques and algorithms), some useful metrics have been applied to compare the effectiveness of different web services TCP techniques. Web services TCP techniques differ in terms of their designs and objectives. Owing to the different goals and designs of TCP approaches, it becomes hard to apply a common estimation mechanism for evaluating the results of TCP approaches. The most commonly-applied evaluation metrics are provided in Table 4. Mei et al. mentioned three web service metrics, namely APFD, 'Harmonic Mean of the rate of Fault Detection' HMGD, and average 'Relative Position' RP [76]. They found that each of the three web service metrics measured different aspects of web services TCP approaches.

**Table 4.** The Distribution of Studies over Evaluation Metrics

Evaluation Metric	References	No
APFD	[46][55][64][7][95][98][67][69][78][97][95][71][79][26][83][84][85][86][88][90][93]	22
APFD-P	[92]	1
NAPFD	[94]	1
FDR	[85]	1
Extended (APFD) <sub>C</sub>	[82]	1
APFDD	[73]	1
HMGD	[68][76]	2
WPFDD	[56]	1
TPFD	[45]	1

MATLAB Tool	[7]	1
Wsrbench	[58]	1
A C program tool	[24]	1
Average RP	[76]	1
Test Cost	[49]	1
Test Efficiency	[49]	1
Harmonic Means	[99]	1
Code Metric	[100]	1

**Table 4** shows the distribution of identified evaluation metrics in primary studies. The top evaluation metric APFD has been reported in 22 primary studies. Wang and Zeng identified the APFD and FDR metrics, indicating that APFD is also discussed with other metrics [85]. In **Table 4**, some primary studies other than those reported in **Table 3** were also used to discuss the APFD and NAPFD metrics. This SLR's findings regarding the metrics proposed for validation of web services TCP have not been mentioned in the previous mapping study by Qiu et al. [2]. As shown in **Table 4**, a total number of 17 metrics have been identified that fill the research gap on the current state of the art on validation of web services techniques. To compare the findings on RQ2, the authors mention 17 metrics, while Catal and Mishra [15] reported only seven metrics in their SLR. Kazmi et al. [31] mentioned 12 metrics, and include those metrics which were stated for test case selection. The lateral mentioned SLR includes precision, recall, efficiency, and F-measure metrics. These metrics are known as inclusiveness metrics. It is noticed that precision, recall, efficiency, and F-measure metrics are best examined in a combination of classifiers [101]. Four lateral mentioned metrics are best explained for fault prediction in traditional programs. Code metric for interface application of web services is used to predict changes produced in web services. Code metrics are widely used for traditional programs. Kumar et al., in their proposed web services TCP, used code metrics on eBay web services, and gained better results than those using a set of other metrics [100]. Askarunisa et al. used APFD, and Harmonic Means (HM) metrics to validate their web services TCP on benchmark weather monitoring services (WMS), and Bible web services [99]. Mei et al. applied the APFD metric to measure the effectiveness of the TCP technique on eight WS-BPEL programs [16].

#### 4.3 (RQ3): Are there Statistical Methods used in web services TCP Techniques?

Since 2016, the trend of using statistical methods in regression testing has increased. The increasing trend of using statistical methods is evident in many primary studies. In a primary study, Jiang et al. integrated statistical methods with the TCP for fault localization [51]. Moreover, they used ANOVA analysis to determine differences between TCP techniques in a group of TCP techniques.

Tahat et al. combined the APFD metric with the HMF metric to determine a fault detection rate by assuming the multiple faults [91]. To investigate the effectiveness of TCP techniques, they determined that detecting defects in regression testing was difficult due to restrictive measures when using the two metrics HMF and APFD in a combined form. Thus, they established the relative position of test cases from the application of varying TCP techniques. Ledru et al. assessed various conditions of their evaluated experiment through the

use of the t-test ANOVA analysis [55]. The t-test ANOVA analysis measured differences between the random and distance-based prioritization of test cases. In some experiments, ANOVA analysis shows its limitations in regards to the size of test cases. For those test experiments, Tukey's HSD test, as an option of ANOVA analysis, is performed to determine differences in multiple mean values.

**Table 5.** Primary Studies Using the Statistical Method

References	Statistical Analysis Used
[51] [85] [90] [91] [94]	p-value
[23] [67] [82]	ANOVA analysis
[102]	t-test
[55]	Tukey's HSD test
[26]	Descriptive statistics

**Table 5** shows that 11 primary studies applied statistical methods regarding TCP techniques for web services. ANOVA analysis is the most-cited statistical method, examined in 15.38% (10/65) x100 of primary studies. Statistical tests, including p-value, t-test, and Tukey's HSD test, also belong to the ANOVA family of analysis tests. Apart from ANOVA analysis, no family or class of statistical method has been observed, which is used for comparing, evaluating, or verifying TCP techniques for web services. Only Marchetto et al. [26] utilized a descriptive analysis. Therefore, a low percentage (1.53%) referred to a statistical analysis other than ANOVA analysis.

#### **4.4 (RQ4): What Issues Related to TCP Exist in Regression Testing, with Respect to Web Services?**

Cloud computing services for 'regression testing' are discussed in various primary studies, within the context of using third party services for testing modified web services. The interface information of web services is merely exposed, and its implementation is also kept in-accessible. In these circumstances, the testing of web services by a third party becomes a challenge [16][97]. Various TCP approaches have been proposed to address the TCP problem, despite mentioning or discussing the challenges related to using third party services. Eight types of potential issues have been identified, that prevail in terms of TCP techniques for web services, and these issues are given as the following.

##### **4.4.1 Scalability**

The scalability challenge is identified during the testing of web services on clouds. The 'Software as a Service' (SaaS) infrastructure of MIDAS is only accessible for the functional testing of web services. The MIDAS platform provides software testing in a controlled manner. To make cloud testing more efficient, large-scale software and hardware distributed systems are planned as a means of troubleshooting and testing complex web systems [6]. In recent research work, Wang et al. revealed that the scalability and stability issues of newly-proposed TCP techniques are not as severe when compared to most previous TCP techniques [103]. From these two primary studies, it can be inferred that scalability remains an issue because



limited research efforts have been made in regards to scalability, in terms of regression testing.

#### **4.4.2 Overhead**

A high overhead issue has been reported to introduce *performance regression testing*. The above factors involve spotting performance regression because several commits have merged from the last testing of web services. The extra efforts and development time has been spent on finding solutions to commits. The overhead issue of TCP techniques has been discussed in the literature. Different researchers have discussed the overhead TCP issue in various ways. Within primary studies, the overhead issue has been identified as a rapid pace of software evolution, which requires regression testing [6][46][70][104][105].

#### **4.4.3 Overfitting Patch**

During test case generation, overfitting becomes an issue when a repair technique produces bad fixes. A bad fix is the human-made modification of source code, used for repairing software. In this regard, repairing the system requires retesting and prioritizing test cases. TCP techniques may suffer from this overfitting issue, while systems are being corrected. Dobolyi et al. have highlighted the issue of over-fitting, which is commonly seen during the validation of approaches [70][106]. To overcome this issue of overfitting patch, Xin and Reiss proposed a technique called DiffTGen for testing patches and for avoiding the overfitting issue during test case generation [107][108]. The DiffTGen technique identified almost 50% of overfitting patches within a few minutes. However, this DiffTGen technique required more sophistication to make it an optimized technique.

#### **4.4.4 Effectiveness**

The effectiveness of regression testing is profoundly affected by many TCP factors for web services. These factors include test case size, fault coverage, fault proneness, and data flow. Additionally, the fault severity and cost of correcting web services determine the effectiveness of a TCP technique. Some of the TCP techniques do not account for fault severity and cost, and both factors are considered equally effective [51][62][64]. Therefore, TCP techniques that ignore these two factors cannot produce an appropriate order of test cases [71][91][109].

#### **4.4.5 Optimization**

In regression testing, optimization of TCP approaches is a severe concern of researchers as they require the ordering of test cases, specifically by executing those with a high priority, before those with a low priority [26][75][111]. The optimization of TCP techniques is a valid reason for seeing the practical implementation of techniques. Many recent attempts have been made to highlight the optimality issue of TCP techniques [82][63][65][105][110]. In these primary studies, the optimization of test case reduction and test case prioritization were attributed as being future research areas for regression testing. Neither of these studies presented a solution for the optimization of TCP approaches. However, in the first study, Anwar et al. [9] concluded that the optimization of TCP techniques was crucial for the cost-effective execution of test suits. The risk of unsafe optimization increased, because some of the vital test cases were skipped and because faults within these skipped test cases remained unaddressed.

#### **4.4.6 Network Fault**

The positioning of faults among networks is a familiar concept used in software testing. A

node position is represented as a fault in a system. In a primary study, Kayes et al. presented the proposal of a network approach to resolving the concerns of researchers, regarding network faults [73]. Dependency among network faults occurs because other faults create some fault consequences. Leading failures need to be addressed before consequent faults are removed.

#### 4.4.7 Robustness

Web services are utilized in complex and critical business systems. The robustness of web services is essential for running a business through these complicated and critical systems [58]. A robustness assessment of critical systems is in itself an issue, which is corroborated by a limited number of primary studies. A comprehensive, robust solution to the TCP technique is mandatory because the security vulnerabilities produced as a result of the behavior and design of web services are only addressed through the use of a robustness approach.

#### 4.4.8 Service Integration

The integration of web services is itself a big challenge for web services providers. It is because a vendor who uses the web services of others does not know the internal behavior of these services. The integration of web services results in a lack of control regarding test case execution and software artifacts, without having complete knowledge about the internal behavior of web services [53].

Fig. 5 shows the number of primary studies that tried to highlight or conquer each issue. Therefore, optimization remains a particular concern to researchers in ten primary studies (15% of the total 65 primary studies), and high effectiveness is a concern in nine primary studies (14% of the total 65 primary studies). “Overhead and over-fitting are prominent in 8% and 5% of publications, respectively. Scalability issue is now a focused area of research in web services TCP as an increased number of web services use the same interface for services. The scalability issue remained less addressed in the existing literature, and two primary studies (3% of total 65 primary studies) revealed this issue for web services TCP. The remaining matters identified, namely robustness, integration, and network fault, all each appear in 2% of the primary studies.

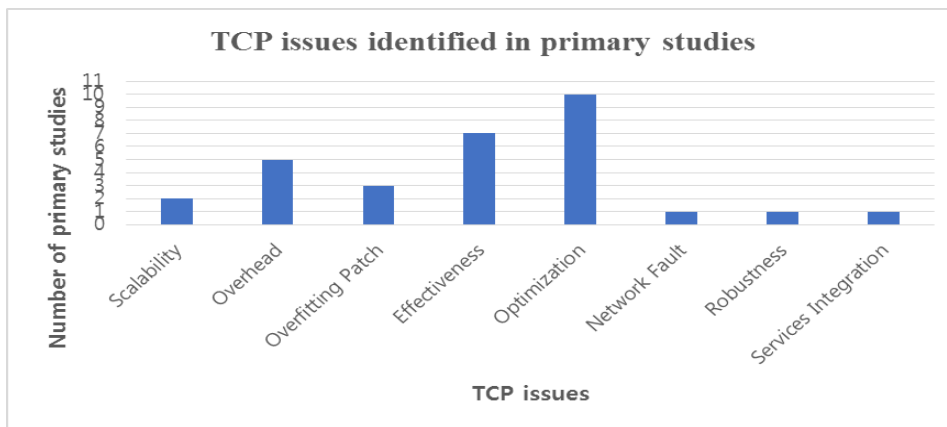


Fig. 5. Identified Issues Impacting TCP techniques

## 5. Threats to Validity

This paper's authors are conscious of threats that may impact the validity of the findings and outcomes of this SLR. First, selection bias could be possible. Although recognized data sources including IEEE Xplore DL, ScienceDirect, ACM DL, Springer, and Wiley online were all used as primary studies, other sources have not been considered due to their scope and time constraints. Thus, it is possible that some relevant studies are missing.

Another possible concern is the need to find adequate search strings. The search strings and repositories used within the research method section have been presented in the method section. The search strings have been refined many times, to determine the maximum number of relevant research studies, and to access only relevant data sources.

Quality assessment may present a problem, leading to inaccurate data results. It was challenging to answer RQ2, RQ3, and RQ4 due to a lack of appropriate measures required for these questions. Therefore, based on the framework proposed for this SLR, the researchers were able to assess primary studies that ensured quality assessment for chosen primary studies.

## 6. Conclusion and Future Work

In this SLR, researchers have classified TCP approaches for criteria-based TCP techniques, models, frameworks, and algorithms. Among the seven criteria-based classes of TCP techniques, coverage-based TCP techniques have been greatly covered in primary studies. Most primary studies have mentioned that APFD is a primarily-applied metric, used to compute or compare the effectiveness of TCP approaches. Three primary studies discussed two tools apart from APFD or extended APFD metrics, which were implemented in the MATLAB and C programming languages. In addition, identified statistical methods were employed in primary studies, including ANOVA analysis (p-value, t-test, and Tukey's HSD test) and descriptive statistics. ANOVA analysis is mainly used to find variances among TCP techniques. The researchers have identified cost-overhead, optimization, and high effectiveness as essential issues related to web services TCP techniques. Besides these issues, overfitting patch, scalability, robustness, network faults, and services integration have also been mentioned, needing additional attention from researchers, within future research work. The current research on TCP techniques for web services contains a few concrete works. However, more issues to TCP techniques for web services have been identified for consideration in future research studies.

## References

- [1] Di Penta, M., Bruno, M., Esposito, G., Mazza, V., & Canfora, G., "Web services regression testing," *Test and Analysis of web Services*, Springer, Berlin, Heidelberg, pp. 205-234, 2007. [Article \(CrossRef Link\)](#).
- [2] Qiu, D., Li, B., Ji, S., & Leung, H., "Regression Testing of Web Service: A Systematic Mapping Study," *ACM Comput. Surv.*, 47(2), 1-46, 2014. [Article \(CrossRef Link\)](#).
- [3] Mei, L., Zhang, Z., Chan, W. K., & Tse, T. H., "Test case prioritization for regression testing of service-oriented business applications," in *Proc. of the 18th international conference on World wide web*, pp. 901-910, 2009. [Article \(CrossRef Link\)](#).
- [4] Chen, L., Wang, Z., Xu, L., Lu, H., & Xu, B., "Test case prioritization for web service regression testing," in *Proc. of Service Oriented System Engineering (SOSE), 2010 Fifth IEEE International Symposium on*, pp. 173-178, 2010. [Article \(CrossRef Link\)](#).

- [5] Elbaum, S., Rothermel, G., & Penix, J, "Techniques for improving regression testing in continuous integration development environments," in *Proc. of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering*, pp. 235-245, 2014. [Article \(CrossRef Link\)](#).
- [6] Hillah, L. M., Maesano, A.-P., Rosa, F. D., Kordon, F., Wuillemin, P.-H., Fontanelli, R., Maesano, L., "Automation and intelligent scheduling of distributed system functional testing," *Int J Softw Tools Technol Transfer*, 19, 281-308, 2017. [Article \(CrossRef Link\)](#).
- [7] Sampath, S., Bryce, R. e., & Memon, A. M, "A Uniform Representation of Hybrid Criteria for Regression Testing," *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING*, 39(10), 1326-1344, 2013. [Article \(CrossRef Link\)](#).
- [8] Do, H., Mirarab, S., Tahvildari, L., & Rothermel, G, "The Effects of Time Constraints on Test Case Prioritization: A Series of Controlled Experiments," *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING*, 36(5), 593-617, 2010. [Article \(CrossRef Link\)](#).
- [9] Anwar, Z., Ahsan, A., & Catal, C, "Neuro-Fuzzy Modeling for Multi-Objective Test Suite Optimization," *J. Intell. Syst*, 25(2), 123-146, 2016.
- [10] Walcott, K. R., Soffa, M. L., Kapfhammer, G. M., & Roos, R. S, "TimeAware Test Suite Prioritization," in *Proc. of the ISSA'06, Portland, Maine, USA*, pp. 1-12, 2006. [Article \(CrossRef Link\)](#).
- [11] Garg, D., & Datta, A, "Parallel Execution of Prioritized Test Cases for Regression Testing of Web Applications," in *Proc. of the ACSC '13 Proceedings of the Thirty-Sixth Australasian Computer Science Conference, Adelaide, Australia*, 2013.
- [12] Liu, C. H., Chen, S. L., & Chen, W.-K, "Cost-Benefit Evaluation on Parallel Execution for Improving Test Efficiency over Cloud," in *Proc. of the Proceedings of the 2017 IEEE International Conference on Applied System Innovation IEEE-ICASI, Sapporo, Japan*, 2017. [Article \(CrossRef Link\)](#).
- [13] Parejoa, J. A., Sánchez, A. B., Seguraa, S., Ruiz-Cortes, A., Lopez-Herrejona, R. E., & Egyed, A, "Multi-objective test case prioritization in highly configurable systems: A case study," *The Journal of Systems and Software*, 122, 287-310, 2016. [Article \(CrossRef Link\)](#).
- [14] Ansari, A., Khan, A., Khan, A., & Mukadam, K, "Optimized Regression Test using Test Case Prioritization," *Procedia Computer Science*, 79, 152-160, 2016. [Article \(CrossRef Link\)](#).
- [15] Catal, C., & Mishra, D, "Test case prioritization: a systematic mapping study," *Software Qual J*, 21, 445-478, 2013. [Article \(CrossRef Link\)](#).
- [16] Mei, L., Chan, W. K., Tse, T. H., & Merkel, R. G, "XML-manipulating test case prioritization for XML-manipulating services," *Journal of Systems and Software*, 84(4), 603-619, 2011. [Article \(CrossRef Link\)](#).
- [17] Farrag, T. A., Saleh, A. I., & Ali, H. A, "Toward SWSs Discovery: Mapping from WSDL to OWL-S Based on Ontology Search and Standardization Engine," *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, 25(5), 1135-1147, 2013. [Article \(CrossRef Link\)](#).
- [18] Simon, B., Goldschmidt, B., & Kondorosi, K, "A Metamodel for the Web Services Standards," *J Grid Computing*, 11, 735-752, 2013. [Article \(CrossRef Link\)](#).
- [19] Petry, F., Ladner, R., Gupta, K. M., Moore, P., Aha, D. W., Lin, B., & Sween, R, "Design of an Integrated Web services Brokering System," *International Journal of Information Technology and Web Engineering*, 4(3), 58-77, 2009. [Article \(CrossRef Link\)](#).
- [20] International Business Machines Corporation, "Patent issued for method and system for ranking services in a web services architecture," *From Journal of Engineering*, 2013.
- [21] Masood, T., Nadeem, A., & Ali, S, "An Automated Approach to Regression Testing of Web Services based on WSDL Operation Changes," in *Proc. of the Emerging Technologies (ICET), 2013 IEEE 9th International Conference, Islamabad, Pakistan*, 2013. [Article \(CrossRef Link\)](#).
- [22] Elbaum, S., Malishevsky, A., & Rothermel, G, "Incorporating varying test costs and fault severities into test case prioritization," in *Proc. of the 23rd International Conference on Software Engineering, IEEE Computer Society*, pp. 329-338, 2001. [Article \(CrossRef Link\)](#).
- [23] Elbaum, S., Malishevsky, A. G., & Rothermel, G, "Test Case Prioritization: A Family of Empirical Studies," *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING*, 28(2), 159-182, 2002. [Article \(CrossRef Link\)](#).

- [24] Rothermel, G., Untch, R. H., & Harrold, M. J., "Prioritizing Test Cases For Regression Testing," *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING*, 27(10), 929-948, 2001. [Article \(CrossRef Link\)](#).
- [25] Zhang, L., Hou, S. S., Guo, C., Xie, T., & Mei, H., "Time-aware test-case prioritization using integer linear programming," in *Proc. of the eighteenth international symposium on Software testing and analysis*, pp. 213-224, 2009, [Article \(CrossRef Link\)](#).
- [26] Marchetto, A., Islam, M. M., Asghar, W., Susi, A., & Scanniello, G., "A Multi-Objective Technique to Prioritize Test Cases," *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING*, 42(10), 918-940, 2016.
- [27] Huang, R., Zhou, Y., Zong, W., Towey, D., & Chen, J., "An Empirical Comparison of Similarity Measures for Abstract Test Case Prioritization," in *Proc. of the Computer Software and Applications Conference (COMPSAC), 2017 IEEE 41st Annual, Turin, Italy, Italy*, 2017. [Article \(CrossRef Link\)](#).
- [28] Kitchenham, B. & Charters, S., "Guidelines for performing systematic literature reviews in software engineering version 2.3," *Technical Report*, 2007.
- [29] Kitchenham, B. A., Budgen, D., & Brereton, P., *Evidence-Based Software Engineering and Systematic Reviews*, Boca Raton, Florida: CRC Press, 2015. [Article \(CrossRef Link\)](#).
- [30] Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M., "Lessons from applying the systematic literature review process within the software engineering domain," *Journal of systems and software*, 80(4), 571-583, 2007. [Article \(CrossRef Link\)](#).
- [31] Kazmi, R., Jawawi, D. N., Mohamad, R., & Ghani, I., "Effective regression test case selection: a systematic literature review," *ACM Computing Surveys (CSUR)*, 50(2), 29, 2017. [Article \(CrossRef Link\)](#).
- [32] Kitchenham, B., Brereton, P., David, B., Mark, T., Bailey, J., & Linkman, S., "Systematic literature reviews in software engineering – A systematic literature review," *Information and Software Technology*, 51(1), 7-15, 2009. [Article \(CrossRef Link\)](#).
- [33] Kitchenham, B. A., Brereton, P., Turner, M., Niazi, M. K., Linkman, S., Pretorius, R., & Budgen, D., "Refining the systematic literature review process two participant-observer case studies," *Empir Software Eng*, 15, 618-653, 2010. [Article \(CrossRef Link\)](#).
- [34] Kitchenham, B., "Procedures for performing systematic reviews," *Keele, UK, Keele University*, 33(2004), 1-26, 2004.
- [35] Zhou, Y., Zhang, H., Huang, X., Yang, S., Babar, M. A., & Tang, H., "Quality Assessment of Systematic Reviews in Software Engineering: A Tertiary Study," in *Proc. of the EASE '15: Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering*, 2015. [Article \(CrossRef Link\)](#).
- [36] Ham-Baloyi, W. t., & Jordan, P., "Systematic review as a research method in postgraduate nursing education," *Health Sa Gesondheid*, 21, 120-128, 2016. [Article \(CrossRef Link\)](#).
- [37] Tsai, W. T., Paul, R., Song, W., & Cao, Z., "Coyote: An xml-based framework for web services testing," in *Proc. of High Assurance Systems Engineering, 2002. Proceedings. 7th IEEE International Symposium on*, pp. 173-174, 2002. [Article \(CrossRef Link\)](#).
- [38] Jones, J. A., & Harrold, M. J., "Test-Suite Reduction and Prioritization for Modified Condition/Decision Coverage," *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING*, 29(3), 195-209, 2003. [Article \(CrossRef Link\)](#).
- [39] Do, H., & Rothermel, G., "On the Use of Mutation Faults in Empirical Assessments of Test Case Prioritization Techniques," *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING*, 32(9), 733-752, 2006. [Article \(CrossRef Link\)](#).
- [40] Bai, X., Dai, G., Xu, D., & Tsai, W. T., "A multi-agent based framework for collaborative testing on web services," in *Proc. of The Fourth IEEE Workshop on Software Technologies for Future Embedded and Ubiquitous Systems, and the Second International Workshop on Collaborative Computing, Integration, and Assurance (SEUS-WCCIA'06)*, pp. 205-210, 2006. [Article \(CrossRef Link\)](#).



- [41] Li, Z., Harman, M., & Hierons, R. M., "Search Algorithms for Regression Test Case Prioritization," *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING*, 23(4), 225-237, 2007. [Article \(CrossRef Link\)](#).
- [42] Zhang, X., Nie, C., Xu, B., & Qu, B., "Test Case Prioritization based on Varying Testing Requirement Priorities and Test Case Costs," in *Proc. of the Seventh International Conference on Quality Software, Portland, OR, USA, 2007*. [Article \(CrossRef Link\)](#).
- [43] Sampath, S., Bryce, R. C., Viswanath, G., Kandimalla, V., & Koru, A. G., "Prioritizing User-Session-Based Test Cases for Web Applications Testing," in *Proc. of the Software Testing, Verification, and Validation, 2008 1st International Conference, Lillehammer, Norway, 2008*. [Article \(CrossRef Link\)](#).
- [44] Roberts, F. S., "Computer science and decision theory," *Ann Oper Res*, 163, 209-253, 2008.
- [45] Krishnamoorthi, R., & Mary, S. A. S. A., "Factor oriented requirement coverage based system test case prioritization of new and regression test cases," *Information and Software Technology*, 51, 799-808, 2009. [Article \(CrossRef Link\)](#).
- [46] Gonzalez-Sanchez, A., Piel, E., Gross, R. A. H.-G., & Gemund, A. J. C. v., "Prioritizing tests for software fault diagnosis," *SOFTWARE – PRACTICE AND EXPERIENCE*, 41, 1105-1129, 2011. [Article \(CrossRef Link\)](#).
- [47] Kundu, D., Sarma, M., Samanta, D., & Mall, R., "System testing for object-oriented systems with test case prioritization," *SOFTWARE TESTING, VERIFICATION AND RELIABILITY*, 19, 297-333, 2009. [Article \(CrossRef Link\)](#).
- [48] Li, J. & Xing, D., "User Session Data based Web Applications Test with Cluster Analysis," in *Proc. of Advanced Research on Computer Science and Information Engineering: International Conference, CSIE 2011, Zhengzhou, China, May 21-22, 2011. Proceedings, Part I, Berlin, Heidelberg: Springer Berlin Heidelberg*, pp. 415-421, 2011. [Article \(CrossRef Link\)](#).
- [49] Bai, X., Kenett, R. S., & Yu, W., "Risk assessment and adaptive group testing of semantic web services," *International Journal of Software Engineering and Knowledge Engineering*, 22(05), 595-620, 2012. [Article \(CrossRef Link\)](#).
- [50] Sampath, S., & Bryce, R. C., "Improving the effectiveness of test suite reduction for user-session-based testing of web applications," *Information and Software Technology*, 54(7), 724-738, 2012. [Article \(CrossRef Link\)](#).
- [51] Jiang, B., Zhang, Z., Chan, W. K., Tse, T. H., & Chen, T. Y., "How well does test case prioritization integrate with statistical fault localization?," *Information and Software Technology*, 54, 739-758, 2012. [Article \(CrossRef Link\)](#).
- [52] Fang, C., Chen, Z., & Xu, B., "Comparing logic coverage criteria on test case prioritization," *Science China Information Sciences*, 55(12), 2826-2840, 2012. [Article \(CrossRef Link\)](#).
- [53] Zhu, H., & Zhang, Y., "Collaborative Testing of Web Services," *IEEE TRANSACTIONS ON SERVICES COMPUTING*, 5(1), 116-130, 2012. [Article \(CrossRef Link\)](#).
- [54] Yoo, S., & Harman, M., "Regression testing minimization, selection and prioritization: a survey," *Softw. Test. Verif. Reliab.*, 22, 67-120, 2012.
- [55] Ledru, Y., Petrenko, A., Boroday, S., & Mandran, N., "Prioritizing test cases with string distances," *Autom Softw Eng.*, 19, 65-95, 2012. [Article \(CrossRef Link\)](#).
- [56] Srikanth, H., & Banerjee, S., "Improving test efficiency through system test prioritization," *The Journal of Systems and Software*, 85, 1176-1187, 2012. [Article \(CrossRef Link\)](#).
- [57] Li, B., Qiua, D., Leungb, H., & Wanga, D., "Automatic test case selection for regression testing of composite service based on extensible BPEL flow graph," *The Journal of Systems and Software*, 85, 1300-1324, 2012. [Article \(CrossRef Link\)](#).
- [58] Laranjeiro, N., Vieira, M., & Madeira, H., "A robustness testing approach for SOAPWeb services," *J Internet Serv Appl*, 3, 215-232, 2012. [Article \(CrossRef Link\)](#).
- [59] Chu, P.-H., Hsueh, N.-L., Chen, H.-H., & Liu, C.-H., "A test case refactoring approach for pattern-based software development," *Software Qual J*, 20, 43-75, 2012. [Article \(CrossRef Link\)](#).
- [60] Tahir, A., Tosi, D., & Morasca, S., "A systematic review on the functional testing of semantic web services," *The Journal of Systems and Software*, 86, 2877-2889, 2013. [Article \(CrossRef Link\)](#).

- [61] Kazi, S., & Stamp, M, "Hidden Markov Models for Software Piracy Detection," *Information Security Journal: A Global Perspective*, 22, 140-149, 2013. [Article \(CrossRef Link\)](#).
- [62] Bozkurt, M., Harman, M., & Hassoun, Y, "Testing and verification in service-oriented architecture: a survey," *SOFTWARE TESTING, VERIFICATION AND RELIABILITY*, 23, 261-313, 2013. [Article \(CrossRef Link\)](#).
- [63] Mei, L., Cai, Y., Jia, C., Jiang, B., & Chan, W. K, "Prioritizing Structurally Complex Test Pairs for Validating WS-BPEL Evolutions," in *Proc. of the Web Services (ICWS), 2013 IEEE 20th International Conference, Santa Clara, CA, USA*, 2013.
- [64] Panigrahi, C. R., & Mall, R, "A heuristic-based regression test case prioritization approach for object-oriented programs," *Innovations Syst Softw Eng*, 10, 155-163, 2013. [Article \(CrossRef Link\)](#).
- [65] Kumar, G., & Bhatia, P. K, "Software testing optimization through test suite reduction using fuzzy clustering," *CSI Transaction on ICT*, 1(3), 253-260, 2013. [Article \(CrossRef Link\)](#).
- [66] Do, H., & Hossain, M, "An efficient regression testing approach for PHP web applications: a controlled experiment," *Softw. Test. Verif. Reliab*, 24, 367-385, 2014. [Article \(CrossRef Link\)](#).
- [67] Fang, C., Chen, Z., Wu, K., & Zhao, Z, "Similarity-based test case prioritization using ordered sequences of program entities," *Software Qual J*, 22, 335-361, 2014. [Article \(CrossRef Link\)](#).
- [68] Zhai, K., Jiang, B., & Chan, W. K, "Prioritizing Test Cases for Regression Testing of Location-Based Services: Metrics, Techniques, and Case Study," *IEEE TRANSACTIONS ON SERVICES COMPUTING*, 7(1), 54-67, 2014. [Article \(CrossRef Link\)](#).
- [69] Thomas, S. W., Hemmati, H., Hassan, A. E., & Blostein, D, "Static test case prioritization using topic models." *Empir Software Eng*, 19, 182-212, 2014. [Article \(CrossRef Link\)](#).
- [70] Huang, P., Ma, X., Shen, D., & Zhou, Y, "Performance regression testing target prioritization via performance risk analysis," in *Proc. of the ICSE 2014 Proceedings of the 36th International Conference on Software Engineering, Hyderabad, India*, pp. 60-71, 2014. [Article \(CrossRef Link\)](#).
- [71] Mei, L., Chan, W. K., Tse, T. H., Jiang, B., & Zhai, K, "Preemptive Regression Testing of Workflow-Based Web Services," *IEEE TRANSACTIONS ON SERVICES COMPUTING*, 8(5), 740-754, 2015. [Article \(CrossRef Link\)](#).
- [72] Wang, X., Jiang, X., & Shi, H, "Prioritization of Test Scenarios using Hybrid Genetic Algorithm Based on UML Activity Diagram," in *Proc. of the Software Engineering and Service Science (ICSESS), 2015 6th IEEE International Conference, Beijing, China*, 2015. [Article \(CrossRef Link\)](#).
- [73] Kayes, I., Islam, S., & Chakareski, J, "The network of faults: a complex network approach to prioritize test cases for regression testing," *Innovations Syst Softw Eng*, 11, 251-275, 2015. [Article \(CrossRef Link\)](#).
- [74] Emam, S. S., & Miller, J, "Test case prioritization using extended digraphs," *ACM Trans. Softw. Eng. Methodol*, 25(1), 1-41, 2015. [Article \(CrossRef Link\)](#).
- [75] Gao, D., Guo, X., & Zhao, L, "Test case prioritization for regression testing based on ant colony optimization," in *Proc. of the Software Engineering and Service Science (ICSESS), 2015 6th IEEE International Conference, Beijing, China*, 2015. [Article \(CrossRef Link\)](#).
- [76] Mei, L., Cai, Y., Jia, C., Jiang, B., Chan, W. K., Zhang, Z., & Tse, T. H, "A subsumption hierarchy of test case prioritization for composite services," *IEEE Transactions on Services Computing*, 8(5), 658-673, 2015. [Article \(CrossRef Link\)](#).
- [77] Nagar, R., Kumar, A., Singh, G. P., & Kumar, S, "Test Case Selection and Prioritization using Cuckoos Search Algorithm," in *Proc. of the 2015 1st International Conference on Futuristic trend in Computational Analysis and Knowledge Management (ABLAZE-2015), Noida, India*, 2015 [Article \(CrossRef Link\)](#).
- [78] Nardo, D. D., Alshahwan, N., Briand, L., & Labiche, Y, "Coverage-based regression test case selection, minimization and prioritization: a case study on an industrial system," *SOFTWARE TESTING, VERIFICATION AND RELIABILITY*, 25, 371-396, 2015. [Article \(CrossRef Link\)](#).
- [79] Hemmati, H., Fang, Z., Mantyla, M. V., & Adams, B, "Prioritizing manual test cases in rapid release environments," *SOFTWARE TESTING, VERIFICATION AND RELIABILITY*, 27(6), 1-25, 2016. [Article \(CrossRef Link\)](#).



- [80] Ryschka, S., Murawski, M., & Bick, M., "Location-Based Services," *Bus Inf Syst Eng*, 58(3), 233-237, 2016. [Article \(CrossRef Link\)](#).
- [81] Hettiarachchi, C., Do, H., & Choi, B., "Risk-based test case prioritization using a fuzzy expert system." *Information and Software Technology*, 69, 1-15. [Article \(CrossRef Link\)](#).
- [82] Hao, D., Zhang, L., Zang, L., Wang, Y., Wu, X., & Xie, T., "To Be Optimal or Not in Test-Case Prioritization," *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING*, 42(5), 490-504, 2016. [Article \(CrossRef Link\)](#).
- [83] Kumar, S., Ranjan, P., & R.Rajesh, "Modified ACO to maintain diversity in Regression Test Optimization," in *Proc. of the 3rd Int'l Conf. on Recent Advances in Information Technology*, 2016. [Article \(CrossRef Link\)](#).
- [84] Ma, T., Zeng, H., & Wang, X., "Test case prioritization based on requirement correlations," in *Proc. of the Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2016 17th IEEE/ACIS International Conference, Shanghai, China*, 2016. [Article \(CrossRef Link\)](#).
- [85] Wang, X., & Zeng, H., "History-Based Dynamic Test Case Prioritization for Requirement Properties in Regression Testing," in *Proc. of the Continuous Software Evolution and Delivery (CSED), IEEE/ACM International Workshop, Austin, TX, USA*, pp. 41-47, 2016. [Article \(CrossRef Link\)](#).
- [86] Panda, S., Munjal, D., & Mohapatra, D. P., "A Slice-Based Change Impact Analysis for Regression Test Case Prioritization of Object-Oriented Programs," *Advances in Software Engineering*, vol. 2016, 1-20, 2016. [Article \(CrossRef Link\)](#).
- [87] Chawla, P., Chana, I., & Rana, A., "Cloud-based automatic test data generation framework," *Journal of Computer and System Sciences*, 82, 712-738, 2016. [Article \(CrossRef Link\)](#).
- [88] Alves, E. L. G., Machado, P. D. L., Massoni, T., & Kim, M., "Prioritizing test cases for early detection of refactoring faults," *SOFTWARE TESTING, VERIFICATION AND RELIABILITY*, 26, 402-426, 2016. [Article \(CrossRef Link\)](#).
- [89] Kaur, M., "Testing in the Cloud: New Challenges," in *Proc. of the Computing, Communication and Automation (ICCCA), 2016 International Conference, Noida, India*, 2016. [Article \(CrossRef Link\)](#).
- [90] Sanchez, A. B., Segura, S., Parejo, J. A., & Ruiz-Cortes, A., "Variability Testing in the wild: the Drupal case study," *Softw Syst Model*, 16, 173-194, 2017. [Article \(CrossRef Link\)](#).
- [91] Tahat, L., Korel, B., Koutsogiannakis, G., & Almasri, N., "State-based models in regression test suite prioritization," *Software Qual J*, 25, 703-742, 2017. [Article \(CrossRef Link\)](#).
- [92] Mostafa, S., Wang, X., & Xie, T., "PerfRanker: Prioritization of Performance Regression Tests for Collection-Intensive Software," in *Proc. of the ISSSTA 2017 Proceedings of the 26th ACM SIGSOFT International Symposium on Software Testing and Analysis, Santa Barbara, CA, USA*, pp. 23-34, 2017. [Article \(CrossRef Link\)](#).
- [93] Rajarathinam, K., & Natarajan, S., "Test suite prioritization using trace events technique," *IET Softw*, 7(2), 85-92, 2013. [Article \(CrossRef Link\)](#).
- [94] Wang, R., Jiang, S., Chen, D., & Zhang, Y., "Empirical Study of the Effects of Different Similarity Measures on Test Case Prioritization," *Mathematical Problems in Engineering*, vol. 2016, 1-19, 2016. [Article \(CrossRef Link\)](#).
- [95] Maheswari, R. U., & JeyaMala, D., "A Novel Approach for Test Case Prioritization," in *Proc. of the Computational Intelligence and Computing Research (ICCIC), 2013 IEEE International Conference, Enathi, India*, 2013. [Article \(CrossRef Link\)](#).
- [96] Felderer, M., & Schieferdecker, I., "A taxonomy of risk-based testing," *Int J Softw Tools Technol Transfer*, 16, 559-568, 2014. [Article \(CrossRef Link\)](#).
- [97] Jiang, B., & Chan, W. K., "Input-based adaptive randomized test case prioritization: A local beam search approach," *Journal of Systems and Software*, 105, 91-106, 2015. [Article \(CrossRef Link\)](#).
- [98] Tyagi, M., & Malhotra, S., "Test case prioritization using multi objective particle swarm optimizer," in *Proc. of the Signal Propagation and Computer Technology (ICSPCT), 2014 International Conference, Ajmer, India*, 2014. [Article \(CrossRef Link\)](#).

- [99] Askarunisa, A., Punitha, K. A. J., & Ramaraj, N, "Test Case Generation and Prioritization for Composite Web Service Based on OWL-S," *Neural Network World*, 21(6), 519, 2011. [Article \(CrossRef Link\)](#).
- [100] Kumar, L., Rath, S. K., & Sureka, A, "Using source code metrics to predict change-prone web services: A case-study on ebay services," in *Proc. of Machine Learning Techniques for Software Quality Evaluation (MaLTeSQuE), IEEE Workshop on*, pp. 1-7, 2017. [Article \(CrossRef Link\)](#).
- [101] He, P., Li, B., Liu, X., Chen, J., & Ma, Y, "An empirical study on software defect prediction with a simplified metric set," *Information and Software Technology*, 59, 170-190, 2015. [Article \(CrossRef Link\)](#).
- [102] Kery, M, *Introduction to WinBUGS for Ecologists A Bayesian Approach to Regression, Anova, Mixed Models, and Related Analyses*, Boston: Elsevier Inc., 2010.
- [103] Wang, Z., Zhao, X., Zou, Y., Yu, X., & Wang, Z, "IMPROVED ANNEALING-GENETIC ALGORITHM FOR TEST CASE PRIORITIZATION," *COMPUTING AND INFORMATICS*, 36(3), 705-732, 2017. [Article \(CrossRef Link\)](#).
- [104] Siddiqui, J. H., & Khurshid, S, "Scaling symbolic execution using staged analysis," *Innovations Syst Softw Eng*, 9, 119-131, 2013. [Article \(CrossRef Link\)](#).
- [105] Zhai, K., Jiang, B., Chan, W. K., & Tse, T. H, "Taking Advantage of Service Selection: A Study on the Testing of Location-Based Web Services through Test Case Prioritization," in *Proc. of the Web Services (ICWS), 2010 IEEE International Conference, Miami, FL, USA*, 2010. [Article \(CrossRef Link\)](#).
- [106] Dobolyi, K., Soechting, E., & Weimer, W, "Automating regression testing using web-based application similarities," *Int J Softw Tools Technol Transfer*, 13, 111-129, 2011. [Article \(CrossRef Link\)](#).
- [107] Xin, Q., & Reiss, S. P, "Identifying Test-Suite-Overfitted Patches through Test Case Generation," in *Proc. of the ISSTA 2017 Proceedings of the 26th ACM SIGSOFT International Symposium on Software Testing and Analysis, Santa Barbara, CA, USA*, pp. 226-236, 2017. [Article \(CrossRef Link\)](#).
- [108] Smith, E. K., Barr, E. T., Goues, C. L., & Brun, Y, "Is the Cure Worse Than the Disease? Overfitting in Automated Program Repair," in *Proc. of the ESEC/FSE 2015 Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering, Bergamo, Italy*, pp. 532-543, 2015. [Article \(CrossRef Link\)](#).
- [109] Park, H., Ryu, H., & Baik, J, "Historical Value-Based Approach for Cost-cognizant Test Case Prioritization to Improve the Effectiveness of Regression Testing," in *Proc. of the The Second International Conference on Secure System Integration and Reliability Improvement, Yokohama, Japan*, 2008. [Article \(CrossRef Link\)](#).
- [110] Ansari, A. S. A., Devadkar, K. K., & Gharpure, P, "Optimization of Test Suite- Test Case in Regression Test," in *Proc. of the Computational Intelligence and Computing Research (ICCIC), 2013 IEEE International Conference, Enathi, India*, 2013. [Article \(CrossRef Link\)](#).
- [111] Athira, B., & Samuel, P, "Web services regression test case prioritization," in *Proc. of the Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference, Krakow, Poland, Poland*, 2010. [Article \(CrossRef Link\)](#).



**Muhammad Hasnain** received his Master Degree (Master of Science in Computer Science) from Abasyn University Islamabad Pakistan in 2016. From 2016-2017, he worked as a Lecturer in Army public college of management sciences Rawalpindi Pakistan. Currently, he is enrolled as a post graduate student in School of Information Technology, Monash University Malaysia. He has published a few conference as well as journal papers. His research interest is focused on regression testing, machine learning, and web services quality enhancement.



**Imran Ghani** is an Associate Professor at Mathematical and Computer Sciences Department, Indiana University of Pennsylvania, United States. He received his Master of Information Technology Degree from UAAR (Pakistan), M.Sc. Computer Science from UTM (Malaysia) and Ph.D. from Kookmin University (South Korea) "His research focus includes agile software development practices, semantics techniques, web services, and software testing.



**Muhammad Fermi Pasha** received his PhD degree in computer science from Universiti Sains Malaysia in 2010. After his PhD degree, he worked as a research fellow at Universiti Sains Malaysia. Currently he is working as a Lecturer in the School of Information Technology in Monash University Malaysia. His research interest is focused on computational neuroimaging, intelligent network security traffic analysis, and healthcare and radiology IT with emphasis on big data. He is an IEEE member and currently supervising PhD students in the latter mentioned research areas.



**Chern Hong Lim** received his Ph.D. and B.Sc.(Hons) degrees in computer science, majoring in Artificial Intelligence from University of Malaya, in 2015 and 2010 respectively. He is currently attached to School of Information Technology at Monash University Malaysia. He is an active member and a committee in IEEE Computational Intelligence Society (CIS) Malaysia Chapter. His research interests include artificial intelligence, machine learning, computer vision, image processing, and fuzzy logic; with focus on image or video content analysis and human motion inference.



**Seung Ryul Jeong** is a Professor in the Graduate School of Business IT at Kookmin University, Korea. He holds a B.A. in Economics from Sogang University, Korea, an M.S. in MIS from University of Wisconsin, and a Ph.D. in MIS from the University of South Carolina, U.S.A. Dr. Jeong has published extensively in the information systems field, with over 100 publications in refereed journals like Journal of MIS, Communications of the ACM, Information and Management, Journal of Systems and Software, among others. Dr. Jeong's areas of interest are Systems Implementation, Process Management, and Machine Learning.