

# The network of faults: a complex network approach to prioritize test cases for regression testing

Imrul Kayes<sup>1</sup> · Shafinaz Islam<sup>2</sup> · Jacob Chakareski<sup>3</sup>

Received: 25 November 2014 / Accepted: 21 July 2015  
© Springer-Verlag London 2015

**Abstract** Regression testing is performed to provide confidence that changes in a part of software do not affect other parts of the software. An execution of all existing test cases is the best way to re-establish this confidence. However, regression testing is an expensive process—there might be insufficient resources (e.g., time, workforce) to allow for the re-execution of all test cases. Regression test prioritization techniques attempt to re-order a regression test suite based on some criteria so that highest priority test cases are executed earlier. In this study, we prioritize test cases for regression testing based on the dependency network of faults. In software testing, it is common that some faults are the consequences of other faults (leading faults). Dependent faults can be removed if and only if the leading faults have been removed. Our goal is to prioritize test cases so that test cases that have exposed the leading faults in the system testing phase, are executed first in regression testing. The leading faults are modeled as the most central faults in the fault dependency network. We present ComReg, a test-case prioritization technique based on the dependency network of faults. We model a fault dependency network as a directed graph and identify leading faults to prioritize test cases for

regression testing. We use a centrality aggregation technique which considers six network representative centrality metrics to identify leading faults in the fault dependency network. We also discuss the use of fault communities to select an arbitrary percentage of the test cases from a prioritized regression test suite. We conduct a case study that evaluates the effectiveness and applicability of the proposed method. We obtain a fault dependency network from the development of a vocabulary learning software. We found that the fault network is a small-world graph with distinguishable community structure. The leading faults are common in all centralities and a re-ordering of test cases is feasible for regression testing based on those leading faults. Our method outperforms traditional regression testing prioritization techniques in detecting fault dependencies. Our modeling of the network of faults provides insights into the requirement of recognizing fault dependencies while re-ordering regression test suites for both research and practice. The dependency model needs further evaluation and improvement considering relevant resources (e.g., man-hours).

**Keywords** Software testing · Regression testing · Test-case prioritization

✉ Imrul Kayes  
imrul@mail.usf.edu

Shafinaz Islam  
shafinazeee10@gmail.com

Jacob Chakareski  
jacob@ua.edu

<sup>1</sup> Computer Science and Engineering, University of South Florida, Tampa, FL, USA

<sup>2</sup> Electrical and Electronic Engineering, Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh

<sup>3</sup> Electrical and Computer Engineering, University of Alabama, Tuscaloosa, AL, USA

## 1 Introduction

Regression testing is performed after a software is modified. The purpose of regression testing is to test the modified software with some test cases to re-establish our confidence that the software will perform according to the modified specification and the newly introduced changes do not hinder the behavior of the unchanged part of the software. In a development cycle, regression testing may begin after the detection and correction of faults in a tested software [1]. Regression

test suite aims to ensure that the evolution of an application does not result in a low-quality software product. However, due to the growing complexity of today's software systems, regression testing has become a challenging task. Moreover, recent trends in software development paradigms introduce new challenges in regression testing. For example, short and iterative "Agile" software development imposes restrictions and constraints on how regression testing can be performed within limited resources [2–4].

Intuitively, the best way to regain confidence from regression testing is to execute all existing test cases from a test suite. Unfortunately, regression testing is often directly associated with high costs. Beizer [5] points out that the regression testing accounts for as much as one-half the cost of software maintenance. One industrial collaborator of Elbaum et al. [6] reports that for one of their products of about 20,000 lines of code, the entire test suite requires seven weeks to run. Some of the most well-studied software failures, for example, the Ariane-5 rocket was blamed on the failure to test changes in a software system [7].

In general, test-case prioritization techniques seek to schedule test cases in an order so that the tester obtains maximum benefit, even if the testing is prematurely halted at some arbitrary point [2]. Regression test prioritization aims to re-order a regression test suite so that those test cases with highest priorities, according to some established criterion, are executed earlier in the process of regression testing than those with lower priorities [8]. Researchers have proposed various techniques for test-case prioritization to re-order the test cases for regression testing. These techniques focus on various aspects of product development, such as coverage-based approaches [8–11], requirement-based approaches [12, 13] and constraint-based approaches [14–16].

However, none of the solutions addressed dependencies among faults while prioritizing. In software testing, it is known that some faults are the consequences of other faults (commonly termed as leading faults). Experience shows that in a software development process, mutually independent faults can be directly detected and removed, but dependent faults can be removed if and only if leading faults have been removed [17]. In worst cases, fault dependencies can create a cascade of faults that can severely affect a software system. For example, in 1990, a fault in the failure recovery code of the AT&T led to cascading faults, which costs 9 h of downtime and at least 60 million in lost revenue [18]. Another example of cascading faults is the escalation of a divide-by-zero exception into a Navy ship's network that left the smart ship dead in the water [19]. Researchers hint that the Internet is also at risk of cascading failures [20]. We argue that a test case that reveals leading faults should be executed first in a regression testing process to get an early confirmation that the software is free from dependent faults.

We attempted a first step to prioritize regression testing based on fault dependency in [21]. We proposed an algorithm to prioritize test cases based on fault dependency. However, in [21], we only considered 1-hop neighborhood or dependencies of faults. This paper uses a fault dependency network to prioritize test cases for regression testing. We leverage the faults' respective positions in the network to determine leading faults (central faults in the network).

The contributions of this work are:

- First, we describe ComReg, which leverages the fault dependency network to prioritize test cases for regression testing. We present a directed graph model for the fault dependency network and identify leading faults (central faults) to prioritize test cases. Our identification of leading faults is based on a centrality aggregation technique. Centralities can represent the position of a fault in a fault network. We propose an aggregation of different representative centrality metrics (indegree, betweenness, closeness, eigenvector, pagerank, and hub centrality) into a final leading score to identify leading faults.
- Second, we discuss the use of fault communities to select  $X\%$  of the test cases from a prioritized regression test suite.
- Finally, we present a case study from the development of a subject software "Tarantula". We discuss the test cases written for the software, the faults it exposed after testing, and the fault network from the exposed faults. We show the identification of leading faults for prioritization and compare the effectiveness with traditional techniques. We also show fault communities for a selection of test cases from the prioritized regression test suite.

The rest of the paper is organized as follows. Section 2 introduces fault dependency-aware test-case prioritization technique, ComReg. Section 3 presents a case study and evaluates ComReg. Section 4 reviews related work and Sect. 5 concludes the paper.

## 2 Fault dependency-aware test-case prioritization

In this section, first, we provide a formal representation of the problem (Sect. 2.1). Then we present our regression testing prioritization technique, ComReg (Sect. 2.2). In Sect. 2.3, we describe how to select and run  $X\%$  of the test cases from a prioritized test suite.

### 2.1 Problem statement

Based on Elbaum et al. [11], we define a prioritization of test cases for regression testing as follows.

Given a test suite  $T$ , the set of permutations of  $T$  as  $PT$  and a function from  $PT$  to the real numbers as  $f$ , a prioritization of test cases for regression testing solution provides an ordered test suite  $T'$  such that for all  $T''$ ,  $f(T') \geq f(T'')$ , where  $T' \in PT$  and  $T'' \in PT$ .

$PT$  represents the set of all possible prioritization (orderings) of  $T$  and  $f$  is a utility function that, applied to any such ordering, yields an award value to that ordering. For example, let us we have  $n$  test cases as  $(T_1, T_2, T_3, \dots, T_n) \in T$ . From those test cases,  $n!$  orderings are possible. Test-case prioritization techniques attempt to find an order from  $n!$  number of orderings such that the order maximizes the utility function  $f$ .

Let us we have  $t$  test cases  $(T_1, T_2, T_3, \dots, T_t)$  in the test suite  $T$ . After running those test cases for a system testing, we get  $n$  faults such as  $(F_1, F_2, \dots, F_n)$  as  $F$ . There exists a relation from  $T$  to  $F$ ,  $R: T \rightarrow F$ , such that for each test case  $t \in T$  there exists none, single or multiple faults  $f \in F$ . Our goal is to prioritize the test suite and select  $X\%$  of the test cases for regression testing.

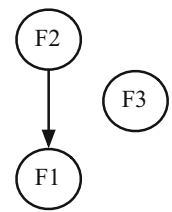
### 2.2 Our approach: ComReg

We propose ComReg, a fault dependency-aware test-case prioritization for regression testing. ComReg is based on the fact that mutually independent faults can be directly detected and removed, but dependent faults can be removed if and only if the leading faults have been removed [17]. A leading fault is the fault that causes dependent faults to occur. For example, consider a simple dictionary program, which has a *load* functionality to read all words and their meanings from text files, a *next word* functionality that allows users to browse words and a *random number generator* for generating a number for an arbitrary selection of a word list. The *next word* functionality is dependent on the *load* functionality in that if the system fails to read words and meanings, there is no way to browse the words. So, consider three following faults that occur.

1. *Fault F1*: words and meanings upload failure.
2. *Fault F2*: does not find the next word.
3. *Fault F3*: random generator does not show a random number.

Figure 1 shows the faults. We can draw an arrow from Fault F2 to Fault F1 to show the dependency of F2 on F1. In this case, F1 is a leading fault and F2 is a dependent fault. However, the Fault F3 is an independent fault (no arrow to or from F3 in Fig. 1). Leading faults tend to represent a relatively small portion of all the faults in a system. For example, Microsoft reports that 80 % of the errors and crashes in Windows and Office are caused by 20 % of the entire pool of faults [22]. We propose to prioritize

Fig. 1 Fault dependency



a regression test suite based on leading faults and to run  $X\%$  of the test cases which contain leading faults. So, in our fault dependency-aware test-case prioritization, an order of the test cases attempts to maximize the utility function  $f$  that determines the number of the leading faults. Fault dependencies could be due to output, data, control or module dependencies.

However, the problem of detecting leading faults is not trivially solvable. The challenge is due to the fact that faults have more than local effects. For example, consider a local effect, Fault A is dependent on Fault B, so A could not be removed before removing B. But if Fault A is dependent on Fault B and Fault B is dependent on Fault C, then Fault A could not be removed before removing Fault B and Fault C. We present various scenarios of fault dependencies considering two Faults F1 and F2 in Fig. 2. Some of the scenarios are listed below.

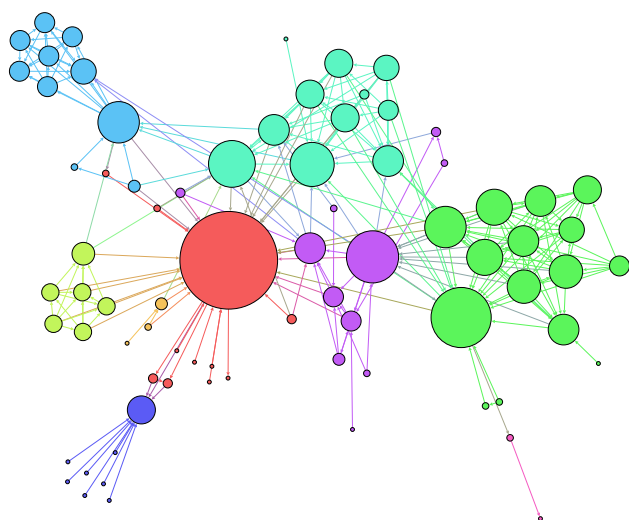
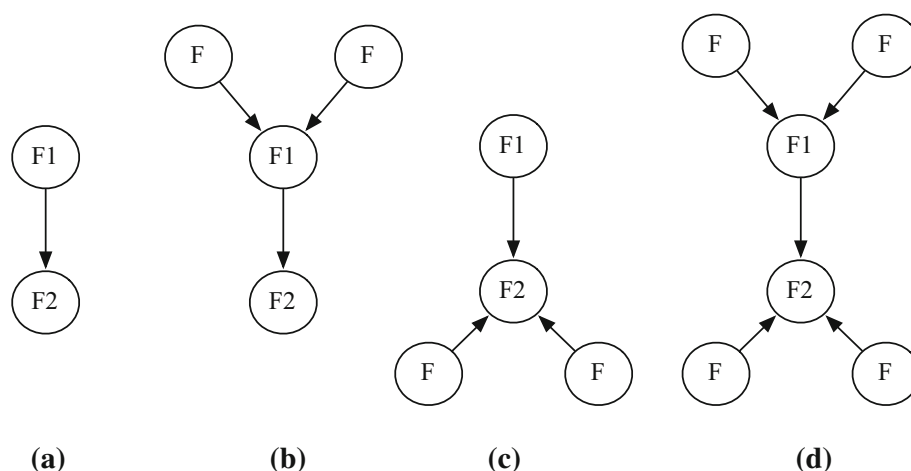
- Fault F1 is dependent on Fault F2 (Fig. 2a)
- Fault F1 is dependent on Fault F2, other faults are dependent on Fault F1 (Fig. 2b)
- Fault F1 is dependent on Fault F2, other faults are dependent on Fault F2 (Fig. 2c)
- Fault F1 is dependent on Fault F2, other faults are dependent on both Fault F1 and Fault F2 (Fig. 2d).

So, it appears that if we consider all faults and their dependencies, the situation becomes very complex. However, this situation can be captured by a complex network as shown in Fig. 3. The network is comprised of 77 faults (shown as nodes) and 254 dependencies (shown as edges). There is an edge from Fault A to Fault B if Fault A is dependent on Fault B. The leading faults in the network are those who occupy central positions.

Formally, we model a fault dependency network as a directed graph  $F = (V, E)$ , where a node  $v \in V$  is a fault and an edge  $e_{ij} \in E$  from  $v_i \in V$  to  $v_j \in V$  denotes that the fault  $v_i$  is dependent on the fault  $v_j$ . The number of nodes and edges are  $|V| = n$  and  $|E| = m$ , respectively. The directed graph can be represented by a  $n * n$  matrix  $F_{n*n}$ , where an entry  $F(i, j)$ :

$$F(i, j) = \begin{cases} 1 & \text{if } e_{ij} \in E \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

**Fig. 2** Examples of various fault dependencies considering two faults: Fault F1 and Fault F2. **a** Fault F1 is dependent on Fault F2. **b** Fault F1 is dependent on Fault F2 and other faults are dependent on Fault F1. **c** Fault F1 is dependent on Fault F2 and other faults are dependent on Fault F2. **d** Fault F1 is dependent on Fault F2 and other faults are dependent on both Faults F1 and F2



**Fig. 3** A fault dependency network of 77 nodes and 254 dependencies. Node size is proportional to indegree

The position of a node (fault) in the network can be represented in network analysis by different centrality metrics. For example, the larger the number connections a fault receives from its direct neighbors, the higher number of other faults depends on the fault. Without removing this fault, all dependent faults could not be removed. Moreover, the larger the number of paths between other pairs of faults a fault is part of, the more it can control the fault propagation between distant faults. Based on this intuition, we conjecture that a fault's 'position' is determined by and manifests via its 'centrality' in the fault network. Next, we will explain how we can use different network centralities to define a fault's position in the network.

We propose to aggregate different representative centrality metrics into a final leading score to identify the leading faults based on [23]. In [23], the authors used centrality aggregation technique to identify influential bloggers in a blogging network. The method has been also used to identify most

retained bloggers [24,25]. We define the leading score of a fault (node) as the average of the positions of that node in decreasing order of centrality scores over various centrality metrics. Specifically, each centrality metric assigns each node a score that can be used to order the nodes in decreasing order of importance (according to that centrality). This allows each fault to receive a rank according to each centrality metric: the first-ranked fault will be the most central one, the last ranked will be the one with the lowest centrality score. Faults having the same centrality score are given the same rank. A fault's final rank is the average rank over all centrality measures. We selected six representative centrality metrics as the focus of our study: indegree, betweenness, closeness, eigenvector, pagerank, and hub centrality.

Degree centrality is defined as the number of links that a node has. In a directed graph like fault dependency graph, two types of degree centralities are possible: indegree and outdegree centrality. For a node, the number of direct incoming connections is characterized as indegree of the node. On the other hand, the number of direct outgoing connections is characterized as outdegree of the node. Although simple, indegree centrality intuitively captures an important aspect of a fault's potential leading position: faults who have many incoming connections from many other faults are those that make other faults to depend. In our fault dependency graph  $F$ , the indegree centrality of a fault  $i$  can be represented by the following equation.

$$\text{indegree}(i) = \sum_{1 \leq j \leq n} F_{ji}. \quad (2)$$

Betweenness centrality, which measures the extent to which a node lies on the shortest paths between other nodes, was introduced as a measure for quantifying the control of a human on the communication between other humans in a social network [26]. Faults with high betweenness centrality may have considerable influence within a fault dependency



network by virtue of their control over fault propagation among other faults. The nodes with the highest betweenness are also the ones whose removal from the network will most disrupt communications between other nodes because they lie on the largest number of paths taken by faults [27]. Formally, the betweenness centrality of a node is the sum of the fraction of all-pairs shortest paths that pass through:

$$C(v) = \sum_{s,t \in V} \frac{\sigma(s, t|v)}{\sigma(s, t)} \tag{3}$$

where  $v$  is the set of nodes,  $\sigma(s, t)$  is the number of shortest  $(s, t)$  paths, and  $\sigma(s, t|v)$  is the number of those paths passing through some nodes  $v$  other than  $s, t$ . If  $s = t$ ,  $\sigma(s, t) = 1$ , and if  $v \in s, t$ ,  $\sigma(s, t|v) = 0$ . Our implementation of betweenness for this research is based on the Brandes algorithm [28].

Closeness centrality measures the mean distance from a node to other nodes, assuming that faults propagate along the shortest paths. Formally, the closeness centrality  $(C(x))$  of a node  $x$  is defined as follows:

$$C(x) = \frac{n - 1}{\sum_{y \in U, y \neq x} d(x, y)} \tag{4}$$

where  $U$  is the set of all nodes;  $n$  is the cardinality of  $U$ ;  $d(x, y)$  is the distance between two nodes, node  $x$  and node  $y$ ;  $d$  is the average distance between  $x$  and the other nodes (there are  $n - 1$  such distances possible and  $d$  is an average of those distances). In our fault dependency network, this centrality measure estimates the amount of faults a fault may have access to compared to other faults. Specifically, a fault with lower mean distance to others can reach others faster.

The centrality of a node does not only depend on the number of its adjacent nodes, but also on their relative importance. Eigenvector centrality allocates relative scores to all nodes in the network such that high-score neighbors contribute more to the score of the node. Formally, Bonacich [29] defines the eigenvector centrality  $C(v)$  of a node  $v$  as the function of the sum of the eigenvector centralities of the adjacent nodes, i.e.,

$$C(v) = 1/\lambda \sum_{(v,t) \in E} c(t) \tag{5}$$

where  $\lambda$  is a constant. This can be rewritten in vector notation, resulting in an eigenvector equation with well-known solutions.

Originally designed as an algorithm to rank web pages [30], PageRank computes a ranking of the nodes in a graph based on the structure of the incoming links. The algorithm assigns a numerical weighting to each node of a network with the purpose of “measuring” its relative importance within the network.

Hubs and authorities are other relevant centralities in the context of the fault network. In a graph, authorities are nodes that contain useful information on a topic of interest; hubs are nodes that know where the best authorities are to be found [27]. A high authority centrality node is pointed to by many hubs, i.e., by many other nodes with high hub centrality. A high hub centrality node points to many nodes with high authority centrality. These two centralities can play a significant role also in our work of finding leading faults. They can infer that the faults that have high hub and authority centrality are not only leading but also they are connected with leading faults.

### 2.3 Fault communities to select X % of test cases

A common approach of regression testing is to select and run X % of the test cases from a prioritized test suite. However, an optimal selection is always challenging. On one hand, a few selections of test cases might remain a significant portion of the software virtually untested. On the other hand, too many selections of test cases will require to test the entire system again. However, in a fault network, fault communities could be leveraged to select an X % of the test cases. Complex networks show communities in them: a community is a subset of nodes within which node to node connections are dense, but between which connections are less dense [31]. Communities are natural outcomes of real-world networks. For example, e-mail network [32], social application network [33], mobile communication network [34], blogging network [35], and yeast protein–protein interaction network [36] revealed community structures. Figure 3 shows communities in a fault network; nodes in a community are colored the same.

Newman proposed a community detection algorithm [37] based on modularity maximization. Modularity is a utility function that computes the quality of a particular division of a network into communities. It is defined as the fraction of the edges that fall within the given community minus the expected such fraction if the edges were distributed at random.

$$Q = (E_1 - E_2) \tag{6}$$

where  $E_1$  is the fraction of edges within communities and  $E_2$  the expected fraction of such edges.

The expected fraction of edges is typically evaluated within a random graph conditioned on the degree sequence of the original network. In that random graph, the probability of an edge between two nodes  $i$  and  $j$  is  $(k_i * k_j) / 2m$ , where  $k_i$  is the degree of node  $i$  and  $m$  is the total number of edges in the network. The modularity can then be written

$$Q = \frac{1}{2m} \sum_{ij} \left( A_{ij} - \frac{k_i * k_j}{2m} \right) \delta(c_i, c_j) \tag{7}$$

$$\delta(c_i, c_j) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ belong to the same community} \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

where  $A_{ij}$  is the matrix representation of the graph,  $\delta$  is the Kronecker delta,  $c_i$  is the label of the community to which node  $i$  is assigned.

The authors describe the modularity for an undirected graph. However, the modularity can be extended for a directed graph such as fault dependency network. In a random directed graph, the probability of an edge from node  $j$  to node  $i$  is  $(k_i^{\text{out}} * k_i^{\text{in}})/m$ . Then for the fault dependency network the above equation could be written as

$$Q = \frac{1}{m} \sum_{ij} \left( F_{ij} - \frac{k_i^{\text{out}} * k_i^{\text{in}}}{m} \right) \delta(c_i, c_j) \quad (9)$$

where  $F$  is a fault dependency matrix and  $F_{ij}$  is 1 if there is an edge from  $j$  to  $i$  and zero otherwise.

We propose to apply the community detection algorithm to uncover communities of the faults. After detecting communities, the faults in the same communities with a leading fault could be identified and corresponding test cases could be executed as a regression test. Moreover, all modules in a software are not the same in terms of fault tolerance. For example, login credential authentication or a module that processes financial transaction are more crucial than a module that prints documents. Furthermore, Pareto principle also (known as 80–20 rule) applies to software systems. The Pareto principle [38] states that for many events, roughly 80 % of the effects come from 20 % of the causes. The Standish Group’s report shows that in a software system, 45 % of the features are never used, 19 % of features are rarely used, 19 % of features are used sometimes, 13 % of features are used often and only 7 % of features are always used [39]. So, as a matter of fact, only 20 % of software features are often and always used. It becomes apparent that ensuring quality of those 20 % of software features is vital. Fault communities could be leveraged to ensure the quality of prioritized features (e.g., 20 % of software features). The leading faults and faults from their communities revealed by the test cases (which target prioritized features) could be used in selecting regression test cases. This way a regression testing can ensure a high customer satisfaction.

### 3 Case study and evaluation

The goal of the case study is to prioritize a test suite of size  $N$  and identify  $X$  % of the test cases from the suite for regression testing. To accomplish the goal, we developed a medium-scale English vocabulary learning software, “Tarantula”. In

the process of development, we wrote test cases, executed the test cases, applied our method ComReg to prioritize the test cases and identified  $X$  % of the test cases for regression testing. In this section, we first give a detailed overview of the functionalities of “Tarantula”. Then we describe the test cases and relevant faults that we have got after applying the test cases in an initial product. We explain the features of the Tarantula, so that readers can understand the underlying goals of the test cases. Finally, we present the leading faults using our centrality aggregation method and compare ComReg with traditional approaches. We also show and discuss the relevance of using community detection techniques in the fault network.

#### 3.1 An overview of Tarantula

The Tarantula is an English vocabulary learning software, which is built targeting high-frequency GRE words. When a user installs and runs the software, she is presented with 50 word lists. Figure 4 shows the first window of Tarantula. The word lists consist of 4054 words and when a user hovers a mouse on a word list icon, it shows the first and the last word in the word list. Users can click on a word list icon to exercise various features of that word list. However, users can use the random button (upper right corner in Fig. 4) to select a random word list (see Fig. 5). When users select a word list, relevant features of the word list are shown in a separate option window as shown in Fig. 6. There are several options available on the option window for users to learn and practice words of a word list. The options are: “Learn Word List”, “Multiple Choice”, “Reverse Challenge”, “Words Jam” and “Flip Words”. Users can click a radiobutton to select an option. The “Learn Word List” feature shows the words and their meanings from the word list serially (see in Fig. 7). Users can use “Next” and “Prev” buttons to view next and previous word, respectively. There is a



Fig. 4 First window of Tarantula

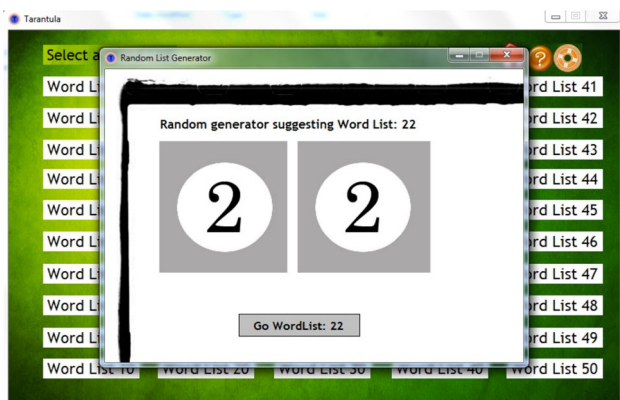


Fig. 5 Random window in Tarantula

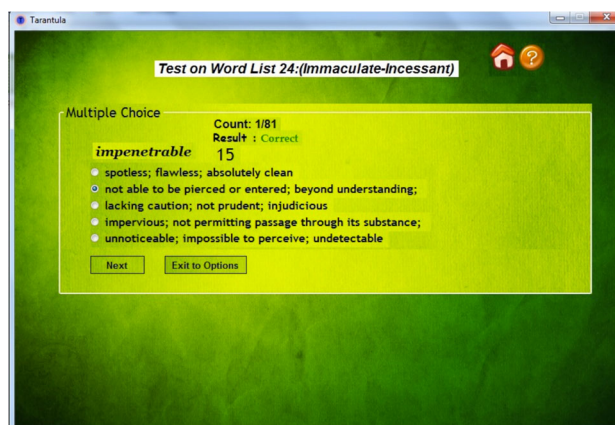


Fig. 8 Multiple Choice feature of Tarantula

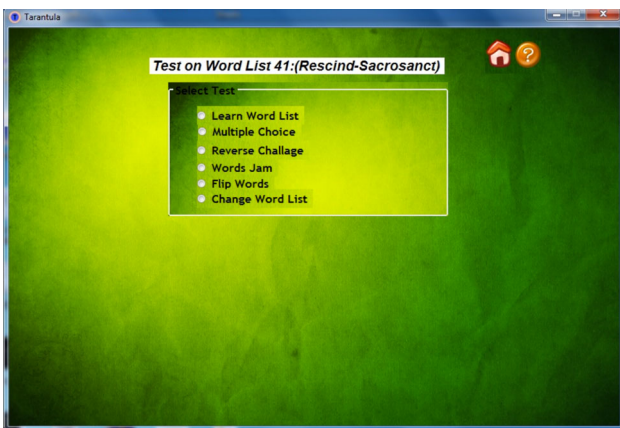


Fig. 6 Option window of Tarantula

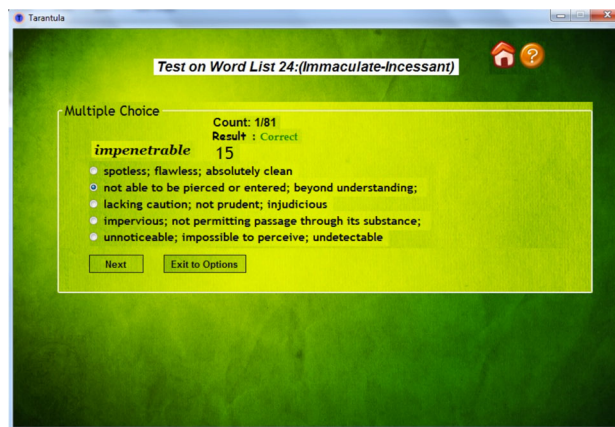


Fig. 9 Reverse Challenge Choice feature of Tarantula

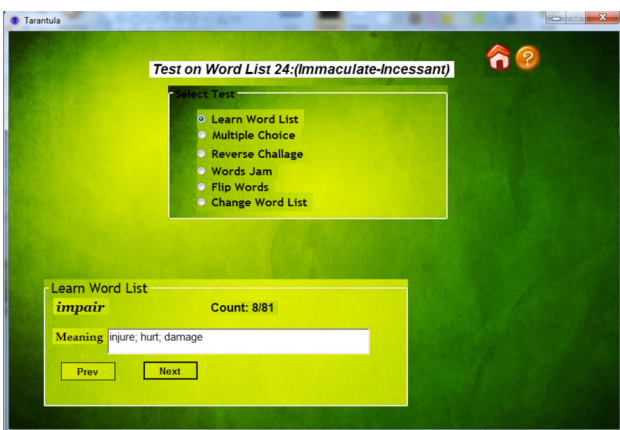


Fig. 7 Learn Word List feature of Tarantula

counter which indicates the serial number or position of the word in the word list.

The “Multiple Choice” feature shows a random word from the word list with five possible meanings (see Fig. 8). Meanings are from the same word list taken randomly, and of them one is appropriate for the word shown. When users click a meaning of the word, the “Result” label shows whether

the selection is right or wrong. There is a timer label which increases in each second to show how much time a user is taking. The “Count” label shows the number of words a user has practiced. Users can click on a “Next” button to get a new word. The “Reverse Challenge” feature is the opposite of “Multiple Choice” feature, where users are expected to match meanings to words (see Fig. 9). The “Words Jam” feature shows ten words from the word list on one side and their meanings from the word list on the other side. Words’ and their meanings’ positions on each side are random (see Fig. 10). Users have to click a word and then the meaning of the word (or vice versa). If the meaning of the word is right then both of them will disappear from the Jam. The “Flip Words” feature shows a random word from the word list to guess (see Fig. 11). Users can click “Flip” button to see the meaning of the word.

### 3.2 Development of Tarantula

The Tarantula is a desktop application, written in C# programming language and we used Microsoft Visual Studio



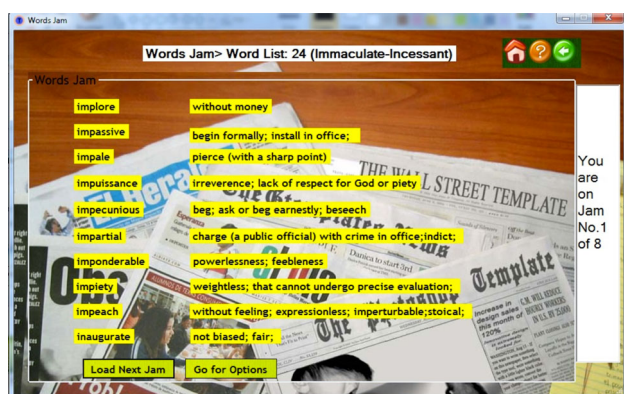


Fig. 10 Words Jam feature in Tarantula

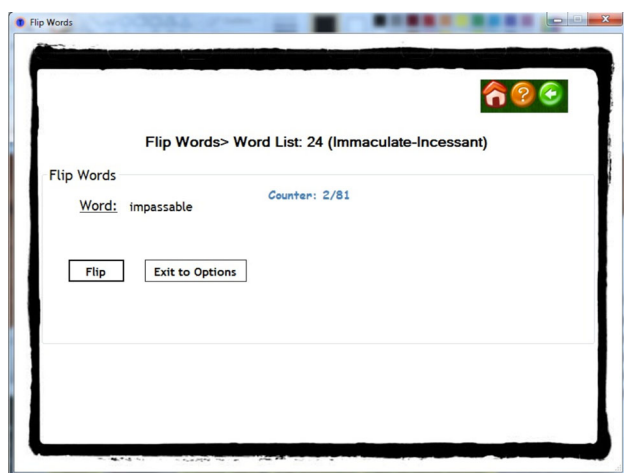


Fig. 11 Flip Words feature in Tarantula

2010 platform. The software consists of 19,390 lines of code. It can be installed from GitHub.<sup>1</sup> The code of Tarantula is also publicly available at GitHub.<sup>2</sup>

We first developed a web crawler using Python programming language to crawl a subset of HTML pages from [40]. These pages have all the words and their meanings. We parsed the crawled HTML pages using a parser (also written in Python). The parser went through all HTML pages and extracted words and meanings using regular expressions. Then we created the repository of fifty word lists (4054 words) from the extracted words and meanings. Finally, we used the repository as a word database for Tarantula.

### 3.3 Test cases and faults

Based on the required features, we wrote fifty test cases before the development. We ran the test cases after finish-

ing an iteration of the development cycle. Sixteen of the test cases revealed 23 faults. The test cases and the faults are below.

- *Test case #1*  
Action: click on a Word List icon to enable the system to load the words of the list with their meanings.  
Expected result: the Word List should be loaded with features.  
Fault #1: the Word List is unavailable due to missing of the file.
- *Test case #2*  
Action: select the *Learn Word List* option from the Radiobuttons of a word list.  
Expected result: a random word from the word list and its meaning should be shown.  
Fault #2: the word in the selected word list is not generated.  
Fault #3: the meaning in the selected world list is not available.
- *Test case #3*  
Action: click on the *Next* button on the *Learn Word List* feature.  
Expected result: a new random word from the word list and its meaning should be shown.  
Fault #4: *Next* button does not generate a random word.  
Fault #5: *Next* button does not generate a meaning.
- *Test case #4*  
Action: click on the *Previous* button event in the *Learn Word List* feature.  
Expected result: a new random word from the word list and its meaning should be shown.  
Fault #6: *Previous* button does not generate a random word.  
Fault #7: *Previous* button does not generate a meaning.
- *Test case #5*  
Action: check the *Count* functionality in the *Learn Word List* feature by clicking on the *Next* and *Previous* buttons.  
Expected result: *Count* should be increased by one on clicking *Next* button and count should be decreased by one on clicking *Previous* button.  
Fault #8: *Count* does not increase after clicking the *Next* button.  
Fault #9: *Count* does not decrease after clicking the *Previous* button.
- *Test case #6*  
Action: select the *Multiple Choice* option from the Radiobuttons of a word list.  
Expected result: a random word from the word list and its possible choices of meanings should be shown. The meanings are also from the same word list.  
Fault #10: the word is not generated.  
Fault #11: meanings are not available.

<sup>1</sup> <https://github.com/ImrulKayes/Tarantula/blob/master/Tarantula1.0.msi>.

<sup>2</sup> <https://github.com/ImrulKayes/Tarantula>.



- *Test case #7*  
Action: verify the functionality of the *Multiple Choice* option. Select the right meaning of the word. Select a wrong meaning of the word.  
Expected result: the system should show “Correct” if the choice is right, otherwise it will show a message saying that the choice is wrong.  
Fault #12: the “wrong” message is not shown.
- *Test case #8*  
Action: verify *Timer* functionality of the *Multiple Choice* option. Select the *Multiple Choice* option from the Radiobuttons of a word list. Then click the *Next* button.  
Expected result: the *Timer* should start from a zero value. It will increase by one after each second. Clicking the *Next* button should set it a zero value.  
Fault #13: the *Counter* does not increase.
- *Test case #9*  
Action: select the *Words Jam* option from the Radiobuttons of a word list.  
Expected result: ten words and their meaning should be shown for matching from the word list.  
Fault #14: words in *Words Jam* are missing.  
Fault #15: meanings in *Words Jam* are missing.
- *Test case #10*  
Action: click a word and then click its meaning in the *Words Jam* feature. Click a meaning and then click its corresponding word in the *Words Jam* feature.  
Expected result: the word and the meaning should be disappeared.  
Fault #16: the word does not disappear.
- *Test case #11*  
Action: in *Words Jam* feature, click a word and click a wrong meaning of the word.  
Expected result: the word and the meaning should not be disappeared.  
Fault #17: the word disappears.
- *Test case #12*  
Action: click *Load Next Jam* in *Words Jam* feature.  
Expected result: ten words and their meanings should be shown to match and *Jam counts* should be increased by one.  
Fault #18: *Jam Count* does not increase.
- *Test case #13*  
Action: select *Flip Words* option from the Radiobuttons of a word list.  
Expected result: a random word should be shown which will allow the users to guess the meaning of the word.  
Fault #19: the word is not generated.
- *Test case #14*  
Action: click the *Flip* button in *Flip Words* feature.  
Expected result: the meaning of the word should be shown and the text “Flip” of the button should be changed as “Next”.

Fault #20: meaning is not available.

Fault #21: text does not change.

- *Test case #15*  
Action: check the random word list generator, click the *Rand* button.  
Expected result: a random word list number should be generated.  
Fault #22: the random generator does not generate a random number.
- *Test case #16*  
Action: click the *Go to Word List* button of the random word list generator.  
Expected result: the word list should be loaded with features.  
Fault #23: the word list is not loaded.

### 3.4 Properties of the fault network

As discussed in Sect. 2.2, a fault dependency network is a directed graph  $F = (V, E)$ , where a node  $v \in V$  is a fault and an edge  $e_{ij} \in E$  from  $v_i \in V$  to  $v_j \in V$  denotes that the fault  $v_i$  is dependent on the fault  $v_j$ . The directed graph can be represented by a  $n * n$  matrix  $F_{n*n}$ , where an entry  $F(i, j)$ :

$$F(i, j) = \begin{cases} 1 & \text{if } e_{ij} \in E \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

The fault dependency matrix can be constructed after the system testing is done. For example, in a Scrum process, a fault review is usually done before the regression testing by examining reported faults on the “Dashboard”. In our case study, running the test cases we have got a fault dependency matrix  $F$  shown in Table 1. The fault dependency matrix has 23 faults and we associated relevant dependencies from Sect. 3.3. Figure 12 shows the largest component of the fault network (22 faults), where node size is proportional to indegree of the node. We used Gephi (<https://gephi.org/>) to visualize and obtain structural properties of the network. The structural properties of the fault network (largest component) are presented in Table 2.

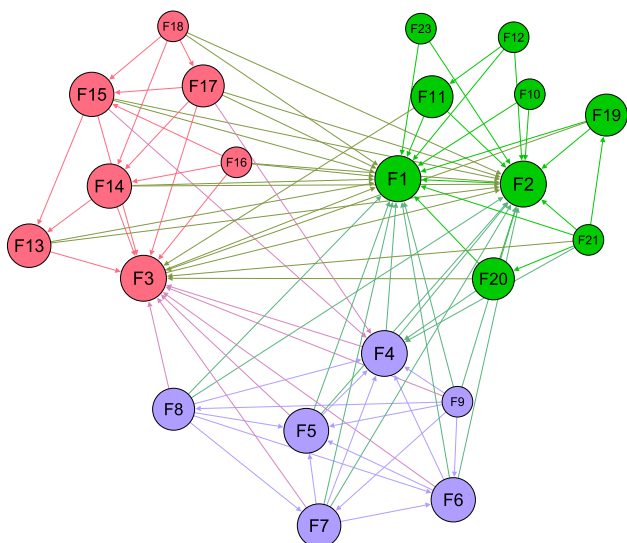
A notable characteristic of the fault network is the high clustering coefficient. Given a network  $G = (V, E)$ , the clustering coefficient  $C_i$  of a node  $i \in V$  is the proportion of all the possible edges between neighbors of the node that actually exist in the network [27]. The clustering coefficient is defined as:

$$C = \frac{3 * \text{Number of triangles}}{\text{Number of connected triples of the nodes.}} \quad (11)$$

In the fault network, for a node  $v_i \in V$ , there could be  $k_i(k_i - 1)$  links exist among the neighborhood of  $v_i$ , where

**Table 1** Fault dependency matrix

Faults ↓→	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
13	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
15	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
16	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
17	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
18	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0
19	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



**Fig. 12** The fault dependency network. Node size is proportional to indegree. Nodes in a community are colored the same (color figure online)

$k_i$  is the number of neighbors of  $v_i$ . So, the local clustering coefficient of fault  $v_i$  in the fault network is:

$$C = \frac{|\{e_{jk} : v_j, v_k \in \text{Neighbor}(V_i), e_{jk} \in E\}|}{k_i * (k_i - 1)} \tag{12}$$

**Table 2** Structural properties of the fault network

Nodes	22
Edges	97
Average indegree	3.95
Average path length	1.074
Average clustering coefficient	0.416

The clustering coefficient for the whole network is the average of the local clustering coefficients of all the nodes  $n$  [41]:

$$C = \frac{1}{n} \sum_{i=1}^n C_i. \tag{13}$$

A high clustering coefficient in the fault networks implies that a faults' connections are interconnected and have a greater effect on one another. The small average path length (1.074), comparable with that of the corresponding random graph of the same size (1.675), together with the high average clustering coefficient (the fault network has average clustering coefficient 0.416, where a same size random graph has

**Table 3** Top 10 faults according to each centrality measurement, sorted in increasing order by rank from left to right (color figure online)

<b>IDC</b>	Fault#1	Fault#2	Fault#3	Fault#4	Fault#5	Fault#14	Fault#15	Fault#6	Fault#13	Fault#7
<b>BC</b>	Fault#1	Fault#2	Fault#3	Fault#4	Fault#15	Fault#14	Fault#17	Fault#21	Fault#21	Fault#13
<b>CC</b>	Fault#1	Fault#2	Fault#3	Fault#4	Fault#15	Fault#5	Fault#6	Fault#7	Fault#8	Fault#9
<b>EC</b>	Fault#1	Fault#2	Fault#3	Fault#4	Fault#5	Fault#6	Fault#7	Fault#8	Fault#9	Fault#15
<b>PC</b>	Fault#1	Fault#2	Fault#3	Fault#4	Fault#5	Fault#13	Fault#14	Fault#15	Fault#6	Fault#11
<b>HC</b>	Fault#1	Fault#2	Fault#3	Fault#4	Fault#5	Fault#6	Fault#7	Fault#8	Fault#9	Fault#15

Faults common to all centralities are colored the same

IDC indegree centrality, BC betweenness centrality, CC closeness centrality, EC eigenvector centrality, PG pagerank centrality, HC hub centrality

**Table 4** Average rank of the top 10 faults

Faults' ID	Average rank
Fault #1	1.00
Fault #2	1.33
Fault #3	2.33
Fault #4	3.33
Fault #15	5.16
Fault #5	5.33
Fault #6	6.00
Fault #14	6.17
Fault #7	6.5
Fault #17	7.00

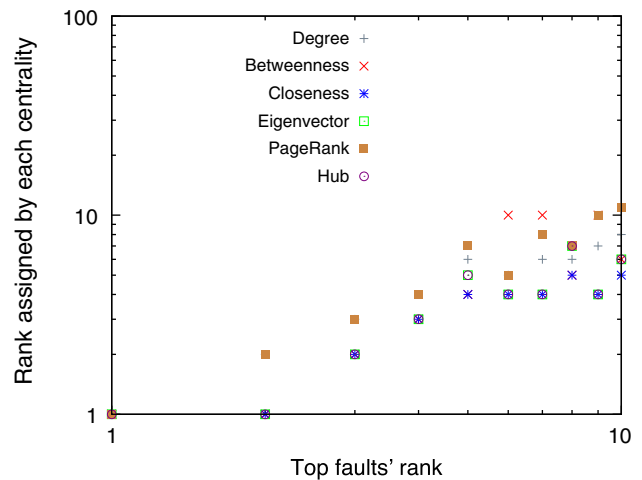
0.295), places the fault network in the category of small-world graphs [41].

### 3.5 Leading faults and prioritized test cases

As described in Sect. 2.2, we use centrality metrics to rank leading faults in the network. To manage the fault network and compute centralities, we used Python 2.7 with the NetworkX<sup>3</sup> library. Faults that appear among the top 10 in multiple centrality metrics are represented in color in Table 3. The average rank of the top 10 leading faults and their average ranks considering all centralities are shown in Table 4. Out of the ten leading faults listed on each centrality, six faults (60 %) are common in all the centralities. To observe more closely, we plot the ranks of the top 10 of the faults assigned by all centralities, shown in Fig. 13. As expected, a fault's assigned ranks from centralities form a cluster and together with all the clusters we can visualize a straight line. This shows that all the centralities tend to rank the same fault in the top.

Note Pareto principle described in Sect. 2.3—for many events, roughly 80 % of the effects come from 20 % of the causes. Pareto principle also holds for the fault dependency network. In the fault dependency network, 78 out of 97 (80.41 %) edges are incident on top 5 nodes out of 23 nodes (21.73 %). It shows that 80.41 % of the fault depen-

<sup>3</sup> <http://networkx.github.io/>.



**Fig. 13** Assigned rank of top 10 most central faults from all centralities

dencies are due to 21.73 % of faults—almost equal figures from Pareto principle.

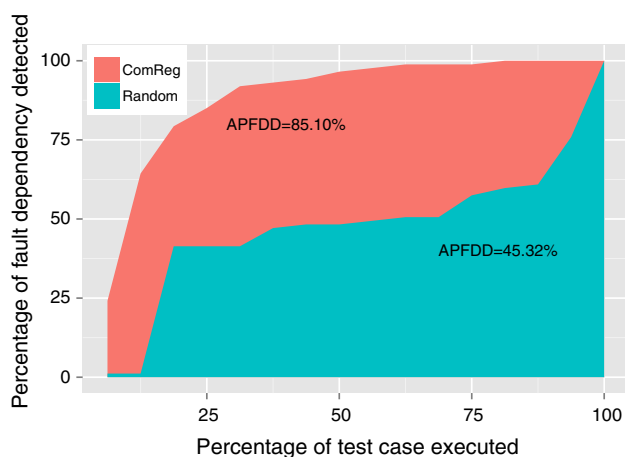
Finally, our prioritized ordering of the test cases for regression testing based on leading faults' exposure in test cases is: T1, T2, T3, T4, T9, T11, T5, T8, T14, T6, T10, T12, T13, T7, T16, T15 (T denotes a test case).

### 3.6 Effectiveness of ComReg

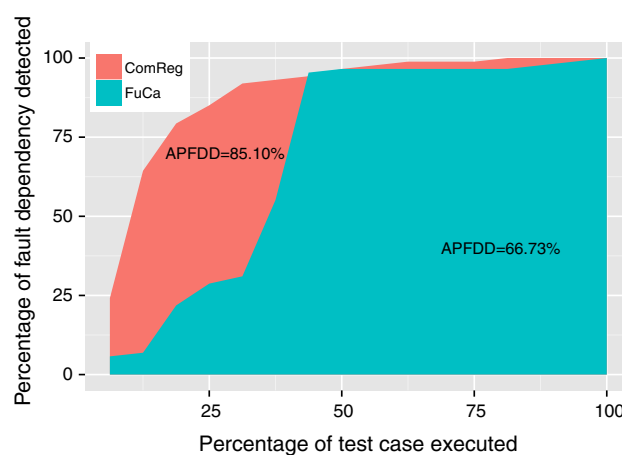
We used three techniques to prioritize our regression test suite and compared them to ComReg. We want to observe which method has faster fault dependency detection rate. The techniques are the following:

1. Prioritization using relevant slices (ReSI): ReSI prioritizes test cases taking into account the coverage requirements present in the relevant slices of the outputs of test cases [42].
2. Prioritization based on function call path (FuCa): FuCa leverages function call-level paths and prioritizes test cases based on those coverage paths [43].
3. Random prioritization (random): random prioritizes test cases based on a randomization algorithm.

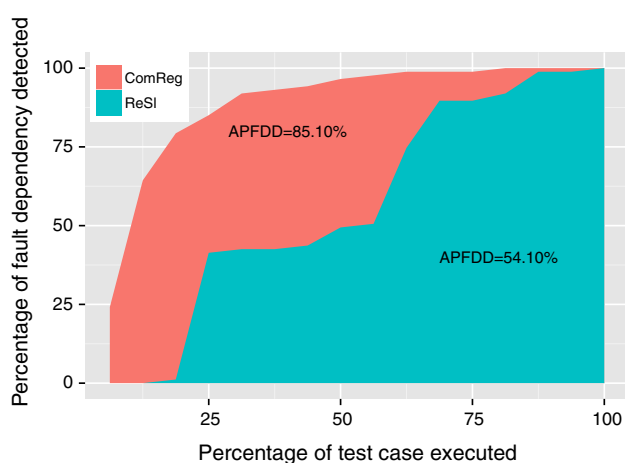




**Fig. 14** APFDD for the prioritized test cases using ComReg and random techniques



**Fig. 16** APFDD for the prioritized test cases using ComReg and function call path techniques [43]



**Fig. 15** APFDD for the prioritized test cases using ComReg and ReSI techniques [42]

We used a metric, average of the percentage fault dependency detected (APFDD) [21], to measure effectiveness of ComReg to the techniques described above. The APFDD quantifies how rapidly a prioritized test suite can detect dependency among faults. The values of the APFDD range from 0 to 100; higher value implies faster fault dependency detection. Figures 14, 15, and 16 show the percentage of test cases executed and the percentage of fault dependency detected for the test cases prioritized by ComReg and other methods (random, ReSI and FuCa, respectively). The areas under the curves represent the weighted APFDD. From the figures, we see that the random prioritization method performed the worst, yielded only 45.32 % APFDD. The ReSI and FuCa methods performed moderately, both were better than the random prioritization method with APFDD 54.10 % and 66.73 %, respectively. Our method ComReg provided the best value of the APFDD (85.10 %), hence outperformed the random, ReSI and FuCa methods in rapidly detecting fault dependencies.

### 3.7 Community detection techniques

We used a popular modularity maximization approach, Louvain method [44], to detect fault communities in the network. Louvain method is a greedy optimization method that attempts to optimize the modularity of a partition of the fault network. The optimization is performed in two steps: modularity maximization and community aggregation. In modularity maximization step, the method looks for “small” communities by optimizing modularity locally. In community aggregation step, it aggregates nodes which belong to the same community and builds a new network whose nodes are the communities. Two steps are repeated iteratively until a maximum of modularity is attained and a hierarchy of communities is produced. Applying the algorithm on the fault network, we have got three fault communities: community #1 (pink color nodes in Fig. 12) has seven faults, community #2 (green color nodes in Fig. 12) has nine faults and community #3 (violet color nodes in Fig. 12) has six faults. Leading faults are distributed among communities. For example, community #1, community #2 and community #3 have one, two and two top leading faults, respectively, out of top 5 leading faults. As discussed in the Sect. 2.3, leading faults and faults from their communities revealed by the test cases (which target prioritized features) could be used in selecting regression test cases. For example, leading fault Fault # 1’s community has eight faults originated from seven test cases (46.66 % of test cases).

## 4 Related work

Different solutions have been proposed to prioritize test cases for regression testing. In this section, we discuss test-case prioritization techniques from the literature.

*Coverage-based prioritization* techniques aim to achieve higher fault detection rates by maximizing early coverage. The solutions are inspired by the intuition that early maximization of structural coverage will also maximize early fault detection. Rothermel et al. proposed a family of techniques [9,10] for test-case prioritization based on several coverage criteria. They considered different types of coverages: branch-total, branch-additional, statement-total, statement-additional, fault exposing potential (FEP)-total, and FEP-additional. A branch-total coverage solution prioritizes test cases according to the number of branches covered by individual test cases. On the other hand, branch-additional prioritizes test cases according to the additional number of branches covered by individual test cases. Statement-total and statement-additional coverage-based solutions are similar to previous two approaches, but rather than considering branches, they consider statements. The FEP-total and FEP-additional are based on program mutation. Program mutation produces a mutant version of the program by introducing modifications to the program source. The prioritization techniques prioritize the test cases such that the test cases can reveal the difference between the original program and the mutant. The authors introduced a metric average percentage of fault detection (APFD) to quantify the success of a prioritization. Elbaum et al. [8,11] further proposed prioritization techniques covering coverage criterion at the function level, while Do et al. [45] considered the coverage criteria at the block level. Korel et al. discussed several model-based test prioritization heuristics in [46,47]. Their coverage criteria is system model; they identified elements of the model related to source-code modifications and applied heuristics to prioritize test cases so that early fault detection in the modified system is maximized. Jones and Harrold described a fine-grain coverage criterion in [48], which considers a modified condition/decision coverage.

*Requirement-based approaches* consider a software's requirement as a basis for prioritization of test cases. Srikanth et al. [12] prioritized test cases based on four factors: requirements volatility, customer priority, implementation complexity, and fault proneness of the requirements. Krishnamoorthi et al. [13] adopted a similar approach. Their prioritization is based on six factors: customer priority, changes in requirement, implementation complexity, completeness, traceability and fault impact. However, a potential weakness of requirement-based approaches is that requirement properties are subjective and thus estimations might be biased.

*Constraint-based approaches* consider different constraints and practical complications in test-case prioritization. Kim et al. [49] consider resource and time constraints. The resource and time constraint do not allow the execution of the entire test suite for a regression testing. They proposed a heuristic that uses historical information to

do test-case prioritization. Alspaugh et al. [16] consider a situation when regression testing is performed in a time-constrained environment. They empirically compared seven Knapsack solvers (e.g., greedy, dynamic programming and the core algorithm) and identified a test suite reordering that rapidly covers the test requirements and always terminates within a specified testing time limit. Walcott et al. [15] proposed a genetic algorithm-based time-aware test-case prioritization technique and empirically compared the approach with the initial ordering, the reverse ordering and two control techniques (random prioritization and fault-aware prioritization). They defined a metric to evaluate the effectiveness of prioritization in a time-constrained environment. Zhang et al. [14] also studied time-aware test-case prioritization problem. Their proposed test case prioritization is based on integer linear programming. They empirically showed that their two proposed techniques outperform genetic algorithm-based time-aware test-case prioritization and four other traditional techniques for test-case prioritization.

Researchers used a number of other criteria to prioritize test cases. Sherriff et al. prioritized test cases based on historical change records in [50]. They proposed a methodology for determining the effect of a software feature change and then prioritized regression test cases by gathering software change records and analyzing them through singular value decomposition. Leon and Podgurski [51] introduced distribution-based filtering and prioritized test cases based on the distribution of the profiles of test cases in the multi-dimensional profile space. Sampath et al. [52] prioritized test cases for web applications. They prioritized test suites by test lengths, frequency of appearance of request sequences and systematic coverage of parameter values and their interactions. Rummel et al. [53] introduced a prioritization technique based on data-flow analysis. They focused on the definition and use of program variables for the data-flow analysis. Jeffrey and Gupta [42] prioritized test cases using relevant slices. Qu et al. [54] prioritized test cases in a black box environment.

However, none of the above solutions considered dependencies among faults in prioritizing test cases for regression testing. In software testing, it is known that some faults are the consequences of other faults (leading faults). So, intuitively, test cases that revealed the leading faults should be executed first in a regression testing in order to get an early confirmation that software is free from dependent faults. In [21], we took the first step to prioritize regression testing based on fault dependency. We proposed an algorithm to prioritize test cases based on fault dependency. We also proposed a metric APFDD to quantify how rapidly a prioritized test suite can detect dependencies among faults. However, that work only considered 1-hop neighborhood or dependencies of faults. This paper leverages a fault network for prioritization.

## 5 Summary and discussions

In this paper, we have presented ComReg, which uses a fault dependency network to prioritize test cases for regression testing. We have modeled a fault dependency network as a directed graph and identified leading faults to prioritize test cases. We have leveraged a network centrality aggregation technique in the fault dependency network to identify leading faults. The centrality aggregation technique considers six representative centrality metrics such as indegree, betweenness, closeness, eigenvector, pagerank and hub centrality and offers a final leading score to identify the leading faults. Our discussions on fault communities shed light on selecting  $X\%$  of the test cases from a prioritized regression test suite. Finally, we have presented a case study and an evaluation of ComReg. In the case study, we have developed an English vocabulary learning software, “Tarantula” and identified leading faults from a fault network after running a set of test cases at the end of the first phase of the development. Evaluation results show that ComReg is better than traditional approaches in detecting leading faults. We have also showed the fault communities in the fault network for test-case selection from a prioritized regression test suite.

The fault dependency network might not be a connected graph. For example, our fault dependency network of Tarantula consists of two components. However, small-world networks tend to have giant components (e.g., [55–58]). A giant component is a connected subgraph that contains a majority of the entire graph’s nodes [59]. The giant component fills most of the network—usually more than half and not infrequently over 90 %—while the rest of the network is divided into a large number of small components disconnected from the rest [27]. Our small-world fault dependency network also has one giant component (22 nodes). So, if a fault network has a large number of nodes and if it shows a large number of connected components, the giant component could be leveraged to detect the leading faults.

Our work has multiple limitations. First, we built a subject software (“Tarantula”) to present a case study and show the effectiveness of our prioritization technique. The Tarantula is a medium-scale software, which lacks the rigorous development cycle of a typical commercial software. This might lead to a higher number of faults in the system testing. Using our proposed method in an industrial software testing setting could provide more insights.

Second, we do not consider the time and resources (e.g., testers) required to identify fault dependencies. If a software is poorly written with a lot of fault cascades, identifications of fault dependencies and their management might be costlier than running the full test suite.

Finally, some centrality algorithms (e.g., betweenness, closeness) used by ComReg are computationally expensive. This was not a major issue for the small fault dependency

graph discussed in this paper. However, for a large-scale fault dependency graph, an approximation algorithm (e.g.,  $k$ -path centrality [60]) with parallel implementation is required for efficiency.

## References

1. Leung HKN, White L (1989) Insights into regression testing [software testing]. In: Proceedings of conference on software maintenance, pp 60–69
2. Yoo S, Harman M (2012) Regression testing minimization, selection and prioritization: a survey. *Softw Test Verif Reliab* 22:67–120
3. Kayes I (2011) Agile testing: introducing PRAT as a metric of testing quality in scrum. *ACM SIGSOFT Softw Eng Notes* 36:1–5
4. Kayes I, Sarker M, Chakareski J (2013) On measuring test quality in scrum: an empirical study. [arXiv:1310.2545](https://arxiv.org/abs/1310.2545)
5. Beizer B (2003) *Software testing techniques*. Dreamtech Press, India
6. Malishevsky AG, Ruthruff JR, Rothermel G, Elbaum S (2006) Cost-cognizant test case prioritization. Department of Computer Science and Engineering, University of Nebraska-Lincoln, Technical Report
7. Hamlet D, Maybee J (2000) *The engineering of software: a technical guide for the individual*, 1st edn. Addison-Wesley Longman Publishing Co., Inc., Boston In:
8. Elbaum S, Malishevsky A, Rothermel G (2002) Test case prioritization: a family of empirical studies. *IEEE Trans Softw Eng* 28:159–182
9. Rothermel G, Untch R, Chu C, Harrold M (1999) Test case prioritization: an empirical study. In: Proceedings of IEEE international conference on software maintenance (ICSM’99), pp 179–188
10. Rothermel G, Untch R, Chu C, Harrold M (2001) Prioritizing test cases for regression testing. *IEEE Trans Softw Eng* 27:929–948
11. Elbaum S, Malishevsky AG, Rothermel G (2000) Prioritizing test cases for regression testing. In: Proceedings of the 2000 ACM SIGSOFT international symposium on software testing and analysis (ISSTA’00). ACM, New York, pp 102–112
12. Srikanth H, Williams L, Osborne J (2005) System test case prioritization of new and regression test cases. In: Proceedings of international symposium on empirical software engineering
13. Krishnamoorthi R, Sahaaya Arul Mary S (2009) Factor oriented requirement coverage based system test case prioritization of new and regression test cases. *Inf Softw Technol* 51:799–808
14. Zhang L, Hou SS, Guo C, Xie T, Mei H (2009) Time-aware test-case prioritization using integer linear programming. In: Proceedings of the eighteenth international symposium on software testing and analysis (ISSTA’09). ACM, New York, pp 213–224
15. Walcott KR, Soffa ML, Kapfhammer GM, Roos RS (2006) Timeaware test suite prioritization. In: Proceedings of the international symposium on software testing and analysis (ISSTA’06). ACM, New York, pp 1–12
16. Alspaugh S, Walcott KR, Belanich M, Kapfhammer GM, Soffa ML (2007) Efficient time-aware prioritization with knapsack solvers. In: Proceedings of the 1st ACM international workshop on empirical assessment of software engineering languages and technologies (WEASEL’07): held in conjunction with the 22nd IEEE/ACM international conference on automated software engineering (ASE’07). ACM, New York, pp 13–18
17. Huang CY, Lin CT (2006) Software reliability analysis by considering fault dependency and debugging time lag. *IEEE Trans Reliab* 55:436–450
18. Neumann PG (1990) Cause of AT&T network failure. *Risks Dig* 9:62



19. Slabodkin G (1998) Software glitches leave navy smart ship dead in the water. *Gov Comput News* 13:33727-1
20. Oppenheimer D, Ganapathi A, Patterson DA (2003) Why do internet services fail, and what can be done about it? In: USENIX symposium on internet technologies and systems, vol 67, Seattle
21. Kayes M (2011) Test case prioritization for regression testing based on fault dependency. In: Proceedings of 3rd international conference on electronics computer technology (ICECT'11), vol 5, pp 48–52
22. Rooney P (2002) Microsoft's CEO: 80-20 rule applies to bugs, not just features. <http://www.crn.com/news/security/18821726/microsofts-ceo-80-20-rule-applies-to-bugs-not-just-features.htm>. Accessed 25 Nov 2014
23. Kayes I, Qian X, Skvoretz J, Iamnitchi A (2012) How influential are you: detecting influential bloggers in a blogging community. In: Proceedings of the 4th international conference on social informatics. Springer, Berlin, pp 29–42
24. Kayes I, Zuo X, Wang D, Chakareski J (2014) To blog or not to blog: characterizing and predicting retention in community blogs. In: Proceedings of the international conference on social computing (SocialCom'14), vol 7. ACM, New York, pp 1–7:8
25. Kayes I, Zuo X, Wang D, Chakareski J (2014) Did you blog yesterday? Retention in community blogs. In: Proceedings of the international conference on social computing (SocialCom'14). ACM, New York, pp 16:1–16:2
26. Freeman L (1977) A set of measures of centrality based upon betweenness. *Sociometry* 40:35–41
27. Newman MEJ (2010) Networks: an introduction. Oxford University Press, USA
28. Brandes U (2008) On variants of shortest-path betweenness centrality and their generic computation. *Soc Netw* 30:136–145
29. Bonacich P (1972) Factoring and weighting approaches to status scores and clique identification. *J Math Sociol* 2:113–120
30. Page L, Brin S, Motwani R, Winograd T (1999) The pagerank citation ranking: bringing order to the web. Stanford InfoLab
31. Girvan M, Newman ME (2002) Community structure in social and biological networks. *Proc Natl Acad Sci* 99:7821–7826
32. Tyler JR, Wilkinson DM, Huberman BA (2005) E-mail as spectroscopy: automated discovery of community structure within organizations. *Inf Soc* 21:143–153
33. Nazir A, Raza S, Chuah CN (2008) Unveiling facebook: a measurement study of social network based applications. In: Proceedings of the 8th ACM SIGCOMM conference on internet measurement (IMC'08). ACM, New York, pp 43–56
34. Onnela JP, Saramäki J, Hyvönen J, Szabó G, Lazer D, Kaski K, Kertész J, Barabási AL (2007) Structure and tie strengths in mobile communication networks. *Proc Natl Acad Sci* 104:7332–7336
35. Kumar R, Novak J, Raghavan P, Tomkins A (2004) Structure and evolution of blogspace. *Commun ACM* 47:35–39
36. Chen J, Yuan B (2006) Detecting functional modules in the yeast protein–protein interaction network. *Bioinformatics* 22:2283–2290
37. Newman ME, Girvan M (2004) Finding and evaluating community structure in networks. *Phys Rev E* 69:026113
38. Pareto V (1971) Manual of political economy. Macmillan, London
39. Duong L (2009) Applying the “80–20 rule” with the standish group's statistics on software usage. <http://bit.ly/beBOfI>. Accessed 25 Nov 2014
40. WordHacker (2006) Gre word list. [http://www.wordhacker.com/en/article/Barron\\_gre\\_list\\_a.htm](http://www.wordhacker.com/en/article/Barron_gre_list_a.htm). Accessed 25 Nov 2014
41. Watts DJ, Strogatz S (1998) Collective dynamics of ‘small-world’ networks. *Nature* 393:440–442
42. Jeffrey D, Gupta R (2006) Test case prioritization using relevant slices. In: Proceedings of 30th annual international computer software and applications conference (COMPSAC'06), vol 1, pp 411–420
43. Zhi-hua Z, Yong-min M, Ying-ai T (2012) Test case prioritization for regression testing based on function call path. In: Fourth international conference on computational and information sciences (ICIS), pp 1372–1375
44. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. *J Stat Mech Theory Exp* 2008:P10008
45. Do H, Rothermel G, Kinneer A (2004) Empirical studies of test case prioritization in a JUnit testing environment. In: 15th international symposium on software reliability engineering (ISSRE'04), pp 113–124
46. Korel B, Tahat LH, Harman M (2005) Test prioritization using system models. In: Proceedings of the 21st IEEE international conference on software maintenance (ICSM'05). IEEE, pp 559–568
47. Korel B, Koutsogiannakis G, Tahat L (2008) Application of system models in regression test suite prioritization. In: IEEE international conference on software maintenance (ICSM'08), pp 247–256
48. Jones J, Harrold M (2003) Test-suite reduction and prioritization for modified condition/decision coverage. *IEEE Trans Softw Eng* 29:195–209
49. Kim JM, Porter A (2002) A history-based test prioritization technique for regression testing in resource constrained environments. In: Proceedings of the 24th international conference on software engineering (ICSE'02), pp 119–129
50. Sherriff M, Lake M, Williams L (2007) Prioritization of regression tests using singular value decomposition with empirical change records. In: The 18th IEEE international symposium on software reliability (ISSRE'07), pp 81–90
51. Leon D, Podgurski A (2003) A comparison of coverage-based and distribution-based techniques for filtering and prioritizing test cases. In: Proceedings of 14th international symposium on software reliability engineering (ISSRE'03), pp 442–453
52. Sampath S, Bryce R, Viswanath G, Kandimalla V, Koru A (2008) Prioritizing user-session-based test cases for web applications testing. In: Proceedings of 1st international conference on software testing, verification, and validation, pp141–150
53. Rummel MJ, Kapfhammer GM, Thall A (2005) Towards the prioritization of regression test suites with data flow information. In: Proceedings of the ACM symposium on applied computing (SAC'05). ACM, New York, pp 1499–1504
54. Qu B, Nie C, Xu B, Zhang X (2007) Test case prioritization for black box testing. In: 31st annual international of computer software and applications conference (COMPSAC'07), vol 1, pp 465–474
55. Faloutsos M, Faloutsos P, Faloutsos C (1999) On power-law relationships of the internet topology. *SIGCOMM Comput. Commun. Rev.* 29:251–262
56. Aiello W, Chung F, Lu L (2001) Random evolution in massive graphs. In: Proceedings of 42nd IEEE symposium on foundations of computer science, pp 510–519
57. Redner S (1998) How popular is your paper? An empirical study of the citation distribution. *Eur Phys J B-Condens Matter Complex Syst* 4:131–134
58. Kayes I, Kourtellis N, Quercia D, Iamnitchi A, Bonchi F (2015) The social world of content abusers in community question answering. In: Proceedings of the 24th international conference on world wide web (WWW'15). International world wide web conferences steering committee, Republic and Canton of Geneva, pp 570–580
59. Newman ME, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *Phys Rev E* 64:026118
60. Jiang K, Ediger D, Bader D (2009) Generalizing  $k$ -betweenness centrality using short paths and a parallel multithreaded implementation. In: international conference on parallel processing (ICPP'09), pp 542–549