

Fuzzy Inference System for Test Case Prioritization in Software Testing

Aron Karatayev , Anna Ogorodova , Pakizar Shamoi*

School of Information Technology and Engineering

Kazakh-British Technical University

Almaty, Kazakhstan

Email: *p.shamoi@kbtu.kz,

Abstract—In the realm of software development, testing is crucial for ensuring software quality and adherence to requirements. However, it can be time-consuming and resource-intensive, especially when dealing with large and complex software systems. Test case prioritization (TCP) is a vital strategy to enhance testing efficiency by identifying the most critical test cases for early execution. This paper introduces a novel fuzzy logic-based approach to automate TCP, using fuzzy linguistic variables and expert-derived fuzzy rules to establish a link between test case characteristics and their prioritization. Our methodology utilizes two fuzzy variables—failure rate and execution time—alongside two crisp parameters: Prerequisite Test Case and Recently Updated Flag. Our findings demonstrate the proposed system’s capacity to rank test cases effectively through experimental validation on a real-world software system. The results affirm the practical applicability of our approach in optimizing the TCP and reducing the resource intensity of software testing.

I. INTRODUCTION

Software testing has become increasingly important in recent years, with regression testing holding significant importance. Its main purpose is to ensure that changes, such as bug fixes or new features, do not introduce any new defects into previously validated code.

However, the regression testing process can often be labor-intensive and costly due to the number of test cases and the need to re-run them frequently. Given these constraints, the industry has turned to prioritization techniques to strategically order test cases, focusing initially on the most critical ones.

Developing an effective prioritization methodology, however, is not a trivial task. It needs to balance different factors, such as test execution time, the likelihood of finding defects, and the historical data of test failures. Additionally, many consistent metrics and evaluation criteria for prioritization can be applicable for certain cases but unsuitable for others. Developing an efficient prioritization strategy that is suitable for different projects is a challenging task. Therefore, a more flexible and convenient test case prioritization (TCP) system is needed to provide an efficient solution.

Currently, TCP systems underscore the significance of this process in software testing and highlight the potential of fuzzy logic, static parameters, and other algorithms like GA to enhance its effectiveness. However, more research is required to address the identified limitations and to explore potential synergies between mentioned approaches.

The current paper provides a methodology for automating the prioritization process by using fuzzy sets and logic. Fuzzy logic is applicable to this problem due to its ability to handle ambiguity and uncertainty, integrate expert knowledge, and combine multiple criteria into a resultant assessment. The main paper’s contribution is capturing experts’ knowledge in constructing the fuzzy inference system, specifically fuzzy sets and fuzzy rules. Additionally, analysis and evaluation were conducted on a real manually collected dataset from real industrial projects.

The paper is structured as follows. This Introduction is Section I. A review of the research on TCP is presented in Section II. Methodology, including fuzzy logic and data collection, are covered in Section III. Section IV presents the study results. Finally, concluding remarks and future works are provided in Section V.

II. RELATED WORKS

TCP is a critical aspect of software testing, ensuring that tests are executed in a sequence that maximizes the early detection of defects and thus optimizes the use of testing resources and time. Various research has explored different methodologies of TCP.

Traditional TCP techniques include basic prioritization strategies involving test adequacy criteria, such as code coverage, to represent different test case behaviours [1]. In one study, a technique for prioritization in regression testing was proposed [2], providing basic concepts of the TCP focused on testing changed code parts as soon as possible. Similarly, the other study introduced an optimized regression testing approach that highlights other possible metrics for prioritization, like the Average Percentage of Fault Detected (APFD), and discusses the need for optimization test case usage [3]. Model transformations in Model-Driven Engineering (MDE) require efficient regression testing due to the high cost of executing many test cases. Techniques focusing on test case rule coverage information have been proposed to prioritize test cases during regression testing of model transformations [4]. Another study introduced an ACO-based approach to optimize the order of test cases in regression testing based on the number of defects detected, execution time, and defect severity [5]. The other study proposed the concept of dynamic test case reordering in which the order of test cases is adjusted during

regression testing in response to events occurring during testing and the most recent test set [6].

Collective intelligence algorithms, particularly the Artificial Fish School Algorithm (AFSA) [7], have been used to solve TCP problems. These algorithms use average test point coverage percentage and adequate execution time to optimize the test case selection process, showing significant improvements in single-target and multi-target TCP problems. Techniques such as Cuckoo-ACO, Epistemic Particle Swarm Optimisation and Leaping Frog Algorithm have been used to address the challenging nature of MOTCP (multiple optimization TCP) [7] by focusing on aspects such as efficient execution time and biological epistasis. Nature-inspired algorithms, such as the hybrid cuckoo search algorithm [8], combine the advantages of different approaches to balance the trade-off between exploration and utilization.

Next, defect-based approaches prioritize test cases based on their ability to detect defects [9]. Mutation testing, a defect-based technique, introduces defects into the original program to create mutants. Test cases are prioritized based on their ability to "kill" these mutants.

In continuous integration (CI) environments, where software changes are often integrated, TCP plays a vital role [10]. Most approaches in CI are history-based, relying on the history of failures and test execution. Highly configurable systems present unique challenges for TCP due to the sheer number of configurations [11]. Approaches in this area often focus on functional and non-functional targets, with results showing that multi-target prioritization typically leads to faster defect detection compared to mono-target prioritization.

The other study presents an approach that prioritizes test cases based on historical failure data, test execution time, and domain-specific heuristics [12]. Another study introduced a new automated TCP technique specifically for web applications [13], focusing on identifying and prioritizing test cases related to database changes. The approach uses a dependency graph of functionality and prioritizes test cases based on their relevance to changed database elements, aiming for early defect detection.

Other methods include the use of Markov chains [14], modified genetic algorithm [15], [16], abstract test case prioritization (ATCP) techniques [17], a new metric based on the probability of error propagation to rank test cases in order of their probabilistic error-finding ability [18].

The application of fuzzy logic in TCP systems is another area of interest due to its potential to handle uncertainty and vagueness [19]. For example, in [20], authors proposed a fuzzy logic-based model for prioritization considering changes in the UML diagram and another based on code changes and compared both. Other studies used software agents and fuzzy logic to prioritize test cases [21], showing the potential benefits of this approach.

The field of test case prioritization has seen a variety of approaches, from traditional strategies focusing on code coverage to more complex and sophisticated techniques using collective intelligence algorithms, hybrid natural algorithms, and error-

based methods. The choice of technique often depends on the specific needs and constraints of the test environment, including factors such as the size and nature of the software, the presence of multiple targets, and the specific characteristics of the test infrastructure.

Integrating fuzzy logic and other advanced computational techniques paves the way for a more robust and reliable software testing framework.

III. METHODOLOGY

The main idea of our research is to develop a convenient TCP system and implement it to increase the time efficiency of the software testing process. The system aims to enhance the effectiveness of software testing by scheduling the execution of the test cases in an optimal order. This section outlines the specific methods used within the proposed methodology.

A. Fuzzy Sets and Logic Theory

This subsection provides an overview of the fundamental terms from fuzzy theory that we use in our research.

1) *Membership Functions and Fuzzy Sets*: A fuzzy set is a class of objects with a continuum of grades of membership ranging between 0 and 1. A fuzzy set A in a universe of discourse X is characterized by a membership function $\mu_A(x)$ which associates with each element x in X a real number in the interval $[0, 1]$, representing the degree of membership of x in A [22], [23]. Mathematically, it is expressed as:

$$\mu_A : X \rightarrow [0, 1]$$

where $\mu_A(x)$ = degree of membership of x in A .

2) *Fuzzy Operations*: The α -cut (Alpha cut) is a crisp set that includes all the members of the given fuzzy subset f whose values are not less than α for $0 \leq \alpha \leq 1$ [22], [23]:

$$f_\alpha = \{x : \mu_f(x) \geq \alpha\}$$

To connect α -cuts and set operations (A and B are fuzzy sets):

$$(A \cup B)_\alpha = A_\alpha \cup B_\alpha, \quad (A \cap B)_\alpha = A_\alpha \cap B_\alpha$$

3) *Fuzzy Rules*: Fuzzy rules are a key component of fuzzy logic, a form of multi-valued logic derived from fuzzy set theory to deal with approximate reasoning. Fuzzy rules express logical relationships between inputs and outputs of a system in a way that mimics human reasoning. These rules are typically expressed in the format: IF x is A THEN y is B , where x is the input variable, A and B are fuzzy sets defined on discourse's input and output universes, respectively [22], [23]. For example, *IF failure rate is high THEN priority is high*. In a fuzzy inference system, the rule base consists of a collection of IF-THEN rules.

B. Dataset collection

Our testing methodology necessitates a diverse and comprehensive pool of test cases. During the extensive exploration process, over 150 test cases from three distinct systems were undertaken. As a result of this work, we focused on only one system. Within this system, we further refined our dataset to

TABLE I: Test cases Dataset (e-commerce system).

#	Test Case	Exec. time	Failure rate	Prerequisite	Recently updated
1	Log in email	10	20	3	false
2	Log in phone	15	28	3	true
3	Register	45	40		false
4	Open profile	5	5	1	false
5	Log Out	5	2	1 / 2	false
6	Retrieve main page	5	60		false
7	Open Category	10	20		false
8	Search item	15	17		false
9	Retrieve item page	10	15	8	true
10	Add to cart	15	5	9	false
11	Open Cart	5	11		false
12	Cart Update	20	11	10, 11	false
13	Remove from Cart	10	5	10	false
14	Add to favorites	20	1		false
15	Retrieve favorites	15	18		false
16	Compare items	40	38		false
17	Checkout	80	46	10	false
18	Track order	30	14	17	false
19	Cancel Order	50	6	17	false
20	Refund order	90	37	17	false

include 48 well-structured and representative test cases. This selection was based on factors such as test case relevance, diversity, and complexity. The chosen test cases were then updated and formatted accordingly to ensure their compatibility with the TCP system. Table I presents a sample of the test cases dataset used in the experiment.

C. Proposed Approach

Fig. 1 presents the proposed fuzzy logic-based approach. Our methodology uses two fuzzy variables—*Failure rate* and *Execution time*—and two crisp parameters: *Prerequisite test case* and *Recently updated* flag.

Fuzzy logic provides a simple framework for dealing with imprecise and ambiguous information. We can assign fuzzy membership values to each test case, indicating their relative importance, by considering various prerequisite parameters like execution time and previous failure history. The test cases are then ranked according to these membership values, with higher values denoting higher priority.

We represent *Execution Time* as an ordered list of terms of the fuzzy variable $X = \text{"Execution Time"}$ using primary linguistic terms $L = \{\text{Short, Medium, High, Very High}\}$, specifying its level. Similarly, the fuzzy variable $Y = \text{"Failure Rate"}$ is represented by terms $L = \{\text{Very low, Low, Medium, High, Very High}\}$. Fig. 2 shows fuzzy input and output variables - the fuzzy sets for the *Execution Time* and *Failure Rate* variables and the output variable *Priority*, represented by $L = \{\text{Very low, Low, Medium, High, Very High}\}$. We use triangular and trapezoidal membership functions representing the linguistic labels expressing the fuzzy variable.

We initially defined fuzzy sets subjectively. One of the advancements we made later was the definition of the fuzzy partition based on the findings of a survey we conducted with QA experts. Three experts participated in this process. Building membership functions based on experts' opinions

TABLE II: Fuzzy rules composed by the experts for the inference system.

Rules	Execution time	Failure rate	Priority
Rule 1	High	Very Low	Very Low
Rule 2	Very High	Very Low	Very Low
Rule 3	Short	Very Low	Low
Rule 4	Medium	Very Low	Low
Rule 5	High	Low	Low
Rule 6	Very High	Low	Low
Rule 7	High	Medium	Low
Rule 8	Very High	Medium	Low
Rule 9	Short	Low	Medium
Rule 10	Medium	Low	Medium
Rule 11	Short	Medium	Medium
Rule 12	Medium	Medium	Medium
Rule 13	High	High	Medium
Rule 14	Very High	High	Medium
Rule 15	Short	High	High
Rule 16	Medium	High	High
Rule 17	High	Very High	High
Rule 18	Very High	Very High	High
Rule 19	Short	Very High	Very High
Rule 20	Medium	Very High	Very High

TABLE III: Parameters used for evaluation.

Parameter	Type	Fuzzy/Crisp	Description
Execution Time	int	Fuzzy	The amount of time it takes for a specific task
Failure Rate	double	Fuzzy	How often the test case fails under normal conditions
Prerequisite Test Case	int	Crisp	A test case that needs to be run first
Recently Updated	boolean	Crisp	A test case with recently modified functions

enabled us to bridge the gap that existed between the system and experts' knowledge.

So, the fuzzy partition was built using the Direct Rating method from [24]. Various methods can be used to question the user to construct fuzzy sets. The Direct Rating method was chosen because it is simple, efficient, and requires fewer converging iterations. In Direct Rating, the user must choose a value from a list of possible linguistic values corresponding to a specific linguistic variable. In our experiment, experts (QA engineers) were asked to categorize different values into one of several categories corresponding to terms. The collected personal judgments were then used to define fuzzy sets. The tailored functions were obtained by limiting the shape of fuzzy sets to triangular or trapezoidal membership functions, as shown in Fig. 2.

Expert-derived fuzzy rules are used to establish a relationship between test case characteristics and prioritization. Table II presents the fuzzy rules. As can be seen, each rule considers the fuzzy sets of *Execution Time* and *Failure Rate* and outputs a *Priority* level for the test case. The fuzzy inference system processes these rules to determine the priority of each test case.

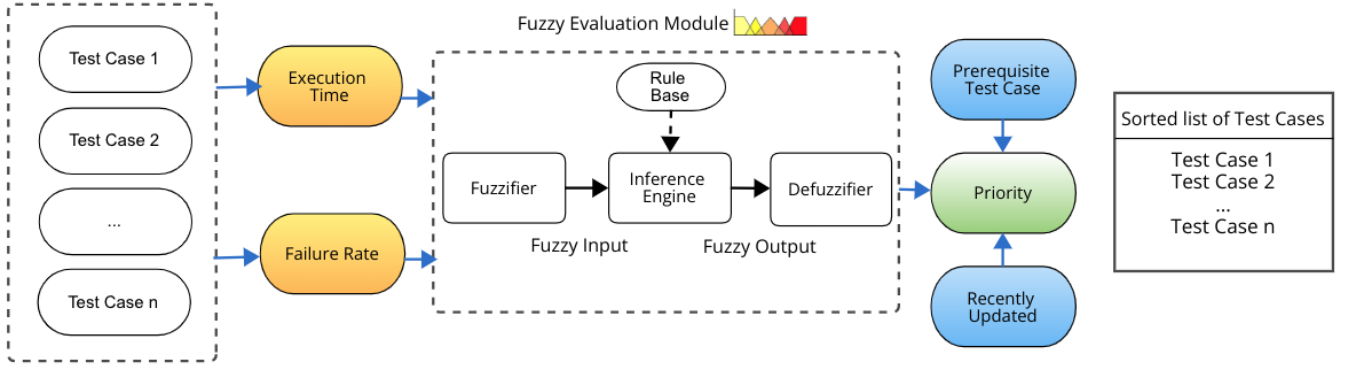


Fig. 1: General Workflow of Test Cases Prioritization System.

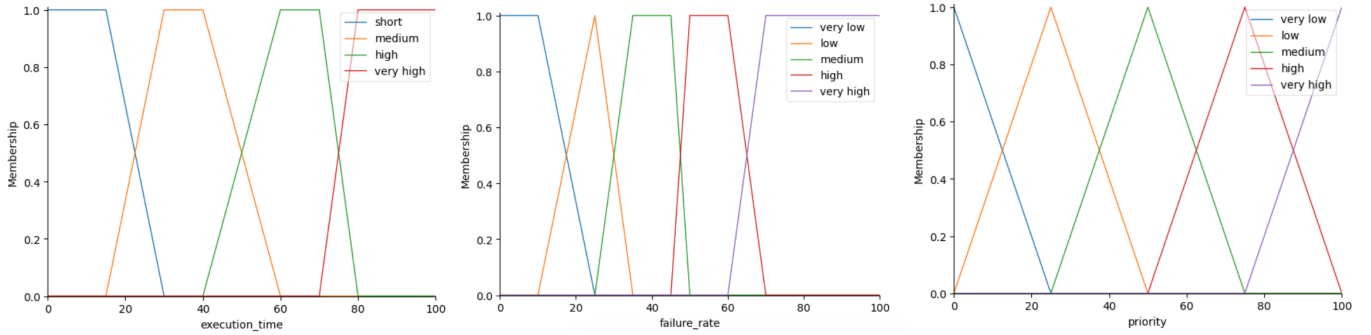


Fig. 2: Input fuzzy sets for *Execution Time*, *Failure rate* and Output fuzzy sets for *Priority*

In addition to the fuzzy logic components, the system also considers two crisp variables: *Prerequisite Test Case* and *Recently Updated Flag* (see Table III). Crisp set specifies value as either 0 or 1, just like in classical binary logic.

The *Prerequisite Test Case* parameter checks whether a test case has a prerequisite test case that needs to be run first. Prerequisite parameters played a crucial role in optimizing the test run flow. We recognized that some test cases had to be run multiple times to test different flows. Therefore, we incorporated the prerequisite parameter to establish a soft hierarchy among the test cases, allowing us to minimize the number of test runs. This approach also enabled us to ensure that all prerequisite test cases were run before the ones that depended on them.

The *Recently Updated Flag* marks test cases that test recently modified functions. This parameter is included based on the assumption that recently changed functions are more likely to contain faults. Test cases marked with this flag are given a higher priority in the system; if the associated function was recently updated, the priority level was raised to the next level.

The workflow of the proposed methodology is as follows:

- 1) **Input.** An array of test cases is received, each defined by a set of parameters. These include the failure rate, execution time, a prerequisites array, and a recently updated flag.

- 2) **Prioritization.** Upon receiving the input, the methodology begins the prioritization process based on fuzzy logic. Each test case's execution time and failure rate are fuzzified and passed to the fuzzy inference system to eventually compute its priority. If a test case has been marked as recently updated, it gets promoted to a higher priority level.

- 3) **Outputs.** Finally, after all test cases have been evaluated and prioritized, we output a sequence of test cases. This output represents the optimized order in which the test cases should be executed.

IV. RESULTS

A. Performance Evaluation

A comparative study was performed to evaluate the effectiveness of the proposed fuzzy logic-based TCP system using three lists of test cases: an unsorted list (the way tests were executed before), a list sorted by a QA engineer, and the list formed based on the proposed methodology. The experiment results are presented in Table IV.

Three key metrics were used to assess the ordering in each list:

- 1) The number of executed test cases. This metric is critical as certain flows necessitate running the same test cases multiple times. Efficient prerequisite mechanics should optimize this number.

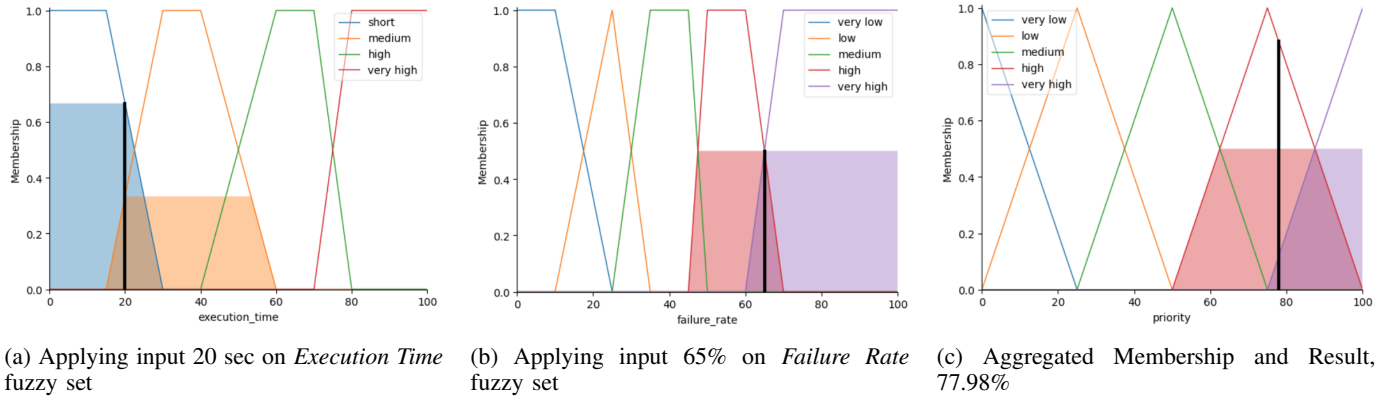


Fig. 3: Simulation Results.

TABLE IV: Evaluation Results

	Unsorted	QA expert	Our methodology
Executed tests amount	32	25	27
Spent time (total)	725 sec	605 sec	630 sec
Failures found	5	5	5

- 2) The total time (in seconds) spent running the tests.
- 3) The number of errors that were identified during testing.

The most important test case was "Retrieve main page," which scored 75, while the one with the lowest priority was "Remove from cart," which scored 20.89. The fuzzy methodology provided satisfactory results compared to manual sorting done by QA specialists closely involved in the project. The proposed methodology significantly reduces the time and effort required by QA engineers.

Despite the room for improvement, the system demonstrated promising results compared to the unsorted list and list formed by an expert. This improvement represents a positive step forward, justifying further investigation and refinement when applying fuzzy logic to TCP.

B. Example

To simulate the fuzzy system, we must first specify the inputs and then use the defuzzification approach. Consider the following example: the *Execution Time* and *Failure Rate* are 20% and 65%, respectively (see Fig. 3). Fuzzy aggregation then aggregates the output membership functions using the maximum operator. Next, we must defuzzify to acquire a priority, which we accomplish using the centroid approach. Aggregation using fuzzy criteria yielded an overall *Priority* of 77.98%. Fig. 3 depicts the visualized simulation results.

V. CONCLUSION

This paper presents a novel fuzzy logic-based approach for automating test case prioritization. It uses fuzzy linguistic variables and expert-derived fuzzy rules to build a link between test case attributes and prioritization. Our methodology combines two fuzzy variables (Failure Rate and Execution Time) with two crisp parameters (Prerequisite Test Case and Recently

Updated Flag). This combination enabled us to benefit from both the fuzzification and fixed hierarchy approaches.

The methodology is highly comparable with manual testing, with the intention of enhancing the efficiency of junior QA specialists in particular. Our experiments on a real-world software system show that the suggested system can efficiently rank test cases.

The key advantage of the methodology compared to similar studies is that real experts participated in knowledge base generation.

The proposed approach has certain limitations. Specifically, two of the main limitations are sensitivity to membership function design and weak generalization across projects. The fuzzy sets and rules specified for one project may not be directly relevant to the other. So, customization is required for each new project or software system, including the need to collect deep domain knowledge and expert insights. This process can be complex and time-consuming. The fuzzy logic-based prioritization system can adapt to different contexts only with the help of expert knowledge.

As for future work, we aim to improve the methodology by including additional prioritization parameters. We plan to integrate this system with the existing automated testing frameworks. This will help streamline the collection of essential parameters such as failure rate and execution time. We plan to extract these parameters directly from test run logs provided by the framework, which will streamline data acquisition and refine our prioritization process.

REFERENCES

- [1] Q. Zhang, C. Fang, W. Sun, S. Yu, Y. Xu, and Y. Liu, "Test case prioritization using partial attention," *Journal of Systems and Software*, vol. 192, p. 111419, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0164121222001285>
- [2] Y. Singh, A. Kaur, and B. Suri, "A new technique for test case prioritization for regression testing," *http://dx.doi.org/10.1177/0971890720050112*, vol. 9, pp. 77–85, 1 2005. [Online]. Available: <https://journals.sagepub.com/doi/pdf/10.1177/0971890720050112>
- [3] A. Ansari, A. Khan, A. Khan, and K. Mukadam, "Optimized regression test using test case prioritization," *Procedia Computer Science*, vol. 79, pp. 152–160, 1 2016.

- [4] S. Iqbal and I. Al-Azzoni, "Test case prioritization for model transformations," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, Part B, pp. 6324–6338, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1319157821002147>
- [5] R. Yan, Y. Chen, H. Gao, and J. Yan, "Test case prioritization with neuron valuation based pattern," *Science of Computer Programming*, vol. 215, p. 102761, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167642321001544>
- [6] N. Kaushik, M. Salehie, L. Tahvildari, S. Li, and M. Moore, "Dynamic prioritization in regression testing," in *2011 IEEE Fourth International Conference on Software Testing, Verification and Validation Workshops*, 2011, pp. 135–138.
- [7] Y. Xing, X. Wang, and Q. Shen, "Test case prioritization based on artificial fish school algorithm," *Computer Communications*, vol. 180, pp. 295–302, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0140366421003510>
- [8] A. Bajaj and O. P. Sangwan, "Discrete cuckoo search algorithms for test case prioritization," *Applied Soft Computing*, vol. 110, p. 107584, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1568494621005056>
- [9] F. Farooq and A. Nadeem, "A fault based approach to test case prioritization," in *2017 International Conference on Frontiers of Information Technology (FIT)*, 2017, pp. 52–57.
- [10] J. A. Prado Lima and S. R. Vergilio, "Test case prioritization in continuous integration environments: A systematic mapping study," *Information and Software Technology*, vol. 121, p. 106268, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0950584920300185>
- [11] J. A. Parejo, A. B. Sánchez, S. Segura, A. Ruiz-Cortés, R. E. Lopez-Herrejon, and A. Egyed, "Multi-objective test case prioritization in highly configurable systems: A case study," *Journal of Systems and Software*, vol. 122, pp. 287–310, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0164121216301935>
- [12] D. Marijan, A. Gotlieb, and S. Sen, "Test case prioritization for continuous regression testing: An industrial case study," in *2013 IEEE International Conference on Software Maintenance*, 2013, pp. 540–543.
- [13] D. Garg and A. Datta, "Test case prioritization due to database changes in web applications," in *2012 IEEE Fifth International Conference on Software Testing, Verification and Validation*, 2012, pp. 726–730.
- [14] G. Barbosa, Érica Ferreira de Souza, L. B. R. dos Santos, M. da Silva, J. M. Balera, and N. L. Vijaykumar, "A systematic literature review on prioritizing software test cases using markov chains," *Information and Software Technology*, vol. 147, p. 106902, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0950584922000623>
- [15] D. K. Yadav and S. Dutta, "Regression test case prioritization technique using genetic algorithm," *Advances in Intelligent Systems and Computing*, vol. 509, pp. 133–140, 2017. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-10-2525-9_13
- [16] P. Konsaard and L. Ramingwong, "Total coverage based regression test case prioritization using genetic algorithm," in *2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2015, pp. 1–6.
- [17] R. Huang, W. Zong, D. Towey, Y. Zhou, and J. Chen, "An empirical examination of abstract test case prioritization techniques," in *2017 IEEE/ACM 39th International Conference on Software Engineering Companion (ICSE-C)*, 2017, pp. 141–143.
- [18] J. Joo, S. Yoo, and M. Park, "Poster: Test case prioritization using error propagation probability," in *2020 IEEE 13th International Conference on Software Testing, Validation and Verification (ICST)*, 2020, pp. 398–401.
- [19] P. Shamo and A. Inoue, "Computing with words for direct marketing support system," in *Proceedings of the 23rd Midwest Artificial Intelligence and Cognitive Science Conference, MAICS 2012*, 2012, pp. 19–26.
- [20] M. Al-Refai, W. Cazzola, and S. Ghosh, "A fuzzy logic based approach for model-based regression test selection," *Proceedings - ACM/IEEE 20th International Conference on Model Driven Engineering Languages and Systems, MODELS 2017*, pp. 55–62, 11 2017.
- [21] C. Malz, N. Jazdi, and P. Gohner, "Prioritization of test cases using software agents and fuzzy logic," in *2012 IEEE Fifth International Conference on Software Testing, Verification and Validation*, 2012, pp. 483–486.
- [22] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 6 1965.
- [23] L. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-iii," *Inf. Sciences*, vol. 9, no. 1, pp. 43–80, 1975.
- [24] J. Cunningham, "Determining an optimal membership function based on community consensus in a fuzzy database system," in *Proceedings of the 44th Annual Southeast Regional Conference*, ser. ACM-SE 44. New York, NY, USA: Association for Computing Machinery, 2006, p. 632–637. [Online]. Available: <https://doi.org/10.1145/1185448.1185587>