Next Release Problem: A Systematic Literature Review

Umer Iqbal[1] Khubaib Amjad Alam[2]

Abstract

Agile software development is a widely used software development methodology which welcomes highly changing customers' requirements. In Agile software development process, the whole software is delivered into a series of small releases. Each release incorporates a subset of whole software requirements. The selection of requirements to be incorporated in the next release is a complex activity. It was first termed as "Next Release Problem" by Bagnal. Many techniques were proposed later on to solve NRP The main objectives of this research are: (1) to classify NRP papers according to four criteria: techniques used, datasets used, objectives (either single or multiple), publication channels and trends; and (2) to analyze these studies from three perspectives: study objectives, optimization techniques to solve NRP and limitations of study. We performed a systematic literature review on NRP studies published in the period 2010-2018 and reviewed them on an automated four electronic databases. We identified a total of 27 studies published between 2010 and 2018 and classified them on predefined classification criteria. Based on the findings of this research it is concluded that multi-objective optimization techniques are the most widely used techniques. Among multi-objective optimization techniques applied in the context of NRP, NSGA-II provides the best solution both in term of convergence speed and solution quality while for the single-objective optimization problem, Simulated Annealing provides promising results. It is also observed that customer's satisfaction is widely used objectives to be maximized in either single objective or multi-objective optimization techniques. Furthermore, 10 real-world datasets were identified during this research. It is observed that the latest optimization techniques are given less attention to solving NRP which have shown promising results in many cases as compared to techniques applied to cater NRP.

Index Terms: Next Release Problem, Next Release Planning, Release Planning, Systematic Literature Review

1 Introduction

The software development process is a process of breaking down software development work into distinct phases to carry out the development process more competently. Many different software process models are proposed until now but the agile software development process gained the most reputation among all. A recent survey showed that almost 70 % of companies use agile sometimes, often and always [1]. In Agile, whole software is delivered in the form of small releases and each release incorporates a small number of whole system requirements. The selection of requirements which will be chosen in the next release is a major challenge.

[1] *National University of Computer and Emerging Sciences, Islamabad | umerpervaiz12@gmail.com*
[2] *National University of Computer and Emerging Sciences Islamabad | khubaib.ajmad@nu.edu.pk*

The selection involves taking care of many criteria's such as customer satisfaction, overall development cost, risk, development time on the basis of which the requirements are chosen for the next release. The Bagnall [2] termed this problem as "Next Release Problem" and proposed a solution based on single objective optimization. He also concluded that NRP is NP-Hard problem and thus the best way to solve the problem is to use heuristic methods. Over time many different techniques and methods are applied on different datasets to solve NRP.

Few studies provide a detailed overview of the techniques proposed for the resolution of NRP. Another problem in the previous studies is that each study focuses on some specific method or technique and does not cover the entire domain. The main goal and purpose of this research is to introduce a broader and precise overview of almost all of the commonly used latest techniques in NRP in the form of a Systematic Literature Review. We followed the guidelines of Kitchenham [3].

The organization of the paper is as follows: The detailed steps and the structured strategy of SLR is described in section 2. Section 3 contains the presentation and discussion of results. Finally, the conclusion and future directions are given in section 4. Section 5 contains the references.

2 Systematic Literature Review

A systematic literature review was conducted by following the guidelines of Kitchenham [3] and the collected data is analyzed in an unbiased and structured fashion. The first and the basic step to start the process of SLR was the formulation of protocol that was designed and structured by Umer Iqbal and reviewed by Dr. Khubaib Amjad Alam. Now the steps performed in SLR are described in the next sections.

A Research Questions

The research questions are given in Table I.

Table 1: Research Questions

RQ #	Research Question	Motivation
RQ 1	What are the existing methods, techniques, and algorithms proposed to cater to the next release problem?	The aim is to identify and compare existing methods and algorithms that are proposed to solve the next release problem.
RQ 3	What datasets are used in the context of NRP?	The aim is to identify different datasets used in the context of NRP and their sources
RQ 4	What is the overall research productivity of next release problem?	The aim is to identify the overall research productivity and to identify different research groups working on next release problems.

B Electronic Databases

Table 2 shows the database and the online link to that database.

Table 2: Electronic Databases

Identifier	Database	URL
ED1	IEEE Xplore	http://ieeexplore.ieee.org/
ED2	ACM	http://dl.acm.org/
ED3	Science Direct	http://sciencedirect.com/
ED4	Springer Link	http://link.springer.com/

The studies that were the part of this research activity were from a time span of 2010 to 2018. The digital libraries that were considered are IEEE, ACM, Science Direct, Springer and Google Scholar databases based on title, abstract, and keywords.

C Search Strategy

The first part of this step is to identify the major key terms and the synonyms and alternatives of these terms. The idea behind the formulation of these terms is to construct a query string that will help to continue the remaining search method. A process consisting of three steps was followed to find the relevant studies to answer the research questions [30]. In the first step, the search string was formed. In the second step, we applied this search string on the selected digital libraries to get the required papers. In the third step, it was made sure that no relevant paper was missed.

Table 3: Search String

Database Name	Search String
IEEE Xplore Digital Library	(Next Release Problem) OR (Next Release Planning)
ACM Digital Library	Content.ftsc: ("Next Release Problem " OR " Next Release Planning ")
Science Direct	("Next Release Problem " OR " Next Release Planning ")
Springer	("Next Release Problem " OR " Next Release Planning")

Table 3 contains the search strings against all the four electronic databases that were used to find the studies for this research. As each database is different from each other so search strings are also different for each database.

D Search Process

In order to make sure that we were not leaving any related study, a two-stage search process [30] was adopted:

- **Initial search stage**

Here, we used the proposed search terms to search for primary candidate studies in the four electronic databases. The retrieved papers were grouped together to form a set of candidate papers.

- **Secondary search stage**

In this step, we reviewed all the studies retrieved after title based search where we read the abstracts of the remaining studies and based on the abstract the studies which were not relevant were excluded and the studies that passed this search qualified for the full-text reading.

E Study Selection Process

This step was designed to get the most relevant studies which were retrieved from five electronic databases in order to answer the research questions. The selection procedure which is given in Figure 1, consists of the following basic steps:

- Initial records
- Title based records
- Abstract based records
- Full article based records

F Inclusion and Exclusion criteria

Based on the above criteria if the study meets the inclusion criteria and none of the exclusion criteria is met then such a study is further moved to the next stage that is quality assessment criteria.

Table 4: Inclusion & Exclusion Criteria

	Inclusion and exclusion criteria
	Inclusion criteria
IC1	Studies related to next release problem
IC2	Articles from peer-reviewed publication venues
IC3	The inclusion of studies from 2010 to 2018
IC4	The inclusion of the most recent article in case of multiple studies on the same theme
	Exclusion criteria
EC1	Articles that are not in the English language
EC2	Editorial, short papers, posters, and extended abstracts will not be included

G Quality Assessment criteria

The quality criteria that is designed for this SLR is given in this section.

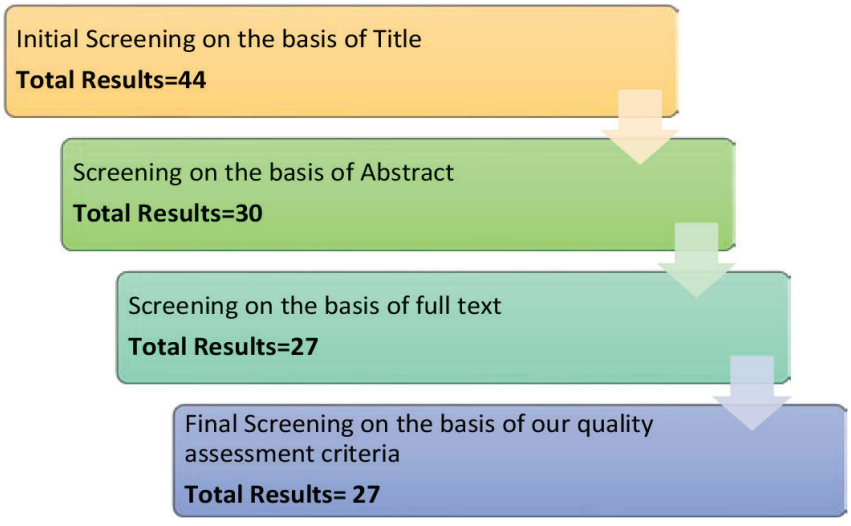

Figure 1: Study Screening Process

Table 5: Quality Assessment Criteria

QC #	Question	Score
QC1	Is the study has clearly defined goals and objectives?	Y|N|P
QC2	Is research in paper assist the aim of next release problem?	Y|N|P
QC3	Is the study propose valid or novelty technique/method?	Y|N|P
QC4	Are limitations of study explicitly stated?	Y|N|P

3 Results & Discussion

This section contains the results and discussion related to the research questions presented in Table. I

RQ1: What are the existing methods, techniques, and algorithms proposed to cater to the next release problem?

The detailed view of the techniques, contribution type that is used in the studies with references and number of studies for each technique are shown in Table 6.

Table 6: Techniques Used to Solve Nrp

Technique	No. of Studies	References of Studies
NSGA-II	10	[3],[5],[6],[16],[18],[23-29]
Ant Colony Optimization (ACO)	4	[4],[13],[17],[26]
Genetic Algorithm (GA)	6	[11],[13],[16],[19],[27],[29]
Simulated Annealing (SA)	8	[4],[11],[16],[18-19], [22],[27],[29]
GRASP	3	[4],[13],[26]
MOEA	4	[3],[5-6],[17]
Hill Climbing	2	[8],[24]
Others	20	[3-16],[18],[20-21],[28-29]

The abbreviations of the techniques used in Table. 6 are given in Table 7.

Table 7: Abbrivation of Techniques

Technique Name	Abbreviation
Non-Dominated Sorting Genetic Algorithm	NSGA-II
Ant Colony Optimization	ACO
Genetic Algorithm	GA
SA	Simulated Annealing
HC	Hill Climbing
GRASP	Greedy Randomized Adaptive Search Procedure
MoCell	Multi-Objective Cellular Genetic Algorithm
MOEA	Multi-objective Evolutionary Algorithm
Others	Others

The graphical representation of the techniques and their usage in percentage is shown in Figure 2.

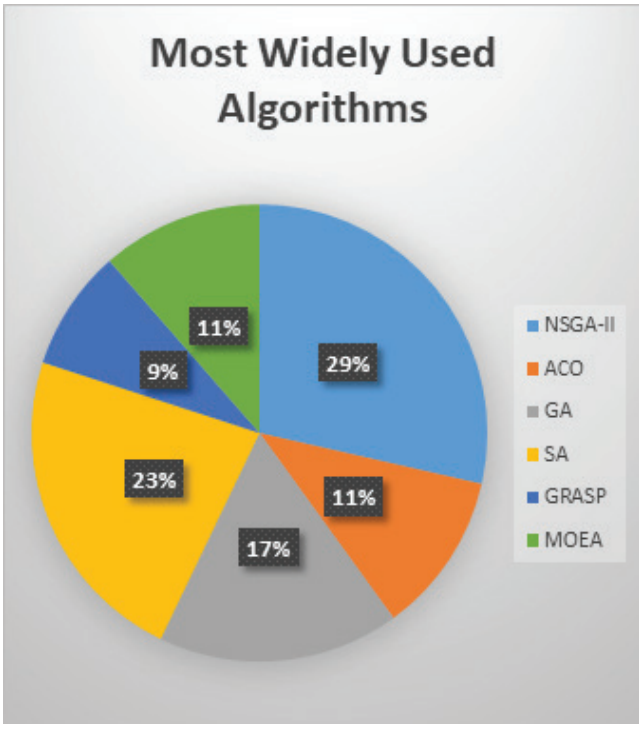

Figure 2: Usage Graph of Techniques

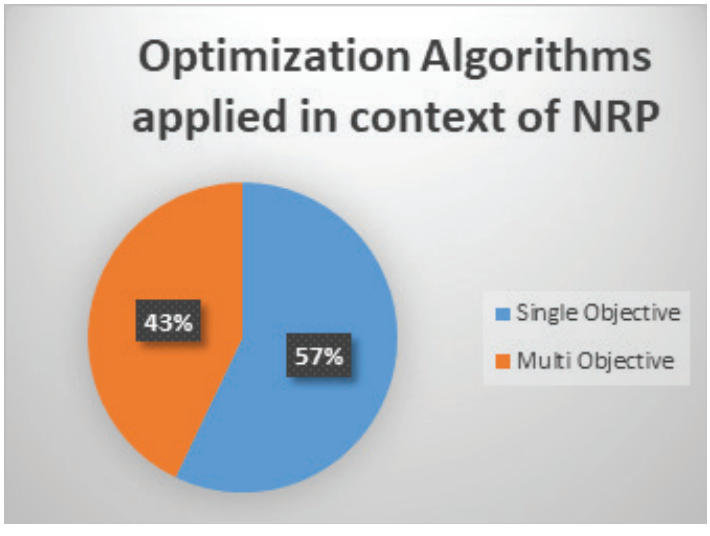

Figure 3: Frequency of Optimization Algorithms

It can be clearly seen in Fig 2 that NSGA-II is most widely used multi-objective optimization technique in the context of NRP while Simulated Annealing (SA) is most widely used in the context of Single Objective NRP. Other techniques like ACO, GA, and MOEA are also used by different authors for comparisons but the overall results are mostly outperformed by NSGA-II and Simulated Annealing. In most of the cases NSGA-II acts as baseline algorithm for comparison. Figure 3 shows that multi-objective optimization techniques have been given more attention as

compared to single objective optimization techniques.

RQ2: What datasets are used in the context of NRP?

All the information about the datasets used in different studies is given in Table 8 with datasets, a number of studies that contains dataset and reference of the studies.

Table 8: Datasets

Dataset	No. of Studies	References of Studies
Synthetic	17	[4-12],[14],[18],[19],[21], [24],[25-28]]
Motorola	6	[12],[14],[18],[27-29]
Mozilla	5	[3],[11],[18],[19],[29]
Eclipse	7	[3],[11],[16],[18],[19],[27],[29]
Gnome	5	[3],[11],[16],[18],[29]
MS-Word	3	[17], [23],[29]
Theme Based RP Dataset	2	[22],[29]
ReleasePlanner	2	[17],[29]
Ralic	2	[18],[29]
Baab	1	[29]

The scenario can be more easily visualized by the statistics presented in Fig 4 which shows the results of the datasets used in studies in the form of percentages. It can be seen that in most of the studies synthetic datasets have been used with the overall percentage of 34 %. Eclipse got the second spot with 14 %. The third position is for Motorola with 12 %. Fourth place is acquired by Gnome and Mozilla with 10% each. Other datasets didn't get much attention because those are not available easily.

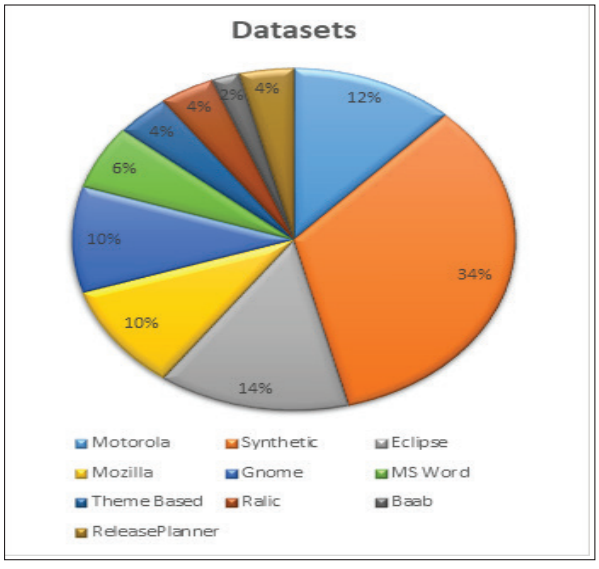

Figure 4: Usage graph of Datasets

RQ3: What is the overall research productivity of next release problem?

To answer this question, we have divided the number of studies into two phases according to years. In the first phase, we analyzed the trend during years 2010-2013 with respect to single and multi-objective techniques. The trend shows that work done on single objective optimization is significant during this time period as compared to multi-objective optimization which can be seen in Figure 5.

Figure 5 Research Trend During Years 2010-2018

The second phase consists of studies from years 2014-2018 and trend was analyzed again with respect to single and multi- objective optimization techniques proposed during this time period. Interestingly multi-objective optimizations techniques were more focused during this time span as compared to single objective optimization techniques which can be visualized through Figure 6

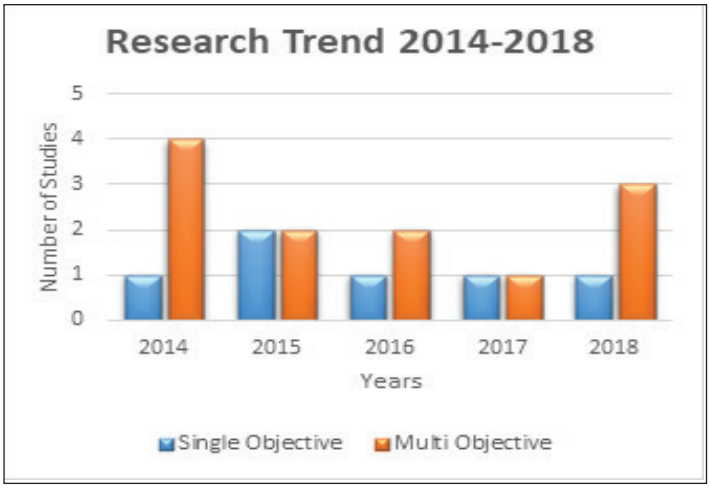

Figure 6 Research Trend During Years 2014-2018

To analyze the overall research productivity, we have considered the research papers that were published in the past years from 2010 to 2018, moreover, we have rated the studies that were able to qualify for the selected studies after quality assessment process and results will be presented. The number of studies according to publication year is shown in Fig 7.

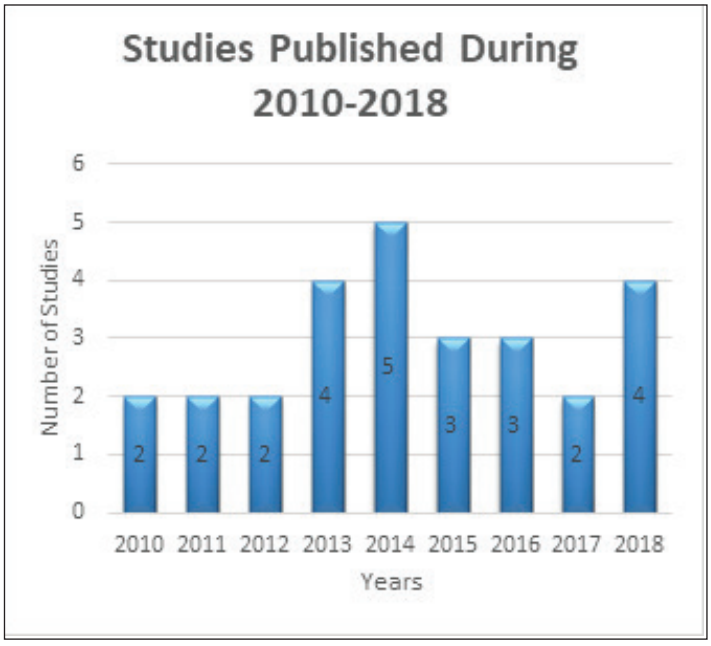

Figure 7: Studies Published During Years 2010-2018

Figure 6 shows that 2014 was the most active research year in which the highest number of studies were published. In 2017 the number was decreased to mere 2 studies but the graph rose again in 2018 with 4 studies published in a single year and more yet to come soon. These numbers are from the span of the last eight years and don't reflect the total number of studies published until yet in the field of NRP.

The quality levels of selected studies are shown in Table 10. The number of studies and the percentage that how many studies falls into a specific group are also shown.

Table 9: Quality Levels of Selected Studies

Quality Level	Number of Studies	Percentage (%)
Very High (score = 4)	12	44.44 %
High (score = 3.5)	8	29.62 %
Medium (score = 3)	7	25.92 %
Total	27	100.00 %

Table 9 reveals that more than 44 % of the studies lies in the highest quality span. While more than 29 % of studies managed to qualify for the high score and only 25.72 % are from medium quality according to our predefined quality assessment criteria.

4 Conclusion

This systematic literature review summarizes the existing literature published in the field of NRP. This paper primarily focused on reviewing the literature on NRP from the span of 2010 to 2018. This study classifies the literature according to different criteria like techniques, contributions, datasets, and quality. 44 studies were identified at the start from different databases and manual sources and after a series of different screening processes only 27 of those qualified for final full-text assessment. The quality of these 27 studies was then analyzed through predefined quality criteria.

The main findings of this research are as follows:

- Multi-Objective Optimization Techniques are the most widely used techniques.
- Among Multi-Objective Optimization Techniques, NSGA-II is widely popular because of its high convergence rate and good quality solutions. Similarly, for Single Objective Optimization, Simulated Annealing is used in most of the cases and showed promising results.
- Synthetic datasets are used in most of the previous studies because of the limited availability of real-world datasets.

NRP is still an emerging field and there is a need to apply the latest state of the art optimization techniques to solve it more efficiently. Furthermore, during the research, it was noted that most of the studies focused on certain common metrics like customer's satisfaction and development cost and other types of metrics like software maintainability, reliability, and traceability are given less attention. These areas should be focused on by researchers in the future.

References

[1] Pmi.org. (2017). 9th Global Project Management Survey. [online] Available at: https://www.pmi.org/-/media/pmi/documents/public/pdf/learning/thoughtleadership/pulse/pulse-of-the-profession-2017.pdf [Accessed 10 Jan. 2019].

[2] Bagnall, A.J., Rayward-Smith, V.J. and Whittley, I.M., 2001. The next release problem. Information and software technology, 43(14), pp.883-890

[3] Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J. and Linkman, S., 2009. Systematic literature reviews in software engineering–a systematic literature review. Information and software technology, 51(1), pp.7-15.

[4] Jiang, H., Zhang, J., Xuan, J., Ren, Z. and Hu, Y., 2010, June. A hybrid ACO algorithm for the next release problem. In Software Engineering and Data Mining (SEDM), 2010 2nd International Conference on (pp. 166-171). IEEE.

[5] Sureka, A., 2014, April. Requirements prioritization and next-release problem under Non-additive value conditions. In Software Engineering Conference (ASWEC), 2014 23rd Australian (pp. 120-123). IEEE..

[6] Cai, X. and Wei, O., 2013, June. A hybrid of decomposition and domination based evolutionary algorithm for multi-objective software next release problem. In Control and Automation (ICCA), 2013 10th IEEE International Conference on (pp. 412-417). IEEE.

[7] Aydemir, F.B., Dalpiaz, F., Brinkkemper, S., Giorgini, P. and Mylopoulos, J., 2018, August. The Next Release Problem Revisited: A New Avenue for Goal Models. In 2018 IEEE 26th International Requirements Engineering Conference (RE) (pp. 5-16). IEEE.

[8] Mauša, G., Grbac, T.G., Bašić, B.D. and Pavˇević, M.O., 2013, July. Hill Climbing and simulated annealing in large scale next release problem. In EUROCON, 2013 IEEE (pp. 452-459). IEEE.

[9] Xuan, J., Jiang, H., Ren, Z. and Luo, Z., 2012. Solving the large scale next release problem with a backbone-based multilevel algorithm. IEEE Transactions on Software Engineering, 38(5), pp.1195-1212.

[10] Li, L., 2016, September. Exact analysis for next release problem. In Requirements Engineering Conference (RE), 2016 IEEE 24th International (pp. 438-443). IEEE.

[11] Paixão, M. and Souza, J., 2013, July. A scenario-based robust model for the next release problem. In Proceedings of the 15th annual conference on Genetic and evolutionary computation (pp. 1469-1476). ACM.

[12] Li, L., Harman, M., Letier, E. and Zhang, Y., 2014, July. Robust next release problem: handling uncertainty during optimization. In Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation (pp. 1247-1254). ACM.

[13] del Sagrado, J., ÁAguila, I.M. and Orellana, F.J., 2011, July. Requirements interaction in the next release problem. In Proceedings of the 13th annual conference companion on Genetic and evolutionary computation (pp. 241-242). ACM.

[14] Harman, M., Krinke, J., Medina-Bulo, I., Palomo-Lozano, F., Ren, J. and Yoo, S., 2014. Exact scalable sensitivity analysis for the next release problem. ACM Transactions on Software Engineering and Methodology (TOSEM), 23(2), p.19.

[15] Jiang, H., Xuan, J. and Ren, Z., 2010, July. Approximate backbone based multilevel algorithm for next release problem. In Proceedings of the 12th annual conference on Genetic and evolutionary computation (pp. 1333-1340). ACM.

[16] Cheng, X., Huang, Y., Cai, X. and Wei, O., 2014, July. An adaptive memetic agorithm based on multiobjecitve optimization for software next release problem. In Proceedings of the Companion Publication of the 2014 Annual Conference on Genetic and Evolutionary Computation (pp. 185-186). ACM.

[17] do Nascimento Ferreira, T., Araújo, A.A., Neto, A.D.B. and de Souza, J.T., 2016. Incorporating user preferences in ant colony optimization for the next release problem. Applied Soft Computing, 49, pp.1283-1296.

[18] Veerapen, N., Ochoa, G., Harman, M. and Burke, E.K., 2015. An integer linear programming approach to the single and bi-objective next release problem. Information and Software Technology, 65, pp.1-13.

[19] Paixao, M. and Souza, J., 2015. A robust optimization approach to the next release problem in the presence of uncertainties. Journal of Systems and Software, 103, pp.281-295.

[20] Chaves-González, J.M. and Pérez-Toledano, M.A., 2015. Differential evolution with Pareto tournament for the multi-objective next release problem. Applied Mathematics and Computation, 252, pp.1-13.

[21] Almeida, J.C., Pereira, F.D.C., Reis, M.V. and Piva, B., 2018, April. The Next Release Problem: Complexity, Exact Algorithms and Computations. In International Symposium on Combinatorial Optimization (pp. 26-38). Springer, Cham.

[22] Araújo, A.A., Paixao, M., Yeltsin, I., Dantas, A. and Souza, J., 2017. An architecture based on interactive optimization and machine learning applied to the next release problem. Automated Software Engineering, 24(3), pp.623-671.

[23] Pitangueira, A.M., Tonella, P., Susi, A., Maciel, R.S.P. and Barros, M., 2017. Minimizing the stakeholder dissatisfaction risk in requirement selection for next release planning. Information and Software Technology, 87, pp.104-118.

[24] Fuchshuber, R. and de Oliveira Barros, M., 2014, August. Improving heuristics for the next release problem through landscape visualization. In International Symposium on Search Based Software Engineering (pp. 222-227). Springer, Cham.

[25] Brasil, M.M.A., da Silva, T.G.N., de Freitas, F.G., de Souza, J.T. and Cortés, M.I., 2011, June. A multiobjective optimization approach to the software release planning with undefined number of releases and interdependent requirements. In International Conference on Enterprise Information Systems (pp. 300-314). Springer, Berlin, Heidelberg.

[26] Del Águila, I.M. and Del Sagrado, J., 2016. Three steps multiobjective decision process for software release planning. Complexity, 21(S1), pp.250-262.

[27] Paixão, M.H.E. and de Souza, J.T., 2013, August. A recoverable robust approach for the next release problem. In International Symposium on Search Based Software Engineering (pp. 172-187). Springer, Berlin, Heidelberg.

[28] Durillo, J.J., Zhang, Y., Alba, E., Harman, M. and Nebro, A.J., 2011. A study of the bi-objective next release problem. Empirical Software Engineering, 16(1), pp.29-60.

[29] Zhang, Y., Harman, M., Ochoa, G., Ruhe, G. and Brinkkemper, S., 2018. An Empirical Study of Meta-and Hyper-Heuristic Search for Multi-Objective Release Planning. ACM Transactions on Software Engineering and Methodology (TOSEM), 27(1), p.3.

[30] Idri, A., azzahra Amazal, F. and Abran, A., 2015. Analogy-based software development effort estimation: A systematic mapping and review. Information and Software Technology, 58, pp.206-230.

[31] Geng, J., Ying, S., Jia, X., Zhang, T., Liu, X., Guo, L. and Xuan, J., 2018. Supporting Many-Objective Software Requirements Decision: An Exploratory Study on the Next Release Problem. IEEE Access, 6, pp.60547-60558..