

Requirement Pattern Extraction using Cluster Based Framework - RAPID

K. Kumar* and R. A. K. Saravanaguru

School of Computing Science and Engineering, VIT University, Vellore - 632014, Tamil Nadu, India;
kkumar@vit.ac.in, sarophd@gmail.com

Abstract

Background/Objectives: In Software Engineering discipline, requirement phase plays vital role in success of a project. To make the requirement analysis process more effective, developers adopt several methods like goal-oriented, aspect-oriented, pattern-oriented etc. In the current scenario recurring requirements can be addressed effectively with requirement patterns rather applying a common procedure to all projects. **Methods/Statistical Analysis:** While adopting common procedure to all project results in increasing cost of quality or even project failure due to the fact that every project is not having similar characteristics. To perform requirement process effectively, it is necessary to select a suitable requirement pattern for a project. Selecting relevant, applicable and reliable pattern is challengeable and prior knowledge is required. **Findings:** The proposed framework groups project characteristics and the pattern sets into k-clusters based on the clustering algorithm. The best suitable pattern will be filtered from the available pattern clusters by applying the filtering algorithm. The suggested patterns along with the situational constraints for a given situation will be stored in the data store. The stack holder can add new pattern details or can update the details of existing patterns **Applications/Improvements:** Finds the analysis patterns that match with the given situational characteristics and requirement of a project. Clustering enhanced with fuzzy to bring the improvement.

Keywords: Clustering, K-Means, Mining, Pattern Extraction, Requirement Pattern, Software Reuse and Filtering

1. Introduction

Nowadays, the project teams which are in process of adopting requirement engineering process needs copying of some of similar requirement functionalities to reduce the requirement engineering process time. But repeated requirement engineering process for software development to single similar applications has a problem includes, repeating previous error or reinventing correct process. To avoid these problems and to make use of similar engineering practices in new projects, requirement patterns are established^{1,2}. The requirement engineering patterns has a well-defined format that describes solution for commonly occurring requirement engineering problems. Pattern has a well-defined format and provides guidelines for engineering procedures. Christopher Alexander stated that "Each pattern describes a problem which occurs over and over again in our environment and

then describes the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice". The existence of patterns aims not only to increase the reusability in solution space but also to leverage the requirement engineering process quality. Realizing the importance of requirement analysis, ACRE (Acquisition of Requirements) framework³ uses commonly followed requirement analysis techniques and established RA (Requirement Analysis) Methods³. The RA Methods acts as guidelines to determine the needs of customers and users. The RA methods aim is to improve communication between stakeholders and to acquire knowledge on environment i.e., domain knowledge. By using the appropriate RE patterns or RA Methods for requirement analysis process could save requirement engineering process time and gather requirements effectively. The Requirement Analysis Pattern Identification (RAPID)

* Author for correspondence

is a framework which identifies patterns that suits for given characteristics of a project. The system suggests suitable patterns which helps the developers to do the requirement engineering process effectively and to speed up the requirement analysis phase.

This paper emphasis on grouping project characteristics and the pattern sets into k-clusters based on the taken clustering algorithm. Using filtering technique the best suitable pattern will be filtered from the available pattern clusters. The suggested patterns along with the situational constraints for a given situation will be stored in the data store. The stack holder can add new pattern details or can update the details of existing patterns.

Patterns are general reusable solutions for repeating configuration issues in an application-autonomous route as recommended by^{2,4,5}. To suit the real standard of basic/ongoing framework, for example, interoperability these configuration pattern supports a great deal⁶. In spite of the fact that the patterns misuse quite a bit of reusability, still it stays to yield poor framework quality because of its ambiguity as pointed by⁷. Likewise these patterns helps in the early assessment of configuration which profoundly fundamental for any framework. The idea of pattern goes for supporting creators and framework modelers in their decision of suitable answers for usually repeating outline issues⁸. The system contains the depiction of the pattern structure and representation that can be utilized to display the specialists in particular application⁹. These configuration design standards are embraced to give answers for usually happening prerequisites building issues.

The usually took after requirement elicitation methods are very little compelling in light of the fact that it doesn't have all around defined workout. This prompts lessening the likelihood of project achievement. Understanding the significance of requirements engineering procedure, ACRE structure has built up requirement analysis strategies for obtaining requirements. It plans to enhance correspondence between stack holders therefore increment in nature of quality requirements. It likewise expects to enhance procuring information on environment¹⁰. Yet, there is a likelihood of task disappointment notwithstanding when the designer took¹¹ after requirement analysis techniques because of the issues that exist in manual choice procedure. It requires more investment for choice procedure, expand many-sided quality of contentions between partners.

2. Proposed System

RAPID Requirement Analysis Pattern Identification finds the requirements analysis patterns that matches with the given situational characteristics of a project¹². It takes project characteristics as input. By applying PCKM algorithm which is based on constrained k-means clustering algorithm¹³, the pattern sets are grouped into k-clusters based on degree of similarity between patterns and project characteristics. And the pattern clusters will be saved in the data store which provides a way for the user to know the related patterns for the given project situation. The best suitable pattern will be filtered from the available pattern clusters by applying BARPC algorithm⁴. The suggested patterns along with the situational constraints for a given situation will be stored in the data store for future reference. It visualize the patterns and its contribution with the given project constraints in the form of pie chart. This functionality made the user to understand the contribution of patterns more easily. Likewise it also suggests requirement engineering patterns based on the problem characteristics. Based on the problem characteristics, the system finds the requirements pattern which describes solutions for the given problem characteristics. The suggested patterns along with the problem constraints for a given situation will be stored in the data store for future reference. It supports the view of detailed description of all patterns by generating report in the form of PDF from data store¹⁴. In addition to this, the framework maintains pattern details in the data store to make the system evolve for changes. The user can add new pattern details or he can update the details of existing patterns. All the discussed above represented with the user level architecture in Figure 1.

2.1 Advantages

- This work helps even the less experienced developers to select suitable requirement analysis patterns.
- It provides a way for the user to know the related patterns for the given project situation.
- Analyses the data sets efficiently. This improves better understanding of data sets.
- It reduces complexity and time taken for selection process.
- It can work well on large data sets.
- It identifies patterns which is suitable for project characteristics.

- Uses best initialization value for k with the help of wrapper search.
- Improves the quality of gathering requirements results in high customer satisfaction.

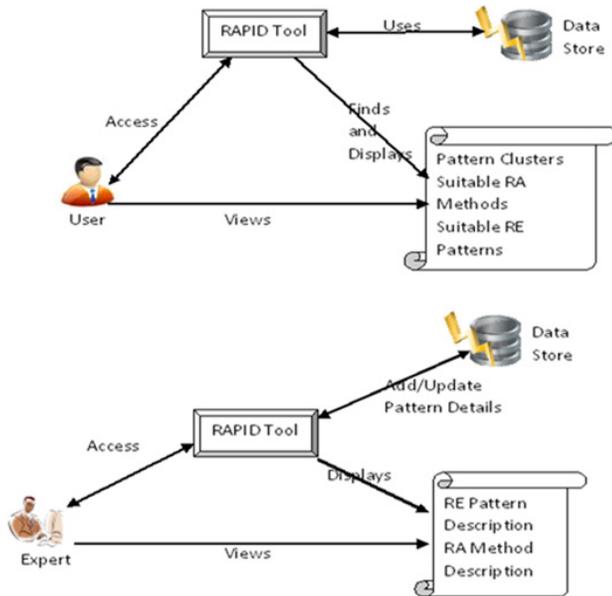


Figure 1. User based architecture.

3. System Architecture

System architecture discussed in the Figure 2 consist of three major modules called requirement analysis, requirement engineering and update. Following sections explain about it:

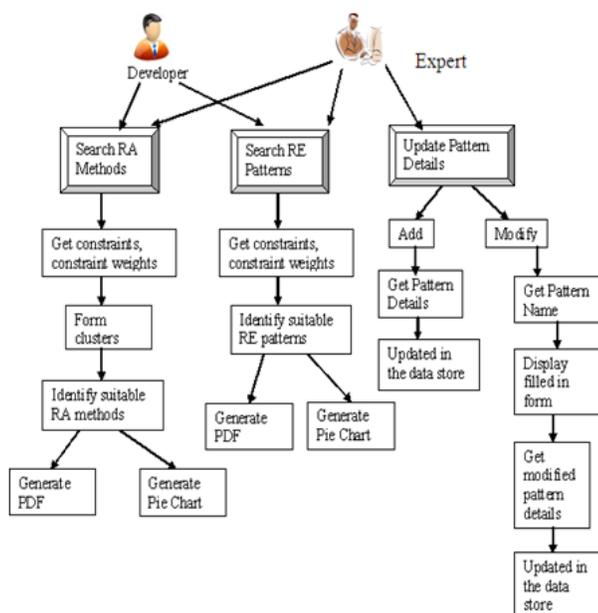


Figure 2. System architecture of RAPID.

3.1 Requirement Analysis (RA) Module

It is for the user who wants to identify suitable Requirements Analysis methods¹⁵ for a project. This module performs the core functionalities of RAPID. User must log into the system for accessing it. This module performs functions like identifying suitable requirement analysis methods, Form RA methods into k-clusters to make the user know the related methods¹⁵, displaying result as report, generate pie chart to make the user to visualize contribution of methods to the constraints, it provides option to view the resulting report in PDF format.

Once the user logs into the system and if the user clicks ‘Search RA Methods’ then the system provides the above mentioned functionalities for the user. To identify suitable requirement analysis methods for a project the user has to give project related details. It includes project name, description, domain, constraints applicable for the project, constraint weights. After giving details if the user clicks ‘Search’ button then the framework first form the RA method sets into k-clusters based on similarity degree between the RA methods using PCKM algorithm. Once the clusters are formed it will be saved in the data store and made available for the user to view the clusters in the form of PDF on the fly. Then the module filters the RA methods that match all the input constraints from the clusters by using BARPC algorithm.

The identified list of methods will be displayed to the user in the form of PDF. The system generates a pie chart to visualize the user that how each RA methods contribute to the given project constraints. Once the user clicks the ‘Log Out’ the system working window will be closed and it redirects to login page.

3.2 Requirement Engineering (RE) Module

This module is for the user who wants to identify suitable Requirements Engineering patterns for a project. This module performs the core functionalities of RAPID. User must log into the system for accessing it. This module performs functions like identifying suitable requirement engineering patterns, displaying result as report, generate pie chart to make the user to visualize contribution of patterns to the constraints, it provides option to view the resulting report in PDF format.

Once the user logs into the system and if the user clicks ‘Search RE Patterns’ then the system provides the above mentioned functionalities for the user. To identify suitable requirement engineering patterns for a project the

user has to give project related details and characteristics of problems that exist in the project. The details include project name, description, domain, characteristics of problems that exist in the project and its weights. After giving details and if the user clicks 'Search' button then the framework identifies the RE patterns that match all the input constraints by using BARPC⁴ algorithm.

The identified list of patterns will be displayed to the user in the form of PDF. The system generates a pie chart to visualize the user that how each RE patterns contribute to the given project constraints. Once the user clicks the 'Log Out' the system working window will be closed and it redirects to login page.

3.3 Update Module

This module is for the experts who want to update pattern details that have been stored in the data store. Expert must log in to access the system. This module performs functions like adding new RA methods and RE patterns into the data store. Updating RA method details and RE pattern details that exist in the data store.

Once the expert logs into the system, the system provides the above mentioned functionalities for the expert. The major role of expert is to add or update RA method details and add or update RE Pattern details. If the expert wants to add new RA Methods and clicks 'Add RA Method' then the system displays a form that contains RA method template. The expert fills and clicks 'Add' button. After validation the details are saved in the data store. The system has the provision to modify the RA method details. If the RA method details which is in the data store needs to be changed, then the expert can update the details. Likewise, If the expert wants to add new RE Patterns and clicks 'Add RE Pattern' then the system displays a form that contains brief details on RE Patterns and the user has to upload a PDF file that has detailed description of RE Pattern.

The expert fills the details, uploads PDF file and clicks 'Add' button. After validation the details are saved in the data store. The system has the provision to modify the RE Pattern details. If the RE pattern details which is in the data store needs to be changed, then the expert can update the details.

4. Implementation of RAPID

Presentation layer provides an interface i.e., client side

interface which acts as intermediate between user and system. The client side interface is developed using java swing. The user access the system through this layer. Also output of process will be displayed to the user through this layer as mentioned in Figure 3.

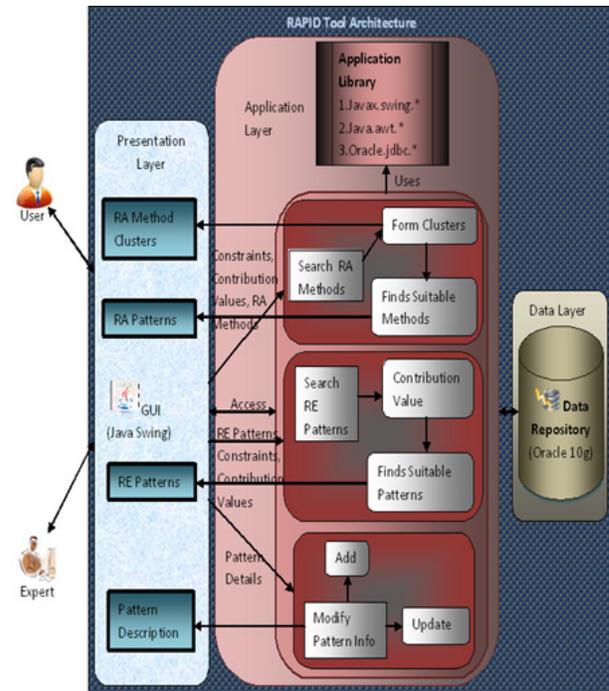


Figure 3. RAPID architecture.

Application layer makes the application to service data requests between a user and a database. It contains RAPID logic which is implemented using java technology. The client library contains the library files and jar files that are required by the processing part to execute.

Finally, data layer uses oracle 10g database for storage. It has RA method details, RE pattern details, characteristics, contribution value between patterns and characteristics in the form of tables. These tables will be globally available to all the internal procedures. JDBC/ODBC Bridge is being used to establish connectivity to the database from the application. The DBMS system will be locally available on the machine where the application is running. For adding, updating, deleting and accessing the data this work using SQL queries built in the internal procedures. Performance of RAPID with other works as depicted in the Figure 4. Here, the ACRE³, Selection method (PCKM), Filtering method (BARPC) are evaluated with the feature parameters and the final assessment presented.

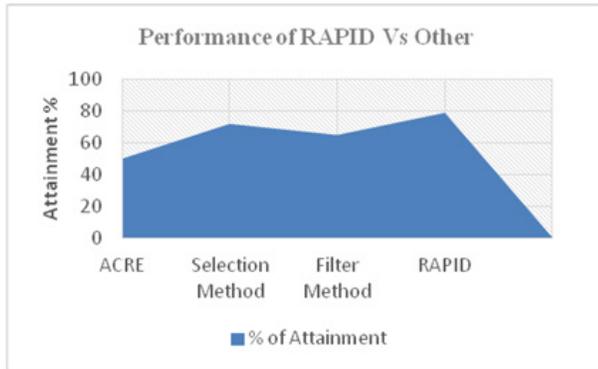


Figure 4. Performance of RAPID.

Comparison of previous work as presented in Table 1 has the parameter based on the features given below:

- It ensures the accuracy and consistency of data identifies suitable requirement analysis methods.
- Produce reports for detailed view of patterns. Identifies suitable Requirements Engineering Patterns.
- Provides summarized view of pattern clusters as a report.
- Provides summarized view of search results as a report and generate Reports in the form of PDF.
- Visualize contribution of patterns to the given characteristics in the form of pie chart.
- Provides a way for updating pattern details.
- It is easy-to-use through its graphical user interface.
- Provides the correctness of the retrieved pattern.

Table 1. Comparison with previous work

Parameter	RAPID	Previous Work	
		Pattern ⁴	Agent ¹²
Technology	Excellent	Average	Average
Accuracy	Yes	Yes	No
Consistency	Yes	Partial	Partial
Easy to Use	Yes	Yes	No
User friendly	Yes	Partial	Yes
Pattern Identification	Partial	Yes	Yes
Analysis method selection	Yes	Yes	No
Tool Development	Yes	No	No

5. Conclusion

This work implemented RAPID to identify requirement analysis patterns that matches the characteristics of a project. This framework successfully finds the suitable Requirement Analysis (RA) Methods, Requirement Engineering Patterns (RE). This framework successfully

finds list of patterns that matches all project characteristics. The identified pattern will be more appropriate to perform requirement analysis process. This framework reduces cost of quality by increasing quality of requirements. It helps the less experienced developer to select best technique for requirement analysis process.

In future, RAPID Requirement Analysis Pattern Identification finds the requirements analysis patterns that matches with the given situational characteristics of a project. It takes project characteristics as input to form fuzzy clustering technique to form clusters and identify the relevant pattern which matches for the given characteristics. It takes project characteristics as input. By applying fuzzy algorithm which is based on constrained k-means clustering algorithm, the pattern sets are grouped into k-clusters based on degree of similarity between patterns and project characteristics. And the pattern clusters will be saved in the data store which provides a way for the user to know the related patterns for the given project situation.

6. References

1. Requirements Engineering Patterns. Working Group on Requirements Engineering Patterns (WGREP) of Section 2.1.6. 2015. Available from: www-library.desy.de/preparch/desy/2004/desy04-233.pdf
2. Hagge L, Lappe K. Sharing requirements engineering experience using patterns. *Software, IEEE*. 2005 Jan-Feb; 22(1):24–31.
3. Chen CH, Khoo LP, Yan W. A strategy for acquiring customer requirement patterns using laddering technique and ART2 neural network. *Advanced Engineering Informatics*. 2002 Jul; 16(3):229–40.
4. Shu-hang G, Yu-qing L, Mao-Zhong J, Jing G, Hong-Juan G. A requirement analysis pattern selection pattern for E-Business project situation. *IEEE International Conference on E-Business Engineering, ICBE'07; Hong Kong*. 2007 Oct 24–26. p. 347–50.
5. Wagstaff K, Cardie C, Rogers S, Schroedl S. Constrained k-means clustering with background knowledge. Brodley C, Danyluk AP, Editors. *Proc of the 18th Int'l Conference on Machine Learning (ICML 2001); San Francisco; Morgan Kaufmann Publishers*: 2001. p. 577–84.
6. Hickey AM, Davis AM. Elicitation technique selection: How do experts do it? *Proceedings of 11th IEEE International Requirements Engineering Conference*; 2003 Sep 8–12. p. 169–78.
7. Issa AA, Al-Ali A. Use case pattern driven requirements engineering. *Second International Conference on Computer Research and Development; Kuala Lumpur*. 2010 May 7–10. p. 307–13.

8. Rapid Application Development Process Model. 2015. Available from: www.careerride.com/pmp-phases-in-rad-model.aspx
9. Fowler M. Analysis patterns: Reusable object models. Addison Wesley; Longman: 1997.
10. Mjeda A, Hinchey M. Requirement-centric reactive testing for safety-related automotive software. 2015 IEEE/ACM 2nd International Workshop on Requirements Engineering and Testing (RET); Florence. 2015. p. 5–8.
11. Gamma E, Helm R, Johnson R, Vlissides J. Design patterns: Elements of reusable object-oriented software. USA; Addison Wesley: 1995.
12. Gaur V, Soni A, Bedi P. An agent oriented approach to requirements engineering. 2010 IEEE 2nd International Advance Computing Conference (IACC); Patiala. 2010 Feb 19-20. p. 449–54.
13. Talpur MAH, Chandio IA, Memon IA, Qureshi TA. Computer-based database system: Rapid response for rural planning agencies to devise transport policies. Indian Journal of Science and Technology. 2014 Oct; 7(10):1603–12.
14. Kannan K, Saravanaguru RA. An approach for Decomposing Requirements into Analysis Pattern using Problem Frames (DRAP-PF). International Conference on Advances in Computing, Communications and Informatics (ICACCI); Kochi. 2015 Aug 10-13. p. 2392–6.
15. Mahmoudi A, Sadi-Nezhad S, Makui A. An extended fuzzy PROMETHEE based on fuzzy rule based system for supplier selection problem. Indian Journal of Science and Technology. 2015 Nov. 8(31):1–11.