

Using Hierarchical Cluster Analysis to Generate Clusters of Agile Practices

Abdullah M. Al-Sakkaf, Nor Laily Hashim, Mazni Omar

*Human-Centered Computing Research Lab, School of Computing, Universiti Utara Malaysia, Malaysia.
Abdullah1@hotmail.com*

Abstract—Agile software development methodologies consist of a set of software practices that can help organizations to produce products faster and deliver what customers want. Despite the benefits they gain from adopting agile practice, organizations could maximize the benefits gained by adopting correlated practices. There is a lack of study on the identification of clusters of independent practices. This paper focuses on identifying clusters of agile practices in software startups in Saudi Arabia. The study was conducted using a questionnaire with 76 software practitioners from software startups in Saudi Arabia. In this paper, 20 agile practices were analyzed using hierarchical cluster analysis. The analysis generated four clusters: Each was associated with a list of practices. These clusters can be used as a guide for agile method tailoring that helps to identify the inter-relationship between different agile practices. The clusters were labeled as project management, quality assurance, team communication, and incremental and iterative clusters. The results can be used to study the co-dependence of agile practices in depth. Moreover, the study can help adopters from similar domains as well as companies with limited resources and experience frequent changes in requirements to adopt these agile practice clusters.

Index Terms—Agile; Agile Practices; Agile Adoption; Agile Practice Clusters.

I. INTRODUCTION

Agile software development methodologies (ASDM) have become very famous on software development methodologies [1]. Each ASDM defines its own processes and practices, but they share in common the same values that are addressed in agile manifesto [2]. From the perspective of software development, choosing and adopting the proper development methodology is a critical task [3]. Each agile development methodology consists of several practices which makes it difficult to select which cluster of practices that fit the needs of a software startup. Software startup is characterized by limited resources, small team, and a product-driven organization.

Many studies show that these methods are adopted partly by selecting a set of agile practices. Therefore, it is difficult for a new adopter to choose proper agile practices sets that fit their organization's needs as ASDM has a big pool of available practices. These practices are useful in its own, but they provide more values when they are working together as a cluster of practices [5]. Agile practices should be selected based on factors that include the organization needs in order to maximize the benefit of adopting a new software development methodology [29].

The aim of this study is to identify codependent agile practices in clusters for software startups. This study used a

quantitative approach to study software startups at the Kingdom of Saudi Arabia (KSA). The results will help the organization to select suitable agile practices cluster based on matching the motivation that correspondingly affects the success of ASDM adoption.

This paper is organized as follows. In Section II, the agile software development methodology is discussed. In Section III, the agile adoption and agile method tailoring are defined. In Section IV the agile practices clusters are presented, as well as, related work that clustering agile practices. In Section V, the research methodology is presented. The analysis results of the hierarchical cluster analysis are also presented in Section VI. Finally, future work and summary related to this work are described in the last section.

II. AGILE SOFTWARE DEVELOPMENT METHODOLOGY

During the last two decades, ASDM has dramatically increased its usage, resulting in a change in the way software development is performed [6]. Unlike traditional development methodologies characterized by sequential phases and heavy upfront planning, agile methodology deals with unpredictability and change by relying on people and close customer collaboration rather than formalized processes [3]. In 2001, the agile manifesto was written, which states that agile development core values [2] are individuals and interactions over processes and tools, working software over comprehensive documentation, customer collaboration over contract negotiation, and responding to change over following a plan. The main attributes of agile are short iterative, collaborative decision-making, quick feedback loops, and continuous integration of code changes into the product [7].

The number of agile methodologies has been growing, leading to the emergence of about 20 different agile or lean methods [6] Agile development methodologies include Scrum, eXtreme Programming (XP), Crystal methodologies, Lean software development, and Feature-driven development [8].

III. AGILE ADOPTION

According to Nerur et al. [3], the adoption of agile methodology is not an easy task, despite the implementation of most of its practices are easy. That happens because the agile adoption represents an organizational change that will affect the company's organizational structure, processes, as well as people's behavior. Likewise, Ayed, Vanderose, and Habra [9] agreed that the adoption of agile software development methods is a wide and complex organizational

change that usually impacts several aspects of the organization (e.g., its structure, culture, management practices, produced artifacts, technologies in use, etc.). In order to successfully handle the several key challenges, it is crucial to understand the context of the organization and carefully study the transformation strategies [9].

Gandomani, Zulzalil, and Ghani [10] indicated that the most important attention within agile adoption process are focusing on people, providing an action plan, transiting challenges identification, providing prerequisites, providing facilitators, and conducting timely assessment. As a consequence, the adoption of agile methodology should be well planned to be successful. However, agile adopters encounter problems due to the lack of guidance and assistance [9]. Particularly, agile methodology is not adopted entirely, but it is adopted as certain practices. Furthermore, some practices are found to be used more frequent on some business domain [6]. It is also found that the success of a project depends on the choice of the agile practices [11]. According to Boehm [12], in practice, the adoption should be only on practices that may be beneficial and applicable in a specific organizational context. Furthermore, every project has its own processes that are tailored for its circumstances and needs [13].

The organization should consider tailored practices based on the current needs. In the field of software development, the tailoring method is the process of adapting the method used to meet the circumstances of use [14]. Further, it could be defined as the adaptation of the method in relation to the aspects, culture, objectives, environment and reality of the organization that is adopting it [15]. When adopting agile development methodology, the organization must consider tailored practices based on the current needs. In summary, the agile adoption should be iterative and the practices should be tailored and selected based on the needs of the organization [16]. This implies that all the practices should not be adopted at once, but rather to find the problem and try to solve it using XP practices [13]. The importance of understanding agile methodology tailoring is to enable companies to select practices to achieve the organization needs, since full agile method adoption can be an overkill for organizations or require a lot of resources [17]. Sometimes this scenario is referred as partial adoption of agile methodology [18].

IV. AGILE PRACTICES CLUSTER

As there is a long list of available agile practices, the agile teams need help in choosing the right combination of practices based on their needs [19]. It was recommended to investigate the practices cluster to determine if the practices are codependent [20].

The term “working set” is used to refer to the set of agile practices that lead to the positive effect in a project [21]. It is defined as “a restricted set of such top important practices, values and goals” [21]. Also, it is known as agile practice clusters [16, 20]. However, evidence of which agile practices working set works well together, and in what contexts, is still lacking [20]. Commonly, an agile adopter needs assistance to choose the proper combination of agile practices [15].

Likewise, another study indicates 15 agile practices clusters [19]. The cluster was generated using principal component analysis with oblique rotation on 58 agile

practices. The new cluster includes agile quality assurance, communication (team), communication (customers), coding standards, etc. In other ways, Trip [5] categorized 12 agile practices into two clusters using qualitative methods. Three experts were interviewed in order to generate new clusters. The new clusters are the project management cluster (include: Release Planning, Iteration Planning, Velocity, daily Stand-up, Retrospectives, and Burndown), and the software development approach (includes the TDD, Refactoring, Continuous Integration, Unit Testing, Coding Standards, and Automated Builds). Furthermore, Melo et al. [22] categorized the agile practices into three clusters which are technical, management, and collective knowledge sharing. Technical practices are pair programming, Burndown chart, and Automated acceptance tests. Management practices are daily meeting, Iteration development, Iteration/release planning, Retrospectives, Checklists, One-on-One meetings, and Timeboxes. Similarly, another study [23] grouped agile practices into three groups, which are the management practices cluster, software process cluster, and software development practices cluster.

V. RESEARCH METHODOLOGY

A survey was designed in order to explore agile practices adoption in software startups. The instrument was developed as a result of an analysis of previous works [5,24]. The questionnaire was sent to about 300 professionals in software startups on KSA. The survey cover letter was sent as the body of the email to briefly describe the questionnaire’s purpose. The questionnaire investigates the adoption of 20 agile practices using a 5-point Likert scale (Never Used - Always Used). Data were collected between March and April of 2016. The survey was sent to CEOs, CIOs, project managers, technical team manager, and developers.

In order to find out the adopted agile practice clusters, hierarchical cluster analysis (HCA) was conducted on the agile practices. The main purpose of the cluster analysis is to find related items in a dataset [25]. Cluster analysis is a convenient method for identifying homogenous groups of objects called clusters that share many characteristics, but are very dissimilar to objects that do not belong to that cluster [25].

Agglomerative hierarchical cluster analysis is used as the clustering procedure. Squared Euclidean distance is used as a measure in HCA, where Euclidean distance is mostly used when variables are in ratio or interval-scaled variables [25]. There are different algorithms used for clustering, and each is used for a different purpose. The method chosen for cluster extraction was Ward’s clustering algorithm as it performs well at recovering clusters [26] and it generates somewhat equally sized clusters [25].

VI. RESULTS

From the total number of sent surveys, approximately 76 were returned, resulting in a rate of return of about 24%. In the questionnaire, the background data were requested from both the respondent and the respondent’s organization. The majority of the respondents were programmers or developers, representing 35% of all respondents. 21% of the respondents were working in an executive level e.g. Chief

Executive Officers (CEOs) and Chief Technology Officers (CTOs). In addition, 35% of the respondents were project managers, system analysts, or IT management. Finally, 6% of the respondents were quality assurance officers or testers. In term of organization size, more than half (51%) of respondents' organizations had 10 employees or less. Also, 17% of the organizations had 10 to 20 employees, while 26% of the organizations had 21 to 40 employees. Only eight percent of the respondents' organizations had more than 40 employees. These can be treated as mature startups.

The HCA was generated using a dendrogram, as shown in Figure 1, where it indicates the possible clusters. A dendrogram was used to determine the number of clusters, which were named as distance-based decision rules. In SPSS, a dendrogram rescales the distances to a range of 0-25 [25].

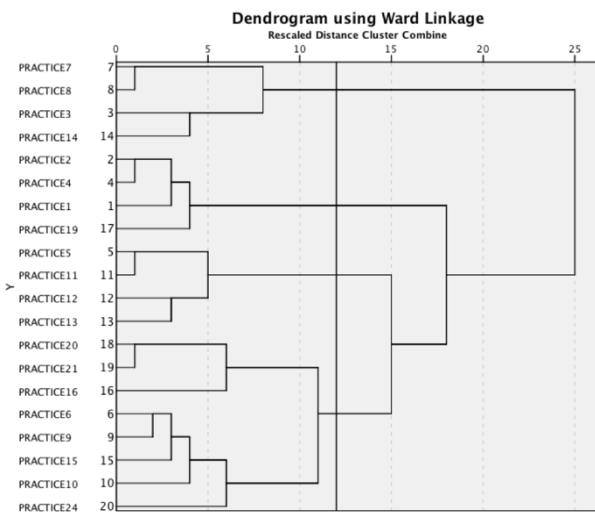


Figure 1: Dendrogram from HCA

In order to perform the cluster analysis, the agile practice variables were analyzed in SPSS 23.0 using the described methods. The analysis produced a total of 19 clusters (count of variables-1) (see Table 1). Further, a dendrogram was generated. As shown in figure 1, the dendrogram at level 12 shows four clusters of agile practices. From the dendrogram, four new clusters were identified.

The new clusters (see Table 2) could be labeled based on several criteria. Cluster 1 was labeled as a project management cluster since three to four practices are similar to those on the project management cluster identified by Tripp [5]. Further, most of their practices support project management activities. As shown in Table 3, the project management cluster was less adopted by software startups because startups implement a loose organizational structure and avoid traditional management [27]. This is caused by the fact that startups work under pressure with limited time and small size team.

The second cluster practices support testing and quality. Cluster 2 was labeled as a quality assurance cluster as shown in Table 2. This cluster shared 75% of the practices in quality assurance cluster defined by Abbas [19]. In continuous integration, “each integration is verified by an automated build (including test) to detect integration errors as quickly as possible” [28], which makes it clear that continuous integration codependent with other testing practices in the same cluster.

Cluster 3 was labeled as a team communication cluster. The practices in this cluster improve the communication inside the development team, hence reflecting the effectiveness of work. Effective communication is ideal for producing high quality products [2]. To clarify, the open work area and single team rules improve the communication between teams, as shown in Table 2. Team communication cluster was the most adopted agile practices cluster by software startups in KSA, as shown in Table 3.

The final cluster was labeled as the incremental and iterative cluster. Cluster 4 was difficult to label because it has eight practices while other clusters have only four practices, as shown in Table 2. Most of these practices support the iterative process like iteration review, story mapping, release planning, and user story. As startups focus on fast movement from idea conception to production by using iterative and incremental approach, these approaches help them to reach their goal [27]. This scenario explains why this cluster is the second most used agile practice clusters (see Table 2).

Table 1
HCA Agglomeration Schedule

Stage	Cluster Combined	Cluster	Coefficients	Stage Cluster First Appears	Next Stage	
	1	2		1	2	
1	7	8	50.000	0	0	15
2	5	11	53.000	0	0	9
3	18	19	61.000	0	0	13
4	2	4	62.000	0	0	8
5	6	9	66.000	0	0	7
6	12	13	79.000	0	0	9
7	6	15	82.000	5	0	11
8	1	2	82.000	0	4	10
9	5	12	86.000	2	6	14
10	1	17	90.667	8	0	17
11	6	10	92.000	7	0	14
12	3	14	100.000	0	0	15
13	16	18	103.500	0	3	17
14	5	6	106.625	9	11	16
15	3	7	112.500	12	1	19
16	5	20	120.750	14	0	18
17	1	16	125.167	10	13	18
18	1	5	127.413	17	16	19
19	1	3	150.969	18	15	0

Table 2
Cluster's Membership

Cluster 1	Cluster 2	Cluster 3	Cluster 4
PRACTICE1 daily meeting	PRACTICE3 unit testing	PRACTICE5 prioritized backlogs	PRACTICE6 team-based estimation
PRACTICE2 short iterations	PRACTICE7 Coding standards	PRACTICE11 single team	PRACTICE9 iteration reviews
PRACTICE4 retrospectives	PRACTICE8 Continuous integration	PRACTICE12 refactoring	PRACTICE10 dedicated product owner
PRACTICE19 iteration planning	PRACTICE14 Test driven development	PRACTICE13 open work area	PRACTICE15 story mapping
			PRACTICE16 collective code ownership
			PRACTICE20 continues deployment
			PRACTICE21 release planning
			PRACTICE24 user story

Table 3
Most Used Agile Practices' Cluster

	N	Mean	Std. Deviation
CLUSTER3	76	3.42	.98
CLUSTER4	76	3.38	1.04
CLUSTER2	76	3.26	1.05
CLUSTER1	76	3.11	1.06

VII. FUTURE WORKS

Although this study found new agile practices clusters, we suggested for further investigation using quantitative research to determine the correlations between the adopting of agile practices clusters and the effectiveness of the individual practices within that cluster. Also, there is a need to discover the relationship between these clusters and other variables, e.g.: quality, project success, customer satisfaction and others. The research recommends to conduct mixed methods (combining qualitative and quantitative) study in order to gain in-depth understanding of the agile practices cluster.

VIII. CONCLUSION

Based on a statistical analysis, this research found four practices clusters. It was constructed using a hierarchical cluster analysis from a total of 20 agile practices. The new clusters are project management cluster, quality assurance cluster, team communication cluster, and incremental and iterative cluster.

In practice, understanding agile practices clusters will help companies to maximize their benefits from adopting these clusters as some agile practices work better when they are adopted together [20]. Software startups need to choose and appreciate a fast process [27], since they do not have available resources to figure out the best way to develop a product. On the contrary, adopting the practices proposed by an ASDM will lead the organization to spend more efforts and resources, hence evaluating each proposed agile practice could help startups to determine which practice bring more value to an organization or otherwise [16, 19].

REFERENCES

- [1] D. West and T. Grant, "Agile development: Mainstream adoption has changed agility," *Forrester Res. Inc.*, vol. 2, (2010).
- [2] K. Beck, M. Beedle, A. Van Bennekum, A. Cockburn, W. Cunningham, M. Fowler, J. Grenning, J. Highsmith, A. Hunt, R. Jeffries, J. Kern, B. Marick, R. C. Martin, S. Mellor, K. Schwaber, J. Sutherland, and D. Thomas, (2001) "Manifesto for Agile Software Development," Agile Alliance.
- [3] S. Nerur, R. Mahapatra, and G. Mangalaraj, "Challenges of migrating to agile methodologies," *Communications of the ACM*, vol. 48, no. 5, ACM, (2005) 72–78.
- [4] L. Williams, "Agile Software Development Methodologies and Practices," in *Advances in Computers*, vol. 80, (2010) 1–44.
- [5] J. F. Tripp and D. J. Armstrong, "Exploring the relationship between organizational adoption motives and the tailoring of agile methods," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, (2014) 4799–4806.
- [6] P. Diebold and M. Dahlem, "Agile Practices in Practice: A Mapping Study," in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, (2014) 30:1–30:10.
- [7] A. Cockburn, *Agile Software Development*. Berlin, Heidelberg: Springer Berlin Heidelberg, (2010).
- [8] T. Dingsøy, T. Dybå, and P. Abrahamsson, "A preliminary roadmap for empirical research on agile software development," in *Proceedings - Agile 2008 Conference*, (2008) 83–94.
- [9] H. Ayed, B. Vanderose, and N. Habra, "Supported approach for agile methods adaptation: an adoption study," in *Proceedings of the 1st International Workshop on Rapid Continuous Software Engineering - RCoSE 2014*, (2014) 36–41.
- [10] T. J. Gandomani, H. Zulzalil, A. A. A. Ghani, A. B. Md. Sultan, and K. Y. Sharif, "An Exploratory Study on Managing Agile Transition and Adoption," vol. 265, (2014) 177–188.
- [11] L. M. R. De Souza Mariz, A. C. C. França, and F. Q. B. Da Silva, "An empirical study on the relationship between the use of agile practices and the success of software projects that use scrum," in *Proceedings - 24th Brazilian Symposium on Software Engineering, SBES 2010*, (2010) 110–117.
- [12] B. Boehm, "Get ready for agile methods, with care," *Computer (Long Beach, Calif.)*, vol. 35, no. 1, (2002) 64–69.
- [13] K. Beck, "Embracing change with extreme programming," *Computer (Long Beach, Calif.)*, vol. 32, no. 10, (1999) 70–77.
- [14] I. Sommerville and J. Ransom, "An empirical study of industrial requirements engineering process assessment and improvement," *ACM Trans. Softw. Eng. Methodol.*, vol. 14, no. 1, (2005) 85–117.
- [15] A. S. Campanelli and F. S. Parreiras, "Agile methods tailoring – A systematic literature review," *J. Syst. Softw.*, vol. 110, (2005) 85–100.
- [16] N. Kurapati, V. S. C. Manyam, and K. Petersen, "Agile Software Development Practice Adoption Survey," in *Agile Processes in Software Engineering and Extreme Programming*, C. Wohlin, Ed. Springer Berlin Heidelberg, (2012) 16–30.
- [17] A. Qumer and B. Henderson-Sellers, "A framework to support the evaluation, adoption and improvement of agile methods in practice," *J. Syst. Softw.*, vol. 81, no. 11, (2008) pp. 1899–1919.
- [18] J. Erickson, K. Lyytinen, and K. Siau, "Agile Modeling, Agile Software Development, and Extreme Programming," *J. Database Manag.*, vol. 16, no. 4, (2005) 88–100.
- [19] N. Abbas, A. M. Gravell, and G. B. Wills, "Using Factor Analysis to Generate Clusters of Agile Practices (A Guide for Agile Process Improvement)," in *2010 Agile Conference*, (2010) 11–20.
- [20] K. Conboy and B. Fitzgerald, "Method and developer characteristics for effective agile method tailoring," *ACM Trans. Softw. Eng. Methodol.*, vol. 20, no. 1, (2010) 1–30.
- [21] L. Krzanik, P. Rodriguez, J. Similä, P. Kuvaja, and A. Rohunen, "Exploring the transient nature of agile project management practices," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, no. c, (2010) 1–8.
- [22] C. de O. Melo, V. Santos, E. Katayama, H. Corbucci, R. Prikladnicki, A. Goldman, and F. Kon, "The evolution of agile software development in Brazil," *J. Brazilian Comput. Soc.*, vol. 19, no. 4, (2013) 523–552.
- [23] S. Lee and H.-S. Yong, "Agile Software Development Framework in a Small Project Environment," *J. Inf. Process. Syst.*, vol. 9, no. 1, (2013) 69–88.
- [24] VersionOne, "10th annual State of Agile Report," (2015).
- [25] E. Mooi and M. Sarstedt, *A Concise Guide to Market Research*. (2001); Berlin, Heidelberg: Springer Berlin Heidelberg.
- [26] H. Finch, "Comparison of Distance Measures in Cluster Analysis with Dichotomous Data," *J. Data Sci.*, vol. 3, (2005) 85–100.
- [27] C. Giardino, M. Unterkalmsteiner, N. Paternoster, T. Gorschek, and P. Abrahamsson, "What Do We Know about Software Development in Startups?" *Software, IEEE*, vol. 31, no. 5, (2014) 28–32.
- [28] S. Stolberg, "Enabling Agile Testing through Continuous Integration," in *2009 Agile Conference*, (2009) 369–3.
- [29] E.O.C. Mkpjoju, N.L.Hashim and A.M.H. Al-Sakkaf, "Motivation for agile adoption among software startups in saudi arabia"