

Software Product Quality in Global Software Development: Finding Groups with Aligned Goals

Panagiota Chatzipetrou, Lefteris Angelis

Department of Informatics
Aristotle University of Thessaloniki
Thessaloniki, Greece
{pchatzip, lef@csd.auth.gr}

Sebastian Barney, Claes Wohlin

School of Computing
Blekinge Institute of Technology
Karlskrona, Sweden
{sebastian.barney, claes.wohlin@bth.se}

Abstract— The development of a software product in an organization involves various groups of stakeholders who may prioritize the qualities of the product differently. This paper presents an empirical study of 65 individuals in different roles and in different locations, including onshoring, outsourcing and offshoring, prioritizing 24 software quality aspects. Hierarchical cluster analysis is applied to the prioritization data, separately for the situation today and the ideal situation, and the composition of the clusters, regarding the distribution of the inherent groupings within each of them, is analyzed. The analysis results in observing that the roles are not that important in the clustering. However, compositions of clusters regarding the onshore-offshore relationships are significantly different, showing that the offshore participants have stronger tendency to cluster together. In conclusion, stakeholders seem to form clusters of aligned understanding of priorities according to personal and cultural views rather than their roles in software development.

Keywords- software; quality; global software development; cluster analysis; goal alignment

I. INTRODUCTION

Global Software Development (GSD) [1] is software work undertaken at geographically separated locations across national boundaries in a coordinated fashion involving real time (synchronous) and asynchronous interaction. In other words, while teams are not co-located, they are still working towards a common goal with a commercially viable product [2]. GSD usually comes with a number of benefits and risks. It is a common practice, which allows companies to improve the product management and the quality of their products

The various groups of stakeholders involved in the software development process in an organization may prioritize software qualities differently. The stakeholders represent various roles, some of them are internal to the organization (insourcing) while others are in outsourcing relationships, and moreover, they may work in different countries, representing onshore and offshore relationships. For value-based software engineering, the need for a common understanding of the prioritization of software quality aspects by key stakeholders is crucial.

The purpose of this paper is to investigate how different people involved in the software development process can be grouped, based on alignment of their views regarding the

prioritization of software quality aspects. For this purpose we use a case study and propose a multivariate statistical methodology, which is able to discover agreements and disagreements between the stakeholders.

The empirical study uses data from a prioritization process, where 65 stakeholders in different roles and groups, working on the development of the same product, prioritize with cumulative voting 24 software quality aspects from their individual perception of two situations: the situation today and their perceived ideal situation.

Hierarchical cluster analysis is applied to the data, separately for each situation and the composition of the clusters, regarding the distribution of the inherent groupings within each of them, is analyzed. Associations between inherent groups and formed clusters are tested using the chi-square test.

In our case study we had stakeholders with various roles and different relationships with the company. However, not all of the combinations were present. Specifically, there are onshore insourcing, onshore outsourcing and offshore insourcing arrangements. There was no stakeholder with offshore outsourcing relationship. The analysis resulted, for both situations, in two clusters. Although some participants from the same role tend to cluster together (e.g. Testers and Architects) in general the compositions of the resulting clusters do not differ significantly. The same is true for the insource-outsource relationships. However, the compositions of clusters regarding the onshore-offshore relationships are significantly different, showing that the offshore participants have stronger tendency to cluster together.

In conclusion, stakeholders seem to form clusters of aligned understanding of priorities according to personal and cultural views, which are not affected much by their roles. The study shows the need for further research to discover: (a) the sources of diversity in prioritization of software qualities among stakeholders and (b) groupings of stakeholders with similar views for efficient team management.

The paper is structured as follows: Section II provides an outline of the related work. Section III describes the dataset. Section IV presents the statistical tools used in the analysis. Section V presents the results from the application of statistical analysis. The conclusions from the analysis are presented in Section VI. Finally, in Section VII we discuss the findings of the paper, the threats to the validity of the study and we provide directions for future work.

II. RELATED WORK

The present paper continues the work presented in [3], [4] and [5]. In [3] the authors have proposed a research methodology to analyze prioritization data of software qualities. The method was applied in a case study concerning specific products developed by a world leading company in telecommunication. Based on Theory-W [6], they identified the success critical stakeholder groups, and by using an ISO 9126-based [7] quality model, they developed a questionnaire [8], which was filled by the representatives of each of the identified success-critical stakeholder groups. At this point they used the method of Hierarchical Cumulative Voting (HCV) [9], [10] to elicit the priorities given to the various aspects of software product quality. Finally, they performed analysis on the data trying to determine the degree to which the key stakeholders are aligned regarding how they perceive software product quality in two different situations (today and ideal).

In [4], the same authors presented a case study that aimed to compare and contrast the priority two groups place on software product quality, i.e. stakeholders internal to the development organization, and stakeholders from outsourcing relationships. Finally, in a recent paper [5], another case study is presented. In this work, the authors explore the levels of alignment between key stakeholder groups on aspects of software quality for two products developed as part of an offshore insourcing arrangement. In all the aforementioned studies, the analysis was based on the calculation of specific numerical metrics of alignment, such as the Spearman correlation coefficient. The importance of stakeholder alignment in the success of a project is discussed in [11] and [12].

As already mentioned, the research of the present paper uses data obtained by the HCV method. The original method of Cumulative Voting (CV) (or the hundred dollar test) is described in [13] as a simple, straightforward and intuitively appealing voting scheme where each stakeholder is given a fixed amount (e.g. 100, 1000 or 10000) of imaginary units (for example monetary) that can use for voting in favor of the most important items (issues, aspects, requirements etc). In this way, the amount of money assigned to an item represents the respondent's relative preference (and therefore prioritization) in relation to the other items. The points can be distributed according to the stakeholder's opinion. Each stakeholder is free to put the whole amount on only one item of dominating importance or to distribute equally the amount to several or even to all of the issues.

The multivariate data obtained from CV are of special type, since the values of all variables have a constant sum. Statistical analysis of CV data was performed in [14] and in [15].

The present paper is an extension of the previous studies. Here, all the stakeholders that prioritize quality aspects of one product are first put together without taking into account their inherent groupings (role, insource- outsource, onshore-offshore) and they are clustered according to their similarities on prioritizations for both situations (today and ideal). Next, the formed clusters are analyzed with respect to

their composition so as to investigate whether similar roles or inherent groupings are clustered together. Furthermore, the distributions of the quality aspects prioritizations in each cluster are visualized graphically by boxplots and compared by a statistical test.

The contribution of the paper is twofold: first we propose a comprehensive statistical methodology that can be used on any prioritization data for discovering groups according to the respondents' views and for analysis of the composition of the groups and second the methodology is applied in a real situation so as to draw interesting results regarding the stakeholder alignment.

III. DESCRIPTION OF THE DATA SET

This paper presents a case study concerning one major product developed by a world leading company in telecommunications, providing a wide range of products and services. These are developed and sold as generic solutions, although customized versions of the product are also developed for key customers. This product is a leading solution for the market and from now on is referred simply as "Product" throughout this paper for reasons of confidentiality.

As mentioned above, stakeholder groups are classified against two criteria:

1) *Insourcing/Outsourcing*

- Insourcing (IN), defines that the work is undertaken by employees internal to the organization.
- Outsourcing (OUT), defines stakeholders employed by a consultancy company or any other subcontractor to work for the organization.

2) *Onshore/Offshore*

- Onshore (ON), defines that the work is undertaken in the home country of the organization (in our case Sweden).
- Offshore (OFF), defines that the work is conducted outside the boundaries of Sweden and specifically in India in the case study.

Furthermore, seven roles were identified according to the responsibilities each stakeholder has in the organization.

- Architects (ARCH) are responsible for the overall product design, assigning requirements to modules of the product.
- Developers (DEV) are responsible for the implementation of requirements.
- Product Support (PS) provides support to users and product owners at customer sites.
- Project Managers (PM) are responsible for planning and executing projects aligned with the priorities of the strategic product management.
- Strategic Product Managers (SPM) have the strategic product responsibility and decide the overall product development direction.
- Tactical Product Managers (TPM) support the strategic product management with expert knowledge of the systems and their architecture. It is also responsible for providing analysis of pre-project

requirements in the form of feasibility, impact and technical dependencies.

- Testers (TSR) are responsible for the verification and validation of requirements.

Stakeholders working onshore are mostly Swedish, while offshore stakeholders are mostly Indian. Outsource stakeholders are the youngest, with the less experience on the product.

The distribution of the stakeholders within the inherent identified groupings is shown in Table I.

Each stakeholder prioritized 24 aspects of software product quality. All 24 aspects are aspects of Software Product Management [16]. The aspects were prioritized using Hierarchical Cumulative Voting (HCV) in two levels. The higher level consists of three aspects and each of them contains a number of aspects of the lower level. Specifically the three aspects of the higher level are [17]:

- A is related to Features (contains 4 aspects on the lower level),
- B is related to System Properties (contains 18 aspects on the lower level),
- C is related to Project Management Constraints (contains 2 aspects on the lower level),

A description of the aspects in both levels (Higher and Lower) is shown in Table II. All the stakeholders were asked to prioritize the software product qualities using HCV from two different perspectives.

Specifically, they prioritized according to their views of how the situation is today and how they believe the priorities should be today, i.e. in their perceived ideal situation since it is important to understand what individuals prioritize, even if it may not represent official opinion. Ultimately these people work on and shape the product, and the misunderstanding of one person can have potentially huge ramifications. The data obtained by HCV were transformed to simple CV results by a procedure described in [9].

IV. THE STATISTICAL METHODOLOGY

In this section we briefly discuss the statistical methods that were used in our analysis. All these techniques comprise a methodology that first discovers clusters of aligned stakeholders and then explores the composition of the discovered clusters with respect to the various inherent groups characterizing the stakeholders. The methodology also involves comparative analysis of the quality aspects within each cluster. So we can distinguish two stages:

TABLE I. GROUPS OF STAKEHOLDERS

Role	Abbrev	TOTAL	IN (<i>Insource</i>)	OUT (<i>Outsource</i>)	ON (<i>Onshore</i>)	OFF (<i>Offshore</i>)
<i>Architects</i>	ARCH	5(7.7%)	5(8.7%)	0	0	5(18.4%)
<i>Developers</i>	DEV	22(33.9%)	15(25.9%)	7(100%)	15(39.5%)	7(26%)
<i>Product Support</i>	PS	2(3%)	2(3.5%)	0	0	2(7.4%)
<i>Project Managers</i>	PM	7(10.8%)	7(12%)	0	4(10.5%)	3(11.1%)
<i>Strategic Product Managers</i>	SPM	9(13.8%)	9(15.5%)	0	7(18.6%)	2(7.4%)
<i>Tactical Product Managers</i>	TPM	10(15.4%)	10(17.2%)	0	9(23.6%)	1(3.7%)
<i>Testers</i>	TSR	10(15.4%)	10(17.2%)	0	3(7.8%)	7(26%)
TOTAL		65	58	7	38	27

TABLE II. QUALITY ASPECTS IN TWO LEVELS

Higher Level	A: <i>Features</i>		B: <i>System Properties</i>		C: <i>Project Management Constraints</i>	
Lower Level	A1	Suitability	B1	Security	C1	Time
	A2	Accuracy	B2	Maturity	C2	Cost
	A3	Interoperability	B3	Recoverability		
	A4	Certification/ Standardization/ Compliance	B4	Containment/ISP/Fault tolerance		
			B5	Learnability		
			B6	Understandability		
			B7	Operability		
			B8	Time behavior		
			B9	Resource behavior		
			B10	Robustness/Stability		
			B11	Performance Management/Statistics		
			B12	Analyzability		
			B13	Changeability		
			B14	Testability		
			B15	Installability		
			B16	Upgradability/Replaceability		
			B17	Configurability/Product Customizability/Adaptability		
			B18	Scalability		

A. Hierarchical Cluster Analysis

Hierarchical cluster analysis (HCA) [18] is a multivariate statistical procedure that attempts to identify relatively homogeneous groups of individual cases, based on their values of a common set of variables, which are used to describe them. In our case study, HCA is used to find groups of stakeholders based on their prioritizations of quality aspects. This analysis uses an agglomerative algorithm. The algorithm performs a stepwise procedure that starts with each stakeholder in a separate cluster. In each step of the algorithm the previously formed clusters are combined into larger clusters, until all stakeholders are members of only one cluster. In the course of the algorithm, the clusters are combined on the basis of the stakeholders' agreement in their prioritizations of quality aspects.

Agreements or disagreements are visualized in a proximity matrix, which contains the values of a measure expressing stakeholders' similarity or dissimilarity. The principal advantage of the hierarchical procedure is that a tree-like hierarchy is structured so as to see in detail the consecutive connections of the stakeholders as the algorithm proceeds.

The agglomerative algorithm in HCA is based on distances between the stakeholders as individuals and between the clusters formed. We used the implementation of HCA in SPSS, which requires definition of distances between individuals and a criterion to combine clusters. In order to obtain a meaningful result, we tried several combinations and we finally decided to use as a distance between stakeholders the chi-square metric [14] and the Ward method [18] as a criterion for combining clusters.

B. Methods to Analyze Clusters

The most important output of HCA is the dendrogram, i.e. a tree diagram, which is a visual representation of the steps of the agglomerative algorithm. It shows how the stakeholders are combined at each step and also visualizes the rescaled distances with connected lines. Dendrograms can be used to assess the cohesiveness of the clusters formed and can signify the appropriate number of clusters. In our analysis we use dendrograms to show how stakeholders belonging to various roles and inherent groups are clustered together.

Another graphical tool we used in our analysis is the box-plot. Specifically, we used it to investigate differences in the prioritized quality aspects between the clusters resulted from the HCA algorithm.

The compositions of the clusters with respect to the different roles and the inherent groupings were represented by contingency tables and they were compared by the chi-square test. Also, the difference of the prioritization of each quality aspect between the clusters was tested by the nonparametric Mann-Whitney test.

V. RESULTS

The prioritized data were analyzed for both the situation as it is understood today and in the perceived ideal situation and separately for: (a) the different roles of stakeholders, (b)

the insource/outsource grouping and (c) the onshore/offshore arrangement.

The results show that although there are some local groupings at the lower levels, we can distinguish two major clusters for both the situation today and the ideal situation. For the situation today, Cluster 1 contains 48 stakeholders and Cluster 2 contains 17 stakeholders. For the ideal situation, Cluster 1 contains 44 stakeholders while Cluster 2 contains 21 stakeholders.

In general the formations of the two clusters under the two situations are not independent, as shown by the chi-square test, which gives $p=0.007$. In Table III we can see the distribution of stakeholders according to the cluster they belong and the perceived situation that they prioritize. As we can see, 37 stakeholders are common in the "larger" clusters while 10 of them are common in the "smaller" clusters. This suggests that the groupings of the stakeholders under the two perspectives have a strong dependence. The result implies that stakeholders who are aligned in their prioritization of the situation today are most probably aligned in their prioritization of the ideal situation, i.e. when a group of people agree in their views of quality aspects of the situation today, their agreement will most probably be maintained when they express the ideal situation.

The entire clustering procedure for the situation today is shown in the dendrogram of Fig. 1. Due to space limitation we omit the corresponding dendrogram for the ideal situation. Moreover, the stakeholders are labeled with their role and inherent groupings in order to show how they are combined in clusters. Table IV shows the distributions of roles within derived clusters.

In the situation today, all Testers (TSR) and both Product Support (PS) stakeholders are members of the large cluster. Generally, there are differences in the compositions of the clusters (for example the role with the largest percentage in Cluster 1 is the Developers (DEV) while in Cluster 2 it is the Tactical Product Managers (TPM). However, there is no strong statistical evidence that the clustering depends on the roles (the chi-square test gives $p>0.05$). This is also true for the ideal situation where even the Testers and the Product Support roles are split.

From Fig. 1 we can see that although some stakeholders with the same role (especially Developers and Testers) are combined at early stages of the algorithm, showing similarity in prioritizations, in general there is no strong connectivity between employees of the same role. This means that if we try to explain the alignment of various stakeholders' in terms of their prioritization, we have to discover other factors or characteristics, apart from their role, which may affect the agreement. The role itself is a very weak criterion for judging the stakeholders' preferences.

TABLE III. CLUSTER ANALYSIS FOR TODAY AND IDEAL SITUATION

		Cluster results for ideal		
		Cluster 1	Cluster 2	Total
Cluster results for today	Cluster 1	37	11	48
	Cluster 2	7	10	17
Total		44	21	65

The compositions of clusters regarding the insource/outsource grouping for the situation today and the ideal situation are presented in Table IV too. We can see that in the situation today the seven outsource stakeholders (which are all in onshore relationship) are clustered together in the first cluster but in general the connections among them are not strong, i.e. they are grouped together at later steps of the algorithm as we can see from Fig 1.

Also the appearance of all of them in the first larger cluster is not statistically significant, i.e. there are no

adequate statistical evidence ($p > 0.05$) that their clustering depends on the outsource characteristic. For the ideal situation, the compositions of the two clusters are quite similar, so it is clear that the clustering does not depend to insourcing/outourcing characteristic. Therefore, the insource/outsource arrangement of stakeholders does not offer any significant information to explain their alignment in prioritization.

TABLE IV. COMPOSITION OF CLUSTERS WITH RESPECT TO THE INITIAL GROUPS

<i>Situation Today</i>	<i>Role</i>	<i>ARCH</i>	<i>DEV</i>	<i>PM</i>	<i>PS</i>	<i>SPM</i>	<i>TPM</i>	<i>TSR</i>	<i>Total</i>
Cluster 1	Count (%)	4(8.3%)	18(37.5%)	4(8.3%)	2(4.2%)	5(10.4%)	5(10.4%)	10(20.8%)	48 (100%)
Cluster 2	Count (%)	1(5.9%)	4(23.5%)	3(17.6%)		4(23.5%)	5(29.4%)		17(100%)
Total	Count (%)	5(7.7%)	22(33.8%)	7(10.8%)	2(3.1%)	9(13.8%)	10(15.4%)	10(15.4%)	65(100%)
	<i>Insource/ Outsource</i>	<i>Insource</i>			<i>Outsource</i>				<i>Total</i>
Cluster 1	Count (%)	41(85.4%)			7(14.6%)				48(100%)
Cluster 2	Count (%)	17(100%)							17(100%)
Total	Count (%)	58(89.2%)			7(10.8%)				65(100%)
	<i>Onshore/ Offshore</i>	<i>Offshore</i>			<i>Onshore</i>				<i>Total</i>
Cluster 1	Count (%)	26(54.2%)			22(45.8%)				48(100%)
Cluster 2	Count (%)	1(5.9%)			16(94.1%)				17(100%)
Total	Count (%)	27(41.5%)			38(58.5%)				65(100%)
<i>Ideal Situation</i>	<i>Role</i>	<i>ARCH</i>	<i>DEV</i>	<i>PM</i>	<i>PS</i>	<i>SPM</i>	<i>TPM</i>	<i>TSR</i>	<i>Total</i>
Cluster 1	Count (%)	4(9.1%)	16(36.4%)	2(4.5%)	1(2.3%)	6(13.6%)	6(13.6%)	9(20.5%)	44 (100%)
Cluster 2	Count (%)	1(4.8%)	6(28.6%)	5(23.8%)	1(4.8%)	3(14.3%)	4(19.0%)	1(4.8%)	21(100%)
Total	Count (%)	5(7.7%)	22(33.8%)	7(10.8%)	2(3.1%)	9(13.8%)	10(15.4%)	10(15.4%)	65(100%)
	<i>Insource/ Outsource</i>	<i>Insource</i>			<i>Outsource</i>				<i>Total</i>
Cluster 1	Count (%)	40(90.9.4%)			4(9.1%)				44(100%)
Cluster 2	Count (%)	18(85.7%)			3(14.3%)				21(100%)
Total	Count (%)	58(89.2%)			7(10.8%)				65(100%)
	<i>Onshore/ Offshore</i>	<i>Offshore</i>			<i>Onshore</i>				<i>Total</i>
Cluster 1	Count (%)	22(50.0%)			22(50.0%)				44(100%)
Cluster 2	Count (%)	5(23.8%)			16(76.2%)				21(100%)
Total	Count (%)	27(41.5%)			38(58.5%)				65(100%)

TABLE V. RESULTS OF THE MANN-WHITNEY TEST FOR THE DIFFERENCES BETWEEN CLUSTERS

Software quality aspects	Today	Ideal		Software quality aspects	Today	Ideal
Suitability (A1)	<0.001	0.004		Resource behavior (B9)	0.976	0.523
Accuracy (A2)	0.008	0.428		Robustness/Stability (B10)	0.289	0.005
Interoperability (A3)	0.003	0.381		Performance Management/Statistics (B11)	0.263	0.483
Compliance (A4)	0.177	0.214		Analyzability (B12)	<0.001	0.340
Security (B1)	0.092	0.833		Changeability (B13)	<0.000	<0.001
Maturity (B2)	0.676	0.001		Testability (B14)	0.002	0.007
Recoverability (B3)	0.003	<0.001		Installability (B15)	<0.001	0.060
Fault tolerance (B4)	0.008	0.710		Upgradability/Replaceability (B16)	0.001	0.408
Learnability (B5)	0.002	0.052		Configurability/Product Customizability/Adaptability (B17)	0.156	0.369
Understandability (B6)	0.002	<0.001		Scalability (B18)	0.030	0.752
Operability (B7)	0.002	0.001		Time (C1)	0.001	0.630
Time Behavior (B8)	0.210	0.041		Cost (C2)	0.893	0.424

A: Features, B: System Properties, C: Project Management Constraints

In Table IV we can see the composition of clusters regarding the initial onshore/offshore grouping in the situation today and ideal situation respectively. It is clear that for today, the compositions of clusters are different. The interesting here is that the second cluster contains mainly onshore stakeholders. Almost all of the offshore stakeholders are grouped in the first cluster. Here, the chi-square statistic shows significant dependence of the clustering on the initial onshore/offshore grouping ($p=0.001$).

The same is true for the ideal situation. Here the chi-square test gives $p=0.045$ showing that the dependence is significant though it seems to be not as strong as for the today situation.

The results show that the alignment of stakeholders is affected by the onshore-offshore relationship and this seems to be an important factor for explaining agreements, which deserves further investigation.

In an effort to identify and compare the profiles of the members forming the derived clusters, we tried to depict how the stakeholders in each one of the clusters prioritize the quality aspects. Having in mind that clusters are essentially groups of people with some degree of coherence in their views, we are interested in comparing the distribution of prioritizations of software quality aspects in each one of them. In Fig. 2, 3 and 4 we can see respectively how the stakeholders in each cluster prioritize the aspects of quality related to Features (A1-A4), System Properties (B1-B18) and Project Management Constraints (C1, C2) for the today situation. The boxplots show the distribution of the prioritization values each aspect received, separately in each cluster. Both clusters are represented in the same figure for comparison.

It is noticeable that the prioritizations of the A1 aspect (suitability) in Cluster 2 have a large variation. In general, all quality aspects of Features are prioritized higher in Cluster 2. On the contrary, the quality aspects of System Priorities receive lower ranking in Cluster 2. Regarding the two aspects of Project Management Constraints, Time seems to be more important for Cluster 2 while Cost seems to be equally important for both clusters. Another interesting point is that all aspects are prioritized more uniformly in Cluster 1 than in Cluster 2 where we generally have larger discrepancies.

In similar Figures (omitted due to space limitation) we can see respectively how the prioritizations of quality aspects related to Features (A1-A4), System Properties (B1-B18) and Project Management Constraints (C1, C2) are distributed within each of the two clusters for the ideal situation. Some of the results are quite similar to the situation today, but there are also some differences. Members of Cluster 1 prioritize the quality aspects in a more uniform manner, the aspect A1 (suitability) seems to be more important for members of Cluster 2 while the other aspects of Features seem to be equally important for both clusters. Some of the aspects of System Properties are most important for Cluster 1 while others are equally important. Both aspects of Project Management Constraints seem to be equally important for both clusters. In general, in Cluster 2 we can observe larger deviations.

Finally, we applied the non-parametric Mann-Whitney test on each one of the quality aspects separately, to see which ones have significance differences between the two clusters. In this way we can characterize the clusters according to their tendency to prioritize more or less certain aspects. The results of the test are shown in Table V where we report the p-value of the test for each quality aspect and for both the today and ideal perspective. The aspects with significant differences between clusters (p-value) are written in bold. These are: for the situation today, aspects A1, A2, A3, B3, B4, B5, B6, B7, B12, B13, B14, B15, B16, B18 and C1 and for the ideal situation, aspects A1, B2, B3, B6, B7, B8, B10, B13 and B14 have statistically significant difference between Clusters.

VI. CONCLUSIONS

The clustering showed that there exist two major clusters in the situation today and the ideal situation. We have shown that the way stakeholders are clustered is not independent from the situation, i.e. there is a large body of common stakeholders that participate in a similar clustering scheme in both situations, today and ideal. This result is consistent with the “garbage can model”, which found that managers prefer to address issues with incremental change, rather than rapid evolution (i.e. people are conservative when it comes to changing things) [19].

The first cluster in both situations is the largest (for convenience we refer to the largest cluster of each scheme as Cluster 1) and is characterized by a more uniform prioritization of quality aspects. The second cluster is much smaller than the first (its size is less than the half size of the first) and is characterized by large variations either within the same aspect or among aspects. The quality aspects under the Features level seem to be in general most important for Cluster 2 while the aspects of the System Properties level seem to be more important for Cluster 1, although some of the results are different for the ideal perspective. Also, Cluster 2 gives significantly larger importance to Time of Project Management Constraints for the situation today.

The clustering does not seem to be significantly affected by the role of the stakeholder or the insource/outsource grouping. The fact that outsourcing stakeholders are not very close to each other but they are all clustered with the insource stakeholders is a positive indication for the company since it shows that this group is mixed with the internal employees and it has not the behavior of a minority which is isolated from the other internal employees. The groups defined by onshore insourcing and onshore outsourcing, work in a co-located environment. Thus it is possible for them to easily ask questions, hold meetings and have impromptu discussions. These factors are possibly the reason for the groups alignment.

On the other hand, it was found that the clustering is associated with the onshore/offshore grouping as almost all of the offshore group are clustered together.

The general conclusion from the analysis is that some stakeholders show agreement in their prioritization and can form groups of homogeneous views.

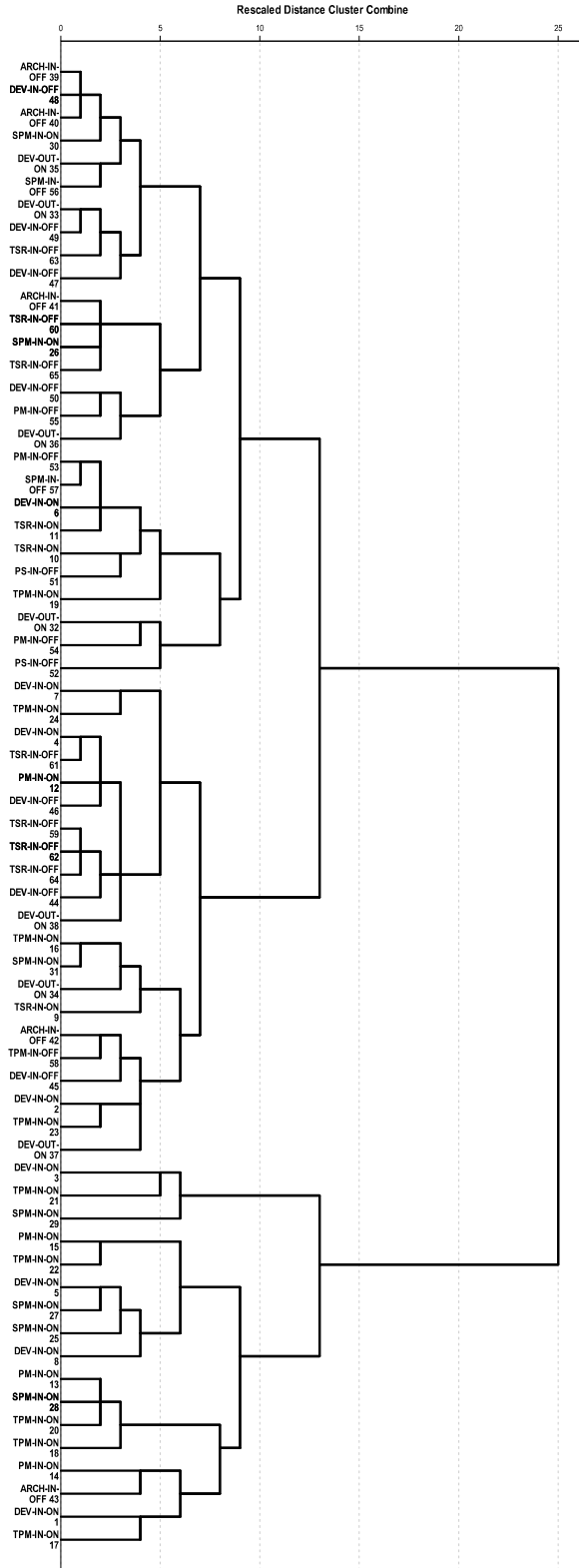


Figure 1. Clustering of stakeholders for the situation today

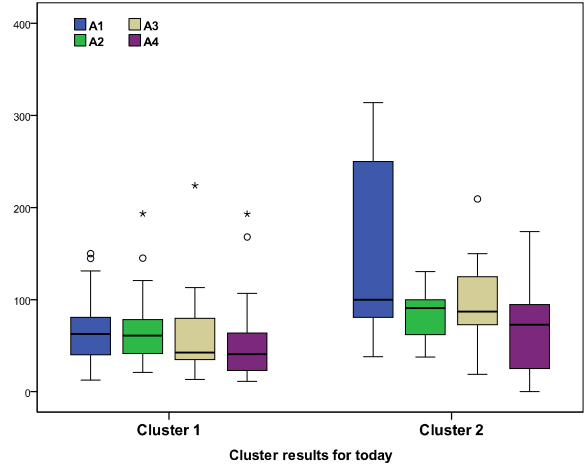


Figure 2. Quality aspects related to Features (A1-A4) for situation today

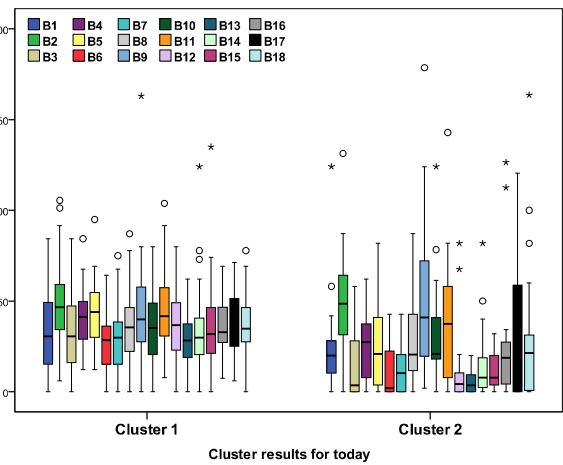


Figure 3. Quality aspects related to System Properties (B1-B18) for situation today

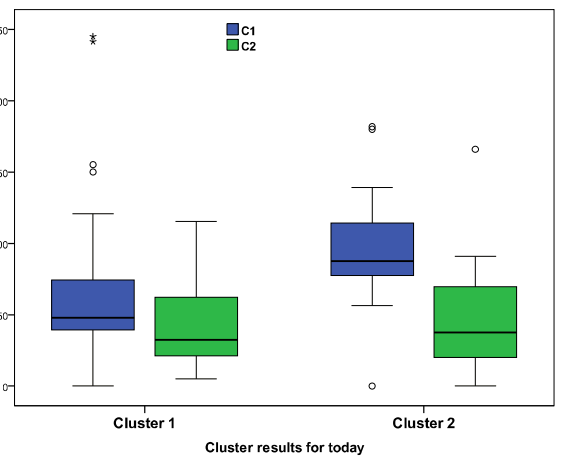


Figure 4. Quality aspects related to Project Management Constraints (C1-C2) for situation today

However, there are factors behind the alignment of stakeholders that remain unseen. The effect of onshore/offshore grouping on the clustering is perhaps an indication that cultural and educational characteristics are very important and that should be recorded and analyzed more.

The methodology presented in this paper is useful for investigating the similarities between stakeholders and the factors that may be responsible for them. Further it may be used as a tool for composing teams of professionals that share similar views on quality aspects and have different roles or belong to different groups internal or external to a company, even from different countries. So the Hierarchical Cluster Analysis and the subsequent analysis of the composition of clusters can be a valuable aid to discover and bring together people that may seem completely irrelevant due to their positions but who in fact have similar views on critical quality aspects.

VII. DISCUSSION – THREATS TO VALIDITY

The paper focuses on issues of alignment between success-critical stakeholders involved in the development of software products. The need for agreement between various groups of stakeholders towards a common goal is very important for the success of a project as emphasized by literature and shown by common practice.

The employees of an organization working in a global environment for the development of a software product may have large differences in their personal characteristics depending on their personality, culture and professional or even social status. Since it is very difficult for an organization to record the views on quality of a large number of employees working on the same project and assess their agreement, the structured and systematic collection of prioritization data through the Cumulative Voting method and the multivariate statistical analysis we present here are valuable aids for studying and understanding the reasons of alignment and possibly the causes of success or failure of a project. Moreover, the present study indicates that there is a need for collection of other attributes (for example cultural, educational, professional etc.) that may reveal the latent factors that affect the alignment.

Regarding the validity threats, we have to remark that our case study is exploratory and is not meant to generalize the results, since the current data are not a representative sample from a population but on the contrary, they have been selected from a specific company and for a very specific product. All methods described here, i.e the collection of data and the statistical techniques, have not been used for studying alignment of the perception of software quality aspects earlier.

The only generalization from the analysis is that if a company wants to investigate in depth the problem of stakeholders' alignment, it needs to take into account more personal characteristics. The study of connecting the prioritization alignment with alignment of other personal attributes is a direction of future research. Also another direction for future research is to go beyond the analysis results and use them to build teams, networks or to perform

other managerial actions that will enhance the joint work towards common goals.

REFERENCES

- [1] D. Šmite, C. Wohlin, T. Gorschek, and R. Feldt, "Empirical evidence in global software engineering: a systematic review," *Empirical Software Engineering*, vol. 15, no. 1, pp. 91–118, February 2010.
- [2] E. Ó. Conchúir, P. J. Ågerfalk, H. H. Olsson, and B. Fitzgerald, "Global software development: Where are the benefits?" *Communications of the ACM*, vol. 52, pp. 127–131, August 2009.
- [3] S. Barney and C. Wohlin, "Software product quality: Ensuring a common goal," *Proc. International Conference on Software Process (ICSP)*, Q. Wang R. Madachy, and D. Pfahl, Eds., May 2009, pp. 256–267.
- [4] S. Barney and C. Wohlin, "Alignment of software product quality goals in two outsourcing relationships," *Proc. 14th International Conference on Evaluation and Assessment in Software Engineering (EASE)*, April 2010.
- [5] S. Barney, C. Wohlin, P. Chatzipetrou and L. Angelis, "Offshore Insourcing: A Case Study on Software Quality Alignment", *Proc. 6th IEEE International Conference on Global Software Engineering*, Helsinki, Finland, August 2011, in press
- [6] Boehm, B., Ross, R.: *Theory-w software project management principles and examples*. *IEEE Transactions on Software Engineering* 15(7) (1989) 902-916
- [7] ISO9126: *Software engineering-product quality-part 1: Quality model*. International Standards Organization (2001)
- [8] S. Barney, C. Wohlin. *Software product quality questionnaire*. <http://sebseb.info/files/icsp2009-questionnaire.xlt> (2008)
- [9] P. Berander and P. Jönsson, "Hierarchical Cumulative Voting (HCV) -Prioritization of requirements in hierarchies," *International Journal of Software Engineering and Knowledge Engineering*, vol. 16, no. 6, pp. 819–849, December 2006
- [10] P. Berander, P. Jönsson. *A goal question metric based approach for efficient measurement framework definition*. *Proc. of the 2006 ACM/IEEE International Symposium on Empirical Software Engineering (ISESE '06)*, New York, NY, USA, ACM (2006) pp. 316-325
- [11] Y. E. Chan, "Why haven't we mastered alignment? the importance of the informal organization structure," *MIS Quarterly Executive*, vol. 1, no. 2, pp. 97–112, June 2002
- [12] Y. E. Chan and B. H. Reich, "IT alignment: An annotated bibliography," *Journal of Information Technology*, vol. 22, no. 4, pp. 316–396, October 2007
- [13] Leffingwell, D. and D. Widrig, *Managing software requirements: A Use Case Approach*, 2nd ed. Addison-Wesley, Boston, 2003
- [14] P. Rovegård, L. Angelis and C. Wohlin, "An Empirical Study on Views of Importance of Change Impact Analysis Issues", *IEEE Transactions on Software Engineering*, vol. 34, No. 4, July/ August 2008, pp. 516-530
- [15] P. Chatzipetrou, L. Angelis, P. Rovegård and C. Wohlin, "Prioritization of Issues and Requirements by Cumulative Voting: A Compositional Data Analysis Framework", *Proc. of EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA)*, 1-3 September 2010, Lille, France, pp. 361-370
- [16] K. Beck, *Extreme Programming Explained: Embrace Change*. Reading, Massachusetts: Addison-Wesley, 2000.
- [17] S. Barney and C. Wohlin, "Software product quality definitions" <http://sebseb.info/files/icsp2009-definitions.pdf>
- [18] R. A. Johnson, D. W. Wichern. *Applied Multivariate Statistical Analysis*, Prentice Hall, 2002
- [19] M. D. Cohen, J. G. March, and J. P. Olsen. *A garbage can model of organizational choice*. *Administrative Science Quarterly*, 17(1):1–25, March 1972.