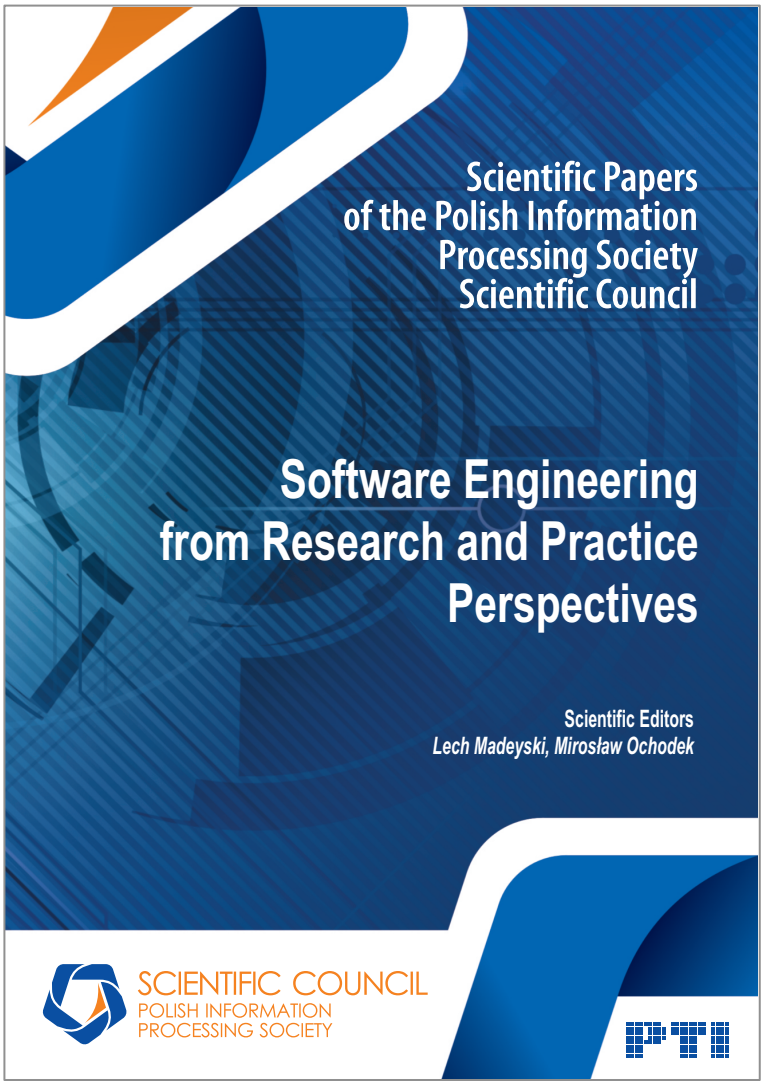
CHAPTER 6

How software development factors influence user satisfaction in meeting business objectives and requirements?

By **Łukasz Radliński**



How to cite:

Łukasz Radliński. "How software development factors influence user satisfaction in meeting business objectives and requirements?." In Software Engineering from Research and Practice Perspective (Eds. Lech Madeyski, M. Ochodek), pp. 101-120, Scientific Papers of the Polish Information Processing Society Scientific Council, ISBN 978-83-63919-16-0, 2014.

Table of Contents

Authors and affiliations

Preface
Lech Madeyski
Wroclaw University of Technology, Faculty of Computer Science and Management, Institute of Informatics
lech.madeyski@pwr.edu.pl
Mirosław Ochodek
Poznan University of Technology, Faculty of Computing
miroslaw.ochodek@cs.put.poznan.pl

CHAPTER 1
Jakub Jurkiewicz
Poznan University of Technology, Faculty of Computing
jakub.jurkiewicz@cs.put.poznan.pl
Piotr Kosiuczenko
Military University of Technology in Warsaw, Department of Cybernetics
pkosoiczenko@wat.edu.pl
Lech Madeyski
Wroclaw University of Technology, Faculty of Computer Science and Management, Institute of Informatics
lech.madeyski@pwr.edu.pl
Mirosław Ochodek
Poznan University of Technology, Faculty of Computing
miroslaw.ochodek@cs.put.poznan.pl
Cezary Orłowski
IBM Center for Advanced Studies on Campus, Gdańsk,
cor@zie.pg.gda.pl
Łukasz Radliński
West Pomeranian University of Technology,
Faculty of Computer Science and Information Technology
lukasz.radlinski@zut.edu.pl

CHAPTER 2
Marek Majchrzak
Capgemini Poland
Wroclaw University of Technology, Faculty of Computer Science and Management, Institute of Informatics
marek.majchrzak@capgemini.com

Łukasz Stilger
Capgemini Poland
lukasz.stilger@capgemini.com
Marek Matczak
Capgemini Poland
marek.matczak@capgemini.com

CHAPTER 3

Tomasz Sitek
Gdańsk University of Technology, Faculty of Management and Economics
tsitek@zie.pg.gda.pl
Artur Ziółkowski
Gdańsk University of Technology, Faculty of Management and Economics
aziolko@zie.pg.gda.pl

CHAPTER 4

Miklós Biró
Software Competence Center Hagenberg,
miklos.biro@scch.at

CHAPTER 5

Stanisław Jerzy Niepostyn
Warsaw University of Technology, Institute of Computer Science
S.Niepostyn@ii.pw.edu.pl
Andrzej Tyrowicz
Agencja Europejska, Andrzej Tyrowicz
a@tyrowicz.eu

CHAPTER 6

Łukasz Radliński
West Pomeranian University of Technology,
Faculty of Computer Science and Information Technology
lukasz.radlinski@zut.edu.pl

CHAPTER 7
Bogumiła Hnatkowska
Wroclaw University of Technology, Faculty of Computer Science and Management, Institute of Informatics
Bogumila.Hnatkowska@pwr.edu.pl
Łukasz Wrona
Wroclaw University of Technology, Faculty of Computer Science and Management, Institute of Informatics
lukwro83@gmail.com

CHAPTER 8
Andrzej Ratkowski
Warsaw University of Technology,
Institute of Control and Computation Engineering
Krzysztof Gawryś
Warsaw University of Technology,
Institute of Control and Computation Engineering
Eliza Świątek
Warsaw University of Technology,
Institute of Control and Computation Engineering

CHAPTER 9
Ilona Bluemke
Warsaw University of Technology, Faculty of Electronics and Information Technology, Institute of Computer Science
I.Bluemke@ii.pw.edu.pl
Anna Stepień
Warsaw University of Technology, Faculty of Electronics and Information Technology, Institute of Computer Science
A.Stepien.1@stud.elka.pw.edu.pl

CHAPTER 10
Michał Ćmil
Poznan University of Technology, Faculty of Computing
cmilmichal@gmail.com
Bartosz Walter
Poznan University of Technology, Faculty of Computing
bartosz.walter@cs.put.poznan.pl

CHAPTER 11

Wojciech Frącz
AGH University of Science and Technology
fracz@iisg.agh.edu.pl
Jacek Dajda
AGH University of Science and Technology
dajda@agh.edu.pl

Chapter 6

How software development factors influence user satisfaction in meeting business objectives and requirements?

User satisfaction is an useful measure of success of software development projects. The goal of this chapter is to analyze if and how individual factors describing software development process and product are related with selected features of user satisfaction. This chapter investigates two features of user satisfaction: meeting stated objectives (MSO) and meeting business requirements (MBR). Achieving such goal involved using visual techniques as well as a range of statistical and data mining techniques. For MSO there are more identified relationships and they are usually stronger than for MBR. Although there are some disagreements in relationships identified with different techniques, there is a common set of explanatory variables identified by most of techniques. Identified relationships can be used to build more complex simulation or predictive models.

User satisfaction is one of the most important features of software quality. In general, information systems are developed to meet the needs of their users. Satisfaction reflects the level of fulfillment of users' needs by a software system, Thus, the inherent problem with user satisfaction is that it is impossible to express it objectively and difficult to empirically prove what influences it.

Nevertheless, this chapter makes an attempt to identify factors that are related with user satisfaction. Typically, the literature on user satisfaction focuses on application and management perspectives, without links to software development context. In contrast, this study focuses on core software engineering factors.

User satisfaction can be treated as an aggregated measure or broken down into a set of detailed characteristics. This analysis uses the extended ISBSG dataset of software projects [I09] where user satisfaction is expressed by eight variables. This study investigates two of these variables that are important from the business perspective, i.e. user satisfaction with the ability of system to **meet stated objectives** (MSO) and to **meet business requirements** (MBR). Other aspects of satisfaction will be investigated in future studies. The main research questions are as follows:

— RQ1. Which software engineering factors influence MSO and MBR?
— RQ2. What is the nature of these relationships?

To answer these questions, we follow a research approach involving the use of a range of statistical and data-mining techniques (explained in Section 6.2). The main contribution of this chapter is a list of software engineering factors

identified by different techniques as related with MSO and MBR. In addition, this chapter discusses the nature of these relationships, pointing out some caution, where appropriate, in interpreting pure quantitative results.

It is important to note that the chapter does not investigate what factors influence if the stated objectives or business requirements are met. Rather, it investigates user satisfaction in these aspects. It is possible that stated objectives are met only in some degree but the user is still generally satisfied with that situation.

6.1 Related work

There are two main related groups of studies, i.e. exploratory studies and prediction studies. The exploratory studies, like the current study, focus on understanding specific phenomena based on analysis of empirical data, expert knowledge, observations, surveys etc. Several such studies investigating user satisfaction have been performed [B05], [MAD12], [P08], [TT10], [WT05]. Studies of user satisfaction with an empirical emphasis have been performed for about 25 years [K93], [W88]. Most of these studies focus on application and management perspectives, typically without strong links to software engineering. In addition, these studies usually involve analysis of very few projects or deeper analysis of just a single project. Thus, while results provided in these studies may be useful in specific context, it is difficult to draw more general conclusions based on them.

In more recent study [SW10] the authors argue that "high rate of developer turnover in projects (due to dissatisfaction) could lead to increasing costs for development firms as well as high user/customer dissatisfaction". The authors observe that with increasing level of user participation the level of developer satisfaction also increases, however the level of user satisfaction slowly decreases. In [RC10] the authors investigate one aspect of user satisfaction, i.e. improving software usability in open-source software. Using a range of statistical methods they analyze factors that might be relevant, i.e., understanding users' requirements, seeking usability experts' opinion by software developers, incremental approach in design, usability testing by managers/developers, and knowledge of user-centered design methods. Since satisfaction with usability will be investigated in future, these results are not relevant for the current study.

The other group of studies aim at building model(s) that could be used to predict future states of certain phenomena based on a set of observations and/or assumed states. While a range of predictive models or frameworks for building them have been proposed in software quality literature [C11], [HB12],

[SBH14], very few studies focus on prediction of user satisfaction [FM04], [PP13]. In contrast with the current study, user satisfaction is defined there as an aggregate measure. In [FM04] the focus is on resource prediction; satisfaction depends here on the combination of software quality (mainly its defectiveness) and specification accuracy. A model developed in [PP13] predicts user satisfaction based on the definition of software requirements.

6.2 Methodology and data

This study uses the extended ISBSG dataset of software projects [I09]. The extension means that it contains additional, usually soft, features, such as user satisfaction. Although the ISBSG dataset has been used in numerous studies, e.g. [FG14], [KJ13], [ML08], such extended version in only very few [R11b], [R12]. After numerous data preprocessing steps, explained later in this section, the subset of the dataset used for the main part of analysis contains data on 89 projects described by a set of variables listed in Table 6.1.

The research methodology contains the following steps:

1. **Basic preprocessing**. This step involved activities such as replacing "don't know" values to "missing"; replacing rare (i.e. with counts close to 1) values of categorical variables by a similar but more common value or marking as "other"; ensuring consistency of values between two variables (e.g. *Client-server* and *Architecture*); removing variables with too many states and very few counts; removing variables with many unclear values (e.g. *Primary programming language*, *1st hardware*, *1st operating system*); transforming variables to their appropriate type (especially logical variables or variables with mixed numeric or interval values); creating logical dummy variables for multi-value categorical variables. This step prepared a dataset to many types of possible future analyses.

Table 6.1. A list of variables used in analysis

Name	Type	N	Notes
Meet stated objectives	logical	89	outcome variable
Meet business requirements	logical	89	outcome variable
Year of project	integer	89	based on project completion date
Adjusted function points	integer	65	also transformed by: ln(x)
Summary work effort	integer	89	in hours; also transformed by: ln(x)
Total defects delivered	integer	72	in the first month after release; also transformed by: ln(x+1)
Development type	nominal	89	new, enhancement, re-development
Architecture	nominal	86	stand-alone, multi-tier/client-server

Name	Type	N	Notes
Client-server	logical	86	
Development platform	nominal	82	PC, Mid-range/multi, mainframe
Language type	nominal	78	3GL, 4GL
Used methodology	logical	81	
Resource level	nominal	89	1-4, the way of recording effort
Debugging tools	logical	72	
Testing tools	logical	72	
Performance monitoring tools	logical	68	
User satisfaction survey	logical	77	
Survey respondent role	nominal	86	customer/user, project manager, sponsor
Project activity scope	logical	77	separate variables for: planning, specification, design, build, test, implement
Organization type	logical	87	separate variables for: computers and software, communications, financial, manufacturing, professional services, other
Application type	logical	88	separate variables for: management information system, network management, transaction/production system
Productivity	numeric	65	in function points per person-hour; also transformed by: ln(x)
Proportion of effort on specification	numeric	67	also transformed by: sqrt(x)
Proportion of effort on build	numeric	81	

2. **Data selection**. This involved filtering the data to the cases with *Data quality rating* set to "A" or "B" as suggested in [I05]; filtering the data so that no attribute describing user satisfaction contains missing values; and filtering the data so that no explanatory variable contains more than 30% of missing data.
3. **Further preprocessing**. This involved repetition of activities as in step 1, but performed on dataset reduced in step 2.
4. **Defining additional variables**. This involved creating variables such as *Productivity*, *Defect rate*, and proportions of effort on specific activities (see Table 6.1). These new variables were then filtered to also meet the criteria for fraction of missing values.

5. **Defining transformed variables**. Since some techniques that were planned to use require normal distribution of variables, this step investigated if such requirement is met and, if necessary, new variables were created after applying commonly used transformations: ln(x), ln(x+1) or sqrt(x) (see Table 6.1).
6. **Defining outcome variables**. The variables describing user satisfaction are originally defined on a 1-4 ranked scale. Since the value "1" and "4" are rare, outcome variables have been defined as logical, i.e. original values "1" and "2" replaced by "false" and original values "3" and "4" replaced by "true" – to indicate if the user satisfaction in particular aspect has been met in a project. Table 6.2 illustrates the distributions for both outcome variables.

Table 6.2. Distributions for two outcome variables

	Meet stated objectives		**Meet business requirements**	
	false	**true**	**false**	**true**
Counts	31	58	19	70

7. **Final variable filtering**. This involved removing variables not suitable for causal analysis, e.g. *Project Id*, *Data quality rating*. Table 6.1 lists variables used in analysis that were kept after this step.
8. **Analysis of correlations for numeric explanatory variables**. This is the first step of the main part of the analysis. Since the outcome variables are dichotomous, this involved the use of a point biserial correlation coefficient which is mathematically equivalent to Pearson product-moment correlation coefficient (assuming encoding values "false"/"true" to "0"/"1") and its interpretation is also the same. Because this step involved using the same statistical test multiple times, i.e. for many explanatory variables, to reduce the risk of reporting false positive results, obtained p values were adjusted according to the false discovery rate (fdr) control [BH95]. This analysis was supported by investigating scatterplots, some of which have been discussed later in this chapter.
9. **Analysis of associations for categorical and logical explanatory variables**. This involved preforming Pearson's chi-squared test or Fisher's exact test for each pair of outcome and categorical/logical explanatory variable. The Pearson's test has been performed when the following conditions were met: a cross-tabulation in a form of z 2x2 table contains at least 5 counts in each cell, in larger tables there are no cells with count of zero and at least 80% of cells have counts of at least 5. Otherwise, the Fisher's test was performed. The p-values obtained in these tests have also been adjusted with "fdr" (as explained in previous step). Statistical-

ly significant relationships, i.e. with p adjusted ≤ 0.05, have been further investigated by analyzing the effect size, i.e. the strength of the relationship, using the following commonly used measures: phi, Cramer's V, Pearson's contingency coefficient, and lambda coefficient.

10. **Analysis of logistic regression models**. The goal of this step was to investigate how each explanatory variable explains the variability of outcome variables. Thus, we built a set of logistic regression models – one for each pair of explanatory and outcome variables. Each of these models contains a single explanatory variable, i.e. it does not take into account any possible interactions between explanatory variables. Such assumption was necessary because adding further variable(s) to the model would result in the need to build the model using fewer cases – most explanatory variables contain missing values and cases with missing values for explanatory variables cannot be used to build the model.
11. **Analysis of data-mining measures of associations**. The goal of this step was to support previous analyses by using other measures of association that are frequently used in data-mining: ReliefF, information gain, gain ratio, and Gini index. The last three measures require variables on non-continuous scale. Thus, to calculate them, each numeric variable has been discretized into five intervals.
12. **Analysis of the CN2 rules** [CN89]. The goal of this step was to learn a set of "if-then" rules that would explain variability in the outcome variables. Such rules can be relatively easily interpreted by human and also used for prediction. What is important, these rules capture the interaction between explanatory variables. An earlier paper [R11b] demonstrated that the CN2 algorithm can produce meaningful and useful rules.

Steps 1-10 were performed with the R statistical software environment [R14] and steps 11-12 using Orange [DCE13].

6.3 Results

6.3.1 Statistical explanatory analysis

The first main step of the analysis was the investigation of relationship between numeric explanatory variables and each outcome variable. Table 6.3 lists the values of point biserial correlation coefficient (r_{pb}) and respective p values (adjusted by "fdr") for each analyzed pair of variables. There are four variables in statistically significant relationship with MSO – one medium-strength positive relationship for the *Year of project* and three medium-strength negative relationships: for *Total defects delivered (ln)*, *Adjusted function points (ln)*, and *Productivity (ln)*.

No statistically significant relationship was found for MBR and any explanatory variable. Without adjusting the p value relationships with *Year of project* and *Adjusted function points (ln)* would be statistically significant.

Table 6.3. Values of measures of association for numeric explanatory variables

Explanatory variable	**N**	**Meet stated objectives**		**Meet business requirements**	
		r_{pb}	**p adj.**	r_{pb}	**p adj.**
Year of project	89	0.40*	0.001	0.25	0.08
Adjusted function points (ln)	65	-0.29*	0.04	-0.28	0.08
Summary work effort (ln)	89	0.14	0.24	0.14	0.34
Total defects delivered (ln)	72	-0.40*	0.002	-0.10	0.59
Productivity (ln)	65	-0.29*	0.04	-0.24	0.13
Prop effort specify (sqrt)	67	-0.19	0.17	0.01	0.94
Prop effort build	81	-0.12	0.28	-0.02	0.94

** - significant at p adjusted. ≤ 0.05*

Figure 6.1 illustrates relationships between strongest numeric explanatory variables and two outcome variables – MSO (parts a and b) and MBR (parts c and d). Parts (a) and (c) confirm that with increasing *Year of project* there are more projects that satisfy both MSO and MBR. However, such straightforward conclusion may be biased by two facts: First, many projects have a *Year of project* with a single common value (2000). Second, there is no project with *Year of project* higher than 2000 for which MSO and MBR were not satisfied. This, based on literature and author's knowledge, cannot be well justified by any theory.

As expected, in projects with fewer defects users were more frequently satisfied in terms of MSO (Figure 6.1, part a). However, no such relationship was found for MBR (Figure 6.1, part c).

Parts (b) and (d) show that with an increased project size (i.e. higher value of *Adjusted function points*) the *Productivity* also increases. Furthermore, in the group of larger projects, the proportion of projects with dissatisfied user increases. However, probably because of imbalanced data, such relationship does not appear to be significant for MBR (Figure 6.1, part d).

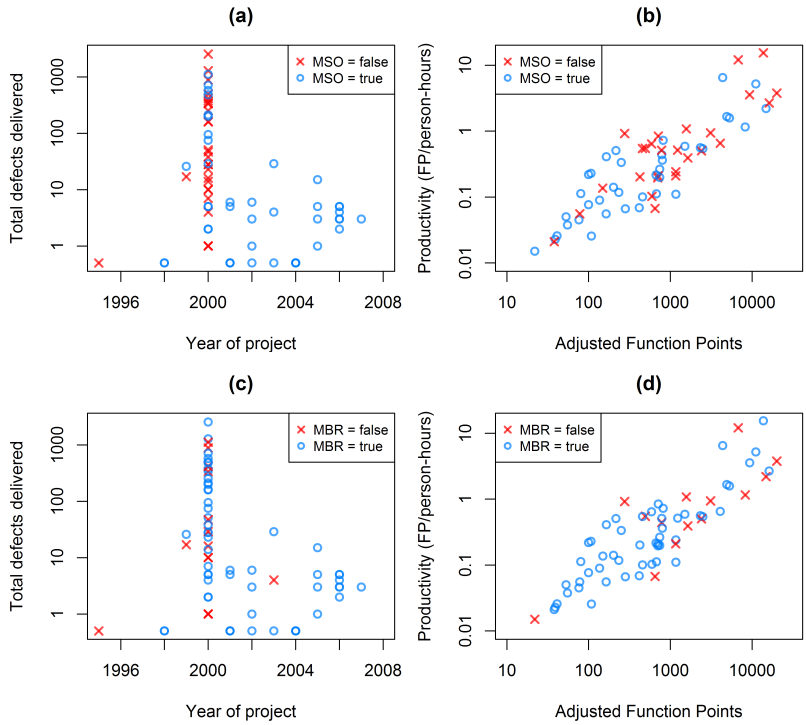


Figure 6.1. Scatterplots for outcome variables and strong explanatory variables

Table 6.4 lists the results of investigating associations between MSO and each explanatory variable. To save space, this table contains only these relationships for which the test of independence confirms the existence of a statistically significant relationship at p adjusted ≤ 0.05. The third column indicates the type of independence test used (either Pearson's chi-squared test or Fisher's exact test), the values of statistic of respective test (where appropriate), and the p value adjusted with "fdr". The last three columns list a range of different measures of effect size. Their values above 0.5 indicate strong relationship, within a range (0.3, 0.5) – moderate, within a range (0.1, 0.3) – weak, and below 0.1 as no relationship. Eight relationships were identified as statistically significant for MSO – at moderate or weak strength. However, in some cases the values of lambda coefficient were unexpectedly low (close to zero) even though other measures of effect size indicate stronger relationship. In these cases, even though there is some level of correlation, these explanatory variables have very

low level of ability in predicting MSO. No relationship was found statistically significant here for MBR.

Table 6.4. Values of measures of association for logical and nominal explanatory variables

Outcome	Explanatory	Method Statistic P. adj.	Phi/ Cramer's V	Pearson's C	Lambda
MSO	Used methodology	Pearson 13.76 <0.01	0.44	0.40	0.41
MSO	User satisfaction survey	Pearson 18.05 <0.01	0.51	0.46	0.47
MSO	Project activity scope design	Fisher ∞ <0.01	0.49	0.44	0.00
MSO	Project activity scope test	Pearson 7.06 0.03	0.33	0.31	0.10
MSO	Application type transaction/ production system	Fisher 4.83 0.03	0.30	0.29	0.00
MSO	Application type management information system	Pearson 8.54 0.02	0.34	0.32	0.24
MSO	Survey respondent role	Fisher – <0.01	0.57	0.50	0.15

6.3.2 Modeling with logistic regression

Table 6.5 provides a list of logistic regression models built for MSO and MBR. Each of these models contains a single explanatory variable and an intercept term (b_0). To save space, this table lists only these models that are statistically significant, i.e. for which p value adjusted with "fdr" is ≤0.05. For MSO ten such models were found statistically significant and only one for MBR. The algorithm for building a logistic regression model provides the values of coefficients in the form of log-odds. To make their interpretation simpler, Table 6.5 provides the values of these coefficients after applying transformation exp(b),

i.e. in the form of odd ratios. For example, according to the model, when *Total defects delivered* = 0, then the odds of reaching user satisfaction in MSO is 4.86. This can further be converted to probability of reaching user satisfaction in MSO as 4.86/(1+4.86) = 0.83, which is quite high value – yet, expected when there are no defects. For one unit increase in *Total defects delivered* we expect to see a decrease in odds of reaching MSO to 0.67 of the odds without such additional defect. In general, the value of $\exp(b_0) \in (0,1)$ indicates a multiplicative decrease in the odds of reaching satisfaction in MSO for one unit increase in value for particular explanatory variable, and the value of $\exp(b_0) \in (1,\infty)$ indicates a multiplicative increase in the odds of reaching satisfaction in MSO.

Table 6.5. List of statistically significant logistic regression single-variable models

Outcome	Explanatory	P adj.	Exp(b_0)	Exp(b_1)
MSO	Year of project	<0.01	0.00	1.80
MSO	Total defects delivered (ln)	<0.01	4.86	0.67
MSO	Used methodology	<0.01	0.70	7.19 (true)
MSO	User satisfaction survey	<0.01	0.64	11.31 (true)
MSO	Project activity scope - design	<0.01	0.87	1.33e+8 (true)
MSO	Project activity scope - test	<0.01	0.92	4.55 (true)
MSO	Application type - transaction / production system	<0.01	1.22	4.91 (true)
MSO	Application type - management information system	<0.01	3.64	0.23 (true)
MSO	Application type - network management	0.01	1.50	8.00 (true)
MSO	Survey respondent role	<0.01	20.00 (cust./user)	0.03 (project man.) 0.37 (sponsor)
MBR	Survey respondent role	<0.01	1.16e+8 (cust./user)	1.80e-8 (project man.) 4.54e-8 (sponsor)

6.3.3 Rankings by data mining measures

Apart from using "traditional" statistical measures of association discussed in earlier subsections, this analysis also investigated a range of measures that are commonly used in data mining. One of the advantages of these measures is that they can all be applied to any type of explanatory variable – however, for information gain, gain ratio, and Gini index, continuous variables need to be discretized into a set of intervals.

Table 6.6 illustrates the ratings for explanatory variables and each outcome variable using mentioned data mining measures. The higher value of each measure indicates a stronger relationship of particular pair of variables. This is additionally visually indicated by the width of the horizontal bar. The list of explanatory variables has been sorted according to the decreasing order of average strength of relationship calculated as an aggregate for both outcome variables and using all four measures. Thus variables at the top are with the strongest relationship with both MSO and MBR. Due to space constraints this figure contains the top 15 explanatory variables according to this overall rating.

Sorting a list of attributes according to each measure produces different order of explanatory variables for each outcome variable. This is caused by the fact that each measure focuses on different aspect of an association. For example, for MSO the values of ReliefF, information gain and Gini index indicate that *Survey respondent role* is the explanatory variable in the strongest relationship with MSO. However, based on gain ratio, for MSO the strongest relationship is with *Project activity scope: design*. According to the average ranking from all four measures, the top six explanatory variables with the strongest relationship with MSO are: *Survey respondent role*, *User satisfaction survey*, *Project activity scope: design*, *Year of project*, *Application type: management information system*, and *Application type: transaction/production system*. Surprisingly, neither any variable indicating project size nor defectiveness was found as strongly related with MSO.

As for MSO, also for MBR the values of ReliefF, information gain and Gini index indicate that *Survey respondent role* is the strongest related variable with MBR. However, based on gain ratio, for MBR the strongest relationship is with *Organization type: computers & software*. According to the average ranking from all four measures, the top six explanatory variables with the strongest relationship with MBR are: *Survey respondent role*, *Productivity (ln)*, *Year of project*, *Organization type: computers and software*, *User satisfaction survey*, and *Productivity*.

Table 6.6. Ratings for explanatory variables with data mining measures

Explanatory	MSO				MBR			
	ReliefF	Inf. gain	Gain ratio	Gini index	ReliefF	Inf. gain	Gain ratio	Gini index
Survey respondent role	0.24	0.26	0.17	0.07	0.38	0.14	0.09	0.03
User satisfaction survey	0.21	0.17	0.17	0.05	0.26	0.03	0.03	0.01
Year of project	0.06	0.26	0.14	0.06	0.11	0.08	0.05	0.02
Project activity scope: design	0.18	0.18	0.21	0.04	0.22	0.02	0.03	0.00
Organisation type: computers & software	0.08	0.05	0.08	0.01	0.12	0.06	0.10	0.01
Application type: transaction/production system	0.12	0.07	0.07	0.02	0.22	0.02	0.02	0.00
Organisation type: financial	0.10	0.04	0.08	0.01	0.07	0.02	0.04	0.01
Organisation type: communications	0.07	0.04	0.06	0.01	0.05	0.03	0.05	0.01
Total defects delivered (ln)	0.00	0.09	0.04	0.03	0.13	0.03	0.01	0.01
Adjusted function points (ln)	0.03	0.03	0.01	0.01	0.09	0.06	0.03	0.01
Organisation type: professional services	0.09	0.02	0.03	0.01	0.16	0.02	0.04	0.01
Summary work effort (ln)	0.07	0.02	0.01	0.01	0.16	0.05	0.02	0.01
Application type: network management	0.08	0.05	0.08	0.01	0.08	0.01	0.02	0.00
Organisation type: other	0.12	0.02	0.02	0.01	0.03	0.03	0.04	0.01
Productivity (ln)	0.03	0.01	0.00	0.00	0.11	0.09	0.04	0.02

6.3.4 Modeling with CN2 rules

Before investigation the details of the learnt CN2 rules, let us analyze the performance of the models represented by these rules. This analysis involves a range of measures: accuracy, F1 score, recall, precision, and Matthews correlation coefficient (MCC). While each measure focuses on different aspect, the interpretation of these measures is straightforward: values closer to 1 indicate more accurate prediction while values closer to zero indicate inaccurate prediction. The values of MCC can be negative to indicate predictions opposite to the actual values.

Table 6.7. Performance measures of generated CN2 rules

Outcome	Validation	Acc.	F1	Recall	Prec.	MCC
MSO	test on train data	0.94	0.95	0.91	0.91	0.86
MBR	test on train data	1.00	1.00	1.00	1.00	1.00
MSO	10-fold CV	0.74	0.81	0.83	0.79	0.42
MBR	10-fold CV	0.78	0.87	0.94	0.80	0.15

Table 6.7 provides the values for performance measures of generated CN2 rules. Rules generated for both outcome variables provide very accurate predictions when tested on the same dataset; for MBR they even perfectly explain the relationships. To investigate the adequacy of the CN2 rule generation algorithm we also analyzed performance achieved in 10-fold cross-validation. Naturally, this yielded in lower values for each measure, but still quite high and comparable with using other techniques, such as k-nearest neighbors, classification trees, naïve Bayes or random forests. Thus, this demonstrates the adequacy of CN2 rules. Yet, the predictive aspect of such analysis will be investigated in future studies.

Table 6.8 lists the rules learnt for MSO. We can observe that these rules use explanatory values already identified as related with MSO: *Total defects delivered*, *Adjusted function points* or *User satisfaction survey*. However, some of these rules also use *Summary work effort*, *Prop. of effort on build* or *Development platform*, i.e. explanatory variables that were earlier not identified as related with MSO. This is caused by the fact that earlier techniques were focused on investigating relationships in pairs, i.e. one explanatory variable for one outcome variable. These rules have the ability to capture information on the interactions between explanatory variables. Unfortunately, the quality of some rules is not high (i.e. closer to zero than to one). Furthermore, some rules cover very few projects – even single one at the extreme.

As for MSO, we also generated a set of CN2 rules for MBR (Table 6.9). Here, we can also see a similar set of explanatory variables that are used in these rules. The most important are: *Survey respondent role*, *Used Methodology*, *Summary work effort*, *Adjusted function points*.

Some rules can be easily interpreted and justified causally by the theory of software engineering. For example, a rule "IF Total Defects Delivered>1121 THEN Meet stated objectives=FALSE" means that if there are large number of defects then we should expect not meeting stated objectives. As another exam-

ple let us analyze a rule "IF Used Methodology=FALSE AND Total Defects Delivered>5 AND Application Type: Network Management=FALSE THEN Meet stated objectives=FALSE". The first two conditions are straightforward – the project does not use any methodology and delivers more than 5 defects (i.e. at least the median of 6 in this dataset). In this case we really should expect no satisfaction in meeting stated objectives. The key is the threshold of tolerable number defects that depends on application type. For network management applications this threshold is typically set to a low value, whereas for other types of application (i.e. perhaps for business use or gaming) it may be higher.

Table 6.8. Rules induced for MSO

Rule quality	Coverage false/true	Rule
0.29	20:5	IF Used Methodology=FALSE AND Total Defects Delivered>5 AND Application Type Network Management=FALSE THEN Meet stated objectives=FALSE
0.46	6:0	IF Survey respondent role=Project manager AND Summary Work Effort<=898 AND Project Activity Scope Planning=TRUE THEN Meet stated objectives=FALSE
0.55	3:0	IF Adjusted Function Points>457 AND Summary Work Effort>4584 AND Client Server=TRUE AND Summary Work Effort<=9653 THEN Meet stated objectives=FALSE
0.48	1:0	IF Total Defects Delivered>1121 THEN Meet stated objectives=FALSE
0.98	1:0	IF Year of Project>2005 AND Prop Effort Build<=0.00 THEN Meet stated objectives=FALSE
0.20	1:36	IF User satisfaction survey=TRUE AND Summary Work Effort>300 THEN Meet stated objectives=TRUE
0.17	2:9	IF Adjusted Function Points<=250 AND Prop Effort Build>0.00 THEN Meet stated objectives=TRUE
0.26	2:7	IF Adjusted Function Points>649 AND Development Platform=PC AND Summary Work Effort<=4167 THEN Meet stated objectives=TRUE
0.43	1:4	IF Summary Work Effort>2102 AND Total Defects Delivered<=290 AND Adjusted Function Points>649 THEN

Rule quality	Coverage false/true	Rule
		Meet stated objectives=TRUE
0.94	0:2	IF Adjusted Function Points>9296 AND Total Defects Delivered>38 THEN Meet stated objectives=TRUE

However, there are some rules that can be hardly explained by a theory. For example, let us analyze a rule "IF Survey respondent role=Project manager AND Summary Work Effort<=898 AND Project Activity Scope Planning=TRUE THEN Meet stated objectives=FALSE". Over 64% of projects meet the condition for effort. For this rule we cannot find a justification that in these projects when a project involves the planning stage and the survey results are provided by a project manager then we should expect no satisfaction in meeting stated objectives.

Table 6.9. Rules induced for MBR

Rule quality	Coverage false/true	Rule
0.25	6:0	IF Survey respondent role=Project manager AND Used Methodology=TRUE AND Summary Work Effort<=1462 THEN Meet business requirements=FALSE
0.32	5:0	IF Adjusted Function Points>594.00 AND Prop Effort Specify<=0.00 AND Adjusted Function Points>738 THEN Meet business requirements=FALSE
0.34	3:0	IF Productivity>0.00 AND Project Activity Scope Planning=FALSE THEN Meet business requirements=FALSE
0.56	3:0	IF Organization type Other=TRUE AND Application Type Management Information System=FALSE AND Year of Project<=2003 THEN Meet business requirements=FALSE
0.97	2:0	IF Summary Work Effort>9076 AND Summary Work Effort<=9653 THEN Meet business requirements=FALSE
0.09	0:29	IF Total Defects Delivered<=7 AND Summary Work Effort>300 AND Organization type Other=FALSE THEN Meet business requirements=TRUE
0.14	0:18	IF Adjusted Function Points<=781 AND Project Activity

Rule quality	Coverage false/true	Rule
		Scope Test=FALSE AND Adjusted Function Points>22 THEN Meet business requirements=TRUE
0.20	0:10	IF Prop Effort Specify>0.00 AND Summary Work Effort>1779 AND Project Activity Scope Planning=TRUE THEN Meet business requirements=TRUE
0.39	1:10	IF Application Type Transaction/Production System=TRUE AND Summary Work Effort>207 THEN Meet business requirements=TRUE
0.86	0:3	IF Summary Work Effort>5541 AND Application Type Management Information System=TRUE AND Total Defects Delivered<=500 THEN Meet business requirements=TRUE

6.4 Limitations and threats to validity

Results obtained in this study are subject to some limitations and threats to validity. First, the dataset used in analysis is not a random sample from population. Initially, ISBSG gathers data from organizations that are willing to share them. Then, this analysis involved the use of a carefully selected subset of the whole ISBSG dataset, as explained in Section 2. Thus, obtained results cannot be generalized to the whole population of projects.

Second, the dataset contains many missing variables. For this reason many variables have not been used at all, while other still have up to 30% of missing values which (1) may bias the results of single-explanatory-variable models/tests and (2) make it difficult to build multi-variable models.

Next, this analysis involved statistical testing of multiple hypotheses. To mitigate the problem of incorrect reporting inflated number of significant results a false discovery rate control was used to adjust obtained p values.

Furthermore, the analysis of pairs of variables involved pairwise removing cases with missing values. Thus, some relationships cover different projects than other relationships.

Finally, the analysis, especially preprocessing steps, involve subjective decisions, for example on grouping states of categorical variables, setting threshold for fraction of missing values, choosing type of variable transformation, etc.

6.5 Conclusions and future work

Obtained results lead to formulating the following conclusions:

1. Although two investigated outcome variables, *satisfaction in meeting stated objectives* and in *meeting business requirements* seem to describe similar phenomena, there are different explanatory variables in relationships with each outcome variable. I.e. some variables that are related with MSO do not seem to be related with MBR, yet the opposite is less likely.
2. *Survey respondent role* is the explanatory variable with the strongest relationship both for MSO and MBR. To put it simple: it matters mostly who you ask about user satisfaction in terms of meeting stated objectives and business requirements.
3. For *user satisfaction in meeting stated objectives* the strongest relationships are with the following variables individually: *Survey respondent role*, *Year of project*, *Total defects delivered (ln)*, *User satisfaction survey*, *Project activity scope: design*, *Application type: management information system*, *Application type: transaction/production system*, and *Adjusted function points (ln)*.
4. For *user satisfaction in meeting business requirements* the strongest relationships are with the following variables individually: *Survey respondent role*, *Productivity (ln)*, *Year of project*, *Organization type: computers and software*, *User satisfaction survey*, *Adjusted function points*, and *Summary work effort*.
5. Analysis of pure quantitative measures of correlation/association may be misleading and should be supported by other techniques, such as scatterplots, that may reveal issues not directly encoded in specific numeric measure.
6. It is difficult to investigate the relationship of interactions between explanatory variables and outcome variable because of relatively high fraction of missing values in explanatory variables.

In future, this analysis will be extended to answer other important questions related to the impact of software development on user satisfaction. It may involve the use of other features of user satisfaction than selected for this analysis. It may also focus on prediction of user satisfaction based on software development characteristics. For example, the previous papers focused on developing a framework [R11a] for building a Bayesian network model [R13] for software quality simulation and prediction, where user satisfaction is one of many features describing software quality. Such model can be updated by the results of this study.

References

[BH95] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B,* vol. 57, no. 1, 1995, pp. 289–300.

[B05] Bokhari, Rahat H.: The relationship between system usage and user satisfaction: a meta-analysis. *Journal of Enterprise Information Management*, vol. 18, no. 2, 2005, pp. 211–234.

[C11] Catal C., Software fault prediction: A literature review and current trends. *Expert Systems with Applications*, vol. 38, no. 4, 2011, pp. 4626–4636.

[CN89] Clark P., Niblett T., The CN2 Induction Algorithm, *Machine Learning*, vol. 3, no. 4, 1989, 261-283.

[DCE13] Demšar, J., Curk, T., & Erjavec, A. Orange: Data Mining Toolbox in Python; *Journal of Machine Learning Research*, vol. 14, 2013, 2349−2353.

[FM04] Fenton N., Marsh W., Neil M., Cates P., Forey S., Tailor M., Making Resource Decisions for Software Projects. In: *Proceedings of the 26th International Conference on Software Engineering*. Washington, DC: IEEE Computer Society, 2004, pp. 397–406.

[FG14] Fernández-Diego M., González-Ladrón-de-Guevara F., Potential and limitations of the ISBSG dataset in enhancing software engineering research: A mapping review, *Information and Software Technology*, vol. 56, no. 6, 2014, pp. 527–544.

[HB12] Hall T., Beecham S., Bowes D., Gray D., Counsell S.,A Systematic Literature Review on Fault Prediction Performance in Software Engineering. In: *IEEE Transactions on Software Engineering*, vol. 38, no. 6, 2012, pp. 1276–1304.

[I05] ISBSG, ISBSG Comparative Estimating Tool V4.0 – User Guide, International Software Benchmarking Standards Group, 2005, www.isbsg.org.

[I09] ISBSG Repository Data Release 11, International Software Benchmarking Standards Group, 2009, www.isbsg.org.

[KJ13] Khatibi Bardsiri V., Jawawi D. N., Hashim S. Z., Khatibi E., A PSO-based Model to Increase the Accuracy of Software Development Effort Estimation, *Software Quality Journal*, vol. 21, no. 3, 2013, pp. 501–526.

[K93] Krishnan M.S., Cost, quality and user satisfaction of software products: an empirical analysis. In: *Proceedings of the 1993 conference of the Centre for Advanced Studies on Collaborative research: software engineering*. vol. 1, IBM Press, 1993, pp. 400–411.

[ML08] Mendes E., Lokan C., Replicating studies on cross- vs single-company effort models using the ISBSG Database, *Empirical Software Engineering*, vol. 13, no. 1, 2008, pp. 3-37.

[MAD12] Mitakos T.N., Almaliotis I.K., Demerouti A.G., *What Factors Influence ERP User Satisfaction?: Perceived Usefulness and Self-Efficacy as Key Directors of the ERP User Satisfaction*, LAP Lambert Academic Publishing, 2012.

[P08] Prasad, V.C.S., Complexities in user satisfaction issues during organisational diffusion of in-house developed new technology tools: the case of an Indian IT company. *International Journal of Information Technology and Management*, vol. 7, no. 3, 2008, pp. 315–326.

[PP13] Proynova R., Paech B., Factors influencing user feedback on predicted satisfaction with software systems. In: Doerr, J. ; Opdahl, A. L. (eds.): *Requirements Engineering:*

Foundation for Software Quality, *Lecture Notes in Computer Science*. vol. 7830. Berlin, Heidelberg: Springer, 2013, pp. 96–111.

[R11a] Radliński Ł., A Framework for Integrated Software Quality Prediction using Bayesian Nets, In: *Proceedings of International Conference on Computational Science and Its Applications*, Santander: Springer, 2011, LNCS vol. 6786, pp. 310-325.

[R11b] Radliński Ł., Factors of Software Quality – Analysis of Extended ISBSG Dataset, *Foundations of Computing and Decision Studies*, vol. 36, no. 3-4, 2011, pp. 293-313.

[R12] Radliński Ł., Empirical Analysis of the Impact of Requirements Engineering on Software Quality, In: *Proc. International Working Conference on Requirements Engineering: Foundation for Software Quality*, Essen: Springer, 2012, LNCS vol. 7195, pp. 232-238.

[R13] Radliński Ł., An expert-driven Bayesian network model for simulating and predicting software quality, In: *Proc. Fifth International Conference on Information, Process, and Knowledge Management*, Nice, France, 2013, pp. 26-31.

[RC10] Raza A., Capretz L.F., Ahmed F., Improvement of Open Source Software Usability: An Empirical Evaluation from Developers' Perspective. *Advances in Software Engineering*, vol. 2010, 2010, pp. 1–12.

[R14] R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

[SBH14] Shepperd M., Bowes D., Hall T, Researcher Bias: The Use of Machine Learning in Software Defect Prediction. *IEEE Transactions on Software Engineering*, vol. 40, no. 6, 2014, 603–616.

[SW10] Subramanyam R., Weisstein F.L., Krishnan M.S., User participation in software development projects. *Communications of the ACM*, vol. 53, no. 3, 2010, p. 137-141.

[TT10] Tarafdar M., Tu Q., Ragu-Nathan T.S., Impact of Technostress on End-User Satisfaction and Performance. *Journal of Management Information Systems*, vol. 27, no. 3, 2010, pp. 303–334.

[W88] White S. D., *An Empirical Study of the User Satisfaction of Accountants with the Software Maintenance Function*, University of Arkansas, 1988.

[WT05] Wixom B.H., Todd P.A., A Theoretical Integration of User Satisfaction and Technology Acceptance. *Information Systems Research*, vol. 16, no. 1, 2005, pp. 85–102.