M. C. Ohlsson and C. Wohlin, "An Empirical Study of Effort Estimation during Project Execution", Proceedings International Symposium on Software Metrics, pp. 91-98, Boca Raton, Florida, USA, 1999.

An Empirical Study of Effort Estimation during Project Execution

Magnus C. Ohlsson and Claes Wohlin
Dept. of Communication Systems
Lund Institute of Technology
Lund University
S - 221 00 Lund
(Magnus_C.Ohlsson, Claes.Wohlin)@tts.lth.se

Abstract

This paper presents an empirical study of effort estimation. In particular, the study is focused on improvements in effort estimations, as more information becomes available. For example, after the requirements phase, the requirements specification is available and the question is whether the knowledge regarding the number of requirements helps in improving the effort estimation of the project. The objective is twofold. First, it is important to find suitable measures that can be used in the re-planning of the projects. Second, the objective is to study how the effort estimations evolve as a software project is performed.

The analysis is based on data from 26 projects. The analysis consists of two main steps: model building based on data from part of the projects, and evaluation of the models for the other projects. No single measure was found to be a particular good measure for an effort prediction model, instead several measures from different phases are used. The prediction models were then evaluated, and it is concluded that it is difficult to improve effort estimations during project execution, at least if the initial estimate is fairly good. It is, however, believed that the prediction models are important to know that the initial estimate is of the right order, i.e. the estimates are needed to ensure that the initial estimate was fairly good. It is concluded that the re-estimation approach will help project managers to stay in control of their projects.

Keywords

Effort estimation, proxy-based estimation, empirical study, experience base, data collection, measurements, education.

1. Introduction

One of the most problematic activities within the software engineering area is project planning and it always includes some uncertainty. There exists a need to make accurate estimates and immature organisations often have problems estimating the required effort. This is especially true if the task is not within their usual domains or some other change occurs [1]. Expert judgement is often used for different types of estimations. Another opportunity is to use one of the cost models that have been developed. Most estimating models are based on historical data from earlier developed projects. Some of the most popular models or methods for size predictions, which is closely related to the effort, are COCOMO [2], Fuzzy-logic [3] and Function Points [4]. Though, there are some problems involved with these methods. For example, Fuzzy-logic requires large amounts of historical data to be able to calculate all the sub-ranges and function points are not directly countable in the final product [5].

The focus of this paper is to evaluate the usefulness of different measures for effort estimation as part of another study for a stepwise introduction of an effort experience base [6], in accordance with the notion of an Experience Factory [7][8], or a corporate memory [9]. The measures believed to be suitable to estimate the effort are hereafter referred to as proxies. A proxy is an indirect measure of the entity of interest. The term proxy is used since it is an approximation or indirect measure of something else. Proxies are further discussed in [10]. An example of a proxy is the number of function points, since it is believed that the effort can be estimated through calculating the function points at an early stage in the development. The selection of a proxy for effort estimation is a crucial task. A proxy should be closely related to the project size and it should be available at the beginning of the project, and new proxies should become available as the project proceeds. Also, it should be customisable, i.e. an organisation develop different projects and they should all be able to take advantage of the historical data generated from different projects.

The relationship between the proxy and the entity, which should be estimated or predicted, should determined using statistical methods. Regression is one suitable statistical method. It could, for example, be a linear, logarithmic or exponential regression. Further, it provides a basis for determining the accuracy of the estimates. It is also possible to calculate prediction intervals i.e. the range of the predictions may be determined.

The paper is organised as follows. The outline of the study is described in Section 2, and the data analysis is described in Section 3. Conclusions are presented in Section 4.

2. Study Outline

2.1. Objectives

The objective of this paper is to evaluate if proxies are useful in the planning process. Effort estimation and planning are closely related to different artefacts created in a project. For example, the more requirements and functionality that is considered to be implemented the more effort it will take. This could be further refined. If a large amount of functionality should be implemented then the design must cover this and it will hence also be related to effort. Therefore, proxies can be useful in the planning process. Also, the needed formalism, in terms of scripts and forms, to improve effort estimation from a managerial viewpoint should be considered. Two questions should be addressed:

- Does the use of proxies improve the effort estimates in comparison to subjective estimates?
- Are the prediction models derived reusable between projects?

2.2. Context

A number of projects from a software development course conducted at the department of Communication Systems, Lund University is used for the study. Each project develops a number of telephone services [11]. The course is read by 100-150 students yearly, which means that 7-12 similar projects are run in parallel since each project is manned with 11-19 students. This implies that all projects produce a large amount of documentation that can be used to evaluate different proxies. The students are aware of that the data from the projects is studied. They are, however, unaware of the exact studies conducted.

2.2.1. Process Model. A somewhat simplified version of the DoD-2167A [12] is used in the course. It is a waterfall model and the objective is to provide a standardised approach, rather than experimenting with a more sophisticated process model. The model, as used in the project, is outlined in Figure 1. In particular, the baselines for different documents in the model are emphasised to provide checkpoints in the process. The study described here is focused around effort needed and used for the different phases in the process model.

2.2.2. Project. The groups consist ideally of 17 students divided into 8 subgroups. One subgroup acts as project leaders. One subgroup is a system group responsible for keeping the system together. Four subgroups are development groups, where each group develops one service. Finally, two subgroups are test groups testing two services each, and with a joint responsibility for the system test (see Figure 2).

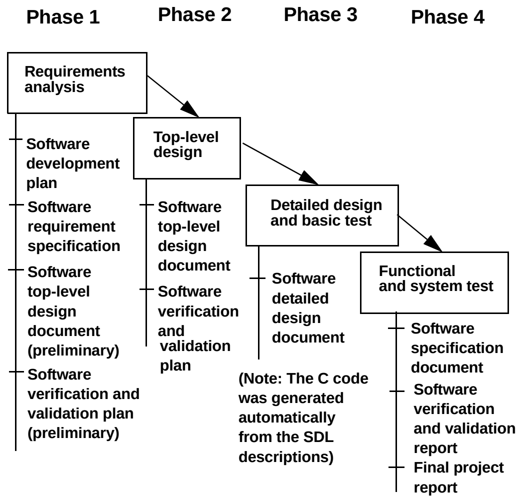

Figure 1: The waterfall development model.

The roles of the subgroups are combined with roles played by department personnel. The department personnel play the roles of customer, external quality assurance personnel and, of course, technical experts, who can help the students when they run into problems. The customer reviews the produced material at two times during the project and also performs an acceptance test at the end of the project. It should be noted that executing the process correctly is viewed as equally important as the final product.

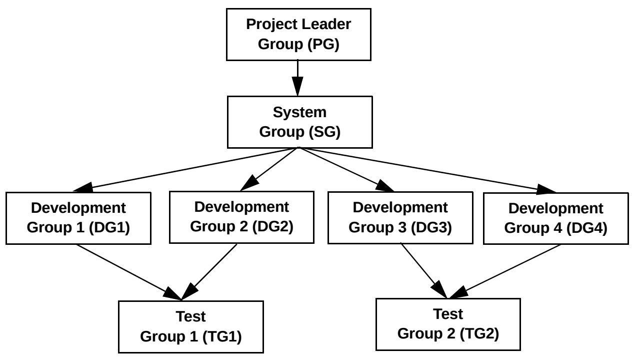

Figure 2: Organisation of the projects.

2.3. Threats

Empirical studies always have threats to their validity. It is always difficult to conduct studies where the results can be interpreted correctly (internal validity), and an even more difficult task is to interpret the generality of the results (external validity). In this particular case the following main threats have been identified.

- Student setting
  Studies in a student setting can always be questioned concerning validity in an industrial environment. In this case, this is not regarded as particular critical as one

objective of the course is to model an industrial environment. In particular it should be noted that the study is based on comparison of different methods for effort estimation.

- Reliability of the data (the students may not be reporting properly)
  The progress reporting weekly is believed to make these risks rather small. Further, the data was collected in retrospect when the final versions of the documents were available.
- Division into project groups
  The students choose which group to work in, and it may lead to that some groups know each other better than others. Random assignment is not an option since the students read different courses in parallel and some of them also commute together.
- Defection from project groups
  Some students have defected but the other project members have mostly divided the undertakings among them to be able to complete the project in time.
- Work expansion
  To fill the available time the students expand the work to fill the available time. This should not be a problem because most of the students take parallel courses and want to finish the work as soon as possible.

2.4. Operation

The data used in the study comes from 26 similar projects from 1996 to 1998. In both 1996 and 1997 seven projects were conducted and twelve projects were conducted in 1998. The effort data used in the calculations come from the projects' final reports. These reports contain information about how much time the projects has spent on different activities and the fault density for different documents, based on faults that were found during inspections and tests.

The data in the reports is a fusion of data from other reports with finer granularity. On a daily basis, the projects' members record the time spent on different activities in the project. This information is gathered in an individual time report and the data form the basis for a weekly report. The weekly report is filled out at the end of every week. The weekly report contains a summation of the person-hours spent in the project for the current week. The weekly report is handed to the project leader at the end of each week. At the end of the project, the weekly reports are gathered and summarised in a final report, from which the data for this study is collected. The data is reported in spreadsheets so that they should be easy to manage.

The proxy data is collected from the requirement specifications, verification and validation plans and the top-level design documents. All projects use the same templates for writing their requirement specifications, test plans and so on, which is a major advantage when comparing the different project. The projects use SDL (Specification and Description Language) [13] in the design phase. SDL is standardised by ITU. This makes the data uniform and possible to compare.

The data has been collected in retrospect from the final reports, and the spreadsheet files delivered at the end of the course.

3. Data Analysis

3.1. Analysis Strategy

The intention of the analysis is to study how the use of proxies changes the accuracy of the effort estimations. This includes finding a suitable proxy that is easily collectable and which can be used by the students when planning their projects. The outcome of this analysis forms the foundation for a new experience package, i.e. a set of lessons learned together with prediction models. The results from the re-planning using proxies are compared with the initial estimations made by the students. The objective is to evaluate if the proxy-based estimates are more accurate than the estimates based on just knowing the approximate size of the projects. The students know the approximate size from previous years, and the size of the projects is also a direct consequence of the size of the course. Thus, the students know the correct order of person-hours the project normally takes. The analysis is summarised in the following steps:

1. Based on the data from 1996 and 1997, different available proxies are evaluated to identify one or a few suitable proxies.
2. Prediction models are built based on the proxies from step 1 and the effort of the projects conducted during 1996 and 1997.
3. New estimates are calculated for the projects in 1998 from the prediction models. Based on the new estimates, it is possible to re-plan or at least check that the initial estimate seems realistic.
4. The accuracy of the estimates from the re-planning is compared with the projects' initial estimates.
5. After finishing the projects in 1998, it is possible to investigate if the proxies are stable across years or if new proxies should be used in the future.
6. Finally, the proxies are compared, the differences and similarities are investigated for the whole data set, i.e. for all 26 projects.

The investigation is carried out using different statistical methods. Suitable proxies are found using correlation and it is then assumed that there exist a linear relationship between the proxy and the effort. It is assumed that this relationship exists within the working interval, i.e. it is not possible to extrapolate these models for larger or smaller projects. The project re-planning includes the use of linear regression.

3.2. Analysis

3.2.1. Deciding proxy. If a proxy should be useful, it should be available early in the development process. A proxy could also be used to re-plan the projects when more information is available. Therefore, proxies that are easy to collect have been chosen. Moreover, it has been decided that the first proxy should be available in the first development phase. The proxies used in this study are described in Table 1. In total, eight proxies are evaluated. The data for the proxies was collected from the following documents:

- Software Requirement Specification (SRS)
- Software Verification and Validation Plan (SVVP)
- Software Top-Level Design Document (STLDD)

The SRS is one of the documents that is completed first. Measurements from this document are therefore very suitable as proxies. Both the SVVP and the STLDD are available as preliminary versions early in the process and could therefore be used for initial estimates and then further refined when the final versions are available. The SVVP provides information about test cases, both system and function tests. The STLDD is a high-level design document and it provides an overview of the system.

Table 1. Description of collected data.

Variable	Description
Requirements	The number of requirements in the requirement specification (SRS).
Tests	The number of test cases the verification and validation plan contains, excluding sub-tests (SVVP).
Sub-tests incl.	The number of test cases the verification and validation plan contains, including sub-tests (SVVP).
Processes	The number of processes in the system (STLDD).
Flowcharts	The number of flowcharts in the design (STLDD).
Outputs	The number of signals from the processes[a] in the design (STLDD).
Inputs	The number of signals to the processes in the design (STLDD).
Total signals	The number of signals in the system (STLDD).

a. SDL is based on processes communicating with signals.

A correlation analysis is conducted to study the relationship between the proxies and the effort in the projects. All values in the calculations are normalised according to the number of project members, i.e. to be able to compare projects with slightly different sizes. The incentive is that there is a correlation between the number of project members and effort of 0.689 with a significance level α of 0.005.

The results from the correlation calculations can be found in Table 2. The problem is that it is difficult to find one proxy that has a high correlation for both parts of the data set. For example, Outputs has a high correlation value in 1997, but it is lower in 1996. An interesting observation is that the number of tests in 1997 has a high negative correlation with the effort, i.e. the more tests the projects wrote, the less effort they had to spend. One explanation may be that the projects' boundaries are well defined and as the projects produce more tests, they increase their understanding of how the system functions and this results in shorter lead times.

Table 2. Correlation between different proxies and project outcome (1996 and 1997).

	1996		1997	
Variable	r	α	r	α
Requirements	0.292	0.547	0.194	0.694
Tests	-0.238	0.630	-0.687	0.092
Sub-tests incl.	0.025	0.960	-0.055	0.912
Processes	0.067	0.983	-0.004	0.994
Flowcharts	0.691	0.089	0.408	0.387
Outputs	0.274	0.574	0.654	0.118
Inputs	0.462	0.317	0.295	0.543
Total signals	0.414	0.378	0.494	0.279

Table 3. PCA for the data from 1996 to 1997.

Variable	Factor 1	Factor 2	Factor 3
Requirements	0.309	0.241	0.782
Tests	-0.261	0.845	0.043
Sub-tests incl.	0.069	0.669	-0.598
Processes	0.797	0.444	-0.163
Flowcharts	0.075	0.640	0.351
Outputs	0.909	-0.047	0.074
Inputs	0.967	-0.108	-0.029
Total signals	0.988	-0.108	-0.029

The relationships between the different variables are investigated using Principal Component Analysis (PCA) [14]. A principal component analysis collects correlated variables into factors and stops extracting factors when they no longer contribute to the explanation of the variance significantly. Before the PCA, the normalised values were also standardised to be able to compare variables with different units. The results can be found in Table 3. They indicate that there is a relationship between the proxies and the documents from which the proxies were collected.

Based on the information above, it was decided to further evaluate three of the proxies. The correlations are fairly low and hence it is not possible to recommend using one particular proxy. Based on the rather close relationship between the factors and the documents, it was decided to use one proxy from each document type. Thus, the proxies are chosen because they come from three different documents and from different phases of the projects. This provides an opportunity to apply multiple linear regression to further refine the estimate, as more information becomes available in the development process. The following proxies are chosen:

- Requirements (from Software Requirement Specification (SRS))
- Tests (from Software Verification and Validation Plan (SVVP))
- Outputs (from Software Top-Level Design Document (STLDD))

Both the number of tests and the number of outputs have correlation figures above 0.5 either 1996 or 1997. Requirements are included although it has a low correlation. It is the only proxy in Factor 3 in Table 3, and also the only proxy related to the requirements specification.

The relationship between the effort and the proxies is studied using multiple linear regression. This allows refinement of the estimates as the projects proceed through the different phases and as more of the proxies become available. First, when only the requirements specification is available, linear regression is used to relate requirements with effort, see Table 4. The regression parameters for the other two cases can be found in Table 5 and Table 6 respectively.

Table 4. Regression parameters for first proxy.

	Requirements
Slope	6.70
Intercept	44.35
R	0.306

Table 5. Regression parameters with first and second proxy.

	Requirements	Tests
Slope	6.42	-1.04
Intercept	48.71	
R	0.322	

Table 6. Regression parameters for all three proxies.

	Requirements	Tests	Outputs
Slope	6.77	-0.97	-1.99
Intercept	54.55		
R	0.344		

The *R* value for the different regression models indicates the goodness of fit. There is an improvement between the first, second and third model, although the *R* values are not impressive. The question is, however, what the *R* value is for the subjective estimates for the individuals in the beginning of the project. The latter is important to remember since this is what the new estimates should be compared with. The *R* value and the regression parameters for the projects' initial plan can be found in Table 7. The *R* value is smaller than any of the proxy models. Further, the inclusion of a third proxy does slightly improve the *R* value but it may affect the results because of correlation between proxies.

Table 7. Regression parameters for the initial plans.

	Initial
Slope	63.27
Intercept	0.093
R	0.217

The models provide the effort for an individual participating in a project. This is a result of the choice to normalise the data. Thus, in order to obtain the total effort for the project, it is necessary to multiply with the number of project members.

3.2.2. Re-plan and Comparison. The next step is to create new estimates for the projects from 1998, i.e. based on the different prediction models derived in the previous section. The new estimates are calculated from the prediction models in Table 4 to Table 6. The new estimates that are calculated using the proxies can be found in Table 8. The first row is the initial effort estimation done by the projects themselves. The following three rows contain the estimates obtained from the three prediction models, i.e. the models using linear regression and multiple linear regression. Finally, the outcome of the projects is also shown in the table. The accuracy of the estimates can be found in Table 9 and it is calculated as *(Estimate-Outcome)/Outcome*.

It should be noted that the planning of the projects is concerned with the total project effort, although proxies available later in the process are used. Another possibility is to use the proxies to plan the remaining project effort. The first alternative is chosen because it provided the possibility to compare the different times of estimation with each other and also with the initial plans.

Some of the estimates based on the proxies are very good while others are not. One example of a poor proxy estimate is the one for project 12 where the estimation error is around 40 percent or higher. Project 7 and project 8 also have very poor estimates. These two projects are characterised by that they are fairly small but they wrote many requirements for their project assignment. Project 12 produced the largest Software Verification and Validation Plan of all projects and a reasonably large Software Requirement Specification and therefore they gained a good understanding of the problem and finished the project early.

On the other hand, the best plans has an absolute error of 0.5 percent. The large variation among these figures depends on the large differences between the projects, i.e. in terms of their approach to the projects. In other words,

Table 8. Estimates for 1998 projects.

Project	1	2	3	4	5	6	7	8	9	10	11	12
Initial	704	908	1200	868	986	956	776	816	1443	1364	1232	796
Reqs	796	863	980	870	1223	1277	923	981	1363	1203	1356	1337
Reqs+Tests	838	925	1030	906	1290	1378	965	1041	1460	1302	1421	1400
Reqs+Tests+ Outputs	798	858	997	883	1263	1287	952	988	1350	1149	1389	1357
Outcome	734	665	895	792	1217	1332	739	736	1500	1017	1189	958

Table 9. Accuracy for 1998 projects (absolute percent values).

Project	1	2	3	4	5	6	7	8	9	10	11	12
Initial	1.4	36.6	34.1	9.6	19.0	28.2	4.9	10.9	3.8	34.1	3.6	16.9
Reqs	8.4	29.8	9.6	9.8	0.5	4.2	24.9	33.3	9.1	18.3	14.1	39.6
Reqs+Tests	14.1	39.1	15.1	14.4	6.0	3.5	30.5	41.4	2.7	28.1	19.5	46.2
Reqs+Tests+ Outputs	8.8	29.0	11.4	11.5	3.8	3.4	28.8	34.3	10.0	13.0	16.8	41.7

although the projects develop the same software, they have major differences in their approach to the problem. An example of the latter is the time spent on writing detailed requirements in comparison with having requirements on a higher abstraction level, which may leave some of the problems to the design phase. It has not been possible to find any pattern in the differences between projects and an ANOVA test did not provide any significant difference between the different estimation models.

Table 10. Mean and standard deviation values for planning accuracy.

Variable	Mean	Std. dev.	Median
Reqs	16.8%	12.4%	11.9%
Reqs+Tests	21.7%	15.1%	17.7%
Reqs+Tests+ Outputs	17.7%	12.6%	12.2%
Initial plans	16.9%	13.3%	13.9%

The means, medians and standard deviations for the estimates can be found in Table 10. The objective is that the use of proxies should improve the effort estimates. A secondary objective is that it may be positive if the new estimate supports the initial estimate. Thus, for the secondary objective it is not so important that the estimates based on the proxies improve; it is more important that they are fairly close to the initial estimate. This provides the project manager with important information for the remaining part of the project, i.e. the project manager knows that the project is under control. From Table 10, it can be seen that the mean values for the use of proxies and the initial plans are fairly close to each other.

It is noticeable that adding the number of tests or the number of outputs to the regression models does not improve the mean value for the accuracy. Though, adding the number of tests increases the median value compared to only using requirements as a proxy. In this case the use of the proxy planning method would not improve the results.

Some minor differences exist between the estimates in Table 10, but it is clear from the table that the most important information for the project manager is that if the proxy estimations support the initial estimate, it is highly likely that the initial estimate is rather good. Thus, a project manager can use the results to increase his/her confidence in the effort estimates. This is an important result, since it means that the project manager obtains feedback as new proxies become available and the project manager can know that the project is running according to plans. This type of continuous feedback is not available for a project manager without suitable proxies. The only opportunity without a proxy is to compare the outcome for one specific phase with the plan. The proxy-based approach discussed here acts as a good complement to this.

3.2.3. Deciding Proxy. A proxy-based approach can hence be an important tool to achieve project control. Thus, it is important to evaluate and re-evaluate the proxies as more projects become available. In other words, the experience should be updated and packaged accordingly. The availability of the projects from 1998 means that a new correlation study has to be conducted to see how the new projects have impacted the choice of suitable proxies. The results from these calculations are presented in Table 11. For the projects from 1998, many of the proxies have very low correlation to the effort of the project. The overall correlation values are very low, i.e. when looking at the whole data set of all 26 projects. One reason may be that the overhead is very different depending on the number of project participants. If this is the case, ways of dealing with individual work should be separated from the time related to the number of individuals in the project. The individual work should be fairly similar between the projects, but the work related to the size of the project is probably non-linear. This may disturb the picture and lower the correlations considerably. This issue is addressed by comparing the correlations without normalisation (see Table 12). These correlations are higher, which seems to support the hypothesis. This is an issue that should be addressed further in the continuation of this study.

Table 11. Correlations between different proxies and project outcome.

	1998		1996-1998	
Variable	r	α	r	α
Requirements	0.206	0.555	0.088	0.677
Tests	0.055	0.876	-0.012	0.955
Sub-tests incl.	-0.314	0.359	0.033	0.876
Processes	-0.167	0.634	-0.245	0.240
Flowcharts	0.246	0.477	0.206	0.326
Outputs	-0.218	0.531	-0.116	0.5856
Inputs	-0.308	0.368	-0.205	0.330
Total signals	-0.315	0.357	-0.179	0.396

Table 12. Correlations between different proxies for non-normalised project outcomes.

	1996-1998 not normalised	
Variable	r	α
Requirements	0.646	< 0.001
Tests	0.378	0.057
Sub-tests incl.	0.365	0.067
Processes	0.283	0.163
Flowcharts	0.374	0.066
Outputs	0.499	0.009
Inputs	0.500	0.008
Total signals	0.516	0.006

The principal component analysis for all the data provides almost the same factors as in Section 3.2.1. The only difference is that flowcharts are related to requirements (see Table 13). In other words, no large differences occurred.

Table 13. PCA for the data from 1996 and 1998.

Variable	Factor 1	Factor 2	Factor 3
Requirements	0.322	0.707	-0.515
Tests	0.308	0.154	0.786
Sub-tests incl.	0.148	0.494	0.599
Processes	0.661	0.412	-0.178
Flowcharts	0.399	0.739	0.039
Outputs	0.720	-0.607	0.046
Inputs	0.932	-0.136	-0.079
Total signals	0.914	-0.385	-0.024

Since all the correlation values are rather low, it is difficult to make any recommendation of which proxies to focus on. A similar approach has to be taken here as was done previously.

It is clear from this study that it is difficult to identify the best proxies. Moreover, the best proxies do not seem to be stable as new projects are added to the experience base. The main positive result of the study is that it may be important to do this in software projects any way, since it provides the project manager with important information that can be used to ensure that the initial estimate is correct. The approach with proxies is hence a valuable tool for project managers although they cannot expect to improve their estimates considerably.

Even though the results did not provide any improved estimates, it is still necessary to continue collect data. It is not possible to conclude that improvements are not possible based on one study. It is necessary to repeat this procedure to be sure on the results.

4. Conclusions

This paper presented a study that was carried out to investigate the usefulness of using proxies for effort estimation as a complement to the initial plans made by project managers prior to the project. Proxies can help in creating mental pictures and provide another level of abstraction when estimating project effort. This will ease and guide the estimation process and people will feel more secure about their estimates. The latter is an important issue supporting the idea of using proxies as an effort estimation tool in software projects.

One problem is to find suitable proxies, which reflect and correlate with project effort. The study included eight proxies from three different documents. The documents relate to different phases, but preliminary versions of the documents are available in the first phase. The results show that it is difficult to find a proxy with good correlation to project effort. One reason for this is that seven of the projects used to build the models had negative correlations between the proxies and the effort. Since, it was not possible to identify one superior proxy, it was decided to investigate three of the proxies. The proxies studied in more detail are from different documents.

The results show that the mean value and standard deviation of the estimation error do not change significantly, as more information becomes available. This is probably due to that the initial estimate is fairly good, and the remaining uncertainty remains throughout the project and it cannot be captured with a proxy only taking product attributes into account. The estimates may improve if process attributes were included in the analysis.

The final recommendation is hence to start using proxies as a means for supporting project managers. The proxies will support managers in their difficult task of controlling software projects. It is important to regularly re-estimate software projects to ensure that they proceed according to plan. The proxy-based approach studied here provides this opportunity.

5. Acknowledgement

This work was partly funded by The Swedish National Board for Industrial and Technical Development (NUTEK), grant 1K1P-97-09673.

6. References

1. W.S. Humphrey, “Managing the Software Process”, Addison-Wesley, 1989.
2. B.W. Boehm, “Software Engineering Economics”, Englewood Cliffs NJ, Prentice Hall, 1981.
3. L.H. Putnam and W. Myers, “Measures for Excellence: Reliable Software on Time, within Budget”, Yourdon Press, 1992.
4. A.J. Albrecht and J.E. Gaffney, “Software Function, Lines of Code and Development Effort Prediction: A Software Science Validation”, IEEE Transactions on Software Engineering, SE-9 (6), pp. 639-647, 1983.
5. B. Kitchenham, “The Problem with Function Points”, IEEE Software, 14(2), pp. 29-31, 1997.
6. M.C. Ohlsson, C. Wohlin and B. Regnell, “A Project Effort Estimation Study”, Journal of Information and Software Technology, 40 (14), pp. 831-839, 1998.
7. V. Basili, G. Caldiera and H.D. Rombach, “Experience Factory”, Encyclopedia of Software Engineering, Vol. 1, edited by J.J. Marciniak, pp. 469-476, John Wiley & Sons, New York, 1994.
8. F. McGarry, “Porting Experience Factory Concepts to New Environments”, Proceedings of the Twentieth Annual Software Engineering Workshop, pp. 71-92, Greenbelt, Maryland, USA, 1995.
9. B. Lennselius, C. Wohlin and C. Vrana, “Software Metrics: Fault Content Estimations and Software Process Control”, Microprocessors and Microsystems 11 (7), pp. 365-375, 1987.
10. W.S. Humphrey, “A Discipline of Software Engineering”, Addison-Wesley, 1995.
11. C. Wohlin, “The Challenge of Large Scale Software Development in an Educational Environment”, Proceedings of the Conference on Software Engineering Education & Training, pp. 40-52, Virginia Beach, Virginia, USA, 1997.
12. DoD-STD-2167A, Defence System Software Development, 1988.
13. F. Belina, D. Hogrefe and A. Sarma, “SDL with Applications from Protocol Specification”, Prentice Hall, United Kingdom, 1991.
14. S.K Kachigan, “Multivariate Statistical Analysis: A Conceptual Introduction”, Radius Press, New York, 1991.