Review of the Doctoral Thesis

Student:	Vo Van Hai
The topic of the Doctoral Thesis:	Optimization of Software Effort Estimation by Improving Functional Points Analysis
Supervisor:	Assoc. Prof. Ing. Zdenka Prokopová, CSc.
Consulting Supervisor:	Assoc. Prof. Ing. Radek Šilhavý, Ph.D.
Study Program:	Engineering Informatics
Workplace:	TBU in Zlin, Faculty of Applied Informatics

The topic of Vo Van Hai's dissertation fulfills the requirement for a dissertation within the "Engineering Informatics" study program.

The doctoral thesis contains 11 parts. Chapter 1 presents the motivation, the work's objectives, and the text's structure. An overview of current software estimation methods and a description of some machine learning algorithms are outlined in the following chapters. Chapter 4 is devoted to the existing knowledge in software effort estimation. The critical part of the doctoral thesis consists of chapters 5 to 7. These chapters describe the proposed method and present the results of the work.
The thesis includes comments on the proposed method validity (chapter 8), a description of the work benefits from the point of view of science and practice (chapter 9), and a conclusion (chapter 10). The final part is the list of references.

The actuality of the dissertation topic

The work deals with the issue of evaluating the development of software systems, especially software effort estimation. This topic is very actual because the correct prediction of software development effort plays a key role in managing large software projects. The accurate forecast then helps to develop software at an optimal price and to meet all requirements.

Fulfillment of the objectives (goals) set in the doctoral thesis

The dissertation's specific objectives and research questions are described in detail in chapter 1.3. The main goal of the dissertation is to propose an approach to improve the accuracy of software effort estimation. The student, in more detail, further specified the primary goal through 4 sub-goals. In my opinion, the sub-goals are not set quite appropriately. In particular, the third and fourth sub-goals, i.e., testing, comparing with other approaches, and contribution of the proposed method, are necessary parts of the work, not sub-goals.
However, all the set sub-goals were fulfilled in work, and above all, the main goal of the work was fulfilled – The new approach to improve the accuracy of software effort estimation was proposed.

The procedure for solving the problem and the results of the dissertation indicate the specific contribution of the doctoral student

The procedure for solving the established research problem is adequate and corresponds to research works of this kind. The thesis contains an overview of the current state of the solved problem. The own proposed solution is based on this summary. However, the chapter "overview of the current state" (chapter 4) should be presented in the front part of the work. Only after that, the chapter with the outline of software evaluation methods and the chapter with the description of machine learning algorithms should be presented. Chapters 2 and 4 summaries state of the art in software estimation methods. From this point of view, chapter 2 should be part of chapter 4. Additionally, the titles of Chapters 2 and 4 are inappropriate, inaccurate, and misleading.
Own research and the work results are presented in chapters 5 to 7. Unfortunately, the structure of this key part of the thesis is unclear and confusing. The core of the work – Proposed method (chapter 5) has only 1 page. Only fig.5-1 presents the proposed framework; the detailed description is missing.
Chapter 6 also seems confusing. This chapter, amongst other things, describes parts of the proposed framework (method) for estimating the software effort. Therefore, this part should have been listed in chapter 5. The title of chapter 6 is unsuitable; the word "experiment" is unacceptable. Chapter 6 describes the methodology proposal for software effort estimation, not an experiment. The experimental part is included in chapter 7.
The presentation of results in chapter 7 is also unclear and confusing; the chapter contains too many tables with minimal description. Included graphs are redundant in this case.

Despite the comments mentioned above, the concrete result of the work is evident. A new methodology (framework) for software effort estimation is proposed. The proposal of this methodology is described and commented on in detail. From the presentation of the results, the definite contribution of the doctoral student in the field of software effort estimation is observable.

Some comments on used methods
The methods used in the work are appropriate to the topic of the work and demonstrate the student's ability to work theoretically and scientifically and to confront the results with the requirements of practice. However, some analyses or statistical analyses are processed inappropriately. These are mainly the analyses that the author carried out in chapter 7. The paired t-test is used to compare the results of the methods. This technique has two competing hypotheses: the null and the alternative. However, the thesis does not include a description of these hypotheses. The definition of test characteristic is also missing. This part of the thesis generally misses the detailed description of using this statistical test.
On the contrary, a systemic approach to designing one's solution and preparing a methodology that modifies existing methods can be positively evaluated.

Significance for practice or the development of a scientific or artistic field

The importance and contribution of the work can be seen both on the theoretical and practical levels. On a theoretical (scientific) level is beneficial:

- the state of the art of software estimation methods, including the overview of software estimation tools used in the software industry
- proposal of a new calibration complexity weight algorithm
- proposal of methodology for the determination of the most suitable clustering algorithm and categorical variable
- proposal a new framework for optimization of software effort estimation

On a practical level, the results of the dissertation can positively influence the process of planning and managing software development while simultaneously being user-friendly and having financial requirements.

Formal editing of the dissertation and its language level

Unfortunately, the structure of the thesis is not good; the work lacks a logical division. The chapters do not have a logical sequence and are inappropriately divided and, above all, inappropriately named. Some criticisms and comments were given in previous parts of the review.
The work contains some other minor formal errors and shortcomings.

- the title of the chapter is given on the last line of the page (pages 17, 55, 56)
- numerical references to equations are missing in the text
- using different types of brackets (tab.2-4)

The quality of some images is low, and the text in the pictures is complicated to read (for example, Fig. 2-1,2-2,2-3. Some images representing software tools (chapter 2.4) are redundant; lacking commentary in the text.
The level of English language is very poor. The text contains many unclear statements and phrases that sometimes lead to an ambiguous interpretation. The same words and phrases are repeated very often in the text. Wrong prepositions and terms are used.

Publication or artistic activity of a doctoral student

The research activity of the doctoral student and the ability of independent scientific work is evidenced by a relatively rich publication activity (a total of 17 professional publications). The composition of research outputs exceeds the requirements placed on students in doctoral study programs. Almost all publications are indexed in the SCOPUS database (15). Three publications are articles in impact journals indexed in the WOS database. For nine publication outputs, the student is listed as the first author.

Questions

1. What does the "relative size" or "functional size" parameter in table 6-3 mean?
2. In Chapter 6.3. you state that two segmentation approaches are considered. However, it is unclear from the description whether both approaches are used in parallel in your proposed method or only one of the mentioned approaches is selected. Please can you explain this obscurity?

3. One part of your proposed method is the "Optimisation Framework". This part of the proposed system is described very faintly in chapter 6.4. Can you describe the optimization process in more detail? Have you used an objective function? Which parameters are optimized?

4. In Chapter 7, you use paired t-tests to compare the results of the CFCW methods with the FPA method. You state two hypotheses. However, I miss the detailed description of this statistical test. The paired sample t-test has two competing hypotheses: the null and the alternative. How did you establish the null and alternative hypotheses? What test characteristic (statistic) did you use?

5. What does the last column in Table 7-33 represent? What does it mean the variable "k-means"?

6. Can you explain why the monitored values for the CFCW-MLP method are different in Tables 7-7 and 7-15? We can also observe the different values for this method in other tables.

Conclusion, recommendation

It can be stated that the doctoral student demonstrated the ability to independent scientific activity. The submitted work meets the minimum requirements for dissertations in terms of formality and substance. I have reservations about the very low level of the final text form of the doctoral thesis. Despite the reservations and remarks summarized above and because the submitted dissertation brings new knowledge for the development of the scientific field and practice

I recommend

the Vo Van Hai's thesis to be accepted by the relevant committee and defended for the award of the Ph.D. degree in the study program "Engineering Informatics".

In Zlin: 11.11.2022

Assoc. Prof. Ing. Bronislav Chramcov, Ph.D.

Review of the Ph.D. thesis
„*Optimisation of Software Effort Estimation by Improving Functional Points Analysis*“
by Vo Van Hai

PhD student Vo Van Hai's dissertation deals with estimating the effort of software projects, which is a non-trivial task given the different applications (ranging from demanding scientific and technical computations to information systems with simpler operations but large amounts of data) and the extent to which ready-made library modules (typically in Java) can be used. However, these estimates are necessary not only for evaluating the complexity of the software work, but also for planning the necessary software development time, its price and the related financing of developers from the knowledge of the solution of "similar" tasks.

It is clear that the earlier evaluation of the effort by the number of lines of source code may give a very distorted information.

The author defines three approaches in estimating the project complexity: 1. *non-algorithmic*, 2. *algorithmic* and 3. *machine learning techniques*. While the first one is based on the experience of a team of experts from solving previous projects and subjective evaluation, the second one already uses mathematical evaluations of multi-criteria functions and the third one works with the "intelligence" of machine learning algorithms. The author's contribution is to combine the advantages of the second and third approaches to create a combined effort estimation algorithm. At the same time, he also considers data segmentation using clustering techniques (a difficult problem with this technique is determining the appropriate criteria for identifying clusters) and the application of machine learning methods to these clusters, which, according to published sources, leads to greater accuracy of the intensity estimates.

The main objectives of the dissertation are to propose a system for calibrating software complexity using the Function Points Analysis (FPA) method and to develop an optimization framework based on regression analysis, machine learning and clustering, followed by a comparison of the results with benchmark data from published papers and an improvement of the FPA method.

It can be concluded that the focus of the thesis is topical, the objectives of the thesis are challenging with a clearly formulated own contribution, and the thesis is **dissertable** if they are met.

After briefly introducing the topic of the thesis and defining the objectives, the author proceeds in Chapters 2, 3 and 4 to a detailed explanation of techniques for estimating the complexity of software projects and software support used for them, information about freely available or paid versions, a description of machine learning algorithms and an overview of the state of the art in the world literature.

From the description, it is evident that algorithmic approaches based on relatively simple formulas (2.1) to (2.6), but with many parameters whose values have been determined by the developers according to the type of software, and are dependent on the size of the source code in thousands of lines, give questionable results in terms of estimation accuracy. Similarly, the determination of correction factors according to relations (2.6) and (2.7) is questionable.

An important part of the explanation is a description of function point analysis with a focus on data and transaction functions and their measurement.

In Chapter 3, dealing with clustering, due to the uncertainty, the author also introduces the use of fuzzy logic tools (Fuzzy C-Mean) and probabilistic expression (Gaussian Mixture Model). He then demonstrates the learning technique with an Artificial Neural Network (ANN).

The core part of the thesis is the design of the own method in Chapter 5 and the experiments that follow. As already mentioned, it consists of a combination of feature point analysis and clustering methods, which entails the need to find a suitable machine learning algorithm, express the clustering criterion and the final calibration of the parameters to estimate the software complexity. The difficulty lies in the fact that there are a large number of machine learning algorithms (the author lists 24 of them in Table 6-1), they are mostly problem-oriented, and there is no universal choice for selecting the "best" one. To find the most suitable one, the author focused on the 5 most commonly used ones and illustrated the selection process with the diagram in Figure 6-2. He also specified 16 key features (denoted by the term "categorical variables") in Table 6-2. The summaries of clustering algorithms in Tables 6-4 to 6-6 by Xu, Andreopoulos, and Gupta and their classification methods are also useful. The author then selected 6 of them, listed in Table 6-7, which is again followed by an illustrative diagram of the selection of the most suitable algorithm.

The experiments used data from the ISBSG repository and some Python libraries in the implementation.

Detailed results and their evaluation for a representative dataset from different areas (finance, communication, government, manufacturing, service industry) are presented in Chapter 7. Pair-wise *t*-tests were used for statistical evaluation.

Of the many possible approaches, the Bayesian Ridge Regression (BRR) algorithm has been shown to consistently achieve the lowest estimation error, and categorical variables and clustering algorithms are best suited for data segmentation. This is also one of the main contributions of working together with the new weighting system to calibrate the complexity of the software.

Finally, the author suggests other possible research directions with respect to new technologies and other methods of complexity estimation.

In terms of content and graphics, the work is at a very good level, as is the language level of the written text in English.

Formal comments:

- On page 33, when interpreting formula (3.1), you state that "R can be seen as the mean distance from member samples to the centroid". More precisely, it is the quadratic mean of the deviations from the centroid x_0, corresponding to the (sample) standard deviation.
- Typographical rules for writing symbols are not always followed, e.g., for relation (3.17) and the following text.

Questions to the dissertant:

1. Is it possible to include in the effort of a software project, besides the actual code and its function points, the effort of obtaining the necessary input data for calculation by the proposed software??
2. You divide the calculation process into 3 phases (shown in Figure 6-19): 1. data preprocessing, 2. calibration, 3. optimization, and then you list a number of criteria (6.1) to (6.6) for evaluation. Which is most telling, since the best results may vary according to the different criteria? Is it not then appropriate to use multi-criteria evaluation? And

what do the actual values of y_i represent in them if this is still a planned project and we only have estimates?

Conclusion:

PhD student Vo Van Hai's dissertation demonstrated the author's overview of software engineering and modelling tools in evaluating software project effort estimates, as well as the author's ability to creatively use existing approaches to formulate his own methodology, combining their advantages.

The PhD student has applied the results in 20 publications in international forums - 3 in impact factor journals and 17 conference papers - with citations. The results can be built upon in further research by the author or his followers. Therefore,

I recommend

Vo Van Hai's Ph.D. thesis to be accepted by the Committee to be presented and defended in the Engineering Informatics study branch

Brno, October 29, 2022

Prof. RNDr. Ing. Miloš Šeda, Ph.D.
Institute of Automation and Computer Science
Faculty of Mechanical Engineering
Brno University of Technology

Opponent: Assoc. Prof. Oldřich Trenz, Ph.D.
Department: Department of Informatics, Faculty of Business and Economics, Mendel University in Brno
Office address: Zemědělská 1, 613 00 Brno
Contact: oldrich.trenz@mendelu.cz; tel.: 545 13 22 67

Opponent's Report – Dissertation Thesis

Author: **Vo Van Hai**
Title of dissertation: **Optimisation of Software Effort Estimation by Improving Functional Points Analysis**

Degree Programme: Engineering Informatics
Degree Course: Software Engineering
Supervisor: Assoc. Prof. Ing. Zdenka Prokopová, CSc.
Consulting Supervisor: Assoc. Prof. Ing. Radek Šilhavý, Ph.D.

General Assessment

The doctoral qualifying paper (dissertation thesis) deals with a new method of software effort estimation using machine learning techniques. The aim of the dissertation is increasing the accuracy in software effort estimation, with the IFPUG FPA method combined with machine learning techniques to allow effort estimation in groups segmented by the Industry Sector categorical variable. The proposed procedures are based on a new calibration of the functional complexity weight algorithm and an effort at the optimization of the framework for the specific clustering.

The dissertation is divided into ten chapters. While the theoretical part is devoted to the introductory chapters (general theory - chapter 2 and chapter 3), the following chapter 4 summarizes the current state of effort estimation. I regard this chapter as crucial for the further direction of the dissertation and I appreciate the author's analysis of the different approaches and limitations.

Chapter 5 deals with the methodological design of the chosen approach. Here I would have expected a more detailed analysis of the situation, but as a basic overview this chapter is sufficient. Chapter Six is devoted to the experimental part, where the author presents the different parts of his research. I positively appreciate the detailed analysis (research) for the ML methods (Table 6-1) and categorical variables (Table 6-2). Overall, this chapter can be evaluated as a dissertable core of the thesis, although some parts would deserve more detail (a more detailed description of the particular facts). This is followed by Chapter Seven, which focuses on the results presented in the thesis and the discussion of the issue. The focus is on the actual algorithm and data segmentation criteria and then on how to apply this to the design of a new weighing system of calibration complexity. The chapter contains a large amount of statistical data, but without detailed comments (interpretations), for example for Chapter 7. 1. 2. Also, the subsequent summary (7. 1. 3) is very brief. The connecting part is Section 7. 3 (New Calibration System and Optimization Framework) where the author uses the most-suitable machine

learning algorithm (the BRR algorithm) and the most-suitable segmentation criterion (the IS categorical variable) and proposes a new calibration complexity weight system. Similarly, here I would appreciate a more detailed discussion of the various facts presented, including a summary.

The last and no less important part of the dissertation is the section devoted to the validity of the presented facts (Chapter 8), the summary of the findings at the scientific and practical level (Chapter 9), and the section devoted to the summary and further possible development of the addressed problem (Chapter 10).

Relevance of the topic

The topic of the dissertation is topical in its content, software effort estimates are a topic still being addressed, and refining these estimates can bring considerable financial savings. Also, an approach based on machine learning methods to improve parameter estimates seems to be a suitable approach.

Meeting the objectives set out in the dissertation

It can be concluded that the stated aim of the dissertation (chapter 1. 3), its division into 4 sub-objectives, has been fulfilled. Although these partial objectives could have been better formulated or moved to the section dedicated to the methodology of the dissertation.

Progress of the solution, results achieved, contribution of the PhD student

Main contribution of the thesis is the proposal of procedures for more accurate software effort estimation by improving the Functional Point Analysis method (FPA). This is done by extending the FPA method in the of machine learning (BRR algorithm), and clustering with categorical variable "IS".

This contribution was achieved by analyzing the currently used machine learning algorithms (Chapter 7. 1) and selecting the best performing algorithm (BRR algorithm), finding the best clustering criterion (Chapter 7. 2), and constructing a new calibration framework (Chapter 7. 3). All of these were complemented by a detailed analysis along with a detailed statistical evaluation.

Contribution to practice and the discipline

The scientific contribution and contribution to practice is mentioned in Chapter 9 (Contribution of the thesis to Science and Practices). This is in particular an innovation of the FPA (Functional Point Analysis) method for more accurate software effort estimation. In this context, implementation costs can be reduced. The experiments conducted showed that the estimates of development complexity using the new calibration weighting algorithm are more accurate than those using the reference IFPUG FPA method.

Formalization of the dissertation

The formal side of the work can be assessed as good, some graphic objects could be published in higher quality or in higher resolution (for example, Fig 2-1, Fig. 2-2, etc). The author often confuses the hyphen and the hyphen in the text. Different number of decimal places for numerical data (e.g. Table 7-59). The literature sources used in the thesis are of relatively high quality, although more up-to-date sources could be found for some areas. These minor comments do not adversely affect the work as a whole.

Publication activity of the PhD student

The author's publishing activity is adequate with regard to their phase of study, the author focuses on both foreign conferences and journal publications. It is necessary to mention that the author of the dissertation has also published the issues elaborated within the dissertation thesis in journals; he has used the outputs within the IEEE, Springer. The author's latest publication is indexed in the periodical in the field of Q02, this can be assessed as very suitable. Overall, the author has 3 publications in journals and 14 conference papers. This number of papers (conferences) can be considered above standard.

Dissertation reminders

Some parts of the dissertation could be more extensive (more annotated content), although this content is available in the referenced publications (publications of the author).

What I missed in the dissertation (conclusion of the thesis) was a discussion of the results achieved (presented) in the form of a comparison with other solutions, research on the topic. However, I appreciate the section devoted to the validity of the results obtained (Chapter 8).

Overall evaluation of the dissertation

Despite some shortcomings, which are mostly of a formal nature, the work is original, interesting and useful, and therefore I evaluate it positively overall.

Since the doctoral student has demonstrated his ability and readiness for independent research and development activities and thus fulfilled the conditions of Section 47, paragraph 4 of Act No. 111/1998 Coll. on Higher Education, I therefore recommend that the doctoral student be awarded the academic title "doktor" after a successful defence.

Questions for discussion

- This thesis deals with an interesting topic, software effort estimation. Where will the results of this thesis be presented, put into practice, besides papers in journals, participation in conferences?
- In the conclusion of the thesis, the author mentions that the topic can be followed up with research considering the productivity factor. Will the author of the dissertation himself follow up on this issue or will this be implemented within the research team?
- The author of the thesis is concerned with the optimization of software efforts, which has resulted in the design of an optimization framework. What optimality criterion (overall) was targeted here?

In Brno, 1. 11. 2022

Assoc. Prof. Oldřich Trenz, Ph.D.
Opponent's signature