EVOLUTIONARY MEASUREMENT-ESTIMATION METHOD FOR MICRO, SMALL AND MEDIUM-SIZED ENTERPRISES BASED ON ESTIMATION OBJECTS

Celar, S.*; Vickovic, L.* & Mudnic, E.*
*University of Split, FESB, R. Boskovica 32, 21000 Split, Croatia
E-mail: stipe.celar@fesb.hr

Abstract:
In the early phase of software project (in the analyses or user requirements specification phase) there are little measurable software artefacts as well as little time for measurements. Reliable software size estimation in such early project phase significantly influence estimation reliability of other project parameters (like effort, duration or cost) and overall project success rate. This paper presents evolutionary method for estimation objects development based on historical measurement data. Estimation objects, recognized from the measurement data in three steps, are classified according their attributes into hierarchy dimensions. Estimation objects are enterprise native and easy for use in early software size estimation. Also, in the paper case study results from the SME applying Function Points for Functional Size Measurement (FSM) are described but the method is independent of FSM or functional size unit.

Key Words: Functional Size Measurement (FSM), Early Software Size Estimation, Function Point, Estimation Objects, Small and Medium-sized Enterprise (SME)

1. INTRODUCTION

We all witness big changes and improvements that software and the whole information and communication technology (ICT) undergone in the last few decades. However, the successful rate of software projects does not follow that rate [1], [2], [3], [4]. Results in Table I show that software projects for almost ten years do not show clear improvements despite the progress of management practices, software development methods, process improvement models and quality standards [5], [6], [7], [8], [9], [10].

Table I: Project success rate (CHAOS report)) [2].

<table>
<thead>
<tr>
<th></th>
<th>'94 %</th>
<th>'96 %</th>
<th>'98 %</th>
<th>'00 %</th>
<th>'02 %</th>
<th>'04 %</th>
<th>'06 %</th>
<th>'09 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Successful</td>
<td>16</td>
<td>27</td>
<td>26</td>
<td>28</td>
<td>34</td>
<td>29</td>
<td>35</td>
<td>32</td>
</tr>
<tr>
<td>Challenged</td>
<td>53</td>
<td>33</td>
<td>46</td>
<td>49</td>
<td>51</td>
<td>53</td>
<td>46</td>
<td>44</td>
</tr>
<tr>
<td>Failed</td>
<td>31</td>
<td>40</td>
<td>28</td>
<td>23</td>
<td>15</td>
<td>18</td>
<td>19</td>
<td>24</td>
</tr>
</tbody>
</table>

Inaccurate estimates of software development effort or project development cost are frequently reported causes of ICT-project failures [11], [12], [13], [14]. Project effort and cost estimation are perhaps the two most crucial issues that a project manager has to make. Majority of the researchers in this field use terms “cost” and “effort” interchangeably [14].


Project managers have to estimate costs and allocate resources to the system development before it is built. Software cost estimation involves the estimation of one or more of the following variables:

- effort (usually in person-months)
- project duration (in calendar time)
- cost (in $ or €).

The above variables are related through the productivity of Software Production Unit (SPU) that depends on a broad number of characteristics of development process and technology, SPU, business process. Some authors mention from several dozen to more than one hundred of different characteristics [7], [10], [15], [16].

The software size is the most important factor that affects the software cost. In the last three decades the software size, e.g. system functionality become more complex and therefore hard to estimate and measure. During that period methods for software functional size measurement were also developed. In fact there are well defined, explored and standardized methods for functional size measuring (like line of code – LOC, function point – FP, use case point – UCP). Further, there are wide spectrum of well described estimation methods – Direct Estimation Methods and Derived Estimation Methods [16], [17].

However, in practice there are some big obstacles for wider and more efficient use of Functionality Size Measurement (FSM) and estimation methods. They could be summarized in the following way:

- It is hard to measure something that is still in development and do not exist fully [17]. It can only be measured based on development documentation or software product itself and such measurement could be time-consuming and expensive [16].
- For efficient application of measurement and estimation techniques the micro and small software development enterprises should improve their overall maturity grade or use adequate software quality models [3].
- In the time of constant market pressure an agile movement become a dominant development paradigm and SPUs are forced to measure and estimate early in the software life cycle and with fewer requirements.

Today, with ISO, CMM, CMMI and other software quality assurance (SQA) models most of micro and small-medium enterprises (or their SPUs) struggle with software quality assurance tasks [18]. They use methods like Program and Evaluation Review Technique (PERT) and earned value to track progress. The problem is with the imprecision and inaccuracy of most software project plans [11].

This paper proposes a small contribution for adoption measurement and estimation methods in micro, small and medium-sized enterprises (MEs and SMEs) environment based on following premises:

- enterprise native SW-objects – estimation objects are produced in enterprise’s ordinary development process and all people are familiar with
- independency of Functional Size Unit (FSU) – functionality size of estimation objects can be measured/estimated in any unit.

The remainder of the paper is organized as follows: Section 2 describes measurement and estimation challenges that micro, small and medium-sized software enterprises are confronting with. Brief overview of software size measurement, estimation metrics and methods is given in Section 3. Section 4 introduces proposed estimation method together with the case study in medium-sized software development enterprise. Section 5 concludes the paper.
2. MEASUREMENT AND ESTIMATION PROBLEMS IN MEs AND SMEs

Micro, small and medium-sized enterprises play a central role in the European economy. They are a major source of entrepreneurial skills, innovation and employment and account for a large proportion of Europe’s economic and professional activity. In practice, 99.8% of businesses in the European Union are SMEs. Between 2002 and 2008, the number of SMEs increased by 24 million (or 13%) whereas the number of large enterprise increased by only 2000 (or 5%). The growth was also reflected in employment figures. In absolute numbers 9.4 million jobs were created in the SME-sector between 2002 and 2008 [19].

Table II: Enterprises in Croatia, December 2011 (SW, ICT and general) [20], [21].

<table>
<thead>
<tr>
<th>Nr. of enterprises</th>
<th>Number of employees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>ALL activities</td>
<td>126.264</td>
</tr>
<tr>
<td>%</td>
<td>100%</td>
</tr>
<tr>
<td>J - Information and communication</td>
<td>4.867</td>
</tr>
<tr>
<td>%</td>
<td>100%</td>
</tr>
<tr>
<td>SW enterprises (J 62 – Comp. programming)</td>
<td>2.755</td>
</tr>
<tr>
<td>%</td>
<td>100%</td>
</tr>
<tr>
<td>SW enterprises in ICT sector (%)</td>
<td>57%</td>
</tr>
<tr>
<td>SW enterprises in general economy (%)</td>
<td>2%</td>
</tr>
</tbody>
</table>

Data about local Croatian and broader USA market (Table II and Table III confirm the above statements. SW enterprises are mainly Micro Enterprises (ME), with 10-20 employees. Survey conducted on 150 software SMEs shows that majority of them very rarely recognised and used basic quality assurance concepts from the internationally recognised quality standards, including ISO and Capability Maturity Model Integration (CMMI) [18].

As an answer on constant market pressure an agile movement becomes a dominant development paradigm during last two decades. In their agility many SPUs reduce significantly time for planning, analysis and design [8]. Further, reducing documentation makes difficulties for applying FSM techniques based on FP that presumed user software requirements specification (SRS) and design documentation. Therefore it is hard to expect that SPUs, especially SMEs, devote enough time to adopt and use relative complex estimation and measurement techniques. In such environment it is hard to make reliable estimation, especially in early project phase, i.e. before design specification or some prototype are done [22], [23]. In fact, estimation failure can be huge [9], [10], [22], [24]. Because of nonexistent measurable artefacts it is not even possible to apply measurement methods as a mean to get reliable estimation. An estimation uncertainty became smaller as project progresses but in the same time its importance decreases significantly (Figure 1).

From those reasons we proposed the method for creating the estimation objects from the enterprise’s historical data. From another side, reliable software size estimation or measurement is only the first step in reliable project cost and effort estimation [3], [4], [14], [25]. It is reason we propose evolutionary three-phased approach for creating native estimating objects and Functional Size Units (FSU) for MEs and SMEs.
3. SOFTWARE SIZE MEASUREMENT AND ESTIMATION METHODS

Software metrics are measures that are used to quantify software artefacts, software development resources, software development process or software related projects. In software engineering more than 300 metrics have been defined [27]. Some of them are only used for scientific purposes and they don’t have practical implementations. Especially for SME estimation methods, or combination of measurements and estimations are tried to be used instead of only measurements. This section gives a short overview of software attributes and methods for their measurement and estimation.

Object-oriented programming languages introduced new measurement objects and measurement attributes. Software complexity described by objects and their attributes can be measured with specialized object-oriented metrics based on use case (Figure 2).
The most measured and estimated attribute is software size. Software size is often difficult to measure and estimate, especially early in a project (prior to the completion of the requirements specification). Later in a project, certainty of an estimate becomes bigger and less important [9], [10], [28] (Figure 1). Certain methods and metrics used in FSM can impact on the measurement’s potential accuracy.

3.1 LOC as functional size unit

Methods for measuring the complexity of software try to identify the measurable software properties and connections between them. Some of the measurable properties are the length, volume, space occupied in memory, the number of lines of code, function points. The Line of Code (LOC, SLOC, KLOC, KSLOC) measure – a count of the number of machine instructions developed, was the first measure applied to software. Its first documented use, the attempt to formally measure software development productivity was in the 70's [29]. LOC method has some big drawbacks. LOC is dependent on the programming language and code size does not represent the functionality of the program. Accurate LOC is to know only after the project is completed, and therefore it is hard to use this method for estimation (except for expert judgments in analogy methods) [7], [16], [30].

Nevertheless, LOC is still popular functional size unit (FSU) and many enterprises express functionality of their products in Source Lines of Code (SLOC) and convert ‘newly’ measures into this one. On this way a measurement process can be automated [16] and SLOC counting for completed projects can be done with minimal effort and without subjectivity [29].

3.2 Function Points as functional size unit

In 1979 Allan Albrecht published Function Points metrics for quantifying the amount of business functionality an information system delivers to its users [15], [16]. It is a synthetic metric derived from enumerating five visible, well-defined external characteristics of software based on 5 Base Functional Components (BFCs) [16], [24], [31]:

1. transactions:
   - external inputs (EI) such as logical transaction inputs or system feeds;
   - external outputs (EO) or external inquiries (EQ) such as online displays, reports or feeds to other systems;
2. data entities:
   - internal logical files (ILF) such as logical groups of user defined data and
   - external interface files (EIF) such as interfaces to other systems.

Function Points related very closely to the types of business applications that IBM was developing at the time. For these types of systems, they are a far superior measure of business value than SLOC and can be much better for an organization that develops these types of systems to use for productivity comparison studies.

Function Points metric has its limitations that make it difficult to start using in software enterprises, especially in small and medium-sized ones. The most important limitations are described below.

FP is relatively complex metric, extensive for learning and requires lots of time to master it. Due to lower limits in calculating function point complexity, the lower limit for normal function point calculations is about 15 FPs [16], [32]. Therefore micro function point approach for smaller changes measurement is new challenging approach. Furthermore, FP cannot be extracted automatically from design documents [33].

FP metric is extensible in counting as well. Due to average price of functionality measurement (4 6 $/FP) the FP are not applicable in practice for SW projects with large
functionality (with more than 15000 FP) [16], [32]. Consequently, the measurement automation for projects with large functionality is challenge as well.

As of 2012 there are more than 20 variants of FP metrics with few conversion rules from one method to other [16], [32] and at least five of them (IFPUG, Mk II, COSMIC, NESMA and FiSMA) were confirmed by the International Organisation for Standardisation (ISO). This diversity of variants additionally complicates practical use of FP. In literature it is possible to find additional function points’ drawbacks (e.g. [34]), depending of author’s specific appliance area or point of view. However, function points are widely accepted and applied concept, and as such they are a basis for many other measurement and estimation methods.

From the estimation point of view there is an important drawback of FP based methods. That is a necessity to analyze system/software under study or more precisely to decompose it (to recognize logical transactions) and make entity and relation model (to recognize data entities) in order to identify the Base Functional Components. It is slow and complicated.

3.3 Object-oriented FSMs

Use cases were first introduced by Ivar Jacobson in the mid 80's. They provide a language for describing the requirements of a software system in a way that facilitates communication between developers and the eventual users of the system. Each use case describes a typical interaction that may occur between a user (human operator or other software system) and the software. The focus is on the functions that a user may want to perform or have performed rather than on how the software will actually perform those functions.

Useful and practical FSUs that are oft used for the most measured software attribute are Use Case Point and Story Point [9], [16], [28]. An Use Case Points method (UCP) were introduced in 1993 by Gustav Karner [29]. Use Case Points count and classify two elements: 1) the actors in the use case and 2) the transactions that are required to make the use case happen. Use case points describe the functionality being delivered rather than the way this functionality is implemented. Unlike Function Points, the Use Case Points can cover a wider spectrum of application types. The problem with using use cases is their lack of standardization across the industry. An organization which has a well defined process for defining use cases could successfully use them for productivity tracking and effort estimation. [16], [29]

A Story Point is usually somewhat larger than a function point perhaps roughly equivalent to two IFPUG function points, or perhaps even more [16]. They are defined within the SPU and are used within that unit to estimate effort and measure productivity and quality.

3.4 Estimation methods

Software size estimation methods can be generally considered as Direct Estimation Methods (or non-algorithmic, Expert Opinion Methods) and Derived Estimation Methods [16], [17]. Direct Estimation Methods (or non-algorithmic) imply the cooperation of one or more experts which directly estimate required elements of the estimation of function points, basing their estimation on experience and intuition.

The basic difference between direct and derived estimation methods is that in the latter, estimation isn’t performed directly on function point values, but rather on different project parameters which are somehow related to function point values.

According to the abovementioned categorization of direct and derived estimation methods, one of the most known direct estimation method is the Delphi according to which the predictions of a number of experts are combined. [17] Decomposition Method is also often used method. By decomposing the project into smaller subprojects, it is possible to conduct the evaluation part by part, in detail. Estimations based on analogy, Parkinson’s Method and ‘Price-to-win’ also fall into the category of direct estimation methods.
Derived Estimation Methods or Algorithmic Model Methods provide with the estimation of complexity as a function of more variables which relate to certain attributes of software development project.

During the late 1970s and during the 1980s, some sophisticated algorithmic software estimating models were developed such as Dr. Boehm’s Constructive Cost Model (COCOMO), PRICE-S (now True S), SLIM, SEER-SEM, and SPQR/20 (now Knowledge Plan).

Each algorithmic model has the form:

\[ \text{Effort} = f(x_1, x_2, \ldots, x_n), \]

where \((x_1, x_2, \ldots, x_n)\) are the cost factors.

The existing algorithmic methods differ in two aspects:
- in the selection of cost factors and
- in the form of function \(f\).

Another important variable relating estimation is the productivity. The relationship between the size of software and the effort required to produce it is named productivity. There are plenty factors influencing the overall software project productivity of SPU – Albrecht mentioned 14 factors [15] while other researchers mention 45 [16] or even more than 100 different productivity factors [7], [10]. These factors have to be taken into account in effort estimation and project planning.

4. PROPOSED METHOD

Software enterprises (or Software Production Unit in software enterprises) can not apply measurement and estimation methods without tailoring them. Learning, tailoring and adopting those methods is not easy process for the ME&SMEs from few reasons:
- FSM and estimation methods itself are quite complex
- lack of trained internal estimators and measurement experts
- internal SME’s processes have to be enough mature to implement FSM and estimation concepts
- lack of a database of historical measurement data.

4.1 Evolutionary approach

CMMI maturity level needed for efficient adopting FSM and estimation methods, of at least core processes, is 3-4 and learning time depends on more factors. Majority of software enterprises are in ME&SME category and struggle with software assurance activities. Measurement process implementation is big challenge for them also. That implementation is not simple and can take several years. As a result in selection of metrics and methods it is necessary to align enterprise’s or SPU’s capabilities with existing metrics and methods. An evolutionary approach, step by step could be useful.

In the first phase called measurement phase as a strategic goal is established application of adequate measurement method. In this phase it is important to:
- create politics of measurements
- identify a person or team to be carrier of measurement politics
- select adequate metrics and methods
- provide internal education and application on real projects
- establish repositories for measurement data
In this phase (Phase A, Figure 3) it is important to recognize objects/entities that could be used for future estimations. It is desirable that objects can be easily recognized in development process and easy connected with measurement data. In addition, metrics closer to large number of developers (not only scientists) have more chances to be widely accepted. For example, UCP is more understandable and simpler for appliance than Unadjusted Function Point (although less researched scientifically).

The objects' granularity can vary from single modules to whole projects, but it is the most relevant to identify those objects that are repeatable in the system (i.e. they can be grouped) but also differ from similar ones by some attribute. Those objects could be for example some class, forms, queries, reports or even some kind of modules.

The next phase is the most intensive phase (Phase B, Figure 3). Measurements and estimations are conduct in parallel. Estimation is done based on measurement data and recognized objects from the Phase A. Strategic goal of this phase is to select and implement adequate estimation metric and method, as well as to establish estimation data repository. Objects recognized from measurement data are classified according their attributes displayed in ontology model (Figure 4) and its dimension \( n \) described in the Table IV. Each dimension consists of some elements \( m \). After classification the objects can be used for software attributes estimation and/or project estimation and compared with measurement data. After reaching these goals, in the Phase B, the SPU can from own classified history data recognise typical produced objects and identify effort/time/cost spent for those objects.

Some external collection of measured data like ISBSG [35], SPQR and others [16], [35] can be used for benchmarking, but their constraints regarding accurateness must be taken into account [16]. Advantage of internal repositories regarding benchmarking repositories is obvious – internal data and internal objects are native for process participants. As a result process participants can easily accept and use such repositories for estimation.

After estimation objects are classified the estimation itself should be validated on same real projects parallel with measurement process. After validation the measurement activities could be replaced with much simpler and cheaper estimation process. It means, the Phase B is completed and begins the final phase, Phase C (Figure 3).

A maximal number of object instances is calculated according to the following formula:

\[
N_{\text{max}} = m^n
\]

where is \( m \) maximum among all dimensions.
For a model with four dimension and two elements in each dimension it comprises 16 estimation elements, what could be quite enough if validated.

### 4.2 Case study

Software Medium-seized Enterprise applied systematically proposed measurement & estimation method. Measurement data are acquired during two years long measurement in development projects. Objects are then classified, validated and used for small and big development projects (Table IV and Figure 5).

An estimation model consists of 4 dimensions with 2 elements at the first three dimensions (Type, Complexity and Development, Figure 5). In the 4th dimensions the model has 4 objects in some branches (Figure 5). The number of estimation objects is 18 (theoretical maximal number would be $4^4$). This number of estimation objects has proved as a sufficient for reliable estimates.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Dimension’s instances recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object Type</td>
<td>It is advisable to distinguish at least two or three different levels according to the objects SPU produces. Business area specificity could be expressed on this level by adding separate vertical.</td>
</tr>
<tr>
<td>Object Complexity</td>
<td>It is advisable to distinguish between two or three different levels relating complexity</td>
</tr>
</tbody>
</table>
| Development perspective    | It is advisable to distinguish between two or three different levels according to the life cycle. Advisable is to distinguish between:  
- new development  
- enhancement  
- re-development of existing software |
| Component perspective      | This is level for the objects. Business area specificity could be expressed on this level by adding separate objects.                                                        |
Table V: Characteristics of 18 final software objects – case of medium-sized enterprise.

<table>
<thead>
<tr>
<th>Object</th>
<th>Description</th>
<th>Measured avg. UFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCN-1</td>
<td>NEW COMPLEX MODULE – 1, it is a new module that links more entities and it does not possess any data processing (e.g. item entry or generating)</td>
<td>44.49</td>
</tr>
<tr>
<td>MCN-2</td>
<td>NEW COMPLEX MODULE – 2, it is a new module that links more entities and it possesses data processing (e.g. item entry or generating)</td>
<td>116.9</td>
</tr>
<tr>
<td>MCN-3</td>
<td>NEW COMPLEX MODULE – 3, it is a new module that possesses complex data processing (e.g. item entry or generating) or it contains more logical units (e.g. write-off record)</td>
<td>318.72</td>
</tr>
<tr>
<td>MCO-1</td>
<td>OLD COMPLEX MODULE – 1, it is just a field adding or removal modification on one tab in complex module (e.g. header modification)</td>
<td>12.66</td>
</tr>
<tr>
<td>MCO-2</td>
<td>OLD COMPLEX MODULE – 2, it is just a field adding or removal modification on several tabs in complex module (e.g. header and items modification)</td>
<td>31.77</td>
</tr>
<tr>
<td>MCO-3</td>
<td>OLD COMPLEX MODULE – 3, it is a modification in one batch in complex module</td>
<td>74.99</td>
</tr>
<tr>
<td>MCO-4</td>
<td>OLD COMPLEX MODULE – 4, it is a modification in several batch in complex module</td>
<td>295.86</td>
</tr>
<tr>
<td>MSN-1</td>
<td>NEW SIMPLE MODULE – 1, it is a new module in a common process category like preview or simple control</td>
<td>4.02</td>
</tr>
<tr>
<td>MSN-2</td>
<td>NEW SIMPLE MODULE – 2, it is a new module that references just one entity. This category also includes complex controls that appear on several places</td>
<td>16.65</td>
</tr>
<tr>
<td>MSO</td>
<td>OLD SIMPLE MODULE; is a modification on simple module (type 1 or 2) that keeps module still simple</td>
<td>3.56</td>
</tr>
<tr>
<td>RCN-1</td>
<td>NEW COMPLEX REPORT – 1, it is a new report that receives data from more entities and does not include any data processing</td>
<td>18.94</td>
</tr>
<tr>
<td>RCN-2</td>
<td>NEW COMPLEX REPORT – 2, it is a new report that receives data from more entities and includes data processing</td>
<td>32.01</td>
</tr>
<tr>
<td>RCO-1</td>
<td>OLD COMPLEX REPORT – 1, it is a modification of the complex report that references less than four entities, or adding or deleting of five or less than five fields is performed without any changes on data processing</td>
<td>6.04</td>
</tr>
<tr>
<td>RCO-2</td>
<td>OLD COMPLEX REPORT – 2, it is a modification of the complex report that references more than four entities, or adding or deleting from five to 15 fields is performed without any changes on data processing</td>
<td>12.75</td>
</tr>
<tr>
<td>RCO-3</td>
<td>OLD COMPLEX REPORT – 3, it is a modification of the complex report that performs adding or deleting more than 15 fields or just one segment of data processing is changed</td>
<td>30.11</td>
</tr>
<tr>
<td>RCO-4</td>
<td>OLD COMPLEX REPORT – 4, it is a modification of the complex report that changes more segments of data processing</td>
<td>89.60</td>
</tr>
<tr>
<td>RSN</td>
<td>NEW SIMPLE REPORT, is a new report that receives data from simple entity</td>
<td>5.09</td>
</tr>
<tr>
<td>RSO</td>
<td>OLD SIMPLE REPORT, it is a modification of simple report that keeps it simple</td>
<td>5.38</td>
</tr>
</tbody>
</table>

Figure 5: 18 final software objects for estimation – case of medium-sized enterprise.

Additional objects' validation followed in large development project lasting for 13 months and size of delivered software is measured by method Mk II FP Index (V1.3) [36] and it was 7.560
Unadjusted Function Points (UFPs). Software size was estimated in the early development phase using the initial and additional user requirements – estimation was 6.773. Overall estimation error was only 11.6% [37].

According to [10] typical software organizations are struggling to avoid estimates that are incorrect by 100% or more while only sophisticated software organizations can achieve results within ±5% of estimated results instead of within ±10%. Similar also states [32] and [35].

5. CONCLUSION

Adoption of measurement and estimation processes belongs under strategic goals for software enterprise. With reliable early software size estimation method an enterprise can improve overall project success rate significantly. Advantage of the proposed evolutionary measurement&estimation method is its independence of measurement and estimation methods, metrics and functional size units. The characteristics of the method could be briefly summarised:

- measurement data from own historical projects are used for estimation objects creating
- created estimation objects are easy to recognise and use for early software size estimation.

This method is appropriate especially for smaller SPU (micro, small and medium-sized enterprises), but it can also be used by large enterprises.

Drawback of this method is its complexity in establishing of estimation objects from the company's own history measurement data.

Future research will focus on reducing effort in data collection through implementing probabilistic methods like Bayesian network.

REFERENCES


