

# A Cost-Benefit Analysis Model for Technical Debt Management Considering Uncertainty and Time

Carlos Fernández-Sánchez, Jessica Díaz, Juan Garbajosa, Jennifer Pérez

Technical University of Madrid (UPM) – CITSEM – Ctra. Valencia, Km. 7, E-28031 Madrid, Spain

{carlos.fernandez, yesica.diaz}@upm.es, {jgs, jenifer.perez}@eui.upm.es

## 1. Introduction

In the last few years, technical debt [1] has been used as a useful means for making the intrinsic cost of the internal software quality weaknesses visible. This visibility is made possible by quantifying this cost. Specifically, technical debt is expressed in terms of two main concepts: *principal* and *interest* [2]. The *principal* is the cost of eliminating—or reducing—the impact of a, so called, technical debt item in a software system; whereas the *interest* is the recurring cost, over a time period, of not eliminating a technical debt item. Previous works about technical debt [2] are mainly focused on estimating principal and interest, and on performing a cost-benefit analysis. This cost-benefit analysis allows one to determine if to remove technical debt is profitable and to prioritize which items incurring in technical debt should be fixed first. Nevertheless, for these previous works technical debt is flat along the time. However the introduction of new factors to estimate technical debt may produce non flat models that allow us to produce more accurate predictions. These factors should be used to estimate principal and interest, and to perform cost-benefit analysis related to technical debt. In this paper, we take a step forward introducing the *uncertainty* about the interest, and the *time frame* factors so that it becomes possible to depict a number of possible future scenarios. Estimations obtained without considering the possible evolution of the interest over time may be less accurate as they consider simplistic scenarios without changes. This assertion is elaborated in the following paragraphs within this section.

Interest *uncertainty* is the probability that no extra cost is derived from technical debt. For example, if the interest for a technical debt item is estimated in terms of extra maintenance cost, but the software system incurring in that technical debt need not be changed over a period of time, then no interest has to be paid during this period. Additionally, the interest does not remain constant over time due to factors external to the project or software system that may take place occasionally. For example, an increase or a decrease in the utilization (e.g., due to seasonal business activity) of a software system could affect the number of emergent problems. Thus, the system could require more or less maintenance at different times.

*Time frame* refers to the time period under study. This time frame could be determined by external constraints or

deadlines such as legal normative issues that an application should be bound to, milestones or, availability of resources or contractual restrictions, among others. The interest is a cost that must be paid continuously over time. That is, the interest over time will cause accumulated costs. Therefore, with enough time ahead, paying off the principal is always profitable because the accumulated interest grows and grows. Hence, determining the time frame is important for a realistic cost-benefit analysis.

This paper presents a technical debt based cost-benefit analysis model that deals with interest uncertainty and considers the time frame under analysis. Cost-benefit analysis about technical debt obtained from this model can be more consistently grounded than when former models are used, since possible future situations that could happen around the system can be systematically included in the analysis.

## 2. Cost-benefit analysis model

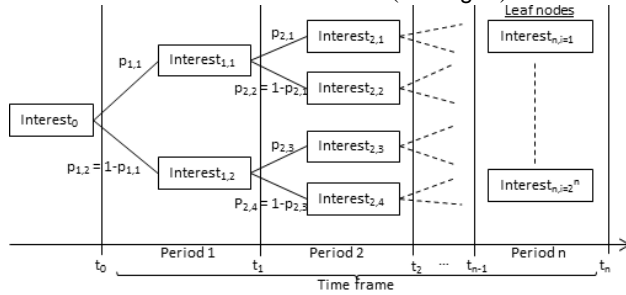
This section presents a model to perform the cost-benefit analysis based on the technical debt concept in which the time frame and the interest uncertainty are model variables. This model takes advantage of binary trees. Binary trees are used to estimate the expected value of the interest. Then, to obtain the net expected value of paying off the principal, the principal is subtracted from the interest expected value.

Binary trees (BT) have been selected because they facilitate the understanding of the technical debt of the system under analysis. This is due to the fact that they can be used to illustrate the possible evolution paths by assigning probabilities to their branches and assigning a weight to each path. That is, tree branches graphically represent several alternatives for the interest evolution. As a result, BT usage facilitates the understanding of the technical debt concept of a system under analysis. More complex trees, with more branches and nodes, would make it harder to understand the technical debt behavior.

BTs are used in the model for estimating the interest uncertainty (see Fig. 1). The tree grows over the time frame under analysis (see Fig. 1). The time frame is divided into periods of time, in such a way the interest can evolve from the current period of time to the next one. Branching coincides with the time frame events ( $t_0 \dots t_n$ ), over which the interest evolves. The root node labeled as  $Interest_0$  represents the current estimated interest at the

moment of the analysis. In *Period 1*, there are two nodes that represent the interest in that moment, one pessimistic labeled as  $Interest_{1,1}$ , and another one optimistic labeled as  $Interest_{1,2}$ . The lines that join the nodes represent the possible paths in the evolution of the interest, and are labeled with the probability of such evolution. For example, there is a line between the node  $Interest_0$  and  $Interest_{1,1}$  that indicates that the interest can evolve following that path, and the probability  $p_{1,1}$  is the probability of such evolution (see Fig. 1). Finally the number of depth levels of the tree is established by the number of time periods defined.

To represent the accumulated interest during the time frame under study is also required. BTs are also useful to represent this. From the tree Fig. 1 it is derived with the accumulated interest ( $AccInt$ ) in their nodes. This new tree has the same probabilities, periods and structure as the one shown in Fig. 1. This tree is named the accumulated interest evolution tree (see Fig. 2).



**Fig. 1. Modeling interest uncertainty using BT**

The accumulated interest evolution tree provides a key data for the study; its leaf nodes represent the possible results of the interest evolution (see Fig. 2). This data is necessary to calculate the expected value (EV) of paying off the principal by calculating the expected interest that would be avoided. As a result, net expected value can be calculated using this data. To calculate the EV, it is necessary to sum all the leaf nodes of the accumulated interest evolution tree ( $AccInt_{n,i}$ ) weighted by their probability. These probabilities are the combined probability (CP) of the whole branch. Let  $n$  be the depth of the tree,  $i$  the number of the node into a level, and  $p_{n,i}$  the probability that the node  $AccInt_{n,i}$  occurs from its parent node, the EV is calculated with the formulas (1), (2) and (3):

$$\text{if } n = 1 \Rightarrow CP_{n,i} = p_{n,i} \quad (1)$$

$$\text{if } n > 1 \Rightarrow CP_{n,i} = p_{n,i} * CP_{n-1, \frac{i+i \bmod 2}{2}} \quad (2)$$

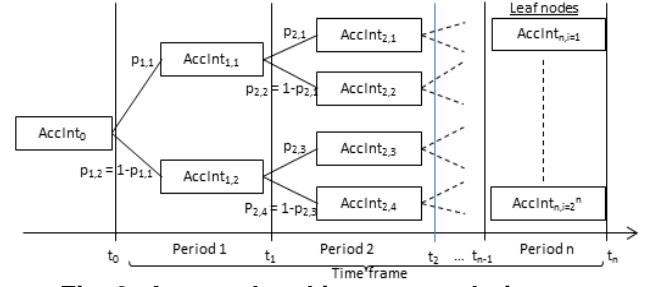
$$EV = \sum_{i=1}^{2^n} LN_i * CP_{n,i} \quad (3)$$

Finally, the net expected value (NEV) of paying off the debt is calculated by subtracting the principal (Principal) to the expected value (EV) (see Formula (4)).

$$NEV = EV - Principal \quad (4)$$

As a result, the cost-benefit analysis model obtains an estimation of the NEV considering the uncertainty, the time frame and the cost of paying off the principal. This is

due to NEV estimation is calculated by considering the possible evolution of the interest, and then the interest uncertainty.



**Fig. 2. Accumulated interest evolution tree**

### 3. Discussion and conclusions

This paper takes a step forward in the cost-benefit analysis based on technical debt by presenting a new model that allows performing cost-benefit analysis dealing with uncertainty of the interest and considering the time frame. In addition, the model helps to reason about the possible situations that can occur around the system and that can affect the evolution of the system interest. The model has been formalized using BTs.

One of the main challenges is to get experience from the use of the model in large projects. As the model uses estimated data as input, it will model realistically the technical debt of the system only if the inputs are correct. It is important to obtain good estimations of such inputs. In this direction, to obtain the input data from models (principal, interest, and probabilities) in a systematic/automatic way would be a landmark. This automation is especially necessary in industrial projects in order not to disturb the normal project development. Another challenge is extending the model with more factors, for example, considering alternative developments to pay off the principal. New factors could introduce new sources of uncertainty, and therefore the model should deal also with them.

Our plans for the future are to cope with the above mentioned challenges. At present issues related to input data are being tackled. One of the final objectives is that this model can be easily used in industrial projects to help technical debt management and to give support in making decisions about when it is necessary to improve a system to solve weaknesses.

### 5. References

- [1] W. Cunningham, "The WyCash Portfolio Management System", OOPSLA '92, 1992, pages 29-30
- [2] I. Ozkaya, P. Kruchten, R. L. Nord, and N. Brown. "Managing technical debt in software development: report on the 2nd international workshop on managing technical debt", ICSE 2011.

This work has been partially sponsored by the Spanish fund (INNOSEP TIN200913849) and UPM under their Researcher Training.