A Proposal of Derivation Methods of FP Transformation Formulas

Masateru Tsunoda, Shinsuke Matsumoto, Akito Monden, Ken-ichi Matsumoto
Graduate School of Information Science, Nara Institute of Science and Technology
{masate-t, shinsuke-m, akito-m, matumoto}@is.naist.jp

Abstract

When FP measurement methods are different between the same functional size projects, FPs are different between them. FP measurement methods are often customized by software vendors. To reinforce accountability for a customer, FP based on a customized method should be transformed to FP based on a standard method. So we proposed two derivation methods of FP transformation formulas. One method is focused on productivity. Productivity of each FP measurement method is transformed by Z-score to derive FP transformation formulas. The other method builds a multiple linear regression model whose objective variable is FP and one of the explanatory variables is effort for each FP measurement method, and makes simultaneous linear equations by the models to derive FP transformation formulas.

1. Introduction

Function point (FP) denotes software functional size, and there are many FP measurement methods such as IFPUG, NESMA, and COSMIC-FFP. Moreover, FP measurement methods are often customized by software vendors. When FP measurement methods are different between the same functional size projects, FPs are different between them. To reinforce accountability for a customer, FP based on a customized method should be transformed to FP based on a standard method. Additionally, projects can be compared precisely with a benchmarking dataset like ISBSG with transforming FP. So we propose two derivation methods of FP transformation formulas. One method is focused on productivity. This method is used when productivity of projects where FPs are measured by different methods seems the to be same. The other method uses a multiple linear regression model. This method is used when it is not obvious that productivity of projects where FPs are measured by different methods is the same and some metrics are able to be used for multiple regression models.

2. Derivation Method Based on Productivity

We assume that the difference of FP measurement methods causes the difference of productivity. Productivity is defined as Effort/FP. Figure 1 shows productivity of projects whose FPs are measured by different FP measurement methods. In this method, we align productivity of each FP measurement method and derive FP transformation formulas as follows.

Step1. Deriving a productivity transformation formula

We derive a productivity transformation formula with Z-score, considering average and variance. μ_A indicates average of productivity of projects in which FPs were measured by method A, and σ_A indicates standard deviation. μ_B indicates average of productivity of projects in which FPs were measured by the method B, and σ_B indicates standard deviation. P_{Ai} indicates productivity of $Proj_i$ (i = 1, 2, …, n) in which FP was measured by method A. That is, P_{Ai} is calculated with FP measured by method A. $\hat{P}_{Bi}$ is converted prod-

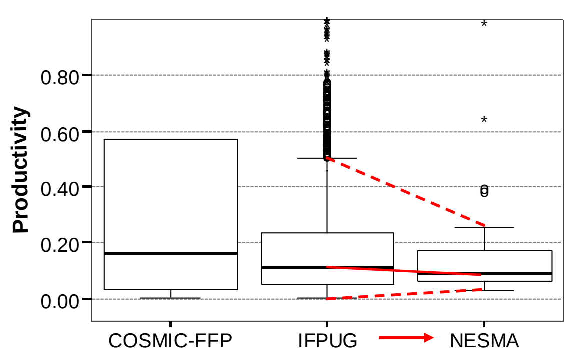

Fig.1 Example of productivity of each FP measurement method

uctivity of $Proj_i$, assuming FP was measured by method B in the $Proj_i$. Namely, $\hat{P}_{\mathrm{B}i}$ is assumed to be calculated with FP measured by method B. $P_{\mathrm{A}i}$ is converted to $\hat{P}_{\mathrm{B}i}$ with the following formula based on Z-score.

$$\hat{P}_{\mathrm{B}i} = \frac{\sigma_{\mathrm{B}}}{\sigma_{\mathrm{A}}}(P_{\mathrm{A}i} - \mu_{\mathrm{A}}) + \mu_{\mathrm{B}} \quad (1)$$

Step 2. Calculating converted FP value of each project

$F_{\mathrm{A}i}$ indicates FP of $Proj_i$ ($F_{\mathrm{A}i}$ is measured by method A), E_i indicates development effort of $Proj_i$, and $\hat{F}_{\mathrm{B}i}$ indicates FP of $Proj_i$, assuming FP is measured by method B in $Proj_i$. $\hat{F}_{\mathrm{B}i}$ is expressed as $\hat{F}_{\mathrm{B}i} = \hat{P}_{\mathrm{B}i}E_i$ because $\hat{P}_{\mathrm{B}i} = \frac{\hat{F}_{\mathrm{B}i}}{E_i}$. With formula (1) and $P_{\mathrm{A}i} = \frac{F_{\mathrm{A}i}}{E_i}$, $\hat{F}_{\mathrm{B}i} = \hat{P}_{\mathrm{B}i}E_i$ is transformed as the following formula.

$$\hat{F}_{\mathrm{B}i} = \frac{\sigma_{\mathrm{B}}}{\sigma_{\mathrm{A}}}F_{\mathrm{A}i} - \left(\frac{\sigma_{\mathrm{B}}}{\sigma_{\mathrm{A}}}\mu_{\mathrm{A}} - \mu_{\mathrm{B}}\right)E_i \quad (2)$$

Each $\hat{F}_{\mathrm{B}i}$ of $Proj_i$ is calculated, using formula (2).

Step 3. Deriving an FP transformation formula

Pairing $F_{\mathrm{A}i}$ with $\hat{F}_{\mathrm{B}i}$, a simple linear regression model is built. The objective variable is FP measured by method B, and the explanatory variable is FP measured by method A. This model changes over from FP measured by method A to FP measured by method B.

In addition to FP measurement methods, some factors affect productivity of software projects [1][2]. To weaken the affect of other factors, projects should be stratified before deriving FP transformation formulas. For example, if team size and programming language have a strong relationship to productivity, only projects that have almost the same team size and same programming language are used for deriving FP transformation formulas.

3. Derivation Method Based on Multiple Linear Regression Models

In this method, a multiple linear regression model whose objective variable is FP and one of the explanatory variables is effort is built for each FP measurement method, and simultaneous linear equations are made by the models to derive FP transformation formulas. Using metrics as explanatory variables of the multiple linear regression model, influence of other metrics can be eliminated.

On a project, $F_{\mathrm{A}i}$ denotes FP measured by method A, $F_{\mathrm{B}i}$ denotes FP measured by method B, E_i denotes development effort, and $X_{i1}, X_{i2}, \ldots, X_{ij}$ ($i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$) denote metrics except for FP and effort. The relationship of $F_{\mathrm{A}i}$ and E_i (the model whose objective value is FP measured by method A) is denoted as the following formula with function f_{A}.

$$F_{\mathrm{A}i} = f_{\mathrm{A}}(E_i, X_{i1}, X_{i2}, \ldots, X_{ij}) \quad (3)$$

Similarly, the relationship of $F_{\mathrm{B}i}$ and Ei (the model whose objective value is FP measured by method B) is denoted as the following formula using function f_{B}.

$$F_{\mathrm{B}i} = f_{\mathrm{B}}(E_i, X_{i1}, X_{i2}, \ldots, X_{ij}) \quad (4)$$

When simultaneous linear equations are made by formulas (3) and (4), values of $F_{\mathrm{A}i}, X_{i1}, X_{i2}, \ldots, X_{ij}$ are given, and unknowns are E_i and $F_{\mathrm{B}i}$, the unknown $F_{\mathrm{B}i}$ (and E_i) can be decided because the number of unknowns and the number of equations are the same. The formula that changes over from FP measured by method A to FP measured by method B is set up by solving $F_{\mathrm{B}i}$ with formulas (3) and (4).

$$F_{\mathrm{B}i} = f'(F_{\mathrm{A}i}, X_{i1}, X_{i2}, \ldots, X_{ij}) \quad (5)$$

4. Conclusions

We proposed two derivation methods of FP transformation formulas, which are based on the difference of productivity and on multiple linear regression models. Our future works are applying the proposed method to the actual project dataset, deriving FP transformation formulas, and confirming validity of the formulas by measuring errors when FPs are transformed from method A to B, and the transformed FPs are transformed from B to A again.

Acknowledgements

This work is being conducted as a part of the StagE Project, the Development of Next Generation IT Infrastructure, supported by the Ministry of Education, Culture, Sports, Science and Technology.]

References

[1] J. Blackburn, G. Scudder, and L. Wassenhove, "Improving Speed and Productivity of Software Development: A Global Survey of Software Developers", *IEEE Trans. on Software Engineering*, Vol.22, No.12, pp.875-885, 1996.

[2] K. Maxwell, L. Wassenhove, and S. Dutta, "Software Development Productivity of European Space, Military, and Industrial Applications", *IEEE Trans. on Software Engineering*, Vol.22, No.10, pp.706-718, 1996.