

# 工数予測支援ツール $e^3$ の実用性評価と機能拡張

古田雄基 肥後芳樹 楠本真二

大阪大学大学院情報科学研究科

{y-furuta,higo,kusumoto}@ist.osaka-u.ac.jp

## 要旨

ソフトウェア開発プロジェクトでの開発工数の見積りは、プロジェクト計画作成において非常に重要な作業である。見積もられた工数を基にして、開発期間、開発人数、開発コスト等が計画される。工数を管理し、プロジェクトを円滑に遂行することにより、納期遅れやコスト超過といったプロジェクトの失敗を未然に防ぐことが可能となる。我々はこれまでに開発予測対象プロジェクトに対して複数の工数予測手法を同時に適用し、妥当な工数予測結果を推薦する工数予測ツール  $e^3$  を開発してきているが、より多くの予測手法の実装やユーザインタフェースの改良、より多くの適用評価が求められていた。そこで、本研究では、 $e^3$  に対して複数の工数予測手法を追加し、ユーザインタフェースの改良を行った。次に、6つのソフトウェア開発データセットを用いた追加実験を行い、これまでの評価結果と同じく、ツールが推薦する予測結果の有用性を確認した。更に、実用性の評価のため、商用見積もりツールである *KnowledgePLAN* との比較実験を行った。比較実験にあたっては、ある企業の63プロジェクトデータを用いた。実験の結果、 $e^3$  の予測精度は *KnowledgePLAN* の予測精度にわずかに劣っていたが、分析を通じて  $e^3$  の予測精度向上のため方向性が確認できた。

## 1. はじめに

ソフトウェア開発において、プロジェクトを円滑に遂行するためには、工数予測が重要である。工数とは、開発期間と要員の積で算出される延べ作業時間を表す数値であり、プロジェクトを管理する上で重要な指標となる。工数を正確に予測し、必要な工数を把握することによっ

て、スケジュールの適切な管理や開発資源の割り当てなどを行うことができる。結果として、納期遅れやコスト超過といったプロジェクトの失敗のリスクを抑えることが可能となる。そのため様々な工数予測手法が提案されている。その中でも、重回帰分析や類似性に基づく手法といった、過去に実施したプロジェクト(過去プロジェクト)の実績データに基づく手法がよく用いられている。実績データに基づく手法では、工数を予測したい進行中のプロジェクト(現行プロジェクト)で収集されたデータと過去プロジェクトの実績データから、現行プロジェクトの工数を予測する。しかしながら、実績データに基づく手法では、予測精度は実績データに大きく依存する[1]ため、1つの手法で高い精度を得ることは難しい。そこで、この問題に対処するために開発されたのが工数予測ツール「 $e^3$ 」である。本稿では、工数予測ツール  $e^3$  に対して行った機能拡張や実用性の評価のための追加実験について述べる。

## 2. 工数予測ツール $e^3$

### 2.1. 特徴

工数予測ツール「 $e^3$ 」プロジェクトの工数の予測値に対して期待される精度に基づいた、工数予測手法の選択を支援するツールである。1件の現行プロジェクトに対して、複数の工数予測手法を適用し、工数予測手法ごとに工数の予測値とその予測値に対して期待される精度を表示する。1度の実行で複数の工数予測手法の予測結果を確認することができ、予測値に対して期待される精度を比較することによって、特定の工数予測手法に依存しない柔軟な工数予測が可能となる。また、出力としては、各工数予測手法の現行プロジェクトの予測値と期待される精度が出力される。厳密に現行プロジェクトの工数の

予測値に対する精度ではないのは、工数予測時には現行プロジェクトの工数の実績値が未知であるため、予測値と実績値の誤差から精度を算出することができないためである。期待される精度の指標としては、ある過去プロジェクトの工数を他の過去プロジェクトを用いて、予測したときの相対誤差の平均値や中央値、相対誤差が25%以内のプロジェクトの割合を採用している。図1は「e<sup>3</sup>」の実際の動作の様子を表している。

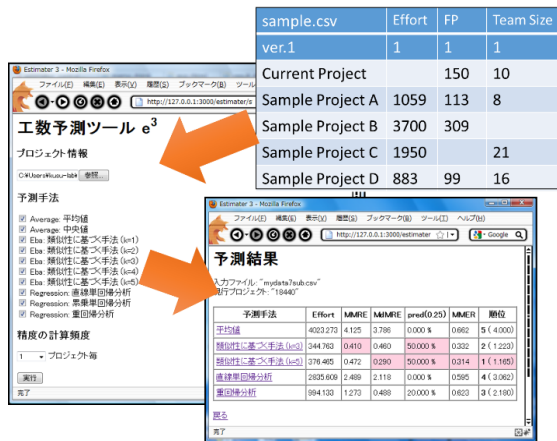


図 1. e<sup>3</sup> の動作

## 2.2. 実装されている工数予測手法

実装されている工数予測手法は、主に類似性に基づく手法と回帰分析に基づく手法の2種類がある。

### 1. 類似性に基づく手法 [2]

過去プロジェクトの中から現行プロジェクトと類似しているプロジェクト(類似プロジェクト)を探し、類似プロジェクトの工数の実績値を基に現行プロジェクトの工数を予測する手法である。これは、プロジェクト情報に含まれるメトリクス値の類似しているプロジェクト同士は工数の値も類似しているという仮説を前提としている。類似プロジェクトのみを用いて予測を行うため、過去プロジェクトの中に特異なプロジェクトが存在したとしても、そのプロジェクトの影響を受けずに予測が可能という長所がある。

### 2. 回帰分析に基づく手法

プロジェクトの工数を目的変数、その他のメトリクスを説明変数として、過去プロジェクト情報を基に回帰式を導出し、現行プロジェクトのメトリクスの値を導出した回帰式に代入することで工数の予測値を得る手法である。

## 2.3 これまでの知見

「e<sup>3</sup>」は文献 [3] で行われた3つのデータセットに対する実験によって、ツールが表示する期待される精度の高い工数予測手法をその都度選択することにより、各データセットにおいて、最も精度の高い手法と同程度の精度で予測を行うことが可能であると確認されている。また、使用するデータセットによって有効な手法も異なり、さまざまな組織から収集されたプロジェクトデータセットに対しては、類似性に基づいた手法が有効であり、同一の組織で収集されたプロジェクトデータセットに対しては、回帰分析を用いた手法が有効であった。

## 2.4 課題

「e<sup>3</sup>」への課題もいくつか指摘されている。一つは、ツールの機能不足の改善である。適用可能な工数予測手法の増大や改良したユーザーインターフェースの提供などが挙げられている。また、ツールの適用が行われたのが、文献 [2] の3つのデータセットに対する適用のみであり、ツールの適用事例が少ないという点も指摘されている。

## 3. e<sup>3</sup> の機能拡張

「e<sup>3</sup>」の機能拡張として、適用可能な工数予測手法を増やすために、類似性に基づく手法において、実装済みのものとは異なる2つの類似度の尺度を用いた手法 [4] とCOCOMO [5] モデルを追加した。実装されている類似性の評価手法は、以下の通りである。

- 調整コサイン類似性

プロジェクトデータを多次元ベクトルとみなし、ベクトルのなす角を類似度の尺度として用いる。ただし、こちらはメトリクスの分散を考慮している。プロジェクト  $p_a$  とプロジェクト  $p_b$  の類似度  $sim(p_a, p_b)$

は以下の式 1 で表される。

$$sim(p_a, p_b) = \frac{\sum_{j \in M_a \cap M_b} \{(v_{a_j'} - \bar{m}_{j'}) \times (v_{b_j'} - \bar{m}_{j'})\}}{\sqrt{\sum_{j \in M_a \cap M_b} (v_{a_j'} - \bar{m}_{j'})^2} \sqrt{\sum_{j \in M_a \cap M_b} (v_{b_j'} - \bar{m}_{j'})^2}} \quad (1)$$

$M_a$  と  $M_b$  はそれぞれ、プロジェクト  $p_a$  と  $p_b$  の欠損していないメトリクスの集合を表している。また  $\bar{m}_{j'}$  はメトリクス  $m_{j'}$  の正規化された特性値の平均値を表す。類似度  $sim(p_a, p_b)$  の値域は  $[-1.0, 1.0]$  である。

- 相関係数

プロジェクトデータを多次元ベクトルとみなし、ベクトルのなす角を類似度の尺度として用いる。ただし、こちらはプロジェクトの分散を考慮している。プロジェクト  $p_a$  とプロジェクト  $p_b$  の類似度  $sim(p_a, p_b)$  は以下の式 2 で表される。

$$sim(p_a, p_b) = \frac{\sum_{j \in M_a \cap M_b} \{(v_{a_j'} - \bar{p}_a) \times (v_{b_j'} - \bar{p}_b)\}}{\sqrt{\sum_{j \in M_a \cap M_b} (v_{a_j'} - \bar{p}_a)^2} \sqrt{\sum_{j \in M_a \cap M_b} (v_{b_j'} - \bar{p}_b)^2}} \quad (2)$$

$M_a$  と  $M_b$  はそれぞれ、プロジェクト  $p_a$  と  $p_b$  の欠損していないメトリクスの集合を表している。また  $\bar{p}_a$ ,  $\bar{p}_b$  はプロジェクト  $p_a$ ,  $p_b$  の正規化された特性値の平均値を表す。類似度  $sim(p_a, p_b)$  の値域は  $[-1.0, 1.0]$  である。

また、ユーザーインターフェースの改良として、入力データのプロジェクトリストから予測に使用しないプロジェクトや、プロジェクトのデータを取り除く機能と、入力データ上で欠損しているプロジェクトのデータを自動補完する方法 [6, 7] を選択できる機能を追加した。具体的には、下記の自動補完方法を選択できるようにした。

- 平均値挿入法

欠損しているプロジェクトデータを全プロジェクトでのそのデータの平均値で補完する

- リストワイズ法

欠損しているプロジェクトデータを持つプロジェクトを除去する。

- K-nn 法

欠損しているプロジェクトに類似しているプロジェクトから、欠損しているメトリクスの値の平均値を算出し、欠損値を補完する。類似度には、ユークリッド距離の逆数を用いる。

- CF 応用法

欠損しているプロジェクトに類似しているプロジェクトから、欠損しているメトリクスの値の加重平均値を算出し、欠損値を補完する。類似度には、調整コサイン類似性を用いる。

以下の図 2 は「 $e^3$ 」の拡張した入力画面である。

## 4. 評価実験

「 $e^3$ 」の評価実験について述べる。評価実験としては、過去に得られた知見と同じ結果が、別のデータセットにおいても得られるかどうかを確認するための実験（評価実験 1）と、「 $e^3$ 」の工数予測ツールとしての実用性を評価するための実験（評価実験 2）を行った。

### 4.1. 評価実験 1

この実験の目的は、ツールが表示する期待される精度の高い工数予測手法をその都度選択することにより、各データセットにおいて、最も精度の高い手法と同程度の精度で予測を行うことが可能であることの確認である。適用対象は PROMISE[8] という見積り研究の Web サイトで公開されている以下の 6 つのデータセットである。

- Desharnais Software Cost Estimation

Desharnais によって収集されたカナダのソフトウェア開発企業における 80 年代のデータセットである

- CM1

NASA のデータを収集し、処理する、宇宙船の計器の開発のデータセット

- JM1

NASA の予測を生成するために、シミュレーションを用いる、実時間地上システムの開発のデータセット

- KC1

NASA の地上データを受信して処理するためのストレージ管理を実装しているシステムのデータセット

- KC2

NASA の KC1 と同じプロジェクトの別の部分のデータセット

- PC1

NASA の地球周回軌道衛星のためのフライトソフトウェアのデータセット

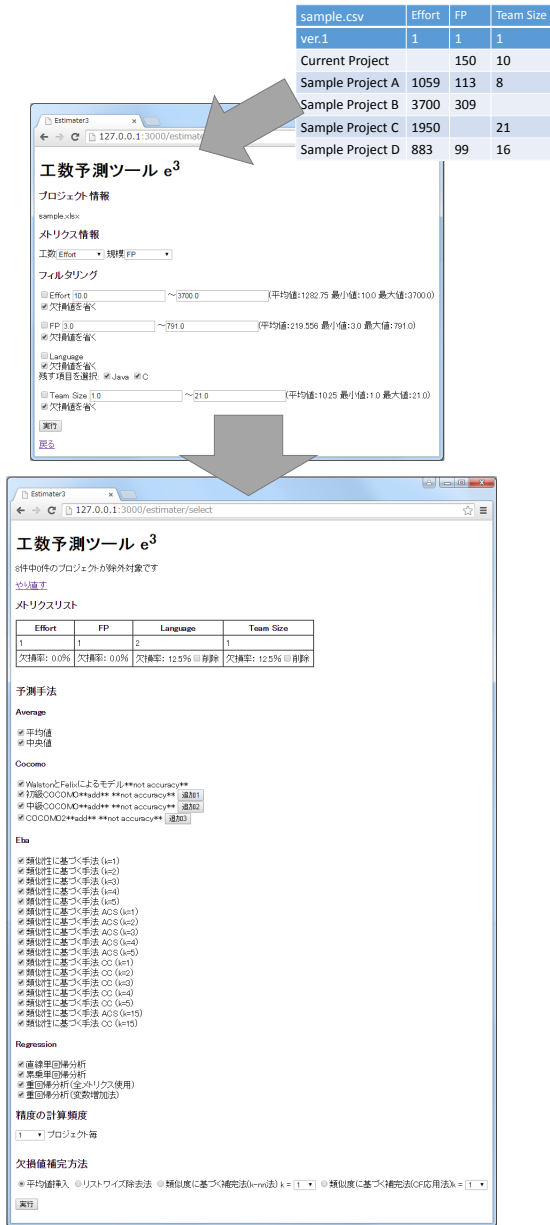


図 2. e<sup>3</sup> の入力画面

評価方法は、ツールが表示する期待される精度の最も高い工数予測手法をその都度選択して使用した場合と、常に同じ工数予測手法を使用し続けた場合とで現行プロジェクトの工数の実測値と予測値の誤差を比較する。実験の結果、適用した全てのデータセットにおいて、ツールが表示する期待される精度の最も高い工数予測手法をその都度選択して使用した場合、それぞれのデータセットで有効な工数予測手法と同程度の誤差で予測が行えることが確認できた。

図 3 に、Dersharnais に適用した結果を示す。この図からデータセットに有効な手法を使い続けた場合の誤差の平均が 0.378、ツールの推薦する手法を選び続けた場合の誤差の平均が 0.382 となっており、同程度の誤差で予測可能であることが確認できる。

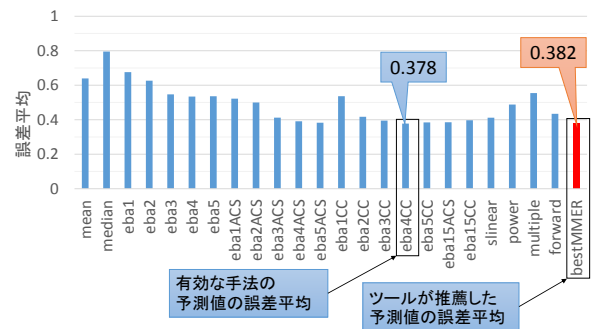


図 3. 評価実験 1 の結果例 (Dersharnais)

## 4.2. 評価実験 2

評価実験 2 として、「e<sup>3</sup>」の工数予測ツールとしての実用性を評価するために、某企業が利用している見積もりツール「KnowledgePLAN」との性能比較実験を行った。

「KnowledgePLAN」はSPR社の開発したソフトウェアプロジェクトの計画を支援するツールである。1万件を超えるプロジェクト実績データを搭載した知識データベースを基に、プロジェクトの所要工数や期間、品質、コスト、成果物の規模などを予測できる。実験には、某企業の63件のプロジェクトを用いた。評価方法は63件のプロジェクトに対して「 $e^3$ 」と「KnowledgePLAN」を適用し、工数の予測値と実測値を比較する。「 $e^3$ 」には、各プロジェクトの工数の予測値を求める際に、入力データとして、予測値を求めるプロジェクト以外の62件のプロジェクトの情報と、予測値を求めるプロジェクトの工数の実測値以外の情報を与えた。「KnowledgePLAN」は、某企業で使用されていた方法を用いて、工数の予測値を求めた。実験の結果、工数の予測値と実測値の誤差をMER (Magnitude of Error Relative)[9]の尺度で評価し、その平均を比べると、「KnowledgePLAN」が0.282、「 $e^3$ 」が0.325となり、「KnowledgePLAN」の方が予測精度は良いという結果になった、しかし、いくつかのプロジェクトに対しては「 $e^3$ 」の方が予測精度がよいということも分かった。結果を以下の表1に示す。しかし、いくつかのプロジェクトに対しては「 $e^3$ 」の方が予測精度がよいということも分かった。

表 1. 実験結果

ツール	平均 MER
$e^3$	0.325
KnowledgePLAN	0.282

#### 4.3. 考察

評価実験2の結果、「 $e^3$ 」の方が予測精度の良いプロジェクトの特徴についてであるが、これらのプロジェクトについては、「 $e^3$ 」の類似性に基づく予測手法が有効に働いていた。今回のデータセットは某企業で収集された、ドメインの限定されているものであるため、過去のこのツールについての知見に基づけば、回帰分析などが有効に働くとされている。そのため、「KnowledgePLAN」の1万件を超える実績データに基づく予測が有効に働いていると考えられる。しかし、プロジェクトデータの中に、ドメインが限定されているものの一部特異なプロジェクトが少数存在し、それらについては「 $e^3$ 」の類似性に基づく手法が有効に働いた可能性が考えられる。

## 5. まとめ

工数予測ツール「 $e^3$ 」の機能拡張と実用性の評価実験について述べた。機能拡張としては、工数見積手法の追加と欠損値補完機能を実装した。また、評価実験としては、文献[8]の6種類のデータセットを用いて、既に得られている知見を確認した。更に、実用性の評価として、KnowledgePLANとの予測精度の比較を行った。今後は、より多くのプロジェクトデータに対する適用実験や実用性の改善に取り組む予定である。

## 6. 謝辞

本研究は、一部日本学術振興会科学研究費補助金基盤研究(S)(課題番号：25220003)の助成を得て行われた。

## 参考文献

- [1] E. Mendes et al. Further comparison of cross-company and within-company effort estimation models for web applications. In *10th IEEE International Symposium on Software Metrics*, pp. 348–357, 2004.
- [2] 角田雅照, 大杉直樹, 門田暁人, 松本健一, 佐藤慎一. 協調フィルタリングを用いたソフトウェア開発工数予測方法. *情報処理学会論文誌*, Vol. 46, No. 5, pp. 1155–1164, May 2005.
- [3] 生方克馬, 柿元健, 楠本真二. 期待される精度の比較による適切な工数予測手法の判別を支援するツール. *電子情報通信学会論文誌 D*, Vol. J95-D, No. 4, pp. 885–894, Apr 2012.
- [4] 大杉直樹, 角田雅照, 門田暁人, 松村知子, 松本健一, 菊池奈穂美. 企業横断的収集データに基づくソフトウェア開発プロジェクトの工数見積もり. *SEC journal*, Vol. 5, No. 1, pp. 16–25, 2006.
- [5] B.W.Boehm, C.Abts, A.W.Brown, S. Chulani, B. K. Clark, E. Horowitz, R. Madachy, D. J. Reifer, and B. Steece. *Software Cost Estimation with Cocomo II*. Prentice Hall, 2000.
- [6] 田村晃一, 柿元健, 戸田航史, 角田雅照, 門田暁人, 松本健一, 大杉直樹. 工数予測における類似性に基づく

欠損値補完法の実験的評価. コンピュータソフトウェア, Vol. 26, No. 3, pp. 44–55, 2009.

- [7] 田村晃一, 門田暁人, 松本健一. 欠損値処理法を用いた工数予測におけるデータ件数と予測精度の関係. コンピュータソフトウェア, Vol. 27, No. 2, pp. 100–105, 2010.
- [8] PROMISE software engineering repository. <http://promise.site.uottawa.ca/SERepository/>.
- [9] T.Foss, E.Stensrud, B.Kitchenham, and I.Myrtveit. A simulation study of the model evaluation criterion mmre. *IEEE Transactions on Software Engineering*, Vol. 29, No. 11, pp. 985–995, Nov 2003.