ソフトウェア開発工数予測における特異プロジェクト除去の効果

The Effect of Removing Noisy Projects in Software Development Effort Estimation

渡邊 瑞穂 † 柿元 健 † 戸田 航史 † 門田 暁人 † 松本 健一 †
Mizuho WATANABE, Takeshi KAKIMOTO, Koji TODA, Akito MONDEN, Ken-ichi MATSUMOTO

† 奈良先端科学技術大学院大学 情報科学研究科
Graduate School of Information Science, Nara Institute of Science and Technology
{mizuho-w,takesi-k,koji-to,akito-m,matumoto}@is.naist.jp

従来，ソフトウェア開発プロジェクトのメトリクス（規模，開発期間など）を説明変数とし，開発工数を目的変数とする工数予測モデルがソフトウェア開発の分野で広く用いられている．ただし，モデル構築に用いるデータセットに特異なプロジェクトが含まれる場合，精度の良いモデルが構築できないことが課題であった．そこで，本稿では，開発規模と開発工数，開発規模と開発期間の関係に着目し，特異なプロジェクトを除去する手法を提案する．ソフトウェア開発企業で収集されたデータを用いた評価実験の結果，提案手法によって特異なプロジェクトを除去することで精度が改善された．

1 はじめに

ソフトウェア開発プロジェクトにおける工数予測は，プロジェクト完遂に必要な資源，及びスケジュール管理を行う上で重要である．必要な工数を過不足なく予測することで，納期遅れ，コスト超過といったプロジェクトの失敗を防ぐことができる．そのため，工数予測に関する数多くの研究が行われている [6] .

例えば，COCOMO[1] や SLIM[7] などの手法では，ソフトウェアのソースコード行数を手法発案者が定義した数式モデルに入力することで，開発工数，開発期間，要員数，生産性を予測する．しかし，これらの手法ではモデル定義の際に組織やプロジェクトに固有の特徴（要因のスキル，顧客の業務分野，顧客との信頼関係など）が考慮されず，予測精度が低いことがある [2].

一方，過去に行われたプロジェクトの実績データを用いて，重回帰分析やニューラルネット [8] などにより組織独自の工数予測モデルの構築も行われている．ここでいう実績データとは，プロジェクト毎に，当該プロジェクトの特徴を表す変数 (工数，開発規模，検出バグ数など) の値を記録，蓄積したものである．この方法により構築されたモデルは組織固有の特徴を反映しているため，COCOMO よりも高い予測精度を得られることが期待される．

しかし，実績データに基づくモデル構築は，実績データ中に含まれる特異なプロジェクトによって，性能の良いモデルが得られない場合がある．ここでいう特異なプロジェクトとは，例えば，設計をやり直したために開発工数が普段の倍以上かかった，ビジネス上の都合のため通常必要とされる半分の期間で開発した，といった「普通でない」プロジェクトのことである．

このような特異なプロジェクトは，実績データから予め除去しておくことが望ましいが，開発当時のプロジェクトの実態（設計をやり直した，など）を知ることは必ずしも容易でない．

そこで，本稿では，プロジェクト実績データに含まれる数値のみを用いて，特異なプロジェクトを特定・除去する 2 つの手法を提案する．一つ目の手法は，開発規模に比べて開発工数が著しく小さいもしくは大きいプロジェクトを除去する手法であり，もう一つは，開発規模に比べて開発期間が著しく短いもしくは長いプロジェクトを除去する手法である．

関連研究として，データセットに含まれるメトリクス値の大小に注目して外れ値（集団から外れた値）を検出する手法が提案されている [5]．しかし，ソフトウェア開発プロジェクトの工数予測においては，開発規模が大きい，開発工数が大きいなどの，個別の値が大きいこと自体は必ずしも特異であるとはいえ

ない．そこで提案手法では，開発規模と開発工数，開発規模と開発期間といった 2 つの値の関係を考慮して特異なプロジェクトを検出する．

さらに，提案手法の有効性を確認するため，ソフトウェア開発企業で収集された実績データを用いた評価実験についても報告する．評価実験では，特異なプロジェクトの除去前と除去後の予測精度を比較し，精度改善の程度を観察した．

以降，2 章では，本稿の実験で用いた工数予測手法であるステップワイズ対数重回帰分析について述べ，3 章では，提案手法である特異なプロジェクトを除去する手法について述べる．4 章では，提案手法の有効性を示すための評価実験の方法として，用いたデータ，実験手順，評価指標を説明するについて説明する．5 章で評価実験の結果について述べ，6 章で結果について考察する．最後に 7 章でまとめと今後の課題について述べる．

2 ステップワイズ対数重回帰分析

本稿では，工数予測手法としてステップワイズ対数重回帰分析を用いる．対数重回帰分析は，多変量解析の一手法である重回帰分析を拡張した見積もり手法であり，ソフトウェア開発に要する工数を予測するために広く用いられている．重回帰分析では，予測対象の変数（目的変数）と，目的変数に影響を与える複数の変数（説明変数）との関係を表した一次式（回帰式）を作成する．回帰式中の各係数と定数は，予測値の絶対誤差（残差）の乗和が最小になるように決定される．作成された回帰式に，現行プロジェクトで計測した説明変数を与えることで，目的変数を予測することが可能となる．対数重回帰分析は，対数変換したデータから回帰式を作成する手法である．(回帰式は，一次式ではなく対数式となる.) 作成された回帰式に対して，対数変化した説明変数を入力し，得られた値を対数逆変換することで予測値が得られる．ステップワイズ対数重回帰分析は，ステップワイズ変数選択法により採用する変数を決定し，対数重回帰分析を行う手法である．ステップワイズ変数選択は次の手順で行われる．

手順 1. 変数を全く含まないモデルを初期モデルとして作成する．

手順 2. 作成されたモデルに対して，各説明変数の係数が 0 でないかの検定を行い，指定した有意水準（本稿の評価実験では，偏 F 値の有意水準を $p_{in} = 0.05$，$p_{out} = 0.1$ とした）で棄却されない場合に変数を採択する．ただし，多重共線性を回避するために，採択する変数の分散拡大要因（VIF）が一定値以上の場合，またはその変数を採択することによって，他の変数の VIF が一定値以上となる場合，その変数は採択しない．

手順 3. 検定により適切な変数が選択されたと判断されるまで手順 2 を繰り返す．

3 提案手法

3.1 概要

本章では，データセットに含まれるメトリクス値に基づいて特異なプロジェクトを除去する手法を提案する．従来の外れ値除去法では，集団から外れている個体を除去する [5]．一方，提案手法では，メトリクス間の関係（開発規模と開発工数，開発規模と開発期間）が，他のプロジェクトと比べて特異な値をとるプロジェクトを特異なプロジェクトとして予測モデル構築用のデータから除去する．

3.2 手法１：開発工数に関する特異プロジェクトの除去

一つ目の手法は，プロジェクトの開発規模と開発工数の関係，すなわちプロジェクトの生産性に基づいて特異なプロジェクトを除去する．開発規模と開発工数の間には，ほぼ一次の相関があると一般に言われている．そこで，提案手法では，開発工数と開発規模の関係から最小二乗法により一次近似式を算出し，算出された一次近似式から大きく外れたプロジェクトを特異なプロジェクトとして除去する．以下にその具体的な手順を示す．

手順 1. 全てのモデル構築用データの開発規模と，同プロジェクトの開発工数から最小二乗法により一次近似式を算出する．開発規模 S を横軸に，開発工数 E を縦軸にとったとき，モデル構築用データは図 1 のように表され，一次近似式は式 (1) で表される．

$$E = aS + b \tag{1}$$

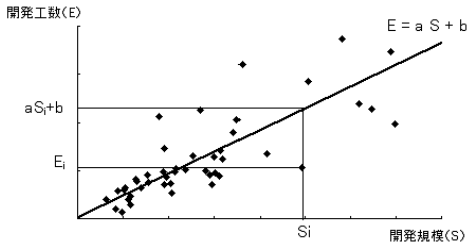

図 1: 開発規模と開発工数の近似直線

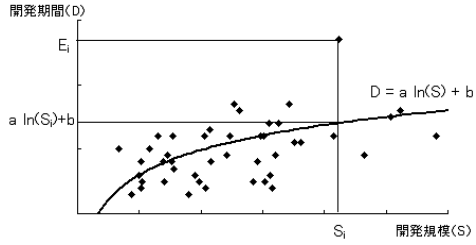

図 2: 開発規模と開発期間の対数近似曲線

$$a = \frac{n\sum_{i=1}^{n} S_i E_i - \sum_{i=1}^{n} S_i \sum_{i=1}^{n} E_i}{n\sum_{i=1}^{n} S_i^2 - \left\{\sum_{i=1}^{n} S_i\right\}^2}$$

$$b = \frac{\sum_{i=1}^{n} S_i^2 \sum_{i=1}^{n} E_i - \sum_{i=1}^{n} S_i E_i \sum_{i=1}^{n} S_i}{n\sum_{i=1}^{n} S_i^2 - \left\{\sum_{i=1}^{n} S_i\right\}^2}$$

ここで，n はプロジェクト件数，a は一次近似式の傾き，b は一次近似式の切片，S_i はプロジェクト i の開発規模，E_i はプロジェクト i の開発工数である．

手順 2. 算出された一次近似式と各プロジェクトとの縦軸方向（開発工数）に関する相対距離を算出する．プロジェクト i と一次近似式との相対距離 d_i は次式（2）で表される．

$$d_i = \frac{|E_i - aS_i - b|}{E_i} \tag{2}$$

手順 3. 算出された相対距離が大きなプロジェクトを特異なプロジェクトとみなし，モデル構築用データから除去する．

3.3 手法２：開発期間に関する特異プロジェクトの除去

二つ目の手法は，プロジェクトの開発規模と開発期間の関係に基づいて特異なプロジェクトを除去する．一般に開発期間は開発工数と三乗根の関係があり [1][7]，3.2 で述べたように開発工数と開発規模の間にはほぼ一次の相関があると言われている．提案手法では開発規模と開発期間の三乗根の関係を対数曲線で近似し，開発規模と開発期間の関係から対数近似式を算出し，算出された対数近似式から大きく外れたプロジェクトを特異なプロジェクトとして除去する．

以下にその具体的な手順を示す．

手順 1. 全てのモデル構築用データの開発規模と，同プロジェクトの開発期間から最小二乗法により対数近似式を算出する．開発規模 S を横軸に，開発工数 E を縦軸にとったとき，モデル構築用データは図 2 のように表され，対数近似式は式 (3) で表される．

$$D = a\ln(S) + b \tag{3}$$

$$a = \frac{n\sum_{i=1}^{n} \ln(S_i) D_i - \sum_{i=1}^{n} \ln(S_i) \sum_{i=1}^{n} D_i}{n\sum_{i=1}^{n} \ln(S_i^2) - \left\{\sum_{i=1}^{n} \ln(S_i)\right\}^2}$$

$$b = \frac{\sum_{i=1}^{n} \ln(S_i^2) \sum_{i=1}^{n} D_i - \sum_{i=1}^{n} \ln(S_i) D_i \sum_{i=1}^{n} \ln(S_i)}{n\sum_{i=1}^{n} \ln(S_i^2) - \left\{\sum_{i=1}^{n} \ln(S_i)\right\}^2}$$

ここで，n はプロジェクト件数，a は対数近似式の係数，b は対数近似式の切片，S_i はプロジェクト i の開発規模，E_i はプロジェクト i の開発工数である．

手順 2. 算出された対数近似式と各プロジェクトとの距離 d_i を算出する．プロジェクト i と対数近似式との距離は次式（4）のように表される．

$$d_i = \frac{|D_i - a\ln(S_i) - b|}{D_i} \tag{4}$$

手順 3. 算出された相対距離が大きなプロジェクトを特異なプロジェクトとみなし，モデル構築用データから除去する．

表 1: 実験に用いたデータセット

変数	平均値	中央値	最大値	最小値	開発工数との相関係数
開発期間	11.30	10.00	36.00	1.00	0.65
トランザクション数	177.47	134.00	886.00	9.00	0.58
開発規模 (調整済 FP)	282.39	247.00	1116.00	62.00	0.73
エンティティ数	120.55	96.00	387.00	7.00	0.50
開発工数	4833.91	3542.00	2394.00	546.00	—

4 評価実験

4.1 目的

提案手法の効果を確認するために評価実験を行った．評価実験では，除去するプロジェクト数を変化させて予測精度の推移も確かめた．

4.2 データセット

評価実験に用いたデータセットは，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 年代のデータである [3] [4]．データセットには 77 件のプロジェクトについて，5 種類の変数が記録されている．データに欠損は含まれない．データの詳細を表 1 に示す．表中の開発期間からエンティティまでを説明変数とし，開発工数を目的変数として予測した．評価実験では，予測を行う工程として，開発対象ソフトウェアのファンクションポイントの計測が終了した時点，すなわち，概要設計もしくは基本設計の完了時を想定している．また，開発期間は，開発初期に予め決定されていることを想定している．

4.3 実験手順

実験の手順は次の通りである．

手順 1. データセットを無作為に 47 件のモデル構築用データ，30 件の評価用データに分割し，これらのペアを 5 組作成した．評価用データのプロジェクトは総工数が未知と仮定され，工数予測の対象となる．

手順 2. モデル構築用データから 2 つの提案手法それぞれで特異なプロジェクトの除去を行った．5 個のモデル構築用データそれぞれについて，除去するプロジェクト数を 1 件から 1 件刻みで除去し，モデル構築用データを除去プロジェクト件数ごとに新たに作成した．

手順 3. 特異なプロジェクトを除去したモデル構築用データを用いてステップワイズ対数重回帰モデルを構築した．

手順 4. 構築したモデルを用いて評価用データの各プロジェクトの開発工数の予測値を算出した．

手順 5. 各プロジェクトについて予測値の絶対誤差平均，絶対誤差中央値，相対誤差平均値，相対誤差中央値を算出した．

4.4 評価指標

評価実験で予測性能の比較に絶対誤差平均値 (MMAE)，絶対誤差中央値 (MdMAE)，相対誤差平均値 (MMRE)，相対誤差中央値 (MdMRE) の 4 種類の精度評価指標を用いた．

それぞれの評価指標は次の式 (5)〜(8) で計算される．ここで，M 件のプロジェクトがあるとする．また，実測値と予測値をそれぞれ X_i，$\hat{X}_i (i = 1 \sim M)$ とし，$A_i = \left| \hat{X}_i - X_i \right|$, $R_i = \frac{\left| \hat{X}_i - X_i \right|}{X_i}$ とおく．

絶対誤差平均値 (MMAE)

$$MMAE = \frac{\sum_{i=1}^{M} A_i}{M} \tag{5}$$

絶対誤差中央値 (MdMAE)

$$MdMAE = \begin{cases} A_n & M = \text{奇数} \\ (A_1 \leq A_2 \leq \cdots \leq A_n \leq \cdots \leq A_{2n-1}) & \\ \dfrac{A_n + A_{n+1}}{2} & M = \text{偶数} \\ (A_1 \leq \cdots \leq A_n \leq A_{n+1} \leq \cdots \leq A_{2n}) & \end{cases} \quad (6)$$

相対誤差平均値 (MMRE)

$$MMRE = \frac{\sum_{i=1}^{M} R_i}{M} \quad (7)$$

相対誤差中央値 (MdMRE)

$$MdMRE = \begin{cases} R_n & M = \text{奇数} \\ (R_1 \leq R_2 \leq \cdots \leq R_n \leq \cdots \leq R_{2n-1}) & \\ \dfrac{R_n + R_{n+1}}{2} & M = \text{偶数} \\ (R_1 \leq \cdots \leq R_n \leq R_{n+1} \leq \cdots \leq R_{2n}) & \end{cases} \quad (8)$$

5 結果

除去するプロジェクト数を変化させて提案手法で特異なプロジェクトを除去し，ステップワイズ対数重回帰分析で予測した時の各評価指標による予測精度のグラフを図 3 に示す．各グラフの横軸は除去プロジェクト数，縦軸はそれぞれの評価指標による予測精度である．除去プロジェクト数 0 件は，特異なプロジェクトの除去を行わない場合の結果となる．図 3 より，いずれの除去手法においても，除去するプロジェクトの数を適切に選ぶことにより，特異なプロジェクトの除去を行わない場合よりも，予測精度が（全ての評価指標において）向上することが確認できる．

ただし，除去するプロジェクトの数によっては，逆に精度が低下することも明らかとなった．特に，除去プロジェクト数を大きくした場合（例えば 15 件以上），ほとんどの評価指標において精度が低下した．

4 つの評価指標（絶対誤差平均値，相対誤差平均値，絶対誤差中央値，相対誤差中央値）を比較すると，絶対誤差平均値と相対誤差中央値については，大きな改善が見られたが，残りの 2 つの指標については，改善が小さいもしくは逆に悪化した．

開発工数に基づいた手法 1 と開発期間に基づいた手法 2 を比較すると，除去プロジェクト数が小さい場合（1〜2 件），手法 1 の方が効果が大きいもしくは手法 2 と同程度であった．ただし，除去プロジェクト数を 9 件にした場合，全ての評価指標で手法 2 の方が効果が大きかった．

各評価指標でそれぞれの除去手法の予測精度が最も高くなった時の除去プロジェクト件数と向上した予測精度の値を表 2 に示す．それぞれの手法および評価指標について，最も精度が向上した除去プロジェクトの数は一定ではない．ただし，それぞれの手法で得られた最高精度の値には大きな差はない．

6 考察

評価実験の結果，いずれの特異なプロジェクト除去手法においても，除去するプロジェクトの数を適切に選ぶことで，予測精度の向上が見られた．ただし，各手法および評価指標ごとに，最も精度が向上した除去プロジェクトの数は一定ではなく，どのようにして除去プロジェクト数を決定するかが重要となることが分かった．

手法 1 については，1 件および 2 件のプロジェクトを除去した場合に，全ての評価指標で改善が見られたことから，全プロジェクト（47 件）の約 4%（2 件）程度を除去するというのが一つの指標となりそうである．ただし，今回の評価実験では一つのデータセットでしか評価していないため，今後，他のデータセットによる実験も行うことが必要である．

手法 2 については，除去プロジェクト数の増加に対する予測精度の変化のばらつきが大きく，現時点では適切な除去プロジェクト数を選ぶことは難しいと考えられる．

7 まとめ

本稿では，特異なプロジェクト除去として，ソフトウェア開発データにおいて特有のメトリクス間の関係を用いて特異なプロジェクトを決定する手法を

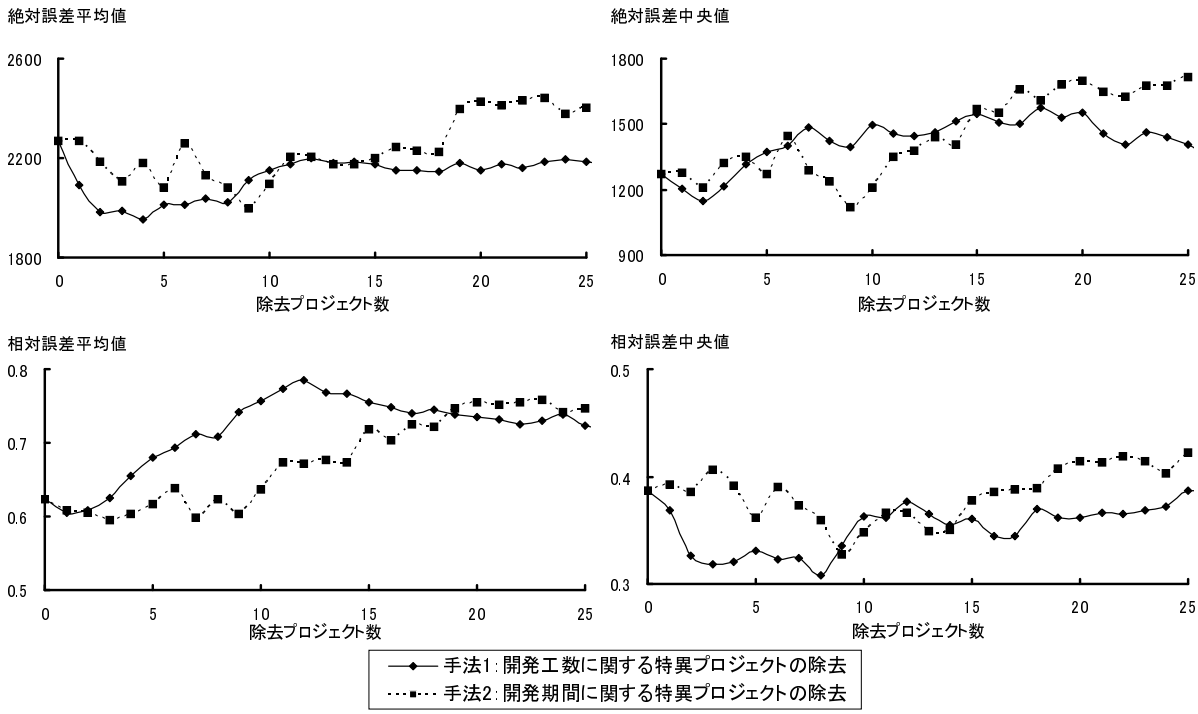

図 3: 除去プロジェクト数を変化させたときの予測精度

表 2: 最も精度が向上した時の除去数と向上値

	手法１（開発工数）		手法２（開発期間）	
	除去数	予測精度向上値	除去数	予測精度向上値
絶対誤差平均値	4	319	9	272
絶対誤差中央値	2	121	9	151
相対誤差平均値	1	0.018	3	0.027
相対誤差中央値	8	0.079	9	0.060

提案し，その効果を実験的に示した．ソフトウェア開発企業で収集された実績データを用いた評価実験の結果，提案手法により特異なプロジェクトを除去することで，特異なプロジェクトを除去しない場合よりも，絶対誤差平均値で 291，相対誤差平均値で 0.027，予測精度が改善された．

今後は，実験結果の信頼性を高めるため，さらに多くの実績データを用いて本稿と同様の実験を行う予定である．また，開発規模，開発工数，開発期間以外のメトリクスに着目した特異なプロジェクト除去手法についても検討してく予定である．

謝辞

本研究の一部は，文部科学省「e-Society 基盤ソフトウェアの総合開発」の委託に基づいて行われた．

参考文献

[1] B. Boehm, Software Engineering Economics, Prentice Hall, 1981.

[2] S. Chulani, B. Boehm, and B. Steece, “ Bayesian Analysis of Empirical Software Engineering Cost Models, ” IEEE Trans. on Software Eng., Vol.25, No.4, pp.573-583, 1999.

[3] J. M. Desharnais, “ Analyse statistique de la productivitie des projets informatique a partie de la technique des point des function, ”Masters Thesis, University of Montreal, 1989.

[4] C. Mair, G. Kadoda, M. Lefley, K. Phalp, C. Schofield, M. Shepperd, and S. Webster, “ An Investigation of Machine Learning Based Prediction Systems, ” J. Systems and Software, Vol.53, Issue 1, pp.23-29, 2000.

[5] 柗本 真佑, 亀井 靖高, 門田 暁人, 松本 健一, “ Fault-Prone モジュール判別モデルに対する外れ値除去法の適用効果 ”, 情報処理学会研究報告, ソフトウェア工学, Vol.2007-SE-155, No.33, pp.49-56, March 2007.

[6] Project Management Institute, A Guide To The Project Management Body Of Knowledge (PMBOK Guides), Project Management Institute, 2004.

[7] L. H. Putnam, “ A General Empirical Solution to the Macro Sizing and Estimating Probrem, ”IEEE Trans. on Software Eng., Vol.4, No.4, pp.345-361, 1978.

[8] K. Srinivasan, and D. Fisher, “ Machine Learning Approaches to Estimating Software Development Effort, ” IEEE Trans. on Software Eng., Vol.21, No.2, pp.126-137, 1995.