

規模あたりの要員数を考慮したソフトウェア開発工数見積モデル

角田雅照[†] 門田暁人[†] 松本健一[†]

奈良先端科学技術大学院大学 情報科学研究科[†]

1. はじめに

ソフトウェア開発プロジェクトにおいて、プロジェクト計画を立案するためには、ソフトウェアの開発工数を見積もる必要があり、そのためにソフトウェア開発工数見積モデルが利用される。ソフトウェア開発工数見積モデルは、開発工数（目的変数）と開発規模など（説明変数）との関係が数学的に表わされたものである。

本稿の目的は、ソフトウェア開発工数見積モデルの精度を高めることである。工数は開発規模に基づいて求められ、生産性は開発規模あたりに必要な工数を示すことから、生産性に強い影響を持つ変数をモデルの説明変数に加えることにより、見積精度の向上が期待される。

以降、まず生産性の分析を行い、規模あたりの開発要員数が生産性に強い影響を持つことを明らかにする。そして、規模あたりの開発要員数を名義尺度に変換し、説明変数として工数見積モデルに用いた場合の見積精度を確認する。

2. 分析対象データ

分析に用いたデータは、財団法人経済調査会によって平成 18 年度に 110 社から収集された、153 プロジェクトの実績値である。プロジェクトごとに、約 140 個の変数が記録されている。ただし、変数によっては欠損値（値が未記録であること）が多く含まれている場合があった。

分析対象のプロジェクトの条件を揃えるため、下記の条件に従ってプロジェクトを抽出した。

- 新規開発である。
- 設計、製造、試験の全工程が実施されている。
- 実績 FP（ファンクションポイント）、開発工数が計測されている（生産性を計算するために必要）。

抽出されたプロジェクトは 43 件であった。プ

ロジェクトのほとんどは事務系のシステムであり、ウォーターフォールで開発が行われていた。

分析対象とした変数を以下に示す。本稿では、各変数をグループ分けし、名称を付けた。

- グループ A: 名義尺度
 - 主開發言語（使用比率が 50% 以上のもの）、システム構成（メインフレームなど）、適用分野（事務系など）、適用業種、クライアント OS、サーバ OS、データベース、開発技法（オブジェクト指向など）、開発作業標準の有無、開発作業標準入手方法（市販品など）
- グループ B: 間隔尺度（5 段階評価）
 - 機能性要求、信頼性要求、開発スケジュール要求、要件の明確度/安定度、ユーザの参画割合、類似システムやライブラリの再利用、プロジェクト管理の経験と能力、アナリストの経験と能力、SE/PG の経験と能力
- グループ C: 比例尺度
 - FP、開発期間、平均要員数（開発工数/開発期間）、開発要員/FP、開発期間/FP

本稿では、FP を開発工数で割った値を生産性と定義した。また、データに含まれていない平均要員数、平均要員数/FP、開発期間/FP を新たに定義した。平均要員数は開発工数を開発期間で除したものである。一般に、開発要員数と開発期間は FP に比例して大きくなるため、開発規模（FP）で標準化した平均要員数/FP と開発期間/FP を定義した。

3. 生産性の分析

生産性に対して影響の強い変数を特定するために、一元配置分散分析を行い、調整済み寄与率を求めた。調整済み寄与率はおおむね 0 から 1 の値をとり、1 に近いほど関連が強いことを示す。分析において全ての変数を名義尺度として扱うため、比例尺度のプロジェクト特性を名義尺度に変換した。各特性の第 1 四分位数以下の値を「低位」カテゴリ、第 3 四分位数以上の値を「高位」カテゴリ、残りの値を「中位」カテゴ

Software Effort Estimation Model Using Team Size per System Size

Masateru Tsunoda[†], Akito Monden[†], and Ken-ichi Matsumoto[†]

[†]Graduate School of Information Science, Nara Institute of Science and Technology

8916-5, Takayama, Ikoma, Nara, JAPAN 630-0192

{masate-t, akito-m, matumoto}@is.naist.jp

りとした。

一元配置分散分析の結果、有意水準 5%で有意な関連が見られた変数は、平均要員数/FP、平均要員数、開発期間/FP であり、寄与率はそれぞれ 0.38, 0.31, 0.19 であった。これまで、開発要員数が生産性に強い影響を与えていることを示した研究がいくつか存在するが[1][2][4]、これらの研究では規模あたりの開発要員数の影響は分析していない。本稿の分析結果は、開発要員数よりも規模あたりの開発要員数のほうが、より生産性に強い影響を与えることを示している。

4. 工数見積モデルの構築

本稿では規模あたりの開発要員数を表す名義尺度の変数を説明変数として用いることを提案する。3章の分析結果より、規模あたりの開発要員数が最も生産性に強い影響を与えていることがわかった。この変数を工数見積モデルの説明変数に加えることにより、見積精度の向上が期待される。ただし、プロジェクトの初期には開発要員数が確定していないことが多いため、モデルの説明変数に規模あたりの開発要員数をそのまま含めると、プロジェクトの初期にモデルを用いることができなくなってしまう。規模あたりの開発要員数を名義尺度に変換して用いることにより、開発要員数が概算のままでもモデルを利用することが可能となる。これまで、(最大) 要員数をそのまま説明変数に用いたモデルしか提案されていない[3]。

名義尺度化した平均要員数/FP を説明変数として用いたモデルと用いないモデルを作成し、見積精度を比較した。目的変数は開発工数であり、平均要員数/FP 以外の説明変数の候補は、プロジェクトの初期に確定していると考えられる変数(2章のグループ A, グループ B, FP) である。Q-Q プロットを描くと FP と開発工数が正規分布していなかったため、それぞれ対数変換を行った。また、FP を説明変数とした単回帰モデルを作成し、Cook の D が 2.73 となったプロジェクト 1 件を外れプロジェクトとみなし、これを除外した。

モデルはステップワイズ重回帰分析により作成した。実験ではデータを学習データとテストデータにランダムに 2 分割することを 20 回繰り返して見積精度を調べた。モデルの精度の評価基準として、一般的に用いられる絶対誤差(見積値-実測値)、相対誤差(絶対誤差/実測値)、MER(絶対誤差/予測値)を用いた。これらの値が小さいほど、モデルの精度が高いことを示す。

予備実験を行い、説明変数の候補をグループ

表 1 モデルの見積精度

	絶対誤差 中央値	相対誤差 中央値	MER 中央値
平均要員数/FP なし	30.5	0.54	0.55
平均要員数/FP あり	19.1	0.37	0.33

B と FP に絞り込んだ。グループ B の変数については、5 段階の値のうち 1, 2 と 4, 5 をそれぞれまとめて 3 段階に変換した。そして、例えば機能性要求が 1, 2 の場合に特に生産性に違いが見られた場合、機能性要求が 1, 2 かそれ以外かを表すダミー変数を作成した。平均要員数/FP は 3 章の方法により名義尺度化した。

名義尺度化した平均要員数/FP を用いた工数見積モデルと用いないモデルの精度を表 1 に示す。それぞれの値は 20 回実験を行った結果の中央値である。表より、名義尺度化した平均要員数/FP を用いたモデルのほうが、精度が高いことがわかる。

5. まとめ

本稿では、ソフトウェア開発工数見積モデルの説明変数に、規模あたりの開発要員数を表す名義尺度の変数を用いることを提案した。実験により、提案方法は見積精度を向上させることを確かめた。

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] Blackburn, J., Scudder, G., and Wassenhove, L.: Improving Speed and Productivity of Software Development: A Global Survey of Software Developers, *IEEE Trans. on Software Eng.*, Vol.22, No.12, pp.875-885 (1996).
- [2] Conte, S. D., Dunsmore, H. E., and Shen, V. Y.: Software Engineering Metrics and Models, p.396, The Benjamin/Cummings Publishing Company, Inc., Menlo Park, California (1986).
- [3] MacDonell, S. G., and Shepperd, M. J.: Comparing Local and Global Software Effort Estimation Models - Reflections on a Systematic Review, Proc. the 1st International Symposium on Empirical Software Engineering and Measurement (ESEM 2007), pp. 401-409, Madrid, Spain (2007).
- [4] Maxwell, K., Wassenhove, L., and Dutta, S.: Software Development Productivity of European Space, Military, and Industrial Applications, *IEEE Trans. on Software Eng.*, Vol.22, No.10, pp.706-718 (1996).