

工数見積もり精度のベースライン確立に向けた データセットの特性分析

角田雅照^{†1} 倉本真美^{†1}

概要: 本研究では、ソフトウェア開発工数見積もりにおける精度のベースライン確立に向け、見積もりモデル構築に用いるデータセットの特性と見積もり精度との関係について議論する。

Analyzing Characteristics of Dataset to Establish the Baseline of Effort Estimation

MASATERU TSUNODA^{†1} MAMI KURAMOTO^{†1}

Abstract: To establish the baseline of the accuracy on software effort estimation, we discuss the relationship between the characteristics of the dataset and the estimation accuracy.

1. はじめに

近年、ソフトウェアはより高機能が求められており、その要求にこたえるためにソフトウェアの規模が大規模化している。大規模なソフトウェアを開発するためには、多くの工数（コスト）が必要となる。このようなソフトウェアを開発するプロジェクトでは、スケジュールやコストに関する管理を行わずにプロジェクトを成功させることは困難である。よって、大規模な開発プロジェクトにおいてプロジェクトの失敗を避けるためには、スケジュールやコストの管理が必要不可欠となる。ソフトウェア開発工数の見積もりは、それらの管理の基礎となる。

スケジュールやコスト管理を正確に行うためには、ソフトウェア開発の工数を高い精度で見積もることは非常に重要であり、そのため、これまでさまざまな定量的工数見積もり手法が提案されてきた。定量的工数見積もり方法として、重回帰分析やプロジェクト類似性に基づく工数見積もりなどがあげられる。定量的に工数を見積もるためには、過去のプロジェクトで収集されたデータを見積もりの根拠データとし、見積もり対象のプロジェクトにおいて既知のデータ、例えば開発言語や開発規模などを用いて、開発工数を見積もる。

ソフトウェア開発の工数見積もり精度は、用いる見積もり方法によっても異なるが、データセットによっても異なる。これは、データセットによって、含まれるデータの品質（データ件数が少ない、必要な説明変数が記録されていない、記録されているデータが不正確であるなど）が異なることが原因である。これまで、いくつかの研究において、データセットの品質の重要性が指摘されており 1), 品質を評価する試みもいくつか行われている 2)。

見積もりモデルの評価を行う場合、評価対象のモデルが仮に性能が高いものであっても、評価に用いるデータセットの品質が低い場合、充分高い見積もり精度を得ることは難しいと考えられる。逆に適切なデータセットを用いれば、かなり見積もり精度が改善される場合もあると考えられる。すなわち、見積もりモデルの性能を評価するためには、データセットの品質も考慮する必要があるといえる。各データセットにおいて、工数見積もり精度に関する何らかのベースラインを決定することができれば、このベースラインに基づいて、見積もりモデルの性能を適切に評価できると考えられる。例えば、あるデータセットを用いる場合、相対誤差 100%がベースラインであるとする、100%に近いモデルは「比較的性能が高く、適切なデータセットを用いることにより、より精度が高まる」などと評価できると期待される。

ベースラインを設定する方法として、広く用いられる見積もり方法の精度、例えば重回帰分析を採用することがある。ただし、ベースラインの設定時に重回帰分析が必ずしも適切に構築されているとは限らず、また、どんなデータセットに対しても重回帰分析を適用可能であるとはいえない。

そこで本研究では、工数見積もり精度のベースライン確立に向け、データセットの特性と見積もり精度との関係について分析を試みる。データセットの特性とは、例えばデータ件数、説明変数の数、各説明変数と目的変数との相関、各変数の分散などが考えられる。その他に、データセットの収集年代や収集期間なども特性の候補として挙げられる。

2. 予備分析

予備分析では、工数見積もりの研究において広く用いられている、重回帰分析とプロジェクト類似性に基づく見積もり方法を用いた。データセットとして、ISBSG データ、

^{†1} 近畿大学
Kindai University

表 1 データセットの特性と見積もり精度との関係

見積もり方法	データセット	BRE 平均値	BRE 中央値	生産性分散	生産性 変動係数
類似性に 基づく手法	ISBSG	258.5%	91.8%	0.076	1.43
	Angel	63.1%	47.0%	0.006	0.87
	Kitchenham	96.6%	54.1%	0.110	1.39
重回帰分析	ISBSG	158.1%	86.0%	0.076	1.43
	Angel	55.3%	37.9%	0.006	0.87
	Kitchenham	78.8%	45.5%	0.110	1.39

Kitchenham データ, Desharnais データを用いた. ISBSG データは, ISBSG が収集した 20 カ国のソフトウェア開発企業のデータである. 不正確なデータを除外した 611 件を分析に用いた. Kitchenham データは, あるソフトウェア開発企業の 1990 年代のデータである. 分析では 135 件のプロジェクトを用いた. Desharnais データは, 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. データセットに含まれるプロジェクトのうち, 欠損値が含まれるものを除いた 77 件を用いた.

データセットの特性として, 予備分析では生産性の分散と変動係数に着目した. 生産性は開発工数と開発規模 (FP などにより計測される) の比として算出される. 生産性の分散に着目した理由は, 工数は開発規模との関係が強いためである. 例えば, 単重回帰分析で予測する場合, 開発規模は説明変数, 工数は目的変数となるが, 生産性の分散が低ければ見積もり精度も高まると考えられる. 分散は平均が大きくなるに従い大きくなるため, 生産性の変動係数についてもデータセットの特性とした.

予測精度評価指標として, BRE (Balanced Relative Error) の平均値と中央値を用いた. 工数の実測値を x , 見積もり値を \hat{x} とするとき, BRE は以下の式で求める.

$$BRE = \begin{cases} \frac{\hat{x} - x}{\hat{x}}, \hat{x} - x \geq 0 \\ \frac{x}{x - \hat{x}}, \hat{x} - x < 0 \end{cases} \quad (1)$$

BRE の値が小さいほど, 工数見積もりの精度が高いことを示す.

分析結果を表 1 に示す. 見積もり方法に関わらず, ISBSG データセットの見積もり精度が最も低く, Desharnais データセットの精度が最も高かった. 生産性の分散は Kitchenham データセットが最も高かったことから, 見積もり精度と生産性の分散はあまり関係が強いとはいえない. 生産性の変動係数については, 差はわずかではあるが ISBSG データセットが最も大きく, Desharnais データセットが最も低かった. このことから, 生産性の変動係数を見積もり精度と何らかの関係がある可能性がある.

3. おわりに

ワークショップでは, 工数見積もり精度のベースライン確立に向け, データセットのどのような特性に着目すべきかについて議論したい.

謝辞 本研究の一部は, 文部科学省科学研究補助費 (基盤 C : 課題番号 16K00113) による助成を受けた.

参考文献

- 1) G. Liebchen and M. Shepperd: Data sets and data quality in software engineering, In Proc. of int. workshop on Predictor models in software engineering (PROMISE), pp.39-44 (2008).
- 2) P. Phannachitta, A. Monden, J. Keung, and K. Matsumoto: Case consistency: a necessary data quality property for software engineering data sets, In Proc. of Int. Conference on Evaluation and Assessment in Software Engineering (EASE), No. 19, 10 pages (2015).