

論文内容の要旨

博士論文題目 プロジェクトの個別性を考慮した定量的ソフトウェア開発工数見積り

氏名 戸田 航史

ソフトウェア開発の初期段階において、開発に必要な労力（開発工数）を高い精度で見積もることが必須である。ただし、ソフトウェア開発は個別性が高いために、あらゆる開発プロジェクトに適合する万能の見積り方法は存在しない。本論文では、従来広く用いられてきた工数見積り方法である線形重回帰モデルによる見積り方法とアナロジーベース法それぞれに対し、開発プロジェクトの個別性を考慮した拡張を行った。これにより平均的な見積り精度の向上と見積りを大きく外すプロジェクトの低減を実現できた。提案する拡張方法の概略は次の通りである。

(1) プロジェクトの特性に応じたフィットデータの選定方法

本方法では、重回帰モデルに基づく工数見積り方法に対し、モデル構築のためのプロジェクトの集合（フィットデータ）を選定する手順を追加する。見積り対象プロジェクトの特性に基づいてフィットデータを選定することにより、個々のプロジェクト向けにカスタマイズされた重回帰モデルの構築を行う。ソフトウェア工学における代表的な実験用データセットである ISBSG データセットを用いた評価実験の結果、フィットデータの選定を行わない従来の重回帰モデルと比較して、相対誤差 MRE (Magnitude of Relative Error) の中央値が 0.552 から 0.383 へ、相対誤差 MER (Magnitude of Error Relative) の中央値が 0.457 から 0.381 となり、見積り精度が向上した。

(2) 類似プロジェクトの特性に応じた見積り回避プロジェクトの選定方法

アナロジーベース法は、開発プロジェクトの個別性を考慮した見積り方法であり、見積り対象の開発プロジェクトに対し、類似するプロジェクト群を過去の実績データから選定し、それらの実績工数に基づいて開発工数を見積もる方法である。ただし、プロジェクトによっては、極端に見積り精度が低くなる（見積りに失敗する）ことが課題であった。提案方法では、アナロジーベース法を拡張し、選定された類似プロジェクト群の特徴に基づいて、見積り精度が低くなるか否かを見積り、見積りを実施すべきか否かを判断する。見積りを実施すべきでないと判断されたプロジェクトは重回帰モデルによる見積りを行い、そうでないプロジェクトでは従来通りアナロジーベース法による見積りを行う。ISBSG データセット、Desharnais データセット、および Kitchenham データセットを用いた評価実験の結果、従来方法と比較して、見積りに失敗したプロジェクトの割合である Pred(100) が最大で 0.254、平均で 0.198 減少し、見積りを大きく外すプロジェクトが減少した。

(論文審査結果の要旨)

本論文では、定量的ソフトウェア開発工数見積りの精度向上を目的として、従来の工数見積り方法に対して、プロジェクトの個別性を考慮した次の2通りの拡張を行っている。(1)重回帰モデルを用いた工数見積りにおいて、モデル構築のためのプロジェクトの集合(フィットデータ)を、見積り対象プロジェクトの特性に基づいて選定する手順を追加する方法、および、(2)アナロジーベース法を用いた工数見積りにおいて、選定された類似プロジェクト群の特性に基づいて、見積り精度が低くなるか否かを判定し、精度が低くなると判定されたプロジェクトは重回帰モデルによる見積りを行い、そうでないプロジェクトでは従来通りアナロジーベース法による見積りを行う方法である。

まず、(1)の提案では、全ての $k=1, \dots, n$ について見積り対象プロジェクトと k 個の特性が一致するプロジェクトの集合を全て求め、各プロジェクト集合を用いて重回帰モデルを作成した時に最も適合度が高くなったものをフィットデータとして選定する。これにより、見積り精度が高くなると期待されるフィットデータを選定できる。従来、系統的なフィットデータ選定方法は提案されておらず、フィットデータの選定は熟練した開発管理者の経験に委ねられていたが、選定の効果は必ずしも高くなかった。提案方法によって、プロジェクトの個別性を考慮した、系統的かつ効果のあるフィットデータ選定が可能となり、ソフトウェア開発現場における高い有用性が見込まれる。

(2)の提案では、アナロジーベース法による工数見積りにおいて、選定された類似プロジェクト群の実績工数の分散が大きいほど、見積りが大きく外れる(相対誤差平均が100%を超える)割合が増えることを実験的に確かめている。そして、そのような見積りを大きく外すと推定されるプロジェクトに重回帰モデルを用いることで、予測精度を改善している。従来、アナロジーベース法は、見積りの根拠である類似プロジェクト群が得られることから、重回帰モデルなどと比べて直感的に分かりやすいという大きな利点があった。ところが、ときおり大きく見積りを外すことが、開発現場に導入する際の大きな問題となっていた。提案方法によって、見積りを大きく外すと推定される場合に限り重回帰モデルを用いることで、アナロジーベース法の特徴を生かしたままその欠点が解決され、開発現場における導入の促進が期待される。

以上のとおり、いずれの提案方法も、研究の位置づけが明確にされた上で、全体にわたって十分具体的に記述されている。ソフトウェア開発企業のデータを用いた評価実験がなされており、信頼性も高い。これらの研究成果は、ソフトウェア開発現場における定量的工数見積りの導入促進に大きな役割を果たし、ソフトウェア業界の発展に大きく貢献するものであり、本論文は博士(工学)論文として価値あるものと認める。