

論文内容の要旨

博士論文題目 ソフトウェア開発管理における予測手法の利用的側面からの評価

氏 名 柿元 健

ソフトウェア開発管理を支援するために、開発工数と信頼性に関して数多くの予測手法が提案されているにも関わらず、その多くは開発現場で利用されていない。本論文では、予測手法を利用的側面から評価することで、その適用の基準や得られる効果を明確にし、開発現場での採用を促進することを目的としている。

まず、第2章、第3章では、開発工数の予測において、予測に必要なデータセットに求められる基準を明らかにするために、データセットに含まれる欠損値、プロジェクトの件数、メトリクス数について、類似性に基づく工数予測手法とステップワイズ重回帰分析の2種類の工数予測手法を評価している。欠損値について、欠損値が生じる3つのメカニズムを想定し、それぞれについて欠損率を変化させたデータセットを多数作成し、各データセットを用いて工数予測を行うことで実験的に評価している。評価実験の結果、類似性に基づく工数予測手法が、ステップワイズ重回帰分析よりもロバスト性が高い、すなわち、欠損のメカニズムに関わらず、欠損率が增大しても予測精度が大きく低下しないことを示している。また、プロジェクト件数とメトリクス数について、それぞれを変化させたデータセットを作成し、各データセットを用いて工数予測を行うことで実験的に評価している。評価実験の結果、プロジェクト件数が50件以上の場合、類似性に基づく工数予測手法がステップワイズ重回帰分析よりも高い精度で予測でき、メトリクス数の増加に伴って予測精度が向上することを示している。

次に、第4章では、信頼性予測の一手法である **fault-prone** モジュール判別において、予測により得られる効果を明らかにするために、各モジュールの **fault** の有無、割り当てられるテスト工数、およびソフトウェアの信頼性 (**fault** 発見率) の関係をモデル化している。費やされたテスト工数に基づいて **fault** 発見率を示すモデルを構築している。構築した **fault** 発見率モデルを用いて、評価指標の1つである F1 値、**fault** モジュール含有率、**fault-prone** モジュール率をパラメータとしてシミュレーションを行っている。シミュレーションの結果、プロジェクト全体の信頼性 (**fault** 発見率) は、F1 値、**fault** モジュール含有率、**fault-prone** と判別されたモジュールの割合によって決定されることを示している。

氏名	柿元 健
----	------

(論文審査結果の要旨)

本論文では、ソフトウェア開発管理における予測手法の開発現場での採用を促進することを目的として、予測手法の利用的側面からの評価によって、その適用の基準や得られる効果を明確にしている。

まず、第2章、第3章では、開発工数の予測において、予測に必要なデータセットに求められる基準を明らかにするために、データセットに含まれる欠損値、プロジェクトの件数、メトリクス数について、類似性に基づく工数予測手法とステップワイズ重回帰分析の2種類の工数予測手法についてコンテキストを明らかにするための評価を行っている。欠損値について、欠損値が生じる3つのメカニズムを想定し、それぞれについて欠損率を変化させたデータセットを多数作成し、各データセットを用いて工数予測を行うことで実験的に評価している。評価実験の結果、類似性に基づく工数予測手法が、ステップワイズ重回帰分析よりもロバスト性が高い、すなわち、欠損のメカニズムに関わらず、欠損率が增大しても予測精度が大きく低下しないことが示されている。また、プロジェクト件数とメトリクス数について、それぞれを変化させたデータセットを作成し、各データセットを用いて工数予測を行うことで実験的に評価している。評価実験の結果、プロジェクト件数が50件以上の場合、類似性に基づく工数予測手法がステップワイズ重回帰分析よりも高い精度で予測でき、メトリクス数の増加に伴って予測精度が向上することが示されている。

次に、第4章では、信頼性予測の一手法である **fault-prone** モジュール判別において、各モジュールの **fault** の有無、割り当てられるテスト工数、およびソフトウェアの信頼性 (**fault** 発見率) の関係をモデル化することで、予測により得られる効果を明らかにしている。構築した **fault** 発見率モデルを用いて、評価指標の1つである F1 値、**fault** モジュール含有率、**fault-prone** モジュール率についてシミュレーションを行っている。シミュレーションの結果、プロジェクト全体の信頼性 (**fault** 発見率) は、F1 値、**fault** モジュール含有率、**fault-prone** モジュール率によって決定されることが示されている。

以上のように、本論文は、ソフトウェア開発管理における予測手法の開発現場での採用を促進するための、予測手法の利用的側面からの評価について、予測に必要なデータセットに求められる基準、および、予測により得られる効果について述べている。どちらの方法も、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おり、ケーススタディや実験により定量的評価がなされており、信頼性が高い。これらの研究成果は、ソフトウェア開発管理の分野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。