

論文内容の要旨

博士論文題目 プロダクトメトリクスと開発者メトリクスを用いた
ソフトウェア信頼性予測モデルの構築方法
氏名 杉本 真佑

大規模化と複雑化が進むソフトウェア開発において、限られたリソースの中で高い信頼性を確保するためには効率的なテストの実施が必須である。その一つの手段は、信頼性の低い、すなわち *fault* (欠陥) を含むソフトウェアモジュールを推定し、テストに費やすリソースを適切に割り当てることである。これまで、過去の開発モジュールの *fault* 検出傾向に基づき、新規開発モジュールに含まれる *fault* の有無を判別する *fault-prone* モジュール判別モデルに関する研究が数多く実施されてきた。

本論文では、*fault-prone* モジュール判別の精度向上を目的として、モデル構築に用いられるフィットデータに関する以下の2つの問題に取り組んだ。

(1) 外れ値除去法適用の効果：一般に *fault-prone* モジュール判別モデル構築のためのフィットデータとしては、過去のプロジェクトで開発されたモジュールの全てが用いられる。しかし、フィットデータ中には他の標本の傾向と異なる外れ値が含まれており、この外れ値が判別モデルの精度を低下させる原因となることが指摘されている。このため、外れ値はあらかじめフィットデータから除去したうえで判別モデルを構築することが望ましい。本論文では、3つの外れ値除去法 (MOA, LOFM, RBM) と新たに提案する外れ値除去法 (CC-MOA) を代表的な3つの *fault-prone* モジュール判別モデルに適用し、その効果を実験的に比較した。実験の結果、LOFMを除く3つの外れ値除去法を用いることで、いずれの判別モデル、データセットに対しても判別精度が改善することを示した。

(2) 開発者メトリクスの提案：従来、モデル構築の説明変数として用いられるメトリクスとしては、ソフトウェアプロダクトの特徴を表すメトリクスが広く用いられている。しかし、*fault* が混入される要因としてはプロダクトの特性のみならず、人的要因、すなわちプロダクトを作成した開発者の特性が強く影響する。*Fault* の有無を目的変数とする判別モデルとしては、*fault* 混入の原因となる要因は可能な限り説明変数として加えることが望ましい。本論文では、開発者に関するメトリクス (開発者メトリクス) を提案し、ソフトウェアの信頼性に対する人的要因に関して分析と実験を行った。実験では、まず *fault* の増加に対して開発者の特性がどの程度関連しているかについて分析を行い、次に *fault-prone* モジュール判別に対して開発者メトリクスを加えた際の効果を確かめた。Eclipse プロジェクトから収集されたモジュールデータを用いた分析の結果、開発者の *fault* 混入率には5倍以上の個人差があること、および、より多くの開発者が変更を加えたモジュールほど *fault* が混入されやすいことが明らかとなった。また開発者メトリクスが判別モデルの精度向上に有効であることを示した。

(論文審査結果の要旨)

本論文では、過去の開発モジュールの `fault` 検出傾向に基づき、新規開発モジュールの `fault` を事前に特定する `fault-prone` モジュール判別モデルの性能向上を目的として、(1) モデルに対してノイズとなる外れ値の除去、および、(2) 人的要因を考慮することによるモデルの表現能力の向上の2つを提案している。

(1) については、判別モデル構築の事前処理として、モデル構築用データ中に含まれる外れ値の除去手法を適用することを提案し、その効果を実験的に確かめている。実験では、NASA/WVU が公開しているデータセットを題材とし、3つの判別モデル(線形判別分析、ロジスティック回帰分析、分類木)に対して、新たに提案する手法を含む4種類の外れ値除去法(MOA, LOFM, RBM, CC-MOA)を適用している。実験の結果、外れ値除去法を適用することで、`fault-prone` モジュール判別モデルの性能向上という効果のみならず、どの判別モデルを用いても安定して高い判別精度が得られることが明らかとなった。これらの成果によって、テスト工数のより適切な割り当てが可能となるのみならず、モデル利用者の利便性が向上するため、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) については、判別モデルの表現能力を補うために、人的要因を考慮に入れた開発者メトリクスの提案を行っている。実験では、Eclipse プロジェクトで開発されたモジュールデータを題材として、信頼性の低下に対する人的要因の影響と、判別モデルに対する開発者メトリクス導入の効果を確かめている。従来の判別モデルはプロダクトの特徴のみに基づいて構築されているためにその表現能力に限界があったが、本論文では、従来考慮されていなかった人的要因を判別モデルに取り入れるとともに、その有効性を実験的に確認しており、高い新規性と信頼性が認められる。また、実験の手順は慎重に設計されており、実験結果の解釈も妥当である。さらに、実験結果に基づいてソフトウェアの信頼性改善活動に対する具体的な提案を行っており、高い有用性が認められる。

以上のとおり、本論文は `fault-prone` モジュール判別モデルの性能向上手段として、モデル構築用データに関する問題の具体的な解決策を提案しており、実験の結果から提案手法の有用性、および、信頼性が十分に示されている。これらの研究成果は `fault-prone` モジュール判別モデルの学術研究の発展に貢献するとともに、ソフトウェアテスト技術の発展に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。