

氏名(本籍・生年月日) 内宮秀明(兵庫県 昭和46年1月31日)  
学位の種類 博士(工学)  
学位記番号 乙 第 235 号  
学位授与の日付 平成27年3月20日  
学位授与の要件 信州大学学位規程 第5条第2項該当  
学位論文題目 Method Mining に基づく error-proneモジュールの予測

論文審査委員 主査 教授 海尻賢二 准教授 小林一樹  
教授 丸山稔 助教 小形真平  
教授 海谷治彦(神奈川大学)

### 論文内容の要旨

ソフトウェア開発では、限られた期間内に運用に耐えうるだけの品質を満たすソフトウェアを開発することが求められる。製作工程ではソースコードを対象としたインスペクションを行うことで品質の向上が期待できる。また、試験工程においては、網羅的な試験項目を消化することにより、納入までの間に誤りを摘出することで品質を向上させることが期待できる。

しかし、インスペクション及びソフトウェアテストは開発規模に応じて、工数も必要となるため、開発の効率化を図るためには、品質を劣化させないことを前提として、実施対象範囲の絞り込みが必要となる。実施対象範囲を絞り込むためには、まだ発見されていない誤りを含んでいる可能性の高い、error-proneモジュールの予測を行うことは有効である。なぜなら、error-proneモジュールと判定したモジュール或いは、error-proneモジュールが関係する機能部全体に対してインスペクションもしくはソフトウェアテストを集中的に行うことで、早期に誤りを摘出するだけでなく、インスペクションやデバッグに費やす工数も削減することが出来るからである。

これまで、数多くのerror-proneモジュールの予測手法が提案されているが、現状、汎用的でかつ最適なerror-proneモジュール予測器は存在しない。また、検査対象プロジェクトに対して、最適(もしくは適切)なerror-proneモジュール予測器をどのようにして選ぶかについては、ほとんど研究が行われていない。

HEはメトリクスの集約値と予測結果の精度の関係を学習させることで、検査対象プロジェクトに対して適切なerror-proneモジュール予測器を見つけ出す手法を提案している。本論文では、検査対象プロジェクトの特徴

を考慮した要素を特徴量として定義し、HEが提唱した手法を応用して、特徴量を考慮したマイニング手法による探索手法を提案する。

上記を踏まえ、本論文では、以下に示す3つの仮説を設定し、各仮説を検証するための実験を行った。

仮説1： 予測閾値に基づいて適切と判定された予測器は一意ではなく、検査対象プロジェクト毎に適切な訓練データと予測アルゴリズムがそれぞれ異なる。

仮説2： 提案するマイニング手法を適用すれば予測閾値による識別の結果から得ることのできるprecisionが十分な値となり、適切な予測器(訓練データと予測アルゴリズム)を識別することができる。

仮説3： 特徴量はマイニング手法の識別精度向上に寄与する。

実証実験の結果、以下の点が判明した。

- (1) 最適な予測器は一意ではなく、対象プロジェクト毎に適切な訓練データと学習アルゴリズムがそれぞれ異なる。
- (2) 同一プロジェクトの過去データが最適な訓練データであると判断することはできず、プロジェクト毎に最適な訓練データは異なるといえる。
- (3) 実証実験の結果、識別精度としてprecisionの値は0.7以上のケースが存在した。よって、適切なerror-proneモジュール予測器を識別できる可能性はあると言える。
- (4) 特徴量を利用することでprecisionが改善したケースが存在した。

これらの実験結果を踏まえ、本論文では、HEの提案手法では考慮されていない特徴量を導入することにより、検査対象プロジェクトに対して最適なerror-proneモジュール予測器を識別できる可能性を示した。