

氏 名	こんどう まさなり 近藤 将成
学位(専攻分野)	博 士 (工 学)
学 位 記 番 号	博 甲 第 9 9 3 号
学位授与の日付	令和 3 年 3 月 25 日
学位授与の要件	学位規則第 4 条第 1 項該当
研 究 科 ・ 専 攻	工芸科学研究科 設計工学専攻
学 位 論 文 題 目	An Empirical Study of Feature Engineering on Software Defect Prediction (ソフトウェア不具合予測における特徴量エンジニアリングの 実証的研究)
審 査 委 員	(主査)教授 水野 修 教授 寶珍輝尚 教授 澁谷 雄

論文内容の要旨

ソフトウェアは我々の生活になくてはならないものである。不具合のあるソフトウェアは破局的な問題を引き起こす可能性があり、そのような事例は数多く存在する。そのため、ソフトウェア開発者は、ソフトウェアテストなどのソフトウェア品質保証活動により、より良い品質のソフトウェアを目指す必要がある。しかし、近年、ソフトウェアの規模は非常に大きくなり、また、依存する環境も複雑になってきている。そのため、限られたソフトウェア品質保証活動のリソースでは、ソフトウェア全体に対してソフトウェア品質保証活動を漏れなく行うことは難しい。

ソフトウェア不具合予測とは、不具合予測モデルを利用して、不具合を持つソフトウェアの構成要素を見つけ出す手法である。不具合を持つソフトウェアの構成要素を見つけることで、開発者は限られたリソースをその構成要素に集中配分し、効果的に不具合を見つけ出すことが期待できる。そのため、不具合予測は開発者及び研究者から注目され、ソフトウェア工学分野において盛んに研究されている。

通常、不具合予測モデルは、過去のソフトウェア構成要素から得られた特徴量を用いて機械学習モデルを学習させることで作成する。そのため、どのような特徴量を利用するかは不具合予測の精度に直結する重要な要素であり、過去の研究では特徴量エンジニアリングに関する研究が行われてきた。特徴量エンジニアリングとは、ドメインに関する知識を用いて特徴量を作成したり、改善したりする活動を指す。しかし、ソフトウェア不具合予測においては、特徴量エンジニアリングによって解決可能だが未解決の問題がいまだに存在する。例えば(1)特徴抽出手法の比較研究が行われていない(2)変更されたソースコードの文脈を考慮した特徴量が提案されていない(3)ソースコードの意味に基づいた不具合予測が行われていない、等である。

本学位論文では、これら 3 つの問題に取り組んでいる。具体的には、(1)複数の特徴抽出手法および特徴選択手法の大規模な比較実験、(2)文脈を考慮した文脈特徴量の提案、(3)ソースコードの意味に基づく深層学習による不具合予測、をそれぞれ 3 章、4 章、5 章において述べている。(1)については、従来の研究から、特徴量、特徴量抽出手法、特徴量選択手法、不具合予測手法、公開データセットを網羅的に収集し、それらをすべて比較する大規模な比較実験を実施し

ている。(2)については、不具合の原因となったとされるソースコードの前後の文脈を抽出して特徴と見なし、不具合予測手法に適用を行っている。(3)については、ソースコードの意味を特徴として捉え、深層学習により学習し、不具合予測を実施する手法の提案を行い、適用実験を行っている。本論文によって示された実験の結果より、(1) 特徴抽出手法および特徴選択手法が不具合予測の精度を改善しつつ、予測に利用する特徴量の数を減らせること、(2) 文脈特徴量が不具合予測の精度を向上させること、(3) ソースコードの意味を利用する手法を提案することで、既存の深層学習を用いた変更に対する不具合予測手法よりも予測精度を改善できることが明らかになった。

本論文の成果により、ソフトウェア不具合の研究における特徴量の抽出の指針が示され、また、効果的な方法を採用することで、不具合予測の精度を高めることが示された。

論文審査の結果の要旨

本学位論文では、ソフトウェアにおける特徴量を用いた不具合の予測に対する複数の観点からのアプローチについて論じている。ソフトウェア不具合の予測は、何らかの特徴量、例えば、ソースコードの複雑さやテストで発見した不具合数等からモデルを作成するものが一般的である。過去に実施された多くの研究において、様々な特徴量が提案されてきた結果、すべての特徴量を網羅することが難しくなっている。そこで本学位論文では、まず、従来提案されてきた多くの特徴量の数の削減について、その効果を確認する実験を行っている。その結果、不具合予測の精度を改善しつつ、利用する特徴量の削減が行えることを確認している。また、次のアプローチとしてソースコードの文脈という特徴に着目し、文脈が不具合予測の精度に与える影響について論じている。具体的にはある不具合の原因となったソースコードの一部分に、その前後の文脈の情報を付け加えた学習を行うことで、不具合予測精度の向上を確認したとしている。また、3つめのアプローチとして、ソースコードの変更に対して、その変更の意味を深層学習により抽出し、不具合予測での特徴として利用することを論じている。実験により、この手法によって、従来の手法よりも高い予測精度を達成することを確認している。

本論文は、ソフトウェアの不具合予測に対してソフトウェアの特徴量が及ぼす影響を多くの実験により明らかにしており、ソフトウェア不具合予測の分野において有用な知見を得ていると評価できる。

本論文は申請者を筆頭著者とする査読を経た以下に示す3編の論文を基礎としている。

- (1) Masanari Kondo, Daniel M. German, Osamu Mizuno, and Eun-Hye Choi, “The impact of context metrics on just-in-time defect prediction,” *Empirical Software Engineering*, Vol. 25, pp.890-939, 2020.
- (2) Masanari Kondo, Cor-Paul Bezemer, Yasutaka Kamei, Ahmed E. Hassan, and Osamu Mizuno, “The impact of feature reduction techniques on defect prediction models,” *Empirical Software Engineering*, Vol. 24, pp.1925-1963, 2019.
- (3) 近藤 将成, 森 啓太, 水野 修, 崔 銀恵, 深層学習によるソースコードコミットからの不具合混入予測, 情報処理学会論文誌, Vol. 59, No. 4, pp.1250-1261, 2018 年 4 月.

また、参考論文として、査読制度のある下記の国際雑誌 2 編にも採録決定、または、掲載され

ている.

- (1) Masanari Kondo, Gustavo A. Oliva, Zhen Ming (Jack) Jiang, Ahmed E. Hassan, and Osamu Mizuno, “Code cloning in smart contracts: a case study on verified contracts from the Ethereum blockchain platform,” Empirical Software Engineering, (accepted, to appear)
- (2) Masanari Kondo, Osamu Mizuno, and Eun-Hye Choi, “Causal-Effect Analysis using Bayesian LiNGAM Comparing with Correlation Analysis in Function Point Metrics and Effort,” International Journal of Mathematical, Engineering and Management Sciences, Vol. 3, No. 2, pp.90-112, 2018.