

論文内容の要旨

博士論文題目

Understanding and Recommending Key Features to Improve Bug Management Process

氏名

Karim Md. Rejaul

(論文内容の要旨)

バグレポートは、開発者がバグを分類・修正するための基本手段であり、その効果的な実施には、開発者にとって重要な記載項目（フィーチャー）がバグレポートに含まれている必要がある。このような重要な記載項目が当初のレポートから欠落しているとバグ修正プロセスの各段階に重要な影響を及ぼすことが既存研究により示されている。本論文では、この観点からバグ修正プロセスを改善するための研究についてまとめたものであり、大きくは二部に分かれる。

前半では、バグレポートにおける重要な記載項目が何であるかを調査した実証的研究を扱う。調査からは以下の事実が確認された：(1) 「再現手順」、「テストケース」、「コード例」、「スタックトレース」、および「予期される動作」は、開発者がバグを修正するための重要な記載項目であるが、当初はレポートに記載されていないことが多い。(2) ハイ・インパクト・バグと呼ばれる深刻度の高いバグでは、その種類によって項目の重要度が異なる。(3) レポートに欠落している項目をバグ修正中に追加質問することは、バグ修正プロセスの効率に大きく影響する。

後半では、バグ報告者を支援することによりバグ修正プロセスを改善する手法を提案し、その評価を行う。まず、報告者がバグレポートの説明にどの項目を重要な記載項目として提供すべきかを4つの機械学習手法を活用して予測する分類モデルを開発し、さらに、バグ修正の履歴に基づいた学習とテキストマイニング手法により重要な記載項目を推奨する手法を開発した。評価実験の結果、以下を確認した：(1) 重要な記載項目の予測モデルからは十分に高い F1 値を得た。(2) 重要な記載項目の予測では、Naïve Bayes Multinomial (NBM)が他の分類手法よりも優れている。(3) 最良モデルは別プロジェクトでの重要な記載項目予測においても有用である。(4) 重要な記載項目の推奨手法は、バグレポートに提供すべき重要な記載項目を適切に推奨できた。

以上

(論文審査結果の要旨)

バグレポートは、開発者がバグを分類・修正するための基本手段であるが、重要な記載項目が当初のレポートから欠落しているとバグ修正プロセスの各段階に重要な影響を及ぼすことが既存研究により示されている。本論文では、この観点からバグ修正プロセスを改善するための研究についてまとめたものである。

まず、バグレポートにおける重要記載項目が何であるかを調査した結果から、以下の事実が報告されている：(1)「再現手順」、「テストケース」、「コード例」、「スタックトレース」、および「予期される動作」は、開発者がバグを修正するための重要記載項目であるが、当初はレポートに記載されていないことが多い。

(2) ハイ・インパクト・バグと呼ばれる深刻度の高いバグでは、その種類によって項目の重要度が異なる。(3) レポートに欠落している項目をバグ修正中に追加質問することは、バグ修正プロセスの効率に大きく影響する。

次に、バグ報告者を支援することによりバグ修正プロセスを改善する手法を提案し、その評価を行っている。まず、報告者がバグレポートの説明にどの項目を重要記載項目として提供すべきかを4つの機械学習手法を活用して予測する分類モデルを開発し、さらに、バグ修正の履歴に基づいた学習とテキストマイニング手法により重要記載項目を推奨する手法を開発した。さらに、評価実験の結果、一定の有効性を示すことができた。

このように、本論文では、バグレポートという重要情報の取り扱いについての実証的な研究に立脚した上で、機械学習やテキストマイニング技術を応用した支援手法を提案することで、ソフトウェア保守をふくむ開発プロセス全般について改善をもたらす取り組みについて報告しており、その有効性についても慎重な評価を行っている。その成果はソフトウェア工学分野に対して明確な貢献を与えるものであり、したがって、本論文は博士(工学)の学位論文として認めるに値すると判断する。

以上