

論文内容の要旨

博士論文題目 OSSの不具合修正プロセス効率化のためのコミッター推薦と不具合修正時期の予測

氏名 伊原 彰紀

本論文は、大規模オープンソースソフトウェア (OSS: Open Source Software) の不具合修正プロセスの効率化を目的として、不具合修正に貢献する有能なコミッターを見つけ出す方法、および、不具合の修正時期を予測する方法を提案する。

OSS プロジェクトにおいてコミッターは、不具合をプロジェクトで修正すべきか否かを審査する役割と、不具合が正しく修正されたか否かを検証する役割を務めており、OSSの不具合を修正する上で必要不可欠な存在となっている。しかしながら、OSSプロジェクトでは、開発者にプロジェクトへの参加/離脱の自由を認めているため、同じコミッターが将来的に不具合修正プロセスを継続的に支えるとは限らない。コミッターの不足は、不具合の修正時間の長期化に繋がり、長期間修正されない不具合はOSS利用企業が修正せざるを得ない状況を引き起こす。

社会に広く普及しているOSSの品質向上、および、修正の効率化のためには、OSSプロジェクトによる不具合修正プロセスの改善だけでなく、OSS利用企業に不具合の修正を促す必要性が高まっている。本論文では、OSSプロジェクト、及び、OSS利用企業の各々の立場における不具合修正プロセスを効率化するために、2つの研究に取り組んだ。

(1) OSSプロジェクトのための有能なコミッターの予測

OSSプロジェクトでは、コミッターの負担を軽減させるために、不具合が正しく修正されたか否かを検証する能力を持つコミッターが求められている。しかし、プロジェクトに参加する一般開発者(大規模プロジェクトでは1万人を超えることも珍しくない)の中からコミッターに昇格する有能な開発者を見つける事は容易ではない。本論文では、不具合修正プロセスにおける開発者の活動量を用いて、多数の一般開発者の中からコミッターを予測するモデルを構築した。実験の結果、再現率が約70%~80%の精度でコミッターに昇格する開発者を予測できることが分かった。

(2) OSS利用企業のための不具合修正時期の予測

OSS利用企業は、修正を要する不具合が次期リリースまでに修正されるか否かを判断し、修正されないのであれば、自社で修正することを検討する必要がある。本論文では、不具合の特徴を示すメトリクス(ベース)に加えて、修正の進捗情報を示す3種類のプロセスメトリクス(状態、期間、参加者)を新たに提案し、OSS利用企業による効率的な不具合修正に向けて、OSS利用企業が次期リリースまでに修正されるか否かを予測するためのモデルを構築した。実験の結果、ベースメトリクスに加えて状態メトリクスを用いることで適合率が約24%~39%、再現率が約14%向上することが分かった。

(論文審査結果の要旨)

本論文では、大規模なオープンソースソフトウェア(OSS: Open Source Software)の開発における不具合修正プロセスを効率化する具体的な方法を提案している。

まず、OSS プロジェクトにおいて、不具合が正しく修正されたか否かを検証する権限を持つ特別な開発者(コミッター)を、一般の開発者の中から選出する(予測する)ためのモデルを構築している。同モデルは、不具合修正プロセスにおける、修正のためのソフトウェア差分(パッチ)の投稿や検証に関する開発者の活動量を数値化し、当該開発者をコミッターとすべきかどうかを判定するものである。代表的な大規模 OSS プロジェクトである Eclipse platform プロジェクト、および、Mozilla Firefox プロジェクトを対象としたモデル評価実験の結果、再現率 70~80%の精度で、コミッターに適した開発者を選ぶことが可能であることを示した。これまで、コミッターの選出に明確な基準は定められておらず、コミッターに適した開発者を選出することは容易ではなかった。結果として、十分な数のコミッターが確保されず、不具合が修正されないまま長期間放置される一因となっていた。本論文の成果は、これまで明らかではなかった「OSS プロジェクトにおけるコミッターの活動」を定量的に表現するとともに、開発者数が1万人を超えることも珍しくない大規模 OSS プロジェクトにおいても、コミッター適格者を客観的、かつ、効率よく選出する方法を提供するものであり、新規性、有用性ともに大である。

次に、OSS プロジェクトにおいて、存在が確認されている不具合の中から、次期リリースまでに修正される可能性の高い不具合を予測するモデルを構築している。同モデルは、不具合の特徴を示すプロダクトメトリクスに、不具合修正の進捗を示す3種類(状態、期間、参加者)のプロセスメトリクスを加えることで、従来よりも高い予測精度を目指すものである。前述と同じ Eclipse platform プロジェクト、および、Mozilla Firefox プロジェクトを対象としたモデル評価実験の結果、3種類のプロセスメトリクスを追加したことで、予測結果の適合率が約24%~39%、再現率が約14%向上することが分かった。OSS は、利用、改変、再頒布が可能のため、最近では、国内外を問わず多くのソフトウェア開発企業がその利用をはじめている。ただし、存在が確認されている不具合を修正する順序や時期は、OSS プロジェクトが決定するものであり、OSS 利用企業の個別の事情や要求に沿ったものとなるとは限らない。同モデルを利用すれば、OSS 利用企業は、どの不具合が次期リリースまでに OSS プロジェクトによって修正されるのか、また、その裏返しとして、どの不具合を自ら修正する必要があるのかが予測可能となる。不具合修正に要する人的資源を効率よく、また、効果的に投入するための新たな指標を開発現場に与えた点で、本研究の成果の新規性、有用性はともに大である。

以上のとおり、本論文は、大規模なオープンソースソフトウェア(OSS)における不具合修正プロセスを、OSS プロジェクト内と OSS 利用企業内の双方において効率化する具体的な方法(モデル)を提案するとともに、その妥当性を、代表的な大規模 OSS プロジェクトから収集されたデータに基づいて示したものである。これらの研究成果は、OSS のさらなる利用を促進するものであり、ソフトウェア工学分野における学術的貢献も大きいことから、本論文は博士(工学)論文として価値あるものと認める。