

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 10 日現在

機関番号：14701

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24500041

研究課題名(和文)リポジトリマイニングに基づく大規模OSS利用支援環境の構築

研究課題名(英文)An environment for mining software repositories to utilize large-scale open source software as a means of productive and quality software development

研究代表者

大平 雅雄(OHIRA, Masao)

和歌山大学・システム工学部・准教授

研究者番号：70379600

交付決定額(研究期間全体)：(直接経費) 4,200,000円

研究成果の概要(和文)：本研究の目的は、ソフトウェア開発企業がオープンソースソフトウェア(OSS)を安心して自社製品へ利用できるようにするために、ボランティア開発者主導で開発されるOSSの品質とプロジェクトの継続性をモニタリングできる支援環境を構築することである。具体的には、(1)OSS開発で利用される各種ソフトウェアリポジトリから自動的にデータを収集・統合するデータベース基盤を構築し、(2)大規模データを目的に応じて分析するためのリポジトリマイニング手法を分類化・プラグイン化し、(3)分析結果をモニタリングするための可視化ツールを開発して、(4)検証実験を通じて支援環境の有用性を評価した。

研究成果の概要(英文)：The goal of the study is to construct an environment for mining software repositories to utilize open source software as a means of productive and quality software development in companies. This study achieved the following results: (1) a database environment for automatic data collection and integration from software repositories, (2) categorization and plug-ins of repository mining techniques for analyzing a large-scale data, (3) visualization tools for monitoring analyzed results and (4) utility evaluation of the constructed environment through verification experiments.

研究分野：ソフトウェア工学

キーワード：オープンソース リポジトリマイニング ソフトウェア工学

1. 研究開始当初の背景

国内外を問わず、企業における OSS (Open Source Software) の利活用が進んでいる。製品開発の効率化と価格競争力の強化、すなわちコスト削減が主な動機付けである。しかし、品質を重視する我が国の開発企業では、未だ OSS の自社製品での利用を躊躇するケースが少なくなく、国際競争への出遅れが目立ち始めている。ソフトウェアが高機能化・複雑化するにつれ、単独の企業がすべてを開発するのではなく、OSS をベースあるいは一部に利用するという開発方式は今後加速していくものと思われる。品質重視の我が国の開発企業が安心して OSS を利用できるよう支援することは喫緊の課題であるといえる。

OSS の利活用の際に代表的な懸念事項としては以下のようなものがある[1]。

例1: 緊急時のサポート対応が迅速に行われるのか(サポート体制)

例2: 発見・報告された不具合が解消されるのか(品質保証)

例3: サポートや不具合対応が継続しておこなわれるのか(OSS プロジェクトの継続性)

これらの懸念を OSS プロジェクト外部の人間が見極めるのは非常に難しい。多くの OSS プロジェクトはボランティア開発者を中心として運営されており、不具合への迅速な対処が見込めない場合や、開発が突然中断・中止される場合などが起こり得るためである。

OSS の利活用の際に生じる懸念を払拭するためには、リポジトリマイニングを適用するのが有効である。リポジトリマイニングとは、バージョン管理システムや不具合管理システムなどに記録された膨大な開発履歴(リポジトリ)に対するデータマイニング手法を総称したものである。専門の国際会議 Mining Software Repository (MSR) が 2004 年から毎年開催されるなど、ソフトウェア工学において現在最も活発な研究分野の 1 つとなっている。前述の例に挙げた懸念を払拭するための手法も既にいくつか提案されている[2]。

しかしながら、OSS プロジェクトが利用しているリポジトリの種類は多種多様であるため、提案手法をそのまま適用できない場合も多く、企業実務者が容易に利用できる状況ではない。また、数多くの様々な手法が提案されているものの手法の体系化、特に分析の目的に沿った手法の分類化が十分ではなく企業実務者が適切な手法を選択するのが難しいという問題があった。

2. 研究の目的

そこで本研究では、ソフトウェア開発企業が OSS を安心して自社製品へ利用できるようにするために、ボランティア開発者主導で開発される OSS の品質とプロジェクトの継

続性をモニタリングできる支援環境を構築することを目的とした。

3. 研究の方法

研究目的を達成するために、以下のサブゴールを設定して研究を進めた。

- (1) OSS 開発で利用される各種ソフトウェアリポジトリから自動的にデータを収集・統合するためのデータベース基盤を構築すること
- (2) 大規模データを目的に応じて分析するためのリポジトリマイニング手法を分類すること、また、手法をプラグイン化すること
- (3) 分析結果をモニタリングするための可視化ツールを開発すること
- (4) 検証実験を通じて支援環境の有用性を評価すること

4. 研究成果

(1) データベース基盤の構築

OSS プロジェクトにおける開発支援システムとして広く利用されているバージョン管理システムのうち CVS, Subversion, Git, 不具合管理システムのうち Bugzilla, Trac, Redmine, メールングリスト管理システムのうち fml, majordomo, mailman から自動的にデータを収集するツールの作成した。また、データフォーマット変換ツールを作成し、様々な OSS プロジェクトのリポジトリデータをマイニングするためのデータベース基盤を整備した。

(2) リポジトリマイニング手法の分類とプラグイン化

まず、研究代表者がこれまで提案した以下のリポジトリマイニング手法をプラグイン形式で支援環境に適用できるよう実装をおこなった。それぞれは、OSS プロジェクトにおけるコミュニケーションの状況や、プロダクト品質を計測・可視化するためのものである。

- タイムラグ分析
- オープンネス分析
- 不具合修正予測
- 社会ネットワーク分析
- コミット予測

次に、これまで多数提案されてきたリポジトリマイニング手法を体系化し、企業実務者の分析用途に合うように整理・分類した。分類された手法の中から以下のものをプラグイン化した。

- バグモジュール予測
- 遅延相関分析
- 変化点検出

(3) モニタリング(可視化)ツールの構築

分析の目的に応じて選択したプラグインの出力(分析結果)をリアルタイムに可視化するためのツールを実装した。実装したツールは、特に分析者(実務者)の利便性を考慮

して Web ブラウザから操作できるように Web ベースのユーザインタフェースを備えている点に特徴がある。

(4) 有用性の評価

構築した支援環境を OSS として広く配布しフィードバックを得るためのアンケートを実施した。また、複数の企業実務者と共に実証実験を通じて支援環境の有用性を評価した。さらに、OSS プロジェクトのプロセス改善に対する支援環境の有用性も評価した。

< 引用論文 >

2010 年 国内オープンソースソフトウェア市場 ユーザ利用実態調査, IDC Japan 株式会社, 2010.

Arie van Deursen (Ed.), Proc. of the 8th working conference on Mining software repositories, ACM, 2011.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 9 件)

吉行 勇人, 大平 雅雄, 戸田 航史, OSS 開発における管理者と修正者の社会的関係を考慮した不具合修正時間予測, コンピュータソフトウェア, 査読有, 2015 (印刷中)

柏 祐太郎, 大平 雅雄, 阿萬 裕久, 亀井 靖高, 大規模 OSS 開発における不具合修正時間の短縮化を目的としたバグトリージ手法, 情報処理学会論文誌, 査読有, Vol.56, No.2, 2015, pp.669-681

<http://id.nii.ac.jp/1001/00102832/>
山谷 陽亮, 大平 雅雄, Passakorn Phannachitta, 伊原 彰紀, OSS システムとコミュニティの共進化の理解を目的としたデータマイニング手法, 情報処理学会論文誌, 査読有, Vol.56, No.1, 2015, pp.59-71

<http://id.nii.ac.jp/1001/00112700/>
Papon Yongpisanpop, Masao Ohira, Akinori Ihara, and Kenichi Matsumoto, Adaptive Search Framework: Better Search Result for Community, Journal of The Infosociomics Society, 査読有, Vol.8, No.2, 2014, pp.29-43

大平 雅雄, 伊原 彰紀, 中野 大輔, 松本 健一, ソフトウェア品質の第三者評価における探索的データ解析ツールの利用とその効果: OSS データを対象とした検証実験, SEC Journal, 査読有, Vol.10, No.1, 2014, pp.24-31

Emad Shihab, Akinori Ihara, Yasutaka Kamei, Walid M. Ibrahim, Masao Ohira, Bram Adams, Ahmed E. Hassan, and Ken-ichi Matsumoto, Studying Re-Opened Bugs in Open Source

Software, Empirical Software Engineering, 査読有, Vol.18, No.5, 2013, pp.1005-1042

DOI: 10.1007/s10664-012-9228-6

正木 仁, 大平 雅雄, 伊原 彰紀, 松本 健一, OSS 開発における不具合割当パターンに着目した不具合修正時間の予測, 情報処理学会論文誌, 査読有, Vol.52, No.2, 2013, pp.933-944

<http://id.nii.ac.jp/1001/00090299/>
Anakorn jongyindee, Masao Ohira, Akinori Ihara, and Ken-ichi Matsumoto, Good Or Bad Committers? - A Case Study of Committers Activities on the Eclipses Bug Fixing Process, IEICE Transactions on Information and Systems, 査読有, Vol.E95-D, No.9, 2012, pp.2202-2210

http://search.ieice.org/bin/summary.php?id=e95-d_9_2202

Passakorn Phannachitta, Akinori Ihara, Pijak Jirapiwong, Masao Ohira, and Ken-ichi Matsumoto, An Algorithm for Gradual Patch Acceptance Detection in Open Source Software Repository Mining, IEICE Transactions on Information and Systems, 査読有, Vol.E95-A, No.9, 2012, pp.1478-1489

http://search.ieice.org/bin/summary.php?id=e95-a_9_1478

[学会発表] (計 10 件)

Yosuke Yamatani and Masao Ohira, An Exploratory Analysis for Studying Software Evolution: Time-Delayed Correlation Analysis, The 6th International Workshop on Empirical Software Engineering in Practice (IWESEP2014), November 2014, Osaka (Japan)

基調講演: 大平 雅雄, ICSME (International Conference on Software Maintenance and Evolution) でのトレンド, ソフトウェア・メンテナンス・シンポジウム 2014, ソフトウェア技術者協会, 2014 年 10 月, 東京.
Akinori Ihara, Yasutaka Kamei, Masao Ohira, Ahmed E. Hassan, Naoyasu Ubayashi, and Kenichi Matsumoto, Early Identification of Future Committers in Open Source Software Projects, The 14th International Conference on Quality Software (QSIC2014), October 2014, Dallas (USA)

Yutaro Kashiwa, Hayato Yoshiyuki, Yusuke Kukita, and Masao Ohira, A Pilot Study of Diversity in High Impact Bugs, The 30th International

Conference on Software Maintenance and Evolution (ICSME2014), October 2014, Victoria (Canada)

招待講演：大平 雅雄，ソフトウェアリポジトリマイニングでできること、できないこと，第9回 SRA 技術セミナー『開発者視点でのビッグデータの活用--ソフトウェア開発を成功に導く魅力的な開発ログデータの利用へ向けて』，2014年2月，東京．

Masao Ohira and Hayato Yoshiyuki, A New Perspective on the Socialness in Bug Triaging: A Case Study of the Eclipse Platform Project, The 5th International Workshop on Social Software Engineering (SES '13), August 2013, Saint Petersburg (Russia)

Akinori Ihara, Yasutaka Kamei, Akito Monden, Masao Ohira, Jacky Wai Keung, Naoyasu Ubayashi, and Ken-ichi Matsumoto, An Investigation on Software Bug Fix Prediction for Open Source Software Projects -A Case Study on the Eclipse Project-, The International Workshop on Software Analysis, Testing and Applications (SATA) 2012, December 2012, Hong Kong (China)

Masao Ohira, Ahmed E. Hassan, Naoya Osawa, and Ken-ichi Matsumoto, The Impact of Bug Management Patterns on Bug Fixing: a Case Study of Eclipse Projects, The 28th IEEE International Conference on Software Maintenance (ICSM2012), September 2012, Trento (Italy)

Anakorn Jongyindee, Pattara Leelaprute, Masao Ohira, and Ken-ichi Matsumoto, Lessons Learned from Collaborative Research in Software Engineering: a Students Perspective, The International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD2012), August 2012, Kyoto (Japan)

Papon Yongpisanpop, Passakorn Phannachitta, Masao Ohira, and Ken-ichi Matsumoto, An Adaptive Search Framework for Supporting Cooperative Work in Organizations, The 10th Asia Pacific Conference on Computer Human Interaction (APCHI2012), August 2012, Matsue (Japan)

6 . 研究組織

(1) 研究代表者

大平 雅雄 (OHIRA, Masao)

和歌山大学・システム工学部・准教授

研究者番号：70379600