

令和 4 年 6 月 24 日現在

機関番号：14301

研究種目：基盤研究(C)（一般）

研究期間：2018～2021

課題番号：18K11241

研究課題名（和文）仮想開発者によるソフトウェア自動修正と進化推薦

研究課題名（英文）Software Automatic Repair and Update Proposal by Virtual Developer

研究代表者

渥美 紀寿（Atsumi, Noritoshi）

京都大学・学術情報メディアセンター・准教授

研究者番号：70397446

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：本研究課題の目的は、ソフトウェアの不具合修正および外部環境の変化に伴う修正にかかる開発コストを削減することである。ソフトウェアの不具合修正では、APIの典型的な呼び出し列を表現するAPI利用パターンに従うように修正することで、APIの誤った利用方法に起因するバグを自動修正するプログラム自動修正手法を提案し、実装した。また、OSSのライブラリのリリースバージョン間における公開APIの変化に関する調査を行い、ライブラリのバージョン更新を支援するための手法を提案した。

研究成果の学術的意義や社会的意義

オープンソースソフトウェアのリポジトリおよびライブラリのリポジトリから得られるすべてのリビジョンのプログラムに対して、プログラム解析技術を用いて利用しているライブラリのメソッドシグネチャと定義しているメソッドシグネチャを抽出するためのフレームワークを構築しており、ソフトウェア開発・保守支援の実現に必要な基盤を確立した点、外部ライブラリの進化に追従するための分析基盤を実現した点、API利用パターンに基づく自動修正手法を確立した点は学術的に意義があるものと考えられる。

研究成果の概要（英文）：The purpose of this research project is to reduce the maintenance cost of repairing software defects and modifying software for the changes in the external environment. We proposed and implemented a method of automatic program repair that automatically fixes bugs caused by incorrect API usage by modifying the software to follow API usage patterns that represent typical call sequences of APIs. We also conducted a survey on changes in public APIs between release versions of OSS libraries, and proposed a method to support version updates of libraries.

研究分野：ソフトウェア工学

キーワード：ソフトウェア自動修正 プログラム解析 ソフトウェア保守

1. 研究開始当初の背景

現在の技術では不具合のないソフトウェアを開発することは不可能であり、日々あらゆるソフトウェアの不具合が報告され、それに対する修正が行われている。しかし、様々なソフトウェアが稼動しており、ソフトウェアの不具合が及ぼす社会的影響は非常に大きい。そのため、ソフトウェアの不具合の早期発見、早期修正が非常に重要である。

ソフトウェアの不具合を修正するためには、その不具合を引き起こす原因となる欠陥箇所を特定し、その欠陥箇所を仕様に基づいて適切に修正する必要がある。近年、これらのことを自動的に行う、ソフトウェアの自動修正に関する研究が進んでおり、GenProg や RSRepair などの手法が提案されている。しかし、これらの手法により生成されたソースコードが適切に不具合を修正する精度はそれほど高くない。そのため、プログラム修正の精度向上が必要である。また、ソフトウェアは進化し続けるため、それに伴って修正方法も変化し続ける。そのため、修正方法の継続的な更新が必要となる。

2. 研究の目的

本研究では、ソフトウェアの修正および外部環境の変化に伴う修正にかかる開発コストを削減することを目的とする。既存のプログラム自動修正手法は、それらが提示する修正内容を自動的に適用できるほど精度が高くなく、発展途上の段階である。また、外部環境の変化に伴うソフトウェアの修正は放置されがちであり、継続的に変化する外部環境に追従していないため、外部環境における不具合が残されている可能性がある。本研究ではソフトウェアの自動修正技術の精度向上および外部環境の変化に伴うソフトウェアの修正支援を実現する。さらに、不具合検出や外部環境の変化を自動的に検出し、プログラムの不具合に対する自動修正や外部環境の変化への追従のための変更までの一連の処理を自動化するためのフレームワークを構築する。この一連の処理における個々の作業をソフトウェア開発 Bot として実現し、その Bot を本研究では仮想開発者と呼ぶ。

ソフトウェア開発のサイクルの中で、継続的インテグレーション (CI) により不具合が検出された時、開発対象のソフトウェアに影響を及ぼす外部環境の変化を検出した時、開発者に通知したソフトウェアの修正リクエストからのフィードバックを得た時のそれぞれの状況に応じて、それぞれ仮想開発者が状況に応じた処理を行うことによって、不具合の修正や外部環境の変化への対応にかかる保守コストを削減する。

3. 研究の方法

本研究では、(1) 仮想開発者のための自動修正フレームワークの構築、(2) 外部環境の変化に伴う修正支援技術の開発、(3) プログラム自動修正手法の精度向上に関して研究を行う。

(1) 仮想開発者のための自動修正フレームワークの構築

プログラムの不具合自動修正や外部環境の変化に伴う修正において、不具合の検出や外部環境の変化の検出から、修正内容を取り入れるまでの一連の処理を仮想開発者が実行し、その結果を開発者に通知するためのフレームワークを構築する。

本フレームワークでは、GitHub や GitLab などのソフトウェアの版管理サービスを中心とした開発支援サービスを用い、CI による自動ビルドやテストにおいて不具合が検出された場合、自動プログラム修正技術を用いて修正を試み、テストに成功した修正をプルリクエストとして送る。開発者はプルリクエストを確認し、問題なければ本流にマージすることによって修正を反映させる。不具合検出からプルリクエストを送るまでの一連の流れが自動化されることにより、保守コストの削減を図る。

(2) 外部環境の変化に伴うソフトウェアの修正支援技術の開発

開発中のソフトウェアが利用しているフレームワークやライブラリ(以降、外部ライブラリと呼ぶ)などの変化を継続的に解析し、開発中のソフトウェアの影響について分析し、利用バージョンの更新や、それに伴うソフトウェアの修正方法を開発者に提示するための仕組みについて研究する。

本研究では、Maven センtralリポジトリに登録されたライブラリを対象に、ライブラリのバージョンごとに提供している API を解析し、そのシグニチャ情報を抽出し、バージョンごとの変化をライブラリの進化情報として記録する。また、ライブラリを利用しているソフトウェアを GitHub から検索し、それらのソフトウェアの変更履歴において、利用しているライブラリのバージョンの特定とバージョン更新時のソースコードの変更内容を抽出し、ソースコードの修正の必要性の有無や修正が必要な場合のソースコードの修正パターンを記録し、開発中のソフトウェアでの修正支援に利用する。

(3) プログラム自動修正手法の精度向上

ソフトウェア自動修正に関する手法として、(a) 欠陥箇所を変異させながら全テストをパスするまで繰り返し、その変異コードを生成する手法や、(b) 意味解析に基づいて欠陥箇所において満たすべき制約を導出し、その制約を満たすプログラムを生成する手法、(c)

機械学習を用いて 大量のバグ修正パッチを学習することでプログラムを自動修正するルールを導出し、そのルールに基づいて修正プログラムを生成する手法など様々な自動修正手法が提案されている。本研究では既存の自動修正手法を調査し、自動修正の精度を向上させるための手法を模索する。

4. 研究成果

仮想開発者のための自動修正フレームワークを構築するために、GitHub における CI の仕組みを用いてビルドやテストの失敗時にプログラム自動修正ツールを実行し、最初に生成されたパッチを自動的に当てたソースコードを自動修正用のブランチにコミットし、プルリクエストを送る環境を構築した。自動修正ツールの中には修正されたと判断されたソースコードを複数生成するツールも存在するが、それらすべてをプルリクエストとして送るようにすることも可能であるが、その場合、開発者はすべてのプルリクエストから適切な修正かどうかを確認することが必要となり、開発者への負担が大きくなるため、本研究では自動修正ツールが最初に生成したパッチのみをプルリクエストとして送信することとした。

外部環境の変化に伴うソフトウェアの修正支援として、外部ライブラリに対する修正支援に関する研究を行った。既存の OSS ライブラリのリリースバージョン間の公開 API の変化に関する調査を行った結果、削除される公開 API の数が少ないライブラリが多数であり、多くのソフトウェアではビルドファイル等で指定している外部ライブラリのバージョン番号を更新するだけで、新しいバージョンへの移行をすることが可能であることがわかった。外部ライブラリを利用したソフトウェアにおいてライブラリの更新時にソースコードの修正が必要な場合、その修正をどの程度支援可能か調査したが、同じライブラリの公開 API に対する更新情報が少なく、支援可能な情報を既存の OSS から得ることはできないことがわかった。

ライブラリリポジトリからライブラリの進化情報を取得し、ライブラリの更新に伴う修正事例に基づいて開発中のソフトウェアに対する修正方法を提示する手法を提案し、既存ソフトウェアにおける外部ライブラリの更新状況の調査とともに口頭発表を行った。

既存のソフトウェア自動修正手法より精度を向上させるために、API 利用パターンを用いた自動プログラム修正手法を提案した。API 利用パターンを用いた自動プログラム修正は、API の典型的な呼び出し列を表現する API 利用パターンに従うように修正することで、API の誤った利用方法に起因するバグを自動修正する手法である。この手法では、事前に準備された API 利用パターンのデータベースから、バグの原因箇所周辺における API の利用方法と類似する API 利用パターンを検索し、見つかったパターンに合うようにソースコードを修正する。本手法における API 利用パターン検索に用いる類似度の定義が適切かを、既存のソフトウェアの修正履歴のうち、メソッド呼び出し 1 つを追加することで対応可能な不具合を利用し、修正後のソースコードから API 利用パターン DB を構築し、修正前のソースコードのバグパターンをキーとして DB を検索した際の修正後ソースコードに対応するパターンの類似度を計測することによって評価した。その結果、正解パターンは類似度が高くなること、正解ではないパターンを誤検出することがわかった。また、正解パターンを含むパターンデータベースを構築することが重要であるが、それを実現することが困難であることがわかった。

API 利用パターンを用いた自動プログラム修正手法および、本手法の有効性を評価するために行った API 利用パターン検索の予備評価について口頭発表を行った。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 長谷川 健太、桑原 寛明、國枝 義敏	4. 巻 37
2. 論文標題 Java Stream APIによるストリーム操作の停止性検査のための型システム	5. 発行年 2020年
3. 雑誌名 コンピュータ ソフトウェア	6. 最初と最後の頁 2_59~2_75
掲載論文のDOI（デジタルオブジェクト識別子） 10.11309/jssst.37.2_59	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 桑原 寛明、國枝 義敏	4. 巻 36
2. 論文標題 情報流解析における制約付き機密度パラメータ	5. 発行年 2019年
3. 雑誌名 コンピュータ ソフトウェア	6. 最初と最後の頁 4_39~4_45
掲載論文のDOI（デジタルオブジェクト識別子） 10.11309/jssst.36.4_39	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計12件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 桑原 寛明、渥美 紀寿
2. 発表標題 API 利用パターンを用いた自動プログラム修正におけるパターン検索の予備評価
3. 学会等名 情報処理学会ソフトウェア工学研究会3月研究会
4. 発表年 2022年

1. 発表者名 荒木 良仁、桑原 寛明、國枝 義敏
2. 発表標題 API利用パターンを用いた自動プログラム修正手法
3. 学会等名 情報処理学会ソフトウェア工学研究会3月研究会
4. 発表年 2021年

1. 発表者名 桑原 寛明
2. 発表標題 情報流解析における機密度ワイルドカードの検討.
3. 学会等名 第27回ソフトウェア工学の基礎ワークショップ
4. 発表年 2020年

1. 発表者名 荒木 良仁, 桑原 寛明, 國枝 義敏
2. 発表標題 自動プログラム修正におけるAPI利用パターンを用いたパッチ生成
3. 学会等名 第27回ソフトウェア工学の基礎ワークショップ
4. 発表年 2020年

1. 発表者名 桑原 寛明, 國枝 義敏
2. 発表標題 機密度パラメータ付き情報流解析のための型検査アルゴリズムとJavaアノテーション
3. 学会等名 日本ソフトウェア科学会 ソフトウェア工学の基礎ワークショップ
4. 発表年 2019年

1. 発表者名 渥美 紀寿, 桑原 寛明
2. 発表標題 ライブラリ進化への追従のためのソフトウェア修正の共有手法の提案
3. 学会等名 電子情報通信学会ソフトウェアサイエンス研究会3月研究会
4. 発表年 2019年

1. 発表者名 荒木 良仁, 桑原 寛明, 國枝 義敏
2. 発表標題 Stream APIを利用するJavaプログラムにおけるストリーム再利用の静的検出手法
3. 学会等名 情報処理学会ソフトウェア工学研究会3月研究会
4. 発表年 2019年

1. 発表者名 吉田 敦, 加藤 大貴, 蜂巢 吉成, 桑原 寛明, 阿草 清滋
2. 発表標題 字句列の共通性に基づく例外処理条件文の抽出手法の提案
3. 学会等名 ソフトウェアエンジニアリングシンポジウム2018
4. 発表年 2018年

1. 発表者名 夏目 雅槻, 相澤 遥也, 渥美 紀寿, 小林 隆志
2. 発表標題 ソースコードの XML表現のための選択例を用いた対話的XPath生成支援
3. 学会等名 電子情報通信学会ソフトウェアサイエンス研究会
4. 発表年 2018年

1. 発表者名 長谷川 健太, 桑原 寛明, 國枝 義敏
2. 発表標題 Java Stream API によるストリーム操作の停止性検査のための型システム
3. 学会等名 第25回ソフトウェア工学の基礎ワークショップ
4. 発表年 2018年

1. 発表者名 桑原 寛明, 國枝 義敏
2. 発表標題 情報流解析における制約付き機密度パラメータ
3. 学会等名 第25回ソフトウェア工学の基礎ワークショップ
4. 発表年 2018年

1. 発表者名 加藤 宗一郎, 吉田 敦, 蜂巢 吉成, 桑原 寛明
2. 発表標題 記述の共通性に着目したプログラムのダイジェスト化手法の提案
3. 学会等名 第25回ソフトウェア工学の基礎ワークショップ
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	桑原 寛明 (Kuwabara Hiroaki) (30432222)	南山大学・理工学部・准教授 (33917)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------