



2026年5月8日

分野：工学系

キーワード：AI、バグ混入コミット、ソースコード、ソフトウェア開発支援

「細かく追うほど誤検出が増える」を解決／
多数決で誤検出を抑える細粒度バグ混入コミット特定技術
 - AI時代の学習データ整備に貢献 -

【研究成果のポイント】

- ◆ ソフトウェアのバグの原因となった変更(バグ混入コミット^{※1})を高精度に特定する新技術「MV-SZZ」を開発
- ◆ トークン^{※2}単位での詳細な変更履歴追跡と多数決に基づく仕組みを用いて最も妥当と思われる変更を選択する仕組みを組み合わせることで、従来手法の課題であった見逃しや誤検出の大幅な削減が可能に
- ◆ ソフトウェア開発コストの低減と品質向上への貢献と高品質なデータセットの構築が可能になることでの AI 時代における学習データ整備への貢献に期待

❖ 概要

大阪大学大学院情報科学研究科の近藤偉成さん(博士前期課程)、肥後芳樹教授(コンピュータサイエンス専攻ソフトウェア工学講座)、九州大学大学院システム情報科学研究科の近藤将成助教、亀井靖高教授(情報理工学専攻)、およびカナダビクトリア大学の Daniel M. German 教授の研究グループは、ソフトウェア開発において、不具合(バグ)の原因になった変更(バグ混入コミット)を高精度に特定する新技術「MV-SZZ」を開発しました。

ソフトウェア開発では、不具合が「いつ」「どの変更によって」混入されたのかを特定することが、品質向上や開発効率の改善において極めて重要です。これまでもバグ混入コミットを特定する技術は提案されてきましたが、精度面に課題がありました。

今回の新技術「MV-SZZ」は、ソフトウェアのソースコード^{※3}をトークンの列として捉え、その変更(開発)履歴を詳細に追跡します。さらに、複数のバグ混入コミットの候補を発見した場合には、多数決に基づく仕組みを用いて、最も妥当と思われる変更をバグ混入コミットとして選択します。これにより、従来手法で問題となっていたバグ混入コミットの見逃しや誤検出を大幅に抑制し、高精度な特定を可能にしました。

この技術は、実用面ではソフトウェア開発現場における開発コスト削減と品質向上に貢献します。また研究面では、バグ混入およびバグ修正に関する高品質なデータセットの構築を可能にし、AI 時代における学習データ整備に大きく貢献します。これにより、バグ予測やバグ自動修正といった研究分野のさらなる

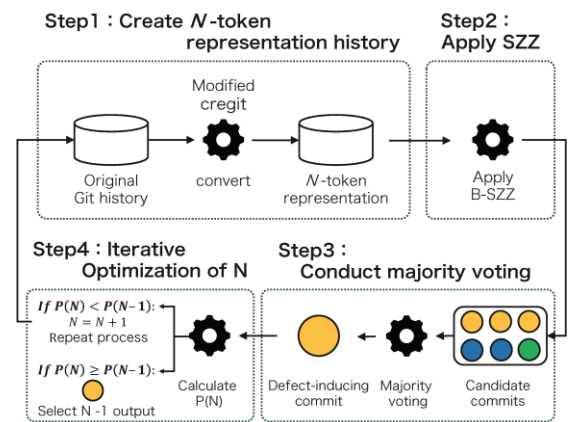


図1 新技術「MV-SZZ」の概要

発展が期待されます。

本研究成果は、米国科学誌「Transactions on Software Engineering」に、2026年5月1日(金)(日本時間)にオンライン公開されました。

【肥後教授のコメント】

バグ混入コミットを特定する最初の論文が米国で発表されてから 20 年あまり、本研究でかなり精度を高めることができたと感じています。しかしまだまだ道半ば。さらに精度をあげる研究と、特定したバグ混入コミットの活用についての研究に引き続き取り組んでいきます。

❖ 研究の背景

ソフトウェアの不具合解析や不具合予測、テスト技術の研究では、「どの変更が不具合を引き起こしたのか」を正確に特定することが不可欠です。この目的のために広く使われてきた手法が SZZ 法^{※4}です。しかし、従来の SZZ 法はソフトウェアのソースコードの開発履歴を「行単位」で追跡するため、行内の細かな修正や無関係な変更(コメントなど)を正しく区別できず、バグ混入コミットの特定精度が十分ではないという課題がありました。近年、ソースコードをトークンの列として捉えて開発履歴を追跡する改良手法も提案されましたが、今度は誤検出が増えるという新たな課題が指摘されていました。

❖ 研究の内容

研究グループでは、バグ混入コミットの見逃しや誤検出が生じる要因を体系的に分析し、以下の 2 つの工夫を組み合わせた MV-SZZ を提案しました。

1. Nトークン表現による履歴追跡

ソースコードを単にトークンに分割するのではなく、複数トークンの連なり(Nトークン)として履歴を解析することで、行内の文脈情報を保ちつつ、バグに関連する変更をより正確に捉えることができます。

2. 多数決によるバグ混入コミットの選択

各 N トークンに対して得られた複数のバグ混入コミットの候補の中から、最も多く支持されたコミットをバグ混入コミットとして選択することで、バグ混入とは無関係な変更に起因する誤検出を効果的に排除できます。

さらに、79 件のオープンソースソフトウェアプロジェクトを対象とした大規模実験を行い、代表的な 6 種類の既存 SZZ 法と比較しました。その結果、MV-SZZ は最も高い性能を示し、本手法の有効性が実証されました。

❖ 本研究成果が社会に与える影響(本研究成果の意義)

本研究成果は、ソフトウェア開発現場およびソフトウェア工学研究の双方において以下の重要な意義を持ちます。

● 開発コストの削減と品質向上に貢献

実際の開発現場において、欠陥の原因特定を迅速かつ正確に行えるようになり、デバッグ作業の効

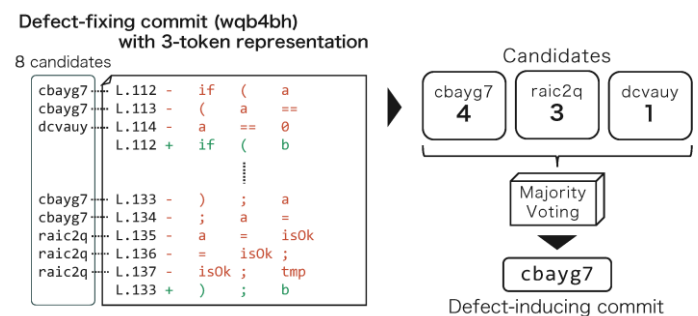


図 2

多数決を用いたバグ混入コミット選択の様子

率化に寄与します。

- 高品質な欠陥データセットの構築を支援

欠陥導入コミットをより正確に特定できることで、欠陥予測や自動修復研究で用いられるデータセットの信頼性が向上します。

- 既存手法全体の改善指針を提示

多数決という汎用的な考え方が、他の欠陥解析手法にも有効であることを示し、ソフトウェアリポジトリマイニング研究の発展に貢献します。

❖ 特記事項

本研究成果は、2026年5月1日(金)(日本時間)に米国科学誌「IEEE Transactions on Software Engineering」(オンライン)に掲載されました。

タイトル: MV-SZZ: An Empirical Study of a Majority Voting-Based SZZ Method

著者名: [Inase Kondo](#), [Masanari Kondo](#), [Daniel M. German](#), [Yasutaka Kamei](#)
and [Yoshiki Higo](#)

DOI: <https://doi.ieeecomputersociety.org/10.1109/TSE.2026.3688089>

❖ 用語説明

※1 バグ混入コミット

ソフトウェアは多数のコミットを積み重ねて進化していくため、不具合の原因を調べる際には、どのコミットが不具合を引き起こしたのかを特定することが重要となる。本研究で扱うバグ混入コミットとは、そのような履歴の中で不具合の原因となったコミットを指す。

コミットとは、ソフトウェア開発において、プログラムの変更内容をひとまとまりとして記録する単位のこと。開発者は、機能の追加や不具合の修正などを行うたびに、その変更内容を「コミット」として履歴に保存する。各コミットには、「いつ」「誰が」「どのファイルを」「どのように変更したか」といった情報が記録されており、ソフトウェアの開発過程を時系列で振り返ることができる。これにより、過去の状態に戻したり、変更の理由を確認したりすることが可能になる。

※2 トークン

トークンとは、ソースコードを構成する最小単位の要素のこと。プログラムの中では、変数名や数値、演算子、記号(かっこやセミコロンなど)がそれぞれ意味を持っており、これら一つ一つがトークンとして扱われる。例えば、1行のプログラムであっても、内部的には複数のトークンに分解できる。ソースコードをトークン単位で捉えることで、「行全体」ではなく、「どの要素がどのように変化したか」をより細かく分析することが可能になる。

※3 ソースコード

ソースコードとは、ソフトウェアの動作内容を人間が理解できる形で記述したプログラムの設計図のこと。ソフトウェア開発者は、プログラミング言語を用いてソースコードを書き、それをコンピュータが実行できる形に変換することでソフトウェアを動かしている。

ソースコードには、計算処理の手順だけでなく、条件分岐やデータの扱い方など、ソフトウェアの

振る舞いに関するすべての情報が含まれている。そのため、ソフトウェアの不具合(バグ)は、ソースコード中の記述ミスや設計上の問題として現れる。

本研究では、このソースコードがどのような変更履歴(コミット)を経て現在の形になったのかに注目し、過去のソースコードの変更を分析することで、不具合の原因となった変更を特定する技術を対象としている。

※4 SZZ 法

ソフトウェア開発において「バグを生み出した変更(バグ導入コミット)」を特定するための分析手法。

❖ 参考 URL

肥後芳樹教授 研究者総覧 URL

<https://rd.iai.osaka-u.ac.jp/ja/74aa4a70cc06fb47.html>

❖ 本件に関する問い合わせ先

<研究に関するお問い合わせ>

大阪大学 大学院情報科学研究科 コンピュータサイエンス専攻ソフトウェア工学講座

教授 肥後芳樹(ひご よしき)

TEL: 06-6879-4108 FAX: 06-6879-4109

E-mail: higo@ist.osaka-u.ac.jp

<広報に関するお問い合わせ>

大阪大学 大学院情報科学研究科 研究戦略企画室

TEL: 06-6105-5907

E-mail: ura-press@ist.osaka-u.ac.jp

❖ 発信先 報道機関

大阪大学から 大阪科学・大学記者クラブ、文部科学記者会、科学記者会

九州大学から 九州大学記者クラブ