

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年 6月20日現在

機関番号：20103

研究種目：挑戦的萌芽研究

研究期間：2009～2011

課題番号：21650008

研究課題名（和文） REMIX:開発者視点のソフトウェア分析ツールキット

研究課題名（英文） REMIX: A Toolkit from Developer's Standpoints

研究代表者

神谷 年洋 (KAMIYA TOSHIHIRO)

公立ほこだて未来大学・システム情報科学部・准教授

研究者番号：70415660

研究成果の概要（和文）： ソフトウェア工学の研究領域で、ソースコード及び実行トレースをリバースエンジニアリングの手法により分析するツールを試作し、ソフトウェア工学に関する国際ワークショップなどで研究発表を行った。一部の機能については、オープンソースの公開レポジトリ上で公開している。本研究において開発したツール群は引き続き開発を行い、より汎用的なツールキットとして実現する予定である

研究成果の概要（英文）： Aiming to researches in software engineering, tools have been developed and presented at workshops on software engineerings. Some features of the tools are hosted in an open-source public repositories. The tools developed in this research projects are going to further development, in order to implement a (more) general purpose toolkit.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,200,000	0	1,200,000
2010年度	600,000	0	600,000
2011年度	600,000	180,000	780,000
年度			
年度			
総計	2,400,000	180,000	2,580,000

研究分野：ソフトウェア

科研費の分科・細目：ソフトウェア工学

キーワード：ツール、プログラム理解、リバースエンジニアリング、保守、検索駆動型開発

1. 研究開始当初の背景

(1) ソフトウェア開発プロセスにおいて開発者が扱うべきドキュメントの種類が増えてきた。（例えば、三階層ウェブフレームワークを用いた場合、プレゼンテーション層はHTMLやJavaScriptによるライブラリ、ロジック層ではJavaなどのプログラミング言

語、データベース層ではSQLなどのデータベースに関する言語などが単一のウェブアプリケーションを実装するために用いられる。）これを受けて、開発者の開発スタイルも、ドキュメントから特定の知識を発見する作業が多くの割合を占めるようになった。

(2) 特定ドキュメントや記述に対する開発、

検索手法は用意されているが、必ずしも相互運用を想定しておらず、相互運用の方法は開発者の手作業によっている。したがって、(1)で上述したような開発プロセスにおける開発者の作業負担は増大した。

2. 研究の目的

本研究プロジェクトでは、開発者が多様なドキュメントを扱うことを要求するソフトウェア開発において、リバースエンジニアリング手法やツールを、開発者が自らの必要に応じて相互運用できるようなツールキットを提供することを目的とした。

3. 研究の方法

既存のリバースエンジニアリングの手法、ツールを相互運用可能な形で提供する手段を確立するため、以下のステップを順に遂行することにより、ツールキットの開発を目指す。(1)既存のリバースエンジニアリングの手法やツールで用いられるモデル(ドキュメント)の関係を定義した上で、相互運用できるようなパーツに分解する。(2)それらのパーツを組みあわせて、相互接続できるようなツールキットを提供する。さらに、研究期間に、Hadoopに代表される大規模並列計算による解析アルゴリズム・手法に著しい発展があったため、それらを取り入れるべくアルゴリズムの工夫などを盛り込んだ。

4. 研究成果

現在までに、ソースコード、コンフィギュレーションファイル(Makefile)、バージョン管理システム(のコミットログ)、コールグラフについてそれらの関係を定義し、分析手法を提案し結果を発表した。

(1) ソースコードとコンフィギュレーションファイルの関係

コンフィギュレーションファイルは、どのソースコードがプロダクト中で利用されるか、利用されないか(したがってデッドコードとなるか)を決めるだけではなく、C言語によって記述されるソースコードにおいては特に、ソースコード中のマクロ置換に影響を与え、実質的にソースコードを変更するなどの目的に利用される。ソースコードに対して行われるコードクローン検出手法の結果に、コンフィギュレーションファイルによって指定されるデッドコードの情報を合わせて解析することで、保守性に影響を与える可能性が高いコードクローンを特定しようと試みた。

(2) バージョン管理システムとソースコード

バージョン管理システムの差し戻しコマン

ド(revert)による影響を取り除いた上で、どの程度のコードが過去のバージョンからコピーされるかを調べた。

(3) コールグラフとソースコード

コールグラフ(関数の呼び出し関係)という、設計ドキュメントレベルの粒度の構造が、重複コード除去という作業に役立つかを調べた。

現在はさらに、実行トレースやHTMLによるユーザーインターフェイスの記述との関係について研究を進めている。

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6件)

① Toshihiro Kamiya, “How Code Skips Over Revisions,” Proc. Fifth International Workshop on Software Clones (IWSC 2011), pp. 69-70 (May 23, 2011). doi 10.1145/1985404.1985420
査読有り

② Toshihiro Kamiya, “Identifying the concepts that are searchable with keywords in code search engines,” Lecture Notes in Artificial Intelligence 6284, pp. 116-123. (2010). 査読有り

③ 神谷年洋, “FreeBSDのソースコードレポジトリにおけるリビジョンを超えたコード断片のコピーの分析”, 電子情報通信学会技術研究報告書, vol. 110, no. 227, pp. 61-65 (2010).
査読なし

④ 神谷年洋, “バージョン管理システムのレポジトリを対象とするコード検索ツールの試作”, 電子情報通信学会技術研究報告書, vol. 110, no. 169, pp. 59-63 (2010).
査読なし

⑤ Toshihiro Kamiya, “Classifying code clones with configuration”, Proc. Forth International Workshop on Software Clones (IWSC2010), pp. 75-76 (2010). doi 10.1145/1808901.1808912
査読有り

⑥ 神谷年洋, “リバースエンジニアリングツールキット Remics の試作”, 電子情報通信学会技術研究報告 vol. 109, no. 170, pp. 7-11 (2009).
査読なし

[学会発表] (計1件)

①神谷年洋, “コードクローン検出技術と応用”, 情報処理学会 第18回組込みシステム研究発表会 (2010/08/09), 函館市

[その他]

①神谷年洋, 肥後 芳樹, 吉田 則裕, “コードクローン検出技術の展開”, コンピュータソフトウェア, Vol.28, No.3, pp.29-42, 2011年8月. 解説記事

② Toshihiro Kamiya, “A code-clone detection tool CCFinder and its successor(s)”, マイクロソフトリサーチ 日本情報学研究賞 受賞者記念講演会, (2010/07/26), 記念講演

③神谷年洋, “コードクローン研究の進化と支援技術の展開”, 第21回 SRA-KTL Technology Seminar, 築地 (2009/10/14), 招待講演

6. 研究組織

(1) 研究代表者

神谷 年洋 (KAMIYA TOSHIHIRO)

公立はこだて未来大学・システム情報科学部・准教授

研究者番号 : 70415660