

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3月31日現在

機関番号：34304

研究種目：若手研究（B）

研究期間：2008 ～ 2011

課題番号：20700029

研究課題名（和文） バースマークの評価技術の確立

研究課題名（英文） Establishment of Evaluation Methods of the Birthmark

研究代表者

玉田 春昭 (TAMADA HARUAKI)

京都産業大学・コンピュータ理工学部・助教

研究者番号：30457139

研究成果の概要（和文）：

実世界でのソフトウェアの盗用を簡単に発見するため、バースマーク技術に着目した。提案されているバースマークを調査し、特に簡易に利用できる使用クラスバースマークに着目した。客観的に評価するため、使用クラスバースマークを用いて、ソフトウェアの分類を試みることにより、当該バースマークの評価手法が確立できた。また、初心者の書いたプログラムの盗用を防ぐための特徴を見つけ出すべく、それらを収集・分析した。

研究成果の概要（英文）：

In this work, we focused on birthmark techniques to detect software theft. We investigated proposed birthmark techniques and we focused used class birthmark because it is most easily in use. A method for objective evaluation of used-class birthmark was established in use of classification method. Furthermore, we investigated the features of program written by novices to prevent program plagiarism.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	900,000	270,000	1,170,000
2009年度	700,000	210,000	910,000
2010年度	700,000	210,000	910,000
2011年度	700,000	210,000	910,000
年度			
総計	3,000,000	900,000	3,900,000

研究分野：総合領域

科研費の分科・細目：情報学・ソフトウェア

キーワード：プログラムの盗用、プログラム改変、ソフトウェアバースマーク、難読化、ソフトウェア保護、耐タンパ

1. 研究開始当初の背景

ソースコードが公開されているソフトウェアが増加している反面、それらを不正に利用してソフトウェアが作成されるソフトウェアの盗用が問題となっている。ソフトウェアが盗用された場合、盗用者はソースコードを公開しない。そのため、ソースコードを対象としたソースコードの類似性計測は利用

できない。

盗まれたソフトウェアを発見するための技術としてバースマークが提案されている。バースマークとは、個々のソフトウェアの動作に不可欠な情報をソフトウェアのバイナリ形式から抽出したものである。その情報を比較することで、ソフトウェア間の類似性を計測できる。

しかし、バースマークの取り出し方は議論されているものの、それぞれのバースマークの発見・立証能力は議論されておらず、不明である。また、バースマークそれぞれの攻撃耐性については述べられておらず、詳細な調査が望まれる。

2. 研究の目的

本研究では、バースマークの客観的評価方法の確立および、盗用を発見するための簡易な手法の確立を目的とする。

3. 研究の方法

- (1) まず、現実世界の攻撃者の行動パターンや持ちうる技術の調査を行う。それら技術により、プログラムがどのように改変されたのか調査を行う。
- (2) 次に、プログラム改変に対するバースマークの耐性を調べる。バースマークは静的解析により得られるものと、動的解析により得られるものがあるため、プログラム改変も、静的改変、動的改変の両方について調査する。
- (3) バースマークの評価方法を導出する。
- (4) バースマークを容易に扱えるよう、試作システムを作成する。試作システムは広範に使ってもらえるよう OSS として公開する。
- (5) また、初心者の書いたプログラムが盗まれないよう、初心者の書くプログラムに特化したバースマークの調査を行う。そのために初心者の書くプログラムがどのような傾向を持つのかを調査する。

4. 研究成果

- (1) 攻撃者の行動パターンや持ちうる技術を調査するため、現実社会で起こっている盗用の事例をまとめた（[その他]③）。また、行動パターンを調査する過程で、従来不可能と言われていた Java 言語を対象とした新たな動的改変を用いた攻撃パターンを開発した。実行時に動的にプログラムを改変しつつ実行する方法であり、その手順をまとめて発表した（[雑誌論文]③）。これを実行した場合、静的解析と動的解析で得られる結果が異なるようになるため、動的バースマークの更なる耐性が求められる。本手法は従来、自己書き換え手法として知られており、プログラムの保護方法として用いられてきた。しかし、Java 言語ではプログラムはロード後書き換えられないという制約のため、不可能とされていた。ただし、Java 5 に動的に書き換える方法であるインスツルメンテーションが導

入されたため、Java 言語でも自己書き換えが可能となった。自己書き換え手法とは、命令の一部をあらかじめ偽の命令に置き換えておき、その命令が実行される直前に本来の命令に書き換え、実行後は再度偽の命令に書き換えるものである。概念を図 1 に示す。

本手法の従来の自己書き換え手法との違いは大きく 2 点ある。1 つは書き換えの単位である。従来手法は 1 命令が書き換えの単位であるが、本手法はクラスファイル全体が書き換えの単位となる。つまり、復帰ルーチン RR で書き換える対象が命令ではなく、クラスである。もう 1 つの違いは、書き換え手段の違いである。従来手法の RR や HR は書き換えのための命令が必要であるが、この命令は書き換え専用の面例ではない。一方の本手法では、書き換え専用のクラスライブラリが必要となる。そのため、従来手法より本手法が使われたことが察知され

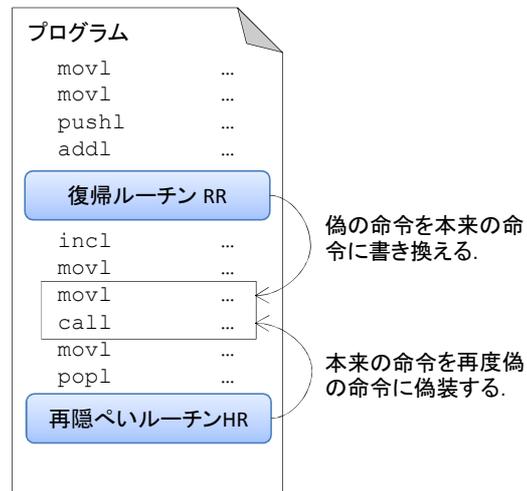


図 1 従来の自己書き換え手法の概念

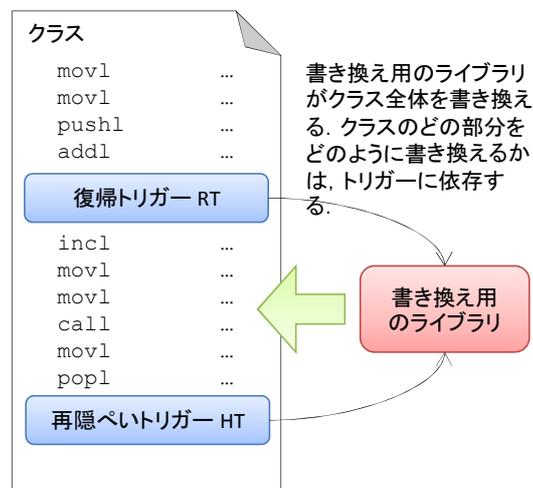


図 2 Java 向け自己書き換え手法の概念

やすい。概念図を図2に示す。書き換えの処理は毎回同じであるが、どの部分を書き換えるのかはどこからのトリガーであるのかにより異なる。この2つの違いは、バースマークに対する攻撃手段として見た場合、攻撃手法が容易に特定できるため、利点である。一方、一般的な保護手法としての観点から見れば、攻撃手法が容易に特定可能であることは欠点であると言える。また、動的にプログラムを書きかえるため、実行時間に100倍程度の大きな遅延がみられた。これらの欠点への対応は、今後の課題である。

- (2) プログラム改変に対するバースマークの耐性を調査するため、プログラムに既存の難読化手法を適用した。それら難読化したプログラムに対し、プログラムの読みやすさを評価するための仮想メンタルシミュレーションモデル (VMSM; Virtual Mental Simulation Model) を用いて難読化されたプログラムの評価を行った。VMSMとは、変数の値を憶えること、忘れることをFIFOキューを用いてシミュレートするモデルである。図3にVMSMの模式図を示す。プログラムを読み進めるために変数名とその値をキューに入れる、すなわち、覚える。そして変数を参照するときにキュー内にその変数が存在しなければ読みにくくなる。VMSMではプログラム理解のコストやプログラムを間違っ理解する可能性を表す動的メトリクスが提案されている。そのメトリクスを用いて、難読化前後の

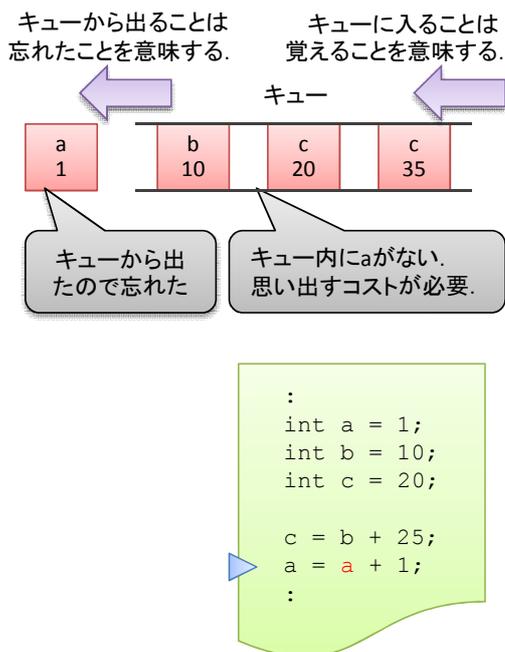


図3 VMSMの模式図

プログラムを比較した。用いた難読化手法は動的名前解決難読化 ([その他] ②) の他, Sandmark に実装されている5種類の計6種類の難読化手法である。その結果、各難読化手法の特徴がメトリクスの増減に現れることがわかった ([雑誌論文] ①)。このことからVMSMのメトリクスの変化から使われた隠ぺい方法が推定できる可能性もあり、バースマークの改変への対応につながると考えられる。

- (3) 提案されているバースマークのうち、使用クラスバースマークに着目して、ソフトウェアの機能面での分類を試みた。クラスの種類と出現頻度に基づいてソフトウェア間の距離を算出し、階層的クラスタリングにより分類する。重み付けにtr-idf法を用いて、ユークリッド距離を用いた場合に最も妥当な分類結果となった。この結果は、似たソフトウェアをまとめるという意味で盗用されたソフトウェアを見つけ出す手段として応用できる。この分類法は、使用クラスバースマークの評価方法として有効である ([雑誌論文] ⑧)。
- (4) 従来多くのバースマーク手法が提案されているが、ツールとして利用できるバースマークは少ない。そこで、Java言語を対象として、バースマークをプラグイン形式で導入可能なツール Stigmata を開発し、オープンソースソフトウェアとして公開した ([その他] ①)。このツールは種々のバースマーク手法をプラグインとして導入できる。また、クラスを単位として指定したバースマークを抽出でき、また、抽出したバースマークを比較できる。そして、その結果である類似度を一覧表示できる。プラグイン開発者は入出力のことは考えず、1つのクラスからバースマークを抽出する部分、また、抽出されたバースマークを比較する部分に集中して開発できる。また、どのようなバースマークでも使い方は同じであるため、このツールの使い方を一度理解すると、どのようなバースマークであっても同じように利用できるようになる。これにより、誰もが容易に盗用の疑いのあるクラスファイルを見つけ出せるようになることを期待できる。
- (5) 従来は、一般的なプログラムを対象としたバースマークが提案されていた。一方、プログラミング演習などでの剽窃が問題となっている。そのため、初心者の書いたプログラムも盗用を防ぐ手段が必要である。その一手段のため、バースマークを適用したい。しかし、プログラムの規模が熟練者の書くプログラムに比べて小さく、従来のバースマークでは偶

然一致してしまう可能性が高い。そのため、初心者の書いたプログラム向けに新たなバースマークを開発すべく、初心者の書いたプログラムの傾向を分析した（〔雑誌論文〕②）。その分析のために、コンパイル単位でのソースコード収集フレームワークを構築した（〔雑誌論文〕④⑤など）。具体的なバースマークの提案は今後の課題であり、その候補として、プログラムの編集履歴に基づいたバースマークが考えられる。

以上のようにバースマークへの攻撃方法や耐性の調査方法、評価方法など、バースマークに関する多岐に渡る研究を行い、国内外に向けて成果を発表した。提案した手法はバースマークの評価やバースマークの耐性に限らず、一般的なプログラムの保護方法などにも応用可能であり、今後のソフトウェア保護技術の一要素として用いられることが期待できる。

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9件）

- ① 玉田 春昭他6名, ``メンタルシミュレーションモデルを利用した動的名前解決難読化手法の読みにくさ評価'', 電子情報通信学会 2012 年総合大会 D-3-3, March 2012. (査読無)
- ② 伏田 享平, 玉田 春昭他3名, ``プログラミング演習における初学者を対象としたコーディング傾向の分析'', 信学技報ソフトウェアサイエンス研究会, Vol. 2012-03-SS, pp. 67--72, March 2012. (査読無)
- ③ 玉田 春昭, 神崎 雄一郎, 門田 暁人, ``Java 言語を対象とした実行時多様化の試み'', 2012 年暗号と情報セキュリティシンポジウム予稿集 (SCIS2012), January 2012. (査読無)
- ④ Haruaki Tamada, Akihiro Ogino, Hirota Ueda, ``A Framework for Programming Process Measurement and Compile Error Interpretation for Novice Programmers'', In Proc. The Joint Conference of the 21th International Workshop on Software Measurement (IWSM) and the 6th International Conference on Software Process and Product Measurement (Mensura) (IWSM/MENSURA 2011), pp. 233-238, 3-4 November 2011. (Nara, Japan) (査 読 有) DOI: <http://doi.ieeecomputersociety.org/1>

0.1109/IWSM-MENSURA.2011.23

- ⑤ Akihiro Ogino, Haruaki Tamada, Hirota Ueda, ``Phynocation: A prototyping of a teaching assistant robot for C language class'', In Proc. HCI International 2011 (HCI2011), 9-14 July 2011. (Orlando, Florida, USA) (査読無)
DOI: 10.1007/978-3-642-21657-2_64
- ⑥ 荻野 晃大, 玉田 春昭, 上田 博唯, ``ロボットをアシスタントとして利用する教育支援システムの研究'', 京都産業大学総合学術研究所所報, 第6号, pp. 45--54, July 2011. (査読無)
- ⑦ 玉田 春昭, 荻野 晃大, 上田 博唯, ``アシスタントロボットを用いたプログラミング教育支援システムの構築'', 信学技報 マルチメディア・仮想環境基礎研究会, Vol. MVE2010-48, pp. 143-148, June 2010. (査読無)
- ⑧ 牛窓 朋義, 門田 暁人, 玉田 春昭, 松本 健一, ``使用クラスに基づくソフトウェアの機能面からの分類'', 信学技報ソフトウェアサイエンス研究会, Vol. SS2009-17, pp. 31-36, August 2009. (査読無)
- ⑨ Haruaki Tamada, Akihiro Ogino, Hirota Ueda, ``Robot Helps Teachers for Education of the C Language Beginners, '' In Proc. HCI International 2009 (HCI2009), 19-24 July 2009. (San Diego, CA, USA) (査読無) DOI: 10.1007/978-3-642-02577-8_41

〔学会発表〕（計2件）

- ① 玉田 春昭, ``ソフトウェアの盗用を防ぐ方法, 発見する方法'', 京阪神地区大学新技術説明会, 22 November, 2011 (大阪).
- ② 玉田 春昭, ``ソフトウェアバースマーク'', 京都産学公連携フォーラム, 5 November, 2008 (京都).

〔産業財産権〕

○出願状況（計1件）

名称：教育支援システムと教育支援方法
発明者：上田 博唯, 荻野 晃大, 玉田 春昭
権利者：大学法人 京都産業大学
種類：特許
番号：特開 2010-231580, 特願 2009-079368
出願年月日：平成 21 年 3 月 27 日
国内外の別：国内

〔その他〕

- ① 静的バースマーク抽出・比較ツール Stigmata.

- <http://stigmata.sourceforge.jp/>
- ② 動的名前解決難読化ツール
DonQuixote.
<http://se.naist.jp/DonQuixote/>
- ③ ソフトウェア盗用の事例 Wiki
<http://cafebabe.jp/piracywiki/>

6. 研究組織

(1) 研究代表者

玉田 春昭 (TAMADA HARUAKI)
京都産業大学・コンピュータ理工学部・助
教
研究者番号：30457139

(2) 研究分担者

なし

(3) 連携研究者

なし