

## 論文内容の要旨

博士論文題目 ソフトウェア盗用を発見するための動的バースマーク

氏名 岡本 圭司

今日、オープンソースソフトウェアの普及に伴い、公開されているソースコードを利用したソフトウェア開発が可能となった。その一方で、公開ソースコードの使用許諾条件（ライセンス）を遵守せず、密かに開発に用いるという新たな形態のソフトウェア盗用が問題となっている。このようなソフトウェア盗用では、盗用者はソースコードを公開しないため、盗用の事実を発見または立証することは非常に困難である。

この問題に対して、バースマークと呼ばれる、個々のソフトウェアに固有の特徴量を抽出し、盗用の発見に役立てようとする研究が行われている。しかし、ソフトウェアの静的な特徴量を用いる従来のバースマークは、Java プログラムには有効であるが、機械語プログラムに対しては利用することが難しい。Java プログラムでは、ソースコードの様々な特徴量が実行プログラムにも残存するため、実行プログラムから抽出したバースマークを用いて、ソースコード盗用の有無を検知できる。一方、一般的な機械語プログラムでは、コンパイラの種類が多く、コンパイラやコンパイルオプションの違いによって生成されるコードが全く異なるものとなる。しかも、コンパイルによってソースコードの特徴量の多くが失われるため、従来のバースマークを盗用の発見に利用することは難しい。

そこで、本論文では、ソフトウェアの実行時に API (Application Program Interface) の呼び出しを観測し、その情報を利用したバースマークを提案する。このようなソフトウェア実行時に得られる情報を用いたバースマークを「動的バースマーク」と呼ぶ。一般に、OS が提供する API は、ファイル操作などの OS の機能を実行するための唯一のインターフェースであるため、あるソースコードをコンパイルして得られる機械語プログラムが呼び出す API の系列や API ごとの呼び出し回数は、用いたコンパイラの種類に関わらず、ほぼ同一のものとなると期待される。そのため、動的バースマークを機械語プログラムから抽出することで、ソースコード盗用の発見に有用である。

本論文では、提案バースマークの有効性を確認する 2 つの実験を行った。1 つめの実験では、提案手法はバースマークが備えるべき性質である保存性と弁別性を備えていることを確認した。2 つめの実験では、機械語プログラムが異なるコンパイラおよびコンパイルオプションによって作成された場合でも、提案バースマークへの影響は軽微であることを確認した。

また、提案バースマークの消去を試みる攻撃に対する耐性の評価を行った。提案手法は、従来から存在する難読化などの機械的変換ツールに対する耐性が期待できるが、人手による攻撃には弱点を持つことが分かった。そこで、編集距離を用いて API の挿入と削除による攻撃を無効にするバースマークの比較方法を提案し、評価実験によりその有効性を確認した。

氏名	岡本 圭司
----	-------

(論文審査結果の要旨)

本論文では、ソフトウェアの盗用を発見するための動的バースマークを提案し、その有効性と攻撃耐性の評価を行っている。さらに、攻撃耐性を向上させるための方法を提案している。

論文中では、まず、オープンソースソフトウェア(OSS)のソースコードを盗用するという新しいタイプの問題が近年発生していることに着目し、現状を調査してその実態を明らかにしている。そして、ソフトウェアの盗用を発見するための従来技術として、ソースコードの類似性を用いる手法、ウォーターマーク、バースマークについて調査し、これらの技術を OSS の盗用に利用する上での問題点を整理した。新しい現実的な問題に着目し、現状の技術とその課題を整理した点には、研究者としての高い問題発見能力、及び、問題整理能力が認められる。

次に、本論文では、ソースコードの入手できない機械語プログラムに対して利用できる動的バースマークを提案している。提案手法では、ソフトウェア実行時の API 呼び出しを記録し、その系列や頻度をバースマークとする。API とは OS がファイル入出力などの機能をアプリケーションに提供するためのインターフェースであり、他に代替手段がないため容易に削除できず、バースマークとして用いる情報として有用である。また、今日普及している Windows プラットフォーム上の機械語プログラムから抽出可能であるため、実用的である。提案手法は、初めて提案された動的バースマークの一つであるのみならず、実用的なプログラムに対して使用できる初めての動的バースマークであり、高い新規性・有用性が認められる。

提案手法の評価においては、オリジナルのソフトウェアとその派生バージョンのソフトウェア、独立に開発されたソフトウェアを用いた実験により、提案手法がバースマークに必要とされる保存性と弁別性を持つことを示している。また、同様に、実験を通じて、提案手法はコンパイラやコンパイルオプションの違いによる影響を受けにくいことを示している。機械語プログラムでは、使用するコンパイラやコンパイルオプションにより大きく異なるバイナリプログラムが作成されるため、従来は、保存性のあるバースマークを構成することが困難であった。提案手法ではこの点が克服されており、高い有用性が認められる。また、実験を通じて保存性や弁別性を評価していることから、提案手法の信頼性は高いといえる。

さらに、本論文では、提案手法の攻撃耐性についても明らかにしている。まず、提案手法は難読化などの等価変換に耐性を持つことを明らかにしている。次に、手作業による攻撃に対しては弱点が認められることを明らかにし、攻撃を受けた場合においても盗用を発見するための新たなバースマーク比較方法を提案している。提案手法を評価してその限界を明らかにした上で、現実的な解を示した点は、提案方法の実用性をさらに高めている。

以上のとおり、本論文は、ソフトウェアの盗用を発見するための実用的な技術を提案するものである。この研究成果により、機械語プログラムのソースコード盗用を発見することが可能となり、盗用の大きな抑止力となることが期待される。よって、これらの研究成果は、ソフトウェアセキュリティの分野に大きく貢献するものであり、本論文は博士(工学)論文として価値あるものと認める。