

ワードやエクセルといったアプリケーションソフト、電子メール、インターネットブラウザから映像、音楽、ゲームソフトといったコンテンツにいたるまで、ソフトウェアは生活のあらゆる場面で欠かすことのできない存在になっています。ソフトウェアを対象とする「ソフトウェア工学」の中でも、ソフトウェアやその内部の秘密情報を保護する「ソフトウェアセキュリティ」の分野で活躍される玉田春昭先生に、盗用発見の新技术「バースマーク」について話していただきました。

コンピュータサイエンス学科
玉田 春昭 助教

バースマーク(プログラム指紋)による盗用発見

増える盗用事例 “プログラム指紋”で盗用を見つけ出す

現在、ソフトウェアの違法コピーや盗用の事例は、意図しないものを含めて、後をたちません。プログラムを暗号化する、簡単に読めないように難読化するという保護手段が多数考え出されてきましたが、新しい防御方法を考え出されては破られるというたちごっこが続いています。また、盗用を発見すること自体も難しくなってきました。

そこで私は、盗用発見の新しい技術として、ソフトウェアの実行に不可欠な特徴に注目した「バースマーク(Birthmark)」を提案しました。“指紋”を応用して考え出した方法で、着想は推理小説から得ました。当初は「プログラム指紋」という名を使っていましたが、ある先生から「これはバースマークだね」と教えられて、Groverが定義していた「バースマーク」を盗用発見の道具として定義し直して使うことになりました※1。

バースマークから ソフトウェアの類似性をはかる

盗まれたことを証明する手段としては、著作権情報などの情報をソフトウェアに密かに埋め込んでおく「電子透かし」の技術がよく知られています。しかし、「盗用発見」の手段としては、壊される可能性やあらかじめ埋め込んでおかなければ

いけない点から不十分です。

ソフトウェアの設計図であるソースコードのコピーを見つけたり類似度を測ったりする方法も提案されていますが、ほとんどのソフトウェアはソースコードをコンピュータ上で実行可能なバイナリレベル※2に変換した形で出回っています。そのため、変換された状態で比較できることが求められているのです。

そこで私は、①難読化も含めたプログラム変換を行っても変化しないこと(保存性)、②全く同じように機能するプログラムであっても、別の人が独立して作った場合、正しく区別できること(弁別性)、という2つの性質を満たすものとして、バースマークを定義しました。具体的な内容は省略しますが、2003年に4種類のバースマークを提案し、上記の性質を満たすことを確かめました。

現在、17のバースマークをJavaクラスファイルから抽出して比較できるツール“stigmata”を、インターネット上で公開しています※3。平均的なサイズのソフトウェアであれば、10秒かからず比較ができます。

ソフトウェアの保守と バースマークの応用

現在、バースマークは、アメリカや韓国、ドイツなどでも提案されていて、合わせて30以上存在しています。バースマークの種類が増えると、互いに弱点を補うことができるものの、比較にかか

る時間は長くなってしまいます。それでも他の手法よりは短い時間でできますが、より高速で無駄なく比較ができるように「どういう種類のソフトウェアには、どういうバースマークを組み合わせるべきか」がわかる、バースマークの評価ツールの開発を目指しています。

盗用発見以外への応用も考えています。ソフトウェアはメンテナンス時に、元のデータのコピーに手を加えて再開発することが少なくありません。これまで、新しい部分の不具合は人の記憶や感覚的なもので確認されていました。しかし、stigmataを使えば、新しくなった部分を特定して、重点的にテストすることが可能になります。その他、プログラミング課題のコピーの発見などにも使えます。

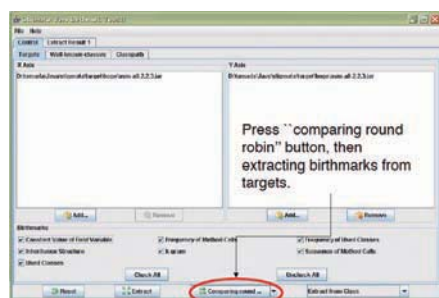
現在、あらゆる分野でソフトウェアが使われています。その数は今後さらに増えるでしょう。こうした中で、他人の作成したソフトウェアを許可なく自分のソフトウェアに組み込んだり(盗用)、内部の秘密情報を改ざんしたり(クラック)することを抑止・防止するソフトウェアセキュリティの分野は、ますます重要な役割を担うようになると思います。

※1 Derrick Groverの「The protection of computer software its Technology and application」(1989年)の中で「ソフトウェアが元来持っている特徴」という意味で定義づけられている。

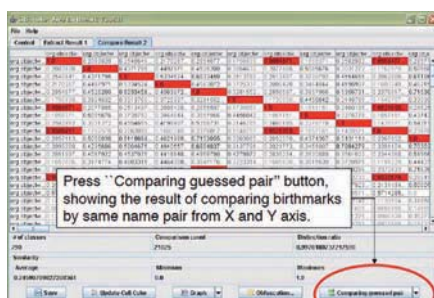
※2 「0」と「1」で構成される、機械(コンピュータ)が実行可能な表現形式。

※3 <http://stigmata.sourceforge.jp>で公開されている。

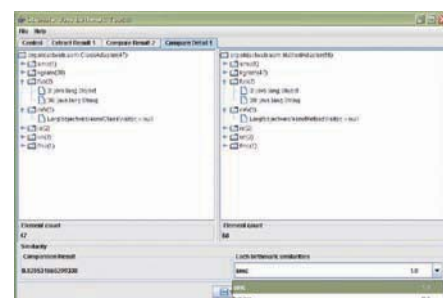
図 Stigmataの比較画面



Stigmataで2つのソフトウェアの類似度を測定するインターフェース



ソフトウェアの部品の総当たり比較結果。赤い部分が高いもの(対角線状に赤くなっているのは同じ部品同士を比べているため)



類似性計算の根拠となる情報を表示する画面