| | |
|---|---|
| 氏名(本籍・生年月日) | 岩 本 一 樹 (静岡県 昭和50年10月12日) |
| 学位の種類 | 博 士 (工 学) |
| 学位記番号 | 甲 第 ６２８ 号 |
| 学位授与の日付 | 平成２７年３月２０日 |
| 学位授与の要件 | 信州大学学位規程 第５条第１項該当 |
| 学位論文題目 | 静的解析に基づくマルウェア分類システムに関する研究 |
| 論文審査委員 | 主査 教 授 和﨑克己 |
| | 教 授 師玉康成 |
| | 教 授 山本博章 |
| | 教 授 丸山 稔 |
| | 教 授 西垣正勝 (静岡大学) |

## 論 文 内 容 の 要 旨

マルウェアはコンピュータセキュリティにおける脅威の1つである．マルウェア対策を行うためには，マルウェアを解析してマルウェアの動作・機能を知る必要がある．

しかし近年，マルウェアは圧縮・暗号化（パック）されマルウェアのコードが隠蔽されている場合が多くあり，パックされた検体は復号しなければ静的解析できない．またパックされた検体は，元の検体がほとんど同じ場合でもバイナリイメージは異なる．それに加えて，ソースコードを少し改変しただけの亜種やコンパイル環境の違いによりバイナリイメージが異なる亜種が大量に作られている．そのため，似たようなマルウェアが大量に存在することになる．

さらにマルウェア感染において，文書ファイルを開くアプリケーションの脆弱性を攻撃する文書型マルウェアが用いられる場合がある．通常，アプリケーションの脆弱性への攻撃が成功すると文書型マルウェアの内部にあるシェルコードが実行される．このシェルコードが文書型マルウェアの内部にあるデータをファイル化する，あるいはファイルをダウンロードする等の手段で実行可能ファイルを作成して開く．文書型マルウェアを動的解析するためには文書型マルウェアが想定する脆弱性をもつアプリケーションが必要になる．しかし文書型マルウェアが実行できるアプリケーションを特定する作業は困難であり，またアプリケーションが準備できるとは限らない．

これらの問題のために，マルウェア解析は困難になっている．文書型マルウェアとしてマルウェアが配布される場合には，文書ファイルが作成する実行可能ファイルを取得する必要がある．実行可能ファイルがパックされているならば，それを復号して元のコードを得る必要がある．さらに大量にあるマルウェアのコードを分類することで，既に解析済みのマルウェアとの類似度から機能を推定したり，解析対象となるマ

ルウェアの優先順位を決める，あるいは解析を行うために適した技術者を選定できれば，マルウェア解析において大いに役立つことが期待できる．よって，本論文では文書型マルウェアからシェルコードを抽出する方法，パックされたマルウェアを展開してオリジナルのコードを抽出する方法，大量のマルウェアを分類する方法を提案する．

文書型マルウェアに対しては，Microsoft Officeの文書ファイルから文書ファイルを開くアプリケーションを用いずにシェルコードを特定する方法について述べる．文書ファイルのエントロピーに基づいて，シェルコードの候補となるバイト列をエミュレータで実行することで効率よくシェルコードを特定することを試みる．特定したシェルコードを動的解析するための実行可能ファイルを作成することで，文書ファイルを開くアプリケーションで脆弱性が攻撃されてシェルコードが実行された状態を再現する．これにより，シェルコードによって作成される実行可能ファイルを得ることができる．本論文では実際の文書型マルウェアを用いて，シェルコードの特定を試みて成否を確認した．また特定したシェルコードを実行して動作を観測した．

本論文では圧縮・暗号化された実行可能ファイルを復号する方法についても述べる．本論文で提案する方法では，独自に作成したエミュレータにより実行可能ファイルの復号コードを実行することで元のバイナリイメージを取得する．このエミュレータは完全にソフトウェアで動作するのでマルウェアが動作する環境を用意する必要がない．復号に合わせてエントリーポイントの特定とインポートテーブルの再構築を行う．エントリーポイントの特定は自身が書き換えたメモリの実行を検知する方法，および既知のエントリーポイントのコードと比較する方法を用いる．本論文では実際のマルウェアの中で圧縮・暗号化された検体に対して復号とエントリーポイントの特定，インポートテーブルの再構築を試みた．また既知のプログラムをパックして，復号が可能であるか調査した．

大量にあるマルウェアを効率よく解析するために，本論文ではマルウェアを分類する方法について述べる．提案する方法ではマルウェアの分類のために，逆アセンブルした結果を制御フロー解析し，あるAPIが呼び出された後に呼び出される可能性のあるAPIの対（API推移）を抽出する．API推移の推移を比較することで，マルウェア間の類似度を算出して階層型クラスタ解析を行いマルウェアを分類する．本論文では実際のマルウェアの分類することでパフォーマンス評価を行い，マルウェアのソースコードをコンパイラおよびコンパイラのオプションを変更してコンパイルしたバイナリから類似度を求めることで性能評価を行った．

最後に本論文では，マルウェアの静的解析に対して本論文で提案する方法の有効性についてまとめ，今後の課題について述べる．