ハニーポットとバイナリコード解析の連携によるネットワーク攻撃の自動防御技術に関する研究（091603006）

Research on Automated Defense Methods against Network Attacks based on Honeypots and Binary Code Analysis

研究代表者
森彰　産業技術総合研究所
Akira Mori　The National Institute of Advanced Industrial Science and Technology (AIST)

研究分担者
高倉弘喜[†]　泉田大宗[††]　橋本政朋[†††]
Hiroki Takakura[†]　Tomonori Izumida[††]　Masatomo Hashimoto[†††]
[†]名古屋大学　[††]産業技術総合研究所　[†††]コーディニュアムソフトウェアラボ
[†]Nagoya University　[††]AIST　[†††]Codinuum Software Lab

研究期間　平成 21 年度～平成 23 年度

概要
基盤ネットワーク上で運用されるハニーポットでの攻撃の捕捉・分析と、二次攻撃のために投入される悪意あるプログラム（マルウェア）の解析の自動化を組み合わせることで、過去に前例のない未知の攻撃に対してもその攻撃の遮断および防御の対策をリアルタイムで同定する技術を確立することを目的とする。囮であることを察知されないようにするためのハニーポット管理・運用技術、および、エミュレーションと静的解析を組み合わせたマルウェア自動解析技術を中心に研究開発を行った。

Abstract
The project aims at establishing technologies for instantly identifying countermeasures against unknown network attacks by integrating automated malware binary code analysis with attack sensing by honeypots deployed in a backbone network. The research focuses on methods of stealth operations of sensitive decoy computers for honeypots and methods of seamless integration of dynamic emulation and static binary code analysis for malware.

１．まえがき
ネットワーク攻撃を発生端緒で同定し、攻撃に関する情報を収集することは、一般への被害を未然に防ぐ上で重要である。本研究では、実際のネットワーク上でハニーポットと呼ばれる囮コンピューターを運用することで、ネットワーク上で発生する攻撃の実態を把握するとともに、投下される悪意あるプログラム（マルウェア）の自動解析と組み合わせることで、それがどのような攻撃であり、どういった対策が有効であるかを早期に知るための、自動防御システムを実現するための技術の確立を目指すものである。

２．研究内容及び成果
①ハニーポット技術に関する研究
実ハードウェア上に直接 OS を搭載し、実ネットワークに接続可能なハニーポットを開発した。特に、攻撃対象のネットワーク環境やサンドボックス環境等を確認してから、システム乗っ取りを試みる不活性型の攻撃に対処するため、環境調査と推定される通信は許可し、それ以外の攻撃にまつわる通信（DoS 攻撃参加やスパムメール送信など）を遮断する通信可否制御モジュールを開発した。外部への攻撃に伴う通信を確認した時点で、ハニーポットの実行イメージと初期起動イメージを比較して、投下されたと推定される攻撃プログラムを自動抽出することができる。ハニーポット上で観測された攻撃の情報は、サニタイズ後、Web サーバを通じて国内外の研究者にベンチマークデータとして公開した。また、迷惑メールを受信し、本文中に埋め込まれたリンクから攻撃プログラムが存在する可能

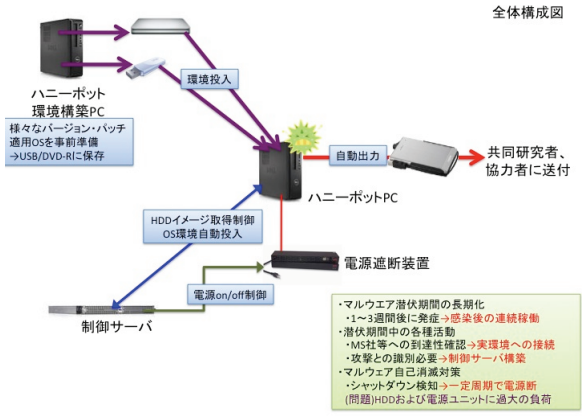

図１：ハニーポットシステム全体構成

性の高い Web サイトへのアクセスを自動的に行うハニーポットシステムを実装した。最大処理能力は、5 分平均で毎分 80 件であり、実運用（平成 23 年度）において、30 分平均で毎分 10 件の迷惑メールを受信し、合計 1,968,593 の URL を抽出して 20,153 件のマルウェアファイル（疑われるものも含む）を収集した。こうして集められる大量の攻撃データを処理するには、攻撃が既知か未知かを判定し、未知の攻撃の分析に注力することが重要である。この点について、これまでより多くの特徴量を考慮することで、従来手法で未知攻撃として検知できなかった 14 種類の攻撃のうち、10 種類を未知攻撃として検知できるようになった。近年多様化しつつある攻撃に対処するために、Linux

OS を標的としたネットワーク攻撃の捕捉を行うハニーポットシステム、および IPv6 で動作するハニーポットシステムの開発も行った。図１に開発したハニーポットシステムの全体構成を示す。研究期間中、当該システムを用いて、ゼロデイ攻撃を含む様々なネットワーク攻撃を捕捉することができた。

②マルウェアバイナリコード解析に関する研究

ネットワーク攻撃の全貌を知るには、対象マシンの制御を奪うために投下される攻撃プログラムの振舞いを解析することが不可欠である。しかし、こうしたマルウェアの解析は、ソースコードが不在のバイナリ形式で行わざるを得ず、しかも暗号化や難読化さらにはデバッガ/エミュレータ検知といった障害を克服しつつ、実行される可能性のある制御パスを漏れなく解析しなければならない。また、近年のマルウェアでは、複数のプロセスやスレッドにまたがって感染・常駐や乗っ取り、外部攻撃が行われることが一般的であるため、実際に実行させながら、実行条件により実行されない制御パスの解析も行う、というように相反する解析手法の併用を迫られる。本研究では、Intel アーキテクチャにおけるハードウェア実行仮想化(VMX)を利用したハイパーバイザーモニター(HVM)と連携して、マルウェアバイナリコードのエミュレーションと静的解析を連動させることにより、実行時の動的情報を参照しながら、実行される可能性のあるコードを網羅的に解析し、マルウェアの全体の振る舞いを自動分析するツールを開発した。ツールは、オープンソースの HVM である Xen を拡張することで実現されている。図２にマルウェアバイナリコード解析部の概要図を示す。

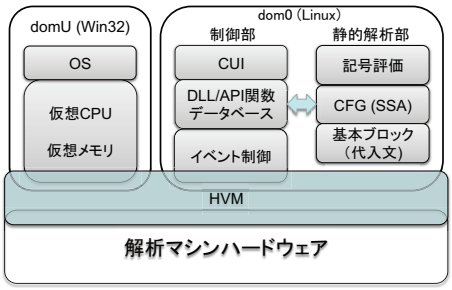

図２：マルウェアバイナリコード解析部

解析対象となるマルウェアを仮想実行環境で起動すると、当該プロセスの実行がそのエントリーアドレスで中断され、そこからバイナリコードの静的解析が行われる。静的解析は、機械命令を解析しながら直接ジャンプ命令を可能な限り追跡し、基本ブロックからなる制御フロー図を生成する。基本ブロック内の機械命令はそれぞれの操作的意味を反映した代入文の列に変換され、さらに静的単一代入（SSA、Static Single Assignment）形式と呼ばれる形に変換される。その上で、残された間接ジャンプ命令の飛び先アドレスを記号的に計算して、可能な限り制御フローを展開していく。そして、これ以上制御フローの展開ができなくなった段階で、仮想環境における実行を再開し、仮想実行がそれまでに作成された制御フローからはみ出した時点で再び静的解析を再開する、というように、仮想実行と静的解析を繰り返しながらバイナリコードの解析を行う。この際、事前に作成された外部 DLL と API 関数のデータベースを参照しながら、API 関数への呼び出しとそこからの復帰を同定し、引数情報とともに制御フロー図の形で解析対象の振る舞いを明らかにする。このように、間接ジャンプを展開しながら制御フローを構築する手法を「展開型静的解析」と呼び、これと仮想実行エミュレーションと組み合わせることで、自己解凍型のマルウェアコードや難読化が施されたコード、および条件分岐等で実行されないコードなども解析することが可能になる。こうした解析を、プロセスごと、スレッドごとに多重化すれば、投下された攻撃プログラムから起動された別プロセス、さらには他のプロセスに注入された遠隔スレッドのコードも自動的に解析することも可能になる。結果として、実際のボットネットワーム等について、それらがアクセスするホストのアドレスやコマンド・コントロールの制御フローを明らかにすることが可能になった。図３に実際にハニーポットで捕捉されたボットネットワームの制御フローの一部を、図４に対応する実行トレースの一部を例示する。図３の灰色ノードは、実際に実行はされなかったものの、静的解析で明らかになったコード部分を表している。

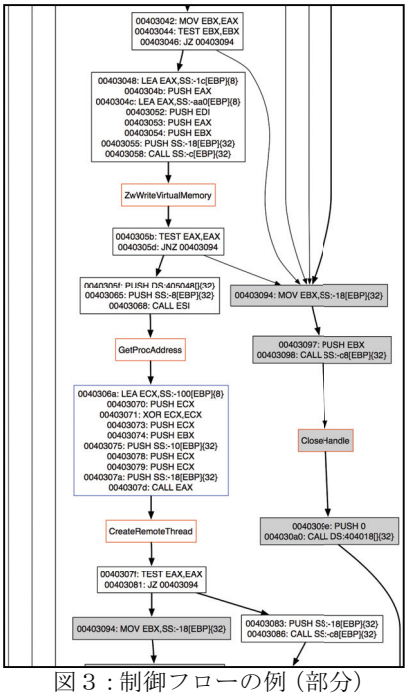

図３：制御フローの例（部分）

```
ZwWriteVirtualMemory(1932,66387968,1242400,2168,1245092)=00000000
GetProcAddress (2088763392,'CreateRemoteThread')=7c810626
CreateRemoteThread(1932,0,0,66060288,66387968,0,1244864)=00000764
CloseHandle(1932)=00000001
ExitProcess(0)
```

図４：実行トレースの例(部分)

３．むすび

本研究では、実環境で稼働する高感度なハニーポット技術を開発し、そこで捕捉された攻撃プログラムのバイナリコードを自動解析することで、どのアドレスから攻撃が行われて、どのような攻撃プログラムが投下され、どのような二次攻撃が行われるか、といった攻撃情報を自動的に分析する技術の開発を行った。要素技術の開発はすでに完了しており、今後は応用へ向けた統合と処理効率の向上、さらなる技術の高度化に取り組む予定である。

【誌上発表リスト】

[1]泉田大宗、森彰、二木厚吉“展開型静的解析と動的解析を連携させたマルウェア解析手法”、コンピューターソフトウェア Vol.29 No.4 （2012 年 10 月、掲載予定）

[2]Moriteru Ishida, Hiroki Takakura, and Yasuo Okabe, “High-Performance Intrusion Detection Using OptiGrid Clustering and Grid-Based Labelling”, Proc. of the 11th IEEE/IPSJ Int. Symp. on Applications and the Internet (SAINT2011), pp11-19, (2011.07.20)

[3]Tomonori Izumida, Kokichi Futatsugi and Akira Mori, “A Generic Binary Analysis Method for Malware”, Proc. of the 5th International Workshop on Security (IWSEC2010), Lecture Notes in Computer Science, Vol.6434 pp119-216 (2010.11.23)