脅威情報とホワイトリストを用いたサイバー攻撃自動対処システムの研究

２０１９年度　先端数理科学研究科

博士学位請求論文（要旨）

脅威情報とホワイトリストを用いたサイバー攻撃自動対処システムの研究

先端メディアサイエンス専攻
重本　倫宏

１　問題意識と目的

近年，民間企業や政府機関，制御システム等の重要インフラを狙ったサイバー攻撃が顕在化しており，個人，企業，国家それぞれの利益や安全性を損なうリスクが高まっている．特に APT（Advanced Persistent Threat）攻撃は，秘密裏に，そして執拗に長期間攻撃を続ける点で従来の脅威とは異なり，マルウェアの侵入を検知して完全に防御することは不可能になりつつある．例えば，2015 年 6 月には，日本年金機構において遠隔操作型マルウェアに感染した職員の端末から基礎年金番号を含む個人情報が約 125 万漏えいし，大きな社会問題となった．

従来のマルウェア対策は，セキュリティベンダなどが提供しているシグネチャ（マルウェア検査パターン）に基づくアンチウイルスソフトにより行われていた．しかし，次々と発生する新種のマルウェアや高度な検知防止機能を持つマルウェアに対して，ベンダ側の解析やシグネチャ作成が追いつかず，即時性のある脅威情報入手ができていない問題があった【問題１】.

このように高度化するサイバー攻撃に対し，組織間で解析結果等の脅威情報を共有して攻撃に備える集団防御の概念が浸透してきた．一部ではマルウェアを解析して得た攻撃に悪用されるサイト情報を共有して，そこに対するインターネット接続を即座に遮断する自動対策技術も実用化されている．しかし，マルウェアの中には，自身がインターネットと通信可能か否かを判断するために，実行初期に正規のサーバに対して疎通確認を行うものが存在する．このようなマルウェアに自動対策技術を適用すると，業務に悪影響を及ぼす可能性があるため，導入に踏み切れず結果的に対策が後手に回る問題があった【問題２】.

本論文では標的型攻撃に対する早期対策の重要性に鑑み，業務影響を考慮した自動対処システムを提案する．ここで前述した２つの問題に対応して以下を目的とする．

【目的１】即時性のある脅威情報を入手が困難であるという問題１に対して，マルウェア解析の高度化と公開情報の構造化により，攻撃者による遠隔操作や情報搾取を防止するために有用となる脅威情報を抽出すること．

【目的２】抽出した脅威情報を用いて対処すると業務に悪影響を与える場合がある問題２に対して，業務への悪影響を抑えつつ，マルウェアによる被害発生リスクを軽減すること．

２　構成及び各章の要約

本論文は 9 章により構成されている．1 章では研究の背景と目的を述べる．2 章では，従来研究を述べる．3 章では高度化するマルウェアの解析技術，4 章では公開情報から脅威情報を収集する技術を提案する．5 章では 3 章，4 章で得られた脅威情報の脅威度を判定する技術を提案する．6 章では脅威情報を活用し，業務悪影響を抑えた自動対処技術を提案する．6 章で述べた対処技術で対応できない脅威に対して，ホワイトリストを活用した自動対処手法を 7 章で提案する．8 章では，ホワイトリストの最適化手法を提案する．9 章では，本論文をまとめる．

1 章要約（研究背景と目的）

高度化するサイバー攻撃に対して現状のセキュリティ対策の問題点を整理し，本研究の目的と課題を設定した．まず目的１に対し，マルウェア解析を高度化すること【課題１】と，公開されている脅威情報を構造化すること【課題２】を設定した．また目的２に対し，脅威情報の脅威度推定精度を向上すること【課題３】と，業務に悪影響を与えずにマルウェアの通信を遮断すること【課題４】を設定した．さらに，本論文で提案する自動対処システムを手動対処，CAPTCHA による対処，淵上らの手法と比較し，従来手法が不審サイトへの接続リスクを軽減するか，業務への影響を軽減するかのどちらかに偏っていたのに対し，提案システムが業務への影響を抑えつつ接続リスクを軽減できる点で新規であることを明らかにした．

2 章要約（基本定義と従来研究）

本論文を通して使用する基本定義を示すとともに，マルウェア解析やインテリジェンス収集，WEB サイトの脅威度推定，自律進化型防御システム，ブラックリストを用いた対処，ホワイトリストを用いた対処の従来研究について述べた．

3 章要約（マルウェア解析高度化）

マルウェア解析を高度化するために，多種環境におけるマルウェアの動的解析結果を用いたマルウェア分類手法を提案した．提案手法は，解析環境間での動作の違いや API コール列等を特徴量とし，深層学習（Recurrent Neural Network）によって解析対象を既知のマルウェアファミリに自動分類するものである．プロトタイプを用いた評価実験により，8,243 検体のマルウェア解析結果を 134 種類のファミリに分類できること，既知のマルウェアを約 64%の精度で分類できることを確認した．また，各マルウェアファミリへ所属する確率を利用することによって，新種マルウェアを抽出できることを確認した．これにより課題１を解決した．

4 章要約（脅威情報収集）

高度化するサイバー攻撃に追従し，適切な対処を行うため，公開されているセキュリティベンダのサイトやブログ等の非構造データから脅威情報を収集して構造化する手法を提案した．セキュリティ分野では，次々と新しい表現（新しいマルウェア名や，新しい脆弱性名等）が現れるため，固有表現を認識することが難しかった．提案手法はこの点に着目し，第１候補が固有表現でないと認識された際に，第２候補を固有表現として抽出するものである．プロトタイプを用いた評価実験により，従来手法より高精度で固有表現を認識できること，従来手法では見逃していた未知語を認識できていることを確認した．また，モデルの再学習時間や１記事に対する固有表現認識に要する時間を計測し，実運用にも耐えうることを示した．これにより課題２を解決した．

5 章要約（脅威度推定）

脅威情報の中には，正規サイトが誤って含まれている場合があり，仮に業務遂行に必要な正規サイトが含まれていると，当該サイトにアクセスできず，業務阻害の要因となってしまう．このため，機械学習をもちいた脅威度の自動推定手法を提案した．提案手法は，URL 文字列や，ドメイン登録に関する情報，DNS 情報などから定めた 35 の特徴量と，3 つの推定器を用いて脅威度の推定を行う．プトトタイプを用いた評価実験では，悪性サイトと良性サイトをそれぞれ 50,000 件用いて学習を行い，脅威度を 92.74%の精度で推定できることを確認した．これにより課題３を解決した．

6 章要約（脅威情報を用いた自動対処）

課題４の解決に向けて脅威情報を用いた自動対処システムを提案した．提案システムは，マルウェアを動的解析することにより脅威情報を自動抽出し，セキュリティベンダの提供する外部データベース（VirusTotal）を参照することで，脅威情報の確信度を算出する．さらに，組織における過去の接続ログから脅威情報を対処した場合の影響を算出し，確信度と影響度に基づき自動対処の適用可否を決定するものである．プロキシと連携したプロトタイプを用いた評価実験により，464 秒でマルウェアの動的解析から対処適用までが完了すること，実マルウェア（732 検体）から 508 件の脅威情報を抽出し，そのうちの 491 件を自動対処できることを確認した．

7 章要約（ホワイトリストを用いた自動対処）

6 章で提案した自動対処システムでは対処できない攻撃へ対応するため，ホワイトリストを用いた自動対処システムを提案した．提案システムは，ホワイトリストに定められた接続先以外には追加認証を要求することによって，人間による意図的な通信は許可するとともに，その認証結果を用いてホワイトリストの精度を高めていくものである．提案システムを用いることで，未知のサイトを利用した攻撃にも対応することが可能となる．プロトタイプを用いた評価実験により，遠隔操作型マルウェアの通信を遮断できること，ユーザの規模が少ないうちは，業務に与える影響が高い（追加認証要求率 1.93%）が，ユーザ規模が 1,000 人程度になると業務に与える影響が抑えられることを確認した．また，プロトタイプの処理性能を測定することで，提案システムが 1,000 人程度のユーザ処理に耐えうることを確認した．6 章及び 7 章で提案したシステムにより，課題４を解決した．

8 章要約（ホワイトリスト最適化）

ホワイトリストを用いた対策では，時間経過とともにホワイトリストの質が低下してしまうため，ホワイトリストの質を最適化する手法を提案した．具体的には，F-measure 評価指標を参考に，誤検知及び検知見逃しリスクの観点から，ホワイトリストの質を定量評価する評価指標を策定した．また，遺伝的アルゴリズムを用いて，与えられたログを基に評価指標の点で最適なホワイトリストを作成する方式を提案した．さらに，運用期間中にホワイトリストの再学習が必要なタイミングを判断する手順を検討した．26 端末の活動ログを用いた評価実験により，誤検知・検知見逃しの両点で比較対象方式より 15%以上優れたホワイトリストを作成できることを確認した．また，本技術により，低負荷で再学習タイミングを判断できることを確認するとともに，10,000 台の端末に対するホワイトリスト運用を実現できる見込みを示した．

9 章要約（結論）

4 つの課題を解決するための手法提案，実装，実験評価を通し，マルウェア解析の高度化と公開されている脅威情報を収集することで目的１を達成し，脅威情報を対処した際の影響度に基づいて自動対処の可否を判断すること，ホワイトリストに定められた接続先以外には追加認証を要求することによって人間による意図的な通信は許可すること，により目的２を達成した．