早稲田大学大学院 基幹理工学研究科

博 士 論 文 概 要

論 文 題 目

Analyzing Hidden Features of Web-based Attacks

Web 感染型攻撃における
潜在的特徴の解析法

申 請 者

Yuta TAKATA

髙田 雄太

情報理工・情報通信専攻 情報システム工学研究

2017 年 7 月

情報化社会の進展に伴い、インターネット上に蓄積された機密情報や金銭を狙うサイバー攻撃の脅威が増加している。攻撃者は、インターネットを通じて他人のクライアントを悪意あるソフトウェア(マルウェア)に感染させることにより、感染したクライアントを不正に操作し、情報の窃取や改ざん、破壊活動を行う。多くのインターネットサービスが Web を通じて提供される現代において、Web は主要なマルウェア拡散経路となっている。

Web を通じたマルウェア感染攻撃は、ドライブバイダウンロード攻撃と呼ばれる。ドライブバイダウンロード攻撃は、攻撃者が用意した悪性 Web サイトへユーザのアクセスを誘導し、Web ブラウザやプラグインの脆弱性を悪用することにより、最終的にクライアントをマルウェアに感染させる。ドライブバイダウンロード攻撃への対策には、ホスト上での対策とネットワーク上での対策がある。ホスト上での対策は、アンチウイルスソフトを用いたマルウェアの検知が挙げられ、ネットワーク上での対策は、URL ブラックリストを用いた悪性通信の検知が挙げられる。これらの対策技術のほとんどは、悪性 URL や攻撃コード、マルウェア等の事前に収集した悪性情報に基づき攻撃を検知する。悪性情報は、悪性 Web サイトへの能動的なアクセスにより収集したり、ネットワークトラヒックの受動的な監視により収集したりする。いずれの方法も悪性情報の収集には有効であるが、受動的な収集方法には観測点が限定的であることやプライバシーの問題があるため、能動的な収集方法が普及している。悪性 Web サイトへの能動的なアクセスには、意図的に攻撃を受ける囮（おとり）のクライアントシステム（ハニークライアント）が用いられる。ハニークライアントは、攻撃を受けることにより発生するシステム上の変化や Web サイトの特徴に基づく機械学習により、悪性 Web サイトを検知する。ただし、攻撃者は対策技術を回避するように行動する。巧妙に細工された攻撃は、既存のハニークライアントによる解析や情報取得を妨害するよう設計されており、悪性情報の露見を防いでいる。このようにハニークライアントにより悪性情報を収集できない場合には、従来の対策技術は有効に機能しない。

本研究は、巧妙化するドライブバイダウンロード攻撃に対応して、ハニークライアントの解析性能を向上させることにより、悪性 Web サイト検知のための情報をより多く収集する方法を提案する。具体的には、悪性 Web サイトにおける４つの巧妙化技術(1) コンテンツ難読化、(2) 環境依存攻撃、(3) 多段転送、(4) Web サイト改ざん、の各々の特徴を捉える解析手法を提案する。提案手法を用いれば、悪性 Web サイトから得られる情報を最大化できる。実働するネットワーク上のデータに基づいて攻撃解析手法の設計と実装を行い、その有効性を実証する。

第１章では、本研究の背景、問題、目標、ならびに研究成果の概要を示す。

第２章では、従来の対策技術ならびに攻撃の巧妙化技術の概要を示す。対策技術は、解析に実環境を使用する高対話型ハニークライアント、およびエミュレー

タを使用する低対話型ハニークライアントの両方を説明する。本研究で実際に用いるのは、攻撃や悪性コンテンツをより詳細に解析できる高い拡張性を持った低対話型ハニークライアントである。また、ハニークライアントで収集した情報に基づき、攻撃を検知する機械学習技術について説明する。攻撃コードの構造等の静的な特徴量や攻撃コードの挙動や多段転送の構造等の動的な特徴量を用いた機械学習による様々な攻撃検知手法を、最新の研究内容を含め記述する。

さらに第2章では、攻撃の巧妙化技術として、コンテンツ難読化や多段転送、環境依存攻撃、Web サイト改ざんについて説明する。攻撃者は、悪性 Web サイトに関連する情報を隠蔽するために、悪性コンテンツに難読化を施したり、多段転送により悪性 URL を頻繁に変更したりする。また、多段転送の過程で、あらかじめ定めた OS やブラウザのみを標的とする環境依存攻撃を仕掛けたり、Web サイト改ざんにより一般の Web サイトを攻撃に組み込んだりする。このような攻撃の巧妙化に対応するためのハニークライアントおよび機械学習技術における課題について記述する。

第3章では、環境依存攻撃に対して、各環境を攻撃するために用意された URL を抽出する手法を提案する。ドライブバイダウンロード攻撃における多段転送は、攻撃の起点となる入口 URL、クライアントのアクセスを中継する踏台 URL、攻撃コードを含む攻撃 URL、そしてマルウェアをホストするマルウェア配布 URL のように、異なる役割を担う複数の URL で構築され、攻撃のスケーラビリティ向上や運用コスト削減に寄与する。環境依存攻撃は、この多段転送の過程で、JavaScript を用いたブラウザフィンガープリンティングによりクライアントの OS やブラウザ等の環境を識別して、その結果に応じて転送先の URL を変更する。ハニークライアントを使用する従来技術は、その環境が攻撃対象に一致しない場合に、攻撃 URL やマルウェア配布 URL 等の悪性 URL へ転送されない。すなわち、悪性 URL へ転送されないハニークライアントは、攻撃コードやマルウェア等の悪性コンテンツを取得することができず、攻撃を検知できない。

そこで第3章では、JavaScript にプログラムスライシングを適用することにより、各環境に用意された転送コードを特定し、それらの実行結果から URL を抽出する手法を提案する。この手法は JavaScript の実行網羅性を向上させることにより、条件分岐文等で本来実行されないコードからも URL を抽出することができる。また、このような転送コードの多くは、難読化が施されている。これに対応するため、URL アクセス時に直接取得できる静的なコードに加え、難読化解除により動的に生成されるコードにも適用できるようブラウザエミュレータを用いて本手法を実装した。本研究では、このブラウザエミュレータがハニークライアントの役割を果たす。事前に収集した悪性 Web サイトの HTTP 通信データを用いて提案手法の評価を行い、通常のアクセスよりも多くの URL を抽出できることを示す。

第4章では、多段転送やJavaScript実行の追跡により、改ざんされたWebサイトにおける悪性コンテンツを特定する手法、ならびに複数のクライアントの環境模擬により攻撃対象を特定する手法を提案する。多段転送における入口URLは、悪性URLへ転送するよう改ざんされたWebサイトが悪用されることが多く、攻撃を疑わないユーザからのアクセスを奪取する。Webサイト改ざんを無害化するためには、そのWebサイト管理者によるインシデント対応が必要不可欠である。しかしながら、改ざんされたWebサイトのURLだけの情報が管理者に報告される従来技術では、多段転送や環境依存攻撃等の巧妙な攻撃に対して、管理者によるインシデント対応のための情報が不足している。

改ざんされたコンテンツは、攻撃コードやマルウェアとは異なり元来は無害なWebサイトに含まれるコンテンツであるため、ハニークライアントを用いて観測できる。改ざんされたコンテンツの特徴やWebサイト改ざんにより発生する転送、攻撃対象のクライアントの環境等の情報は、迅速なインシデント対応のために重要であるとともに、攻撃検知にも有用な情報である。

そこで第4章では、攻撃者により改ざんされたコンテンツを特定するとともに、当該コンテンツにより攻撃の脅威に晒されるクライアントの環境を特定する手法を提案する。具体的には、ブラウザエミュレータにより多段転送やJavaScript実行を詳細に追跡して、どのコンテンツがどのURLへ転送するかを特定する機能を実装する。本手法は、複数のクライアントの環境を模擬し、Webサイトを解析することにより、どのようなクライアントの環境が脅威に晒されるのかを特定する。本手法を評価するために、前章と同様の実験環境を用いて悪性WebサイトのHTTP通信データを収集して解析した。提案手法により、改ざんされたコンテンツや攻撃対象のクライアントの環境情報を効果的に特定できることを示す。

以上に述べたとおり、本研究では悪性Webサイトにおける4つの攻撃巧妙化技術の特徴を明らかにした。ハニークライアントを用いてそれらを解析する手法を提案して実装した。いずれの手法も巧妙化するドライブバイダウンロード攻撃が普遍的に持つ特徴を捉えることができる実用的な手法であり、実データを利用した評価によりその有効性を確認した。本研究の提案手法により、ハニークライアントの解析性能を飛躍的に向上することができる。さらに、本研究の成果に基づき改良したハニークライアントで収集した情報は、アンチウイルスや機械学習等の既存技術における悪性Webサイトの検知精度向上にも寄与する。本研究の成果は、よりセキュアなWebサイトを実現するために有用である。

早稲田大学　博士（工学）　学位申請　研究業績書

氏 名　髙田 雄太　　印

(2017 年 7 月 現在)

種 類 別	題名、　発表・発行掲載誌名、　発表・発行年月、　連名者（申請者含む）
○論文	Yuta Takata, Mitsuaki Akiyama, Takeshi Yagi, Takeshi Yada, and Shigeki Goto, ``Fine-grained Analysis of Compromised Websites with Redirection Graphs and JavaScript Traces,'' IEICE Transaction on Information and Systems, Vol.E100--D, No.8, pp.1714--1728, August 2017.
○論文	Yuta Takata, Mitsuaki Akiyama, Takeshi Yagi, Takeo Hariu, and Shigeki Goto, ``MineSpider: Extracting Hidden URLs Behind Evasive Drive-by Download Attacks,'' IEICE Transaction on Information and Systems, Vol.E99--D, No.4, pp.860--872, April 2016.
○国際会議	Yuta Takata, Mitsuaki Akiyama, Takeshi Yagi, Takeshi Yada, and Shigeki Goto, ``Website Forensic Investigation to Identify Evidence and Impact of Compromise,'' in Proceedings of the 12th EAI International Conference on Security and Privacy in Communication Networks (SecureComm), pp.431--453, Guangzhou, China, October 2016.
○国際会議	Yuta Takata, Mitsuaki Akiyama, Takeshi Yagi, Takeo Hariu, and Shigeki Goto, ``MineSpider: Extracting URLs from Environment-dependent Drive-by Download Attacks,'' in Proceedings of the 39th Annual International Computers, Software & Applications Conference (COMPSAC), pp.444--449, Taichung, Taiwan, July 2015.
○国際会議	Yuta Takata, Shigeki Goto, and Tatsuya Mori, ``Analysis of Redirection Caused by Web-based Malware,'' in Proceedings of the Asia Pacific Advanced Network (APAN) 32nd Meeting Network Research Workshop, pp.53--62, New Delhi, India, August 2011.
講演	高田雄太, 秋山満昭, 八木毅, 矢田健, 後藤滋樹, ``Web ブラウザ実装差異を悪用する解析回避コードの抽出と分類,'' 電子情報通信学会 暗号と情報セキュリティシンポジウム（SCIS）, 2017 年 1 月.
講演	高田雄太, 寺田真敏, 村上純一, 笠間貴弘, 吉岡克成, 畑田充弘, ``マルウェア対策のための研究用データセット ~MWS Datasets 2016~,'' 情報処理学会 研究報告コンピュータセキュリティ（CSEC）, 2016-CSEC-74, vol.17, pp.1--8, 2016 年 7 月.
講演	高田雄太, 秋山満昭, 八木毅, 針生剛男, ``プログラムスライシングを用いた環境依存コードの実行網羅性向上による潜在的URL の抽出,''情報処理学会 コンピュータセキュリティシンポジウム（CSS）2014 論文集, vol.2014, no.2, pp.17--24, 2014 年 10 月.
講演	Yuta Takata, Mitsuaki Akiyama, and Takeo Hariu, ``Extracting Redirect-Chain Variations in Drive-by Download Attacks Using Emulation of Various Client Environments,'' USENIX Security Symposium Poster Session, August 2014.

早稲田大学　博士（工学）　学位申請　研究業績書

種 類 別	題名、　発表・発行掲載誌名、　発表・発行年月、　連名者（申請者含む）
講演	高田雄太, 秋山満昭, 針生剛男, ``ドライブバイダウンロード攻撃に使用される悪性なJavaScriptの実態調査,'' 信学技報, vol.113, no.502, ICSS2013-72, pp.59--64, 2014年3月.
講演	高田雄太, 森達哉, 後藤滋樹, ``Web感染型マルウェアのリダイレクト解析,'' 情報処理学会 第73回全国大会講演論文集, vol.2011, no.1, pp.497--498, 2011年3月.
書籍	八木毅, 青木一史, 秋山満昭, 幾世知範, 高田雄太, 千葉大紀, ``実践サイバーセキュリティモニタリング,'' コロナ社, 2016.
その他（論文）	Yumehisa Haga, Yuta Takata, Mitsuaki Akiyama, and Tatsuya Mori, ``Building a Scalable Web Tracking Detection System: Implementation and the Empirical Study,'' IEICE Transaction on Information and Systems, Vol.E100--D, No.8, pp.1663--1670, August 2017.
その他（国際会議）	Toshiki Shibahara, Yuta Takata, Mitsuaki Akiyama, Takeshi Yagi, and Takeshi Yada, ``Detecting Malicious Websites by Integrating Malicious, Benign, and Compromised Redirection Subgraph Similarities,'' in Proceedings of the 41st Annual IEEE Computer Software and Applications Conference (COMPSAC), Turin, Italy, July 2017.
その他（国際会議）	Toshiki Shibahara, Kohei Yamanishi, Yuta Takata, Daiki Chiba, Mitsuaki Akiyama, Takeshi Yagi, Yuichi Ohsita, and Masayuki Murata, ``Malicious URL Sequence Detection using Event De-noising Convolutional Neural Network,'' in Proceedings of the IEEE International Conference on Communications (ICC), Paris, France, May 2017.
その他（国際会議）	Takuya Watanabe, Mitsuaki Akiyama, Fumihiro Kanei, Eitaro Shioji, Yuta Takata, Bo Sun, Yuta Ishii, Toshiki Shibahara, Takeshi Yagi, and Tatsuya Mori, ``Understanding the Origins of Mobile App Vulnerabilities: A Large-scale Measurement Study of Free and Paid Apps,'' in Proceedings of the 14th International Conference on Mining Software Repositories (MSR), pp.14--24, Buenos Aires, Argentina, May 2017.
その他（講演）	Fumihiro Kanei, Yuta Takata, Mitsuaki Akiyama, Takeshi Yagi, and Takeshi Yada, ``Protecting Android Apps from Repackaging by Self-Protection Code,'' Network and Distributed System Security Symposium (NDSS) Poster Session, February 2017.
その他（講演）	石井悠太, 渡邉卓弥, 金井文宏, 高田雄太, 塩治榮太朗, 秋山満昭, 八木毅, 森達哉, ``Androidサードパーティーマーケットの大規模調査,'' 電子情報通信学会 暗号と情報セキュリティシンポジウム（SCIS）, 2017年1月.

早稲田大学　博士（工学）　学位申請　研究業績書

種 類 別	題名、　発表・発行掲載誌名、　発表・発行年月、　連名者（申請者含む）
その他（講演）	山西宏平, 芝原俊樹, 高田雄太, 千葉大紀, 秋山満昭, 八木毅, 大下裕一, 村田正幸, ``畳み込みニューラルネットワークを用いた URL 系列に基づくドライブバイダウンロード攻撃検知,'' 情報処理学会 コンピュータセキュリティシンポジウム（CSS）2016 論文集, vol.2016, no.2, pp.811--818, 2016 年 10 月.
その他（講演）	芳賀夢久, 高田雄太, 秋山満昭, 森達哉, ``Web トラッキング検知システムの構築とサードパーティトラッキングサイトの調査,'' 情報処理学会 コンピュータセキュリティシンポジウム（CSS）2016 論文集, vol.2016, no.2, pp.1079--1086, 2016 年 10 月.
その他（講演）	Yumehisa Haga, Yuta Takata, Mitsuaki Akiyama, Tatsuya Mori, and Shigeki Goto, ``Canvas Fingerprinting in the Wild: A Large-scale Measurement and Evaluation,'' International Symposium on Research in Attacks, Intrusions and Defenses (RAID) Poster Session, November 2015.
その他（講演）	Fumihiro Kanei, Mitsuaki Akiyama, Yuta Takata, and Takeshi Yada, ``Observing Interaction between Java and JavaScript for privacy leakage detection in Android,'' International Symposium on Research in Attacks, Intrusions and Defenses (RAID) Poster Session, November 2015.
その他（講演）	Toshiki Shibahara, Takeshi Yagi, Mitsuaki Akiyama, Yuta Takata, and Takeshi Yada, ``Detecting Malicious Web Pages based on Structural Similarity of Redirection Chains,'' ACM Conference on Computer and Communications Security (CCS) Poster Session, October 2015.
その他（講演）	芝原俊樹, 八木毅, 秋山満昭, 高田雄太, 矢田健, ``リダイレクトの構造的類似性に基づく悪性 Web ページ検知手法,'' 情報処理学会 コンピュータセキュリティシンポジウム（CSS） 2015 論文集, vol.2015, no.3, pp.496--503, 2015 年 10 月.
その他（講演）	芳賀夢久, 高田雄太, 秋山満昭, 森達哉, 後藤滋樹, "Canvas Fingerprinting を用いた Web トラッキングの検証と実態調査,'' 情報処理学会 コンピュータセキュリティシンポジウム（CSS） 2015 論文集, vol.2015, no.3, pp.686--693, 2015 年 10 月.
その他（講演）	Takeo Hariu, Keiichi Yokoyama, Mitsuhiro Hatada, Takeshi Yada, Takeshi Yagi, Mitsuaki Akiyama, Tomonori Ikuse, Yuta Takata, Daiki Chiba, and Yasuyuki Tanaka, ``Security Intelligence for Malware Countermeasures to Support NTT Group's Security Business,'' NTT Technical Review, vol.13, no.12, pp.1--7, December 2015.
その他（講演）	針生剛男, 横山恵一, 畑田充弘, 矢田健, 八木毅, 秋山満昭, 幾世知範, 高田雄太, 千葉大紀, 田中恭之, ``NTT グループのセキュリティビジネスを支えるマルウェア対策用セキュリティインテリジェンス,'' NTT 技術ジャーナル, vol.27, no.10, pp.18--22, 2015 年 10 月.
	その他 特許４件