様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　６ 年　 ６ 月 １１ 日現在

機関番号：１７１０４
研究種目：基盤研究(C)（一般）
研究期間：2021～2023
課題番号：２１Ｋ１１８４８

研究課題名（和文）キャンパスBYODを見据えた効率的な暗号化通信の分析に関する研究

研究課題名（英文）A Study on Analyzing Encrypted Communications for Campus BYOD

研究代表者
佐藤　彰洋（Satoh, Akihiro）

九州工業大学・情報基盤センター・准教授

研究者番号：３０６０９３７６

交付決定額（研究期間全体）：（直接経費）　　3,000,000円

研究成果の概要（和文）：大学におけるBYODでは，マルウェアに感染済みの端末をネットワークに持ち込まれることが課題となる．マルウェアによる通信の検出はブラックリストやレピュテーションに頼ることになるが，それらは誤検出を伴うため管理者による検出原因の調査と特定が必須である．しかしながら，その作業は暗号化通信の普及により困難を極めることとなる．本研究では，悪性と判別された膨大な量の暗号化通信において，その原因の効率的な特定を実現するものである．

研究成果の学術的意義や社会的意義
大学におけるBYODのみならず，公衆無線LANなど自身が所有する端末を外出先のネットワークに接続する利用形態は，今後増加するものと想像できる．その一方，マルウェアに感染した端末をネットワークに持ち込まれる可能性はより高まることになる．本研究の核心を成す学術的問いは「暗号化により通信内容が隠蔽される状況下において，悪性と判別された通信の原因を特定するに十分な根拠を見出せるか」に集約される．この解の探究は，マルウェアの感染が疑われる端末に対する迅速且つ的確な措置を可能とすること，延いてはネットワークの堅牢性の向上に大きく寄与するものとなる．

研究成果の概要（英文）：Some of the most serious security threats facing computer networks involve malware. One common way to detect infected machines in a network is by monitoring communications based on blacklists. Administrators should pursue their detection causes by investigating the communications themselves. However, the investigation extremely difficult due to their encryption. In this study, we realize an approach for improving the cause investigation of malicious communications detected through blacklists.

研究分野：ネットワークセキュリティ

キーワード：マルウェア　暗号化通信　ドメイン名　機械学習

１．研究開始当初の背景

コロナ禍におけるリモートワークの導入を背景に，私物情報端末の業務利用，すなわち BYOD (Bring Your Own Devices)が急速に浸透しつつある．高等教育の現場でも，対面と遠隔の両講義を円滑に実施するため，学生個人の端末を必携とする BYOD 体制への移行が必須となっている．その一方で，マルウェアに感染済みの端末をキャンパスネットワークに持ち込まれることが大きな課題となる．マルウェアに感染した端末は，攻撃者の指令を受けることで，ランサムウェアの拡散，フィッシング詐欺や標的型攻撃への利用など，様々な犯罪活動に加担する．大学における BYOD では，学生の専攻や研究，または出身国により用途が多岐に渡るため，端末側のみでの防御には限界がある[1]．故に，ネットワーク側で観測される通信からマルウェアの感染を検出することが求められる．

これまではマルウェアの通信を補足するために，パケットのペイロードを参照する DPI (Deep Packet Inspection)が用いられてきた．一方，米国 Cisco Systems 社によると，2017 年の時点でインターネットにおける暗号化通信の割合は 50%を超えること，それと併せて約 70%のマルウェアが通信を暗号化することが報告されている．この暗号化通信の普及に伴い DPI の適用範囲が極僅かとなることから，ブラックリストやレピュテーションなど，通信先の良悪に基づく検出に頼らざるを得ないのが現状である．しかしながら，ブラックリストやレピュテーションは更新時機に起因して大量の誤りが生じること，それよる検出結果は正誤の判断が困難であることが問題となる[2]．故に，管理者による検出原因の調査と特定が必須となるが，それら一連の作業は通信の暗号化により困難を極めることになる．

本研究の核心を成す学術的問いは「暗号化により通信内容が隠蔽される状況下において，悪性と判別された通信の原因を特定するに十分な根拠を見出せるか」に集約される．この解の探究は，マルウェアの感染が疑われる端末に対する迅速且つ的確な措置を可能とすること，延いてはネットワークの堅牢性の向上に大きく寄与するものとなる．

[1] F. L. Lévesque et al., “Technological and Human Factors of Malware Attacks: A Computer Security Clinical Trial Approach,” ACM Transactions on Privacy and Security, 2018.
[2] M. Kührer et al., “Paint It Black: Evaluating the Effectiveness of Malware Blacklists,” International Workshop on Recent Advances in Intrusion Detection, 2014.

２．研究の目的

本研究の目的は，悪性と判別された膨大な量の暗号化通信において，その原因の効率的な特定を実現することである．これにより，マルウェアの感染が疑われる端末に対する迅速且つ的確な措置が可能となるため，キャンパスネットワークの堅牢性の向上が期待できる．

学術的独自性と創造性は，ブラックリストやレピュテーションにより悪性と判別された通信の前後には，その原因の特定を助ける通信群が存在することに着目した点にある．これら一連の通信群の特徴に基づくことで原因ごとの分類を，その分類結果と教師データとして保持するマルウェアの悪性通信との類似性を比較することで原因の特定を可能とする．特に，教師データの不足が精度の低下に直結することを踏まえ，その不足を擬似データの自動生成で補う点は，他に類を見ない挑戦的な取り組みである．

３．研究の方法

本研究の核となる，(1)不審セッション導出技術，(2)不審セッション分類技術，(3)不審セッション特定技術の確立を目指す．以降，それらの詳細を述べる．

(1)不審セッション導出技術

悪性と判別された通信とそれに付随する通信群を，不審セッションとして集約する技術の確立を目指す．本技術の着想は次の知見，(a)マルウェアによる一連の通信は幾つかのタスクに細分化されること，(b)悪性と判別された通信がマルウェアのタスクの一部によるものである場合，その通信の前後には同一タスクに起因する通信群が存在することから得たものである．故に，それら通信群を考慮に入れることで，原因特定の確度を飛躍的に向上させることが可能となる．

具体的なアプローチは，DNS (Domain Name System)の名前解決におけるドメイン間の共起関係を利用することで，悪性と判別された通信と結び付きの強い通信群，すなわち同一タスクに起因する通信群を選択した．DNS に着目した理由は，マルウェアによる通信に先んじて名前解決が

生じるため両者は密接な関係性を有することに加え，その名前解決先が学内 DNS となるためクエリログから通信内容を取得可能であるが故である．また，通信におけるドメイン間の共起を文章における単語間の共起に帰着できるため，その導出には自然言語処理の分野における深層学習や機械学習を応用した．

(3-2)不審セッション分類技術

不審セッションを原因ごとに分類する技術の確立を目指す．本技術の着想は次の知見，(a)不審セッションはマルウェアの活動の一端を成すタスクと対応付けられること，(b)タスクが担う役割の差異はパケット交換の差異として表れることから得たものである．故に，不審セッションにおけるパケットの外観的特徴を比較することで，その原因の同異を推定することが可能となる．

具体的なアプローチは，不審セッションを成す通信群において，そのパケットのサイズ，送信方向，送信間隔などに加え，暗号化のためのネゴシエーションの特徴を利用した．また，それら通信群を画一的に扱うのではなく，前述の「不審セッション導出技術」における共起関係，すなわち悪性と判別された通信との結び付きの強さを重みとして付与した．パケットの外観的特徴に基づく分類では精度の担保が懸念事項となっていたが，ブラックリストやレピュテーションにより悪性と判別された通信のみに分類対象を限定しているため，その問題を大幅に緩和することに成功した．

(3)不審セッション特定技術

不審セッションの分類結果に対して，その原因を特定する技術の確立を目指す．本技術の着想は次の知見，(a)マルウェアは狙いによりボット，クリプトマイナー，バンキング，エクスプロイトなどに大別できること，(b)未知のマルウェアによる通信だとしても，その活動を細分化したタスク単位の比較により，既知のマルウェアによる通信との類似性を見出せることから得たものである．故に，多様なマルウェアの通信を予め教師データとして保持すること，それとの類似性を比較することにより不審セッションの分類結果に検出原因を付加することが可能となる．

具体的なアプローチとして，教師データには，学術研究のために公開されている悪性通信のデータセットなどを用いた[3]．しかしながら，暗号化の普及に伴い，マルウェアの通信を十分量確保することが困難であることが問題であった．教師データの不足が精度の低下に直結することを踏まえ，その不足を GAN (Generative Adversarial Network)による擬似データの自動生成で補うことを試みた．

[3] H. Hindy et al.,“A Taxonomy of Network Threats and the Effect of Current Datasets on Intrusion Detection Systems,” IEEE Access, 2020.

４．研究成果

本研究は核となる 3 つの技術，(1)不審セッション導出技術，(2)不審セッション分類技術，(3)不審セッション特定技術の総合評価を実施した．具体的には，プロトタイプシステムの構築[4]，九州工業大学のキャンパスネットワークの調査[5]，そのキャンパスネットワークにおける実証実験を通じた有効性の検証である[6,7]．その結果，1 ヶ月間に観測された DNS クエリから，6520 の感染が疑われる悪性クエリを検出できることが明らかになった．その 1 ヶ月分のデータを処理するのに必要とする計算時間は約 12000 秒であり，実用の範囲内に収まるであろうことを確認した．この結果は，暗号化の有無を問わずマルウェアの通信を特定することに成功したことを示している．すなわち，本研究の学術的問い「暗号化により通信内容が隠蔽される状況下において，悪性と判別された通信の原因を特定するに十分な根拠を見出せるか」に一定の解を示し得たと言える．

[4] 佐藤彰洋 他，“DNS クエリに基づくドメインの分散表現法”，情報処理学会インターネットと運用技術シンポジウム，Dec 2022.
[5] 佐藤彰洋 他，“学外公開アドレス管理システム”，学術情報処理研究，No.27，pp.167-173，Nov 2023.
[6] 佐藤彰洋 他，“DGA マルウェアにより生成された悪性ドメインの検出”，九州工業大学情報基盤センター年報，Vol.4，pp.31-44，Mar 2024.
[7] 佐藤彰洋 他，“辞書に基づく DGA マルウェアにより生成された悪性ドメインの検出”，九州工業大学情報基盤センター年報，Vol.4，pp.45-58，Mar 2024.

総務省はこれまで公衆無線 LAN の整備を推進してきた．実際，公衆無線 LAN の利用者数は増加傾向にあり，2018 年度末時点で約 5600 万人，2020 年度末時点で約 6400 万人と予想されている[8]．一方，高等教育の現場では，学生個人の端末を必携とする BYOD 体制を検討する動きが盛ん

になってきている．文部科学省の協力を得た大学 ICT 推進協議会の調査によると，BYOD を導入している大学は 30%を超えていることが報告されている[9]．このように，自身が所有する端末を外出先のネットワークに接続する利用形態は，今後増加するものと容易に想像できる．その一方，マルウェアに感染した端末をネットワークに持ち込まれる可能性はより高まることになる．

本研究の実現により，ネットワークに内在する感染端末を迅速に排除することが可能となる．その成果は，公衆無線 LAN やキャンパスネットワークなど，端末の持ち込みを前提としたネットワークにおいて，セキュリティを向上するための要素技術と成り得る．故に，本研究は，日本における情報通信基盤の整備の方向性とも合致するものである．

[8] サイバーセキュリティタスクフォース 公衆無線 LAN セキュリティ分科会，“公衆無線 LAN セキュリティ分科会 報告書，” 2018．
[9] 大学 ICT 推進協議会 ICT 利活用調査部会，“BYOD を活用した教育改善に関する調査研究結果報告書，” 2018．

５．主な発表論文等

〔雑誌論文〕　計3件（うち査読付論文　1件／うち国際共著　0件／うちオープンアクセス　1件）

１．著者名 佐藤彰洋，戸田哲也，和田数字郎，福田豊，中村豊	４．巻 27
２．論文標題 学外公開アドレス管理システム	５．発行年 2023年
３．雑誌名 学術情報処理研究	６．最初と最後の頁 167-173
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

１．著者名 佐藤彰洋，福田豊，中村豊	４．巻 4
２．論文標題 DGAマルウェアにより生成された悪性ドメインの検出	５．発行年 2024年
３．雑誌名 九州工業大学情報基盤センター年報	６．最初と最後の頁 31-44
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 佐藤彰洋，福田豊，中村豊	４．巻 4
２．論文標題 辞書に基づくDGAマルウェアにより生成された悪性ドメインの検出	５．発行年 2024年
３．雑誌名 九州工業大学情報基盤センター年報	６．最初と最後の頁 45-58
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

〔学会発表〕　計2件（うち招待講演　0件／うち国際学会　0件）

１．発表者名 佐藤彰洋，福田豊，中村豊
２．発表標題 DNSクエリに基づくドメインの分散表現法
３．学会等名 情報処理学会インターネットと運用技術シンポジウム
４．発表年 2022年

１．発表者名 Akihiro Satoh, Gen Kitagata, Yutaka Nakamura
２．発表標題 Numerical Representation of DNS Queries for Cybersecurity
３．学会等名 RIEC Annual Meeting on Cooperative Research Projects
４．発表年 2022年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関