A-017

オープンデータを活用したアタックサーフェスマネジメント支援
Support for attack surface management using open data

松田 健[1)2)]　木村 良太[1)]　王 芸[1)]　園田 道夫[2)]
Takeshi Matsuda　Ryota Kimura　WANG YUN　Michio Sonoda

1　はじめに

近年のサイバー攻撃の目的は多くは金銭の搾取であり，その標的は国家や企業だけでなく，個人にも広がっており，個人情報やシステムのログイン情報など，攻撃の被害から防御すべき情報も多岐に渡っている．本研究では，ネットワークからアクセス可能な IT 資産を守るためのアタックサーフェスマネジメントの考え方を支援する仕組みを，ネット上から収集可能なオープンデータを用いて構築し，セキュリティ対策の評価を実現するモデリングの手法について検討する．本稿では，実際の攻撃で使用されたソフトウェアとその種類，具体的な活用方法に関する情報を収集し，LLM の手法を活用したアタックサーフェスマネジメント支援の方法を提案する．

2　準備

2.1　アタックサーフェス

アタックサーフェスとは，企業や組織が持つ情報技術を活用するためのすべての情報資産に対し，外部からアクセス可能なものやそこに到達するための経路などを指す用語であり，攻撃対象領域とも呼ばれている．つまり，アタックサーフェスとはどこが攻撃対象であるかということを特定するものである．一方，攻撃者がターゲットの目標を把握し，それにアクセスする際に使用される手法やツールのことを攻撃ベクトルといい，防御側の視点から考えるとどのような攻撃が可能であるかということについて考えるものである．アタックサーフェスマネジメントでは，アタックサーフェスと攻撃ベクトルの両方について考慮し，情報資産のすべてを守る手法を考えたり，起こり得るリスクを考えそれを最小化する手法を考えたりすることが重要である．

2.2　従来研究

関連研究について整理する．[1] は，多層防御などの強制アクセス制御技術におけるアクセス制御ポリシーを確実に保護すべきであることをアタックサーフェスと考え，そのポリシーを保護する方法について議論している．[2] は，車載システムにおける脅威分析から導出したアタックサーフェスに基づいたテストに対して，脅威モデリングとアタックテストを連携させた脅威モデリング連携型アタックテストが提案されている．[3] は，異種コンポーネントを含む複雑なシステムの記述に対処するために，コンポーネント指向のビューでシステムのデータ相互作用とリソース配分をモデル化する 2 層階層型攻撃対象領域ネットワークを提案している．[4] は，衛星インターネットや航空インターネットなどもアタックサーフェスを拡張させ宇宙空地ネットワーク向けの攻撃対象領域管理技術の実装戦略を提案している．

3　提案手法

アタックサーフェスマネジメントの実現には，具体的な攻撃事例を分析することが重要である．そのような情報は，国家安全保障，サイバーセキュリティ，ヘルスケアなどの研究開発および技術支援を行っているアメリカの Miter 社がまとめている MITRE ATT&CK [5] のページにかなり詳細に記載されている．

具体的には，攻撃が成立するまでの戦略と手法，攻撃成立後の戦略と手法だけでなく，攻撃に使用されたことがあるツールとどのようなグループがどのような目的で使用したかなどの付加価値の高い情報を取得することができる．

しかしながら，[5] の情報は非常に膨大であり，アタックサーフェスマネジメントへの活用には情報の集約が必要である．

そこで，本研究は，過去に実際の攻撃で使用されたソフトウェアの情報を収集し，企業や組織において情報を管理する立場である管理者側の注意だけで防ぎきれるものであるか，それとも企業や組織において従業員側 (以下，ユーザ側という）の行動によってその脅威が変化するものであるかの情報を自動的に判断する手法を提案する．

その手順は以下の通りである．

【提案アルゴリズム】

1. https://attack.mitre.org/software/からソフトウェア名とその詳細情報を取得
2. ソフトウェアの種類と各ソフトウェアの説明文章を取得
3. ソフトウェアの説明文章から文章埋め込み表現を取得
4. それぞれのソフトウェアに対して管理者側の対策だけで防ぎきれる脅威であるかどうか判断するラベルを付与
5. 文章の埋め込みベクトルを求めて，前のステップで与えたラベルを自動付与

Web サイト [5] から収集できる攻撃ツール（ソフトウェア）のデータ例を表 1 に示す．ID はサイトに掲載されているものをそのまま利用し，それ以降はツール名，関連するツール名，詳細情報となっている．なお，2025 年 6 月 6 日現在では，877 個の攻撃ツールの情報が掲載されていた．

一方，Electronic Transactions Development Agency のサイトからも同様の情報を収集可能であるが，より詳細な説明やソフトウェアの種類など情報を収集することができ，2025 年 6 月 6 日現在では，2180 個の攻撃ツールの情報掲載が確認された．

上述のデータから，攻撃ツールとその活用方法，影響を与える範囲，驚異の種別など，様々な情報の入手が可能である．

本研究では，企業や組織が持つ情報資産に関する脅威について，組織構成員が使用するシステムの開発や運用を行う管理者の対策だけで防ぎきれるものか否かを表 1

1）　阪南大学大学院企業情報研究科
2）　国立研究開発法人 情報通信研究機構

ID	ソフトウェア名	関連ソフトウェア	説明
S0045	ADVSTORESHELL	AZZY, EVILTOSS, NETUI, Sedreco	ADVSTORESHELL is a spying backdoor that has been used by APT28 from at least 2012 to 2016. It is generally used for long-term espionage and is deployed on targets deemed interesting after a reconnaissance phase.
S0440	Agent Smith		Agent Smith is mobile malware that generates financial gain by replacing legitimate applications on devices with malicious versions that include fraudulent ads. As of July 2019, Agent Smith had infected around 25 million devices, primarily targeting India though effects had been observed in other Asian countries as well as Saudi Arabia, the United Kingdom, and the United States.
S0066	3PARA RAT		3PARA RAT is a remote access tool (RAT) programmed in C++ that has been used by Putter Panda.

表 1　攻撃ツールの情報の例 [5] からの引用

の情報から判断可能なモデルの構築法を提案する．

4　実験

本研究は [5] のデータを用いて提案手法を用いた実験を行う．

4.1　ラベル付与の方針

本研究を進める上で，現時点において，本研究の目的を遂行するためのラベル付与の一般的な指針，つまり，攻撃ツールの説明文に対して，管理者の対策だけで脅威が防げるものであるかどうか適切に判断する方法は存在していないものと考えられる．

そのため，本研究はそのための指針の 1 つを提案するものでもある．

例えば，DualToy という攻撃ツールは，USB 経由で接続された Android および iOS デバイスに悪意のあるアプリケーションをインストールする Windows マルウェアであるという説明がなされているが，これはシステムの管理者だけが注意するだけでは脅威を取り除くことは難しいものの 1 つであると考えられる．

一方で，Fgdump というツールは Windows パスワードハッシュ ダンパーであるという説明がされており，この類のものは管理者側の脅威対策であると考えられる．

このような方針でラベル付与をし続けることは，新しい攻撃ツールの開発が今後も続くことを考えると人が付与していくことは困難である．

なるべく，少ない教師データから適切なラベル付与を実現する教師あり学習の方法を構築可能であるかどうかを本研究では調査する．

以下，管理者が対策を施せばユーザの脅威に対するリスクを低下できる可能性があるものを 1，それ以外は 0 のラベルを付与する．例えば，ユーザがマルウェアに感染する場合に，管理者が用意している範囲でそれが起こり得るかどうかを考え，その上で対策の施しようがないものが 0 である．

4.2　モデルの構築

本研究では，Sentence-BERT（SBERT）を用いて 768 次元の文章の埋め込みベクトルを得る方法を採用した．得られたベクトルに，4.1 の手法で与えたラベルを付与し，ランダムフォレストを用いてラベルが与えられていないデータに対し，ラベルを付与する実験を行った．なお，本研究では，miter att&ck のサイトから収集した 877 個のデータのうち，100 個の攻撃ツールにラベルを付与した．

4.3　データ

付与されたラベルを持つ 100 個のデータを用いて，その中から 10 個 (1:3 個)，20 個 (1:5 個)，30 個 (1:6 個)，40 個 (1:8 個)，50 個 (1:8 個) の学習用データセット $T_{10}, T_{20}, T_{30}, T_{40}, T_{50}$ をそれぞれ準備した．なお，データ数の多いデータセットは，個数が少ないデータを完全に含んでいるように構成している．データセットの構成に使用しなかった残り 50 個のデータ (1:10 個) を検証用データとして使用した．

4.4　結果

どの学習データに対しても，検証データのラベルはすべて 0 と判定された．しかし，ラベルの付与をしなかった残りの 777 個のデータに対しては，以下のような結果が得られている．T_{10}, T_{20} は 1 と予測しなかった．なお，

学習	1 と予測	特記事項
T_{30}	4	Dump password hashes, Active Directory
T_{40}	8	Command tools, Frameworks
T_{50}	3	上述と同様

表 2　未知データに対する予測の一部

この 777 個のデータには正解ラベルが存在しないため，FP, FN についての確認はできず，一般的な機械学習の性能評価指標は使用できない．

5　考察

実験に使用した T_{40} と T_{50} を比較すると，ラベル 1 を持つデータは全く同じで，ラベル 0 のデータ数が t_{50} は 10 個ほど T_{40} より多い状態である．これにより，表 2 のように T_{40} で学習した結果の方が未知のデータに対して効果的であるような結果が得られたと考えられる．しかし，この結果は 0 ラベルに対応する情報量が増えたことで本来得たい情報であるラベル 1 のデータ，つまり管理者がアタックサーフェスマネジメントを実施する上で重要な情報が削ぎ落されたことになっている．これは全体の特徴を捉えられていないことを意味しており，今後は本研究を遂行するために必要なデータの個数を増やす方法を考えることが重要な課題である．

参考文献

[1] 井上 洋樹, 辻 秀典, 橋本 正樹: Access Control Policy Protection Empowered by the TrustZone Technology. コンピュータセキュリティシンポジウム 2018 論文集 2018 (2), 941-948 (2018)

[2] 西尾 泰彦, 城間 政司, 井上 博之: 脅威モデリング連携型アタックテストによる 車載ネットワーク脅威分析手法, 情報処理学会論文誌 58 (12), 1943-1953, 2017-12-15 (2017)

[3] Kangyu Huang, Lin Yang, Renfang Fu, Shengli Zhou, Zheng Hong: HASN: A hierarchical attack surface network for system security analysis. China Communications, Volume: 16, Issue:5 (2019)

[4] Haiyi Wang, Xin Heng, Weichen Li: Research on Attack Surface Management Technologies and Telecom Operators' Strategies for Space-Air-Ground Networks. 2024 IEEE/CIC International Conference on Communications in China (ICCC Workshops)

[5] miter att&ck. https://attack.mitre.org/

[6] Electronic Transactions Development Agency. https://apt.etda.or.th/cgi-bin/listtools.cgi