不審活動の端末間伝搬に着目した標的型攻撃検知方式の提案と評価

川口 信隆† 築地原 護‡ 井手口 恒太† 谷川 嘉伸† 冨村 英勤†

†株式会社日立製作所
nobutaka.kawaguchi.ue@hitachi.com

‡株式会社日立アドバンストシステムズ

あらまし:本稿では，拡散活動に着目することで標的型攻撃を検知する方式を提案する．提案方式では，攻撃者が，侵入先ネットワークに拡散するに伴い，被攻撃端末における不審性が上昇することに着目する．そして，不審性が向上した端末間の関係をグラフとして抽出・評価することで，攻撃を検知する．本方式は，個々の端末やプロセスに着目するだけでは検知できない，高度な攻撃に対応できる．ある組織の同一部署に設置された30台の端末を対象に2か月間実施した評価実験を通じ，提案方式は既存方式と比べ誤検知を1/10に低減しながら，攻撃を97%の確度で検知できることを示した．

Detection of Advanced Persistent Threat based on Cascade of Suspicious Activities over Multiple Internal Hosts

Nobutaka Kawaguchi† Mamoru Tsuichihara‡ Kota Ideguchi†
Yoshinobu Tanigawa† Hideyuki Tomimura†

†Hitachi, Ltd.
nobutaka.kawaguchi.ue@hitachi.com
‡Hitachi Advanced Systems, Ltd.

Abstract In this paper, we propose a novel detection scheme for advanced persistent threat (APT). The scheme focuses on the phenomena that as an attacker intrudes into multiple hosts via lateral movement, the compromised hosts in turn exhibit a few suspicious activities. Then, the scheme extracts the lateral movement as a graph with the suspicious hosts as the nodes, and detects the attack based on the graph. Evaluation experiments with two month monitoring of 30 hosts in a same department of an organization show our scheme achieves more precise detection performance than an existing scheme.

1 はじめに

近年，標的型攻撃を仕掛ける攻撃者の技術力が向上し，様々な攻撃回避技術を用いるようになるに従い，攻撃を構成する要素（被攻撃端末・プロセスなど）を個々に分析するのみでは検知が難しくなりつつある[1] [2][4][13].

上記の問題を解決するために，我々は，標的型攻撃の 1 ステップである拡散活動に着目した検知方式を提案する．拡散活動では，攻撃者は，遠隔操作ツール（e.g. Psexec[9]）やシステムの脆弱性を悪用して，ある端末から別の端末に渡り歩く[3][11]．本方式は，複数端末の様々な活動ログに対して時空間分析を行い，検知を実現する．

本方式は，攻撃に起因する可能性がある活動の発生頻度を基に端末の不審性を分析する．そして，拡散活動に伴い不審性が高い端末が連鎖的に現れる現象を，被攻撃端末をノードとするグラフ構造として抽出し，グラフがある基準を満たすとき，標的型攻撃が発生していると判断してアラートを上げる．

本方式の特徴は，複数の端末で発生する，不審性が僅かでもある様々な種類の活動を収集・相関分析し，拡散活動を表すグラフ構造として抽出する点にある．これにより，個々の端末で行う活動の不審性が低い標的型攻撃の検知が可能になる．また，既存技術と比べて，誤検知の発生を抑えられる．

我々は，3 種類の攻撃シナリオと，ある組織の同一部署に属する30台の端末の2か月間の活動ログを用いた評価実験を通じ，検知精度及び，プロセスやファイルシステムに痕跡を残さない高度な攻撃に対する有効性の両面で，本方式は既存方式より優れていることを確認した．

以下，第 2 章では関連研究と本方式の位置付けについて説明する．第 3 章・第 4 章では本方式の詳細及び実装について述べる．第 5 章では性能評価及び実験の詳細について述べる．第 6 章を本稿のまとめとする．

2 背景

既存研究の多くは，標的型攻撃を構成する単一要素（個々の被攻撃端末・プロセス，マルウェア，C&C 通信，エクスプロイト）を基に攻撃検知を行っている．

動的解析・静的解析を基にマルウェアを検知・分類する技術に関しては，これまでに多数の研究発表がなされている[5][6][7]．これらの研究では，ルールや振舞いモデルに従って，マルウェアの検知或いはファミリーの分類を行う．攻撃の高度化に伴い，正規ツールやフットプリントが小さいマルウェアが利用され，発見しやすい既知脆弱性が用いられなくなるようにつれ，これらの方式で攻撃を効果的に検知するのは難しくなりつつある[1][2][4][13]．

文献[8]で提案されているプロセスホワイトリスト機構は，あるホストで実行可能なプログラムの種類を制限する．このアプローチはマルウェアが含まれる実行ファイルの起動を防ぐことができるが，ユーザやシステムが頻繁に新規アプリケーションのインストールを行うような環境では，多数の誤検知が発生するという問題がある．また，このアプローチではプロセスやファイルシステムに痕跡を残さない攻撃に対応できない．

文献[9]では，C&C サーバに対するトラヒックと拡散活動に起因する通信を関連付けることで標的型攻撃を検知する方式が提案されている．このアプローチは，拡散活動・C&C 通信に用いられる特定のプロトコル（Microsoft SMB）の特徴に強く依存しているため，既知の攻撃に対しては有効であるものの，異なるタイプの通信が用いられる攻撃を検知するのは難しい．

3 標的型攻撃検知方式

3.1 コンセプト

提案する標的型攻撃検知方式の目的は，標的型攻撃により組織ネットワークに侵入したマルウェア／攻撃者による拡散活動を複数のネットワーク内端末の活動の不審性を基に早期検知することで攻撃被害を最小化することにある．本方式のアプローチは，標的型攻撃に関する以下の事実・仮説に基づく．

- 事実：標的型攻撃では，攻撃者にとっては，最終目的を達成する前の検知を回避することが重要である．このため，攻撃者は，Windows 標準コマンドの活用などを通じて，アンチウイルスなどによる検知を避ける傾向にある．このため，個々の端末やプロセスのみを分析して検知を行うのは困難になりつつある[1] [2][4][13]
- 仮説1：一方で，端末の通常オペレーションと攻撃活動とでは目的が異なるため，攻撃に伴い，「やや不審な活動（以下，不審活動）」が端末内で発生すると考えられる．例としては，通常オペレーションでは使用頻度が低い Windows コマンドの実行などがある
- 仮説2：これらの不審活動は，通常オペレーションでも発生しうる活動であるため，この点のみに着目して検知を行おうとすると，誤検知が大量に発生する恐れがある．しかし，攻撃者が拡散活動で複数端末を渡り歩く場合，不審活動を行う端末（不審端末）が連鎖的に発生する．この現象は，通常オペレーション時には発生しづらく，高精度な検知に活用可能である

本方式は，個々の端末の活動を個別に分析するのではなく，不審活動・不審端末の発生タイミングの相関性を基に，複数の端末を跨った拡散活動の痕跡を「不審活動グラフ」として構築し，攻撃を検知する．これにより，個々の活動・端末を監視するだけでは発見が困難な，不審性が低い高度な標的型攻撃の検知が可能になる．

3.2 ユースケース

本方式のユースケースを図 1 に示す．本方式は，ネットワーク内に設置された拡散活動検知装置と Web プロクシ，各端末にインストールされている端末監視機能(TM)から構成される．

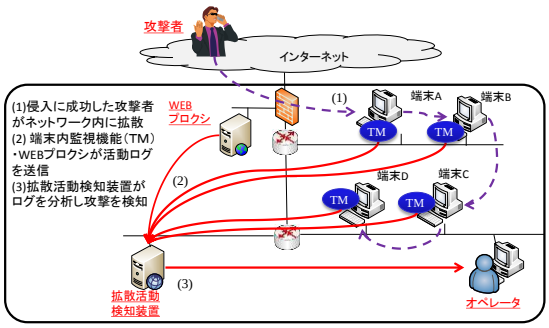

図 1 本方式のユースケース

ユースケースでは先ず，インターネットから組織ネットワークへの侵入に成功した攻撃者が，端末 A →端末 B→端

末C→端末Dの順で不正アクセスを行い，活動を広げていく．その際には，サーバからの機密情報の窃取，サービスに対する脆弱性攻撃，Psexec などの遠隔実行ツールを用いた遠隔端末操作などが行われる．

これに対して本方式では，先ず，各端末内にインストールされた端末監視機能及びネットワーク内に設置されたWeb プロクシが，端末の活動（ファイルアクセスやネットワーク通信など）のログを逐一拡散活動検知装置に送信する．拡散活動検知装置はログをリアルタイムに分析し，不審活動を特定する．そして，不審活動の不審度がある基準を満たす端末を不審端末として特定する．さらに，複数の不審端末から構成される不審活動グラフを検知する．不審活動グラフがある基準を満たす場合，ネットワーク内で拡散活動が発生していると判断し，アラートを送信する．

3.3 検知手順

図 2，図 3 に本方式の検知手順の概要及び，アーキテクチャを示す．拡散活動検知装置は，各端末の不審性を「不審活動度」として評価する．不審活動度は，時間経過に伴い，不審活動の内容・頻度に応じて増減する．各端末は不審活動度に応じた「状態」をとる．不審活動度が閾値未満である端末は「正常状態」にあり，閾値以上である端末は「不審状態」にあるとする．本方式では，攻撃者が端末で悪意ある活動を行うに従い，端末の不審活動度は徐々に上昇し，ある時刻に正常状態から不審状態に転換する．同様に，不審活動度が閾値を下回ると，端末の状態は正常状態に再転換する．

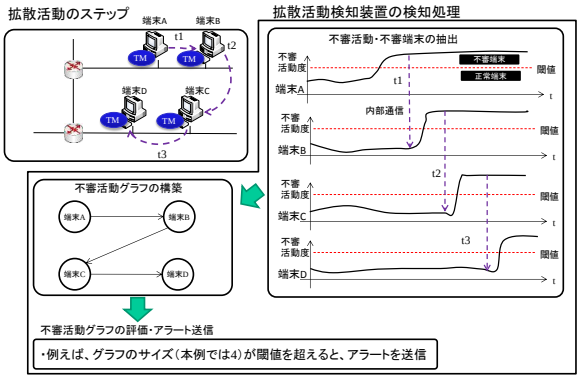

図 2 検知手順概要

拡散活動検知装置は，検知された不審端末を基に不審活動グラフを構築する．不審活動グラフは，不審端末をノード，不審端末間で発生する TCP/IP 通信（内部通信）を有向辺とする有向グラフである．本方式では，拡散活動に伴い，正常状態であった端末が不審端末に次々に転換していく現象を捉える． このため，概略を示すと，2 つの端末（端末 A 及び端末 B）は，

（1）不審状態にある端末Aから，正常状態である端末Bに対して内部通信が発生する，（2）内部通信発生後一定内に，端末 B の状態が不審状態に転換する，場合に限り，A→Bというグラフ構造を取る．

最後に，不審活動グラフの形状（例えばサイズ）を評価し，評価値が閾値を超えた場合にアラートを送信する．

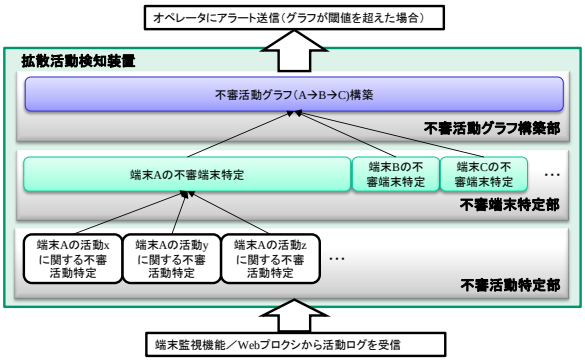

図 3 拡散活動検知装置アーキテクチャ

3.4 不審活動特定

標的型攻撃に伴い発生する場合があると考えられる不審活動を，以下の 2 種類に分類する．

- アノマリ型：攻撃が発生していない一定の通常オペレーション期間で観測されなかった或いは観測頻度が低い活動（例：普段実行しないプロセスの起動）
- ルール型：通常オペレーション中でも発生しうるが，経験的に，攻撃中に発生する頻度が高いことが分かっている活動（例：実行ファイルの作成）

ここで，アノマリ型不審活動の特定には通常状態を定義した通常プロファイルが必要である．このため，本方式は，通常プロファイル作成のための事前学習を行う学習フェイズ，学習結果を用いて検知を行う検知フェイズという2つのフェイズを持つ．このため，本方式は機械学習型の検知手法の一種といえる．

どういった活動を不審活動として捉えるかは，標的型攻撃やマルウェアに関する知見（[13][14]など）を基に決める必要がある．我々は，以下の不審活動を定義した．

3.4.1 アノマリ型不審活動及び特定方法

（1）プロセス起動

端末ごとに，学習フェイズ中に一回以上発生したプロセス

のリストをプロセスに関する通常プロファイルとして構築する. 検知フェイズ中で通常プロファイルに含まれないプロセスが起動すると, 不審活動として特定される.

(2)ポートオープン

端末ごとに, 学習フェイズ中に一回以上 TCP/UDP のリスニングポートを開いたプロセスのリストをポートオープンに関する通常プロファイルとして構築する. 検知フェイズ中で通常プロファイルに含まれないプロセスがポートを開こうとすると, 不審活動として特定される.

(3)内部通信

端末ごとに, 学習フェイズ中に一回以上 TCP/IP 通信を行った同一組織ネットワーク内の端末のリストを, 内部通信に関する通常プロファイルとして構築する. 検知フェイズ中で通常プロファイルに含まれない通信先端末に対して通信が発生すると, 不審活動として特定される.

(4)Web 通信

端末ごとに, 学習フェイズ中に一回以上, CONNECT メソッドで 10KB 以上のデータ送信を伴う通信を行った Web サーバのドメインを, Web 通信に関する通常プロファイルとして構築する. 検知フェイズ中で, 通常プロファイルに含まれないドメインに対して CONNECT メソッドで一定サイズ以上のデータ送信が発生すると, 不審活動として特定される.

3.4.2 ルール型不審活動及び特定方法

(1)実行ファイル作成

端末のローカルドライブに対する実行ファイルの作成は不審活動として特定される. 標的型攻撃発生の際は, 攻撃ツールのダウンロードが行われるため, 実行ファイル作成が発生する場合があると考えられる.

60秒内に複数の実行ファイルが作成された場合は, 1つの不審活動としてカウントされる. これは, 通常オペレーション時に, 複数の実行ファイルから構成される新規アプリケーションがインストールされる際に, 多数の不審活動として特定されないようにするためである.

(2)ICMP Echo Request

端末が ICMP Echo Request を 3 回以上送信し, 受信した ICMP Echo Reply 数が送信数の 1/3 以下になるとき, 不審活動としてカウントされる. 標的型攻撃発生の際は拡散先端末を探索するアドレススキャンにより Reply を伴わない ICMP Echo Request が発生する場合があると考える.

3.5 不審端末特定

拡散活動検知装置は, 端末監視機能・Web プロクシから取得した不審活動を基に, 各端末の過去 window 時間内における不審活動度を求める.

Si(x,t1,t2)が時刻 t1 から t2 に端末 x で発生した活動 i の集合であり, 監視対象の不審活動が n 種類ある場合, 時刻 t における端末 x の不審活動度 *SUS (x,t)*は, 以下の数式(1)で求める.

$$
\begin{aligned}
&\mathrm{SUS}(\mathrm{x},\mathrm{t}) = \\
&W_{Proc} * \left|\mathrm{Uniq}\left(S_{Proc}(x, t-window, t)\right)\right| + \\
&W_{Port} * \left|\mathrm{Uniq}\left(S_{Port}(x, t-window, t)\right)\right| + \\
&W_{WEB} * \left|\mathrm{Uniq}\left(S_{WEB}(x, t-window, t)\right)\right| + \\
&W_{Comm} * \left|\mathrm{Uniq}\left(S_{Comm}(x, t-window, t)\right)\right| + \\
&W_{File} * \left|S_{File}(x, t-window, t)\right| + \\
&W_{ICMP} * \mathrm{Min}\left(1, \left|S_{ICMP}(x, t-window, t)\right|\right) \qquad (1)
\end{aligned}
$$

ここで Wi は活動 i に関する重み付け, |.|関数は集合内の要素数を返す関数である. Uniq 関数は発生時刻以外の属性が異なる要素の集合を返す関数である. このため, 例えばプロセス起動に関する不審活動では, 同一プロセスを window 時間内に何度も起動しても不審活動度は上昇しない. Min 関数は引数のうち値が小さいものを返す関数である. このため, ICMP Echo Request に関する不審活動は, 「有り」又は「無し」の 2 値で評価される.

上式が示すように, 不審活動度は, 重み付けされた各不審活動の発生頻度の合計値として求められる. このため, 攻撃者が 6 種類の不審活動を全ては行わない場合, 或いは監視が困難な場合であっても, 不審活動度を算出できる. 攻撃者が各端末内でどのような活動を行うかを事前に予想することは難しい. また, 攻撃者のスキルによっては不審活動の一部は観測が難しくなる可能性がある. 例えば, 攻撃が, 起動中の他プロセスを奪取する場合や OS カーネルに手を加える場合は, プロセス起動やポートオープン, 実行ファイル作成を正しく監視できない可能性がある.

一方で, ネットワークに関する活動である, ICMP Echo Request, 内部通信, Web 通信はネットワーク上からの監視が可能なため監視見逃しが少なくなると考えられる. このため, ネットワークに関する不審活動度のみを用いて攻撃検知を行いたい場合, もう1つの不審活動度 SUS_{Net}を用いる. SUS_{Net} は数式(1)から Web 通信, 内部通信, ICMP Echo Request に関わる項だけを取り出したものである.

最後に，不審活動度が閾値 TH_{SUS} を超えた場合，当該端末は不審端末と判断される．ある端末が不審端末であるかどうかはリアルタイムに判断される．このため，ある時刻に不審端末に転換した端末において，その後の不審活動の発生が無いと不審活動度は減少し，閾値を下回ると正常端末に再転換する．

3.6 不審活動グラフ構築

不審活動グラフ G を<V,E>と表記する．V は不審端末に対応するノードの集合，E はノードを繋ぐ内部通信に対応するエッジの集合である．不審活動グラフ G は，攻撃者が E に含まれる内部通信を通じて V に含まれる端末間を渡り歩くことを示している．ノードは，端末が不審端末に転換する度に生成される．

ノード v∈V は，（host, generation_time, last_activity_time）で表記される．host は，ノードの基となる端末の識別子(IP address)，generation_time は，端末が不審端末に転換した時刻である．last_activity_time は，端末が正常端末に再転換する前に，最後に行った不審活動の発生時刻（最終不審活動時刻）である．

エッジ e∈E は(src, dst generation_time)で表記される．src は内部通信を開始した端末に対応するノードの識別子，dst は内部通信を受信した端末に対応するノードの識別子，generation_time は内部通信が開始され，エッジが生成された時刻である．ここで，内部通信には不審活動に加えて正常活動も含まれる．

新しく生成されたノード v2 は，以下の 4 条件を満たすとき，既存グラフ G<V,E>に含まれるノード v1 に，エッジ e を介して連結される．

1. ¬∃v∈V (v.host == v2.host)
2. (v1.host == e.src.host and v2.host == e.dst.host) and (v1.generation_time <= e.generation_time <= v1.last_activity_time+gap1)
3. ¬IsSuspicious(v2.host,e.generation_time)
4. e.generation_time <= v2.generation_time <= e.generation_time+gap2

条件 1 は，v2 が示す端末と同じ端末を示すノードがグラフ G に存在しないことを規定する．

条件 2 は，エッジ e は v1 が示す端末から v2 が示す端末に対する内部通信を表現し，開始時間は，v1 の最終不審活動時刻から一定時間（*gap1*）内であることを規定する．

条件 3 は，内部通信の受信端末は，受信時点では正常端末であることを規定する．*IsSuspicious(x,t)*は端末 x が時刻 t の時点で不審端末であるか否かを返す関数である．

条件 4 は，内部通信受信後，受信端末が一定時間内（*gap2*）に不審端末に転換することを規定する．

v2 が v1 と連結される，即ち v1→v2 となるとき，G<V,E>は，V=V∪v1, E=E∪e のとおり更新される．

v2 を連結可能なノードを持つグラフが存在しない場合，v2 は新規グラフ G’<V{v2}, E{φ}>を構成する．

不審活動グラフは時々刻々と変化・成長していく．不審活動グラフに含まれるノード数をグラフスコアと呼称する．拡散活動検知装置は，グラフスコアが閾値 TH_{GRAPH} 以上になった場合に標的型攻撃の発生を検知する．

4 実装

端末監視機能は C#で実装した．各不審活動の監視には，Windows が提供する標準 API を用いた．プロセスの正確な識別には起動元実行ファイルのハッシュ値などを用いる必要があるが，今回は実装の簡便さの点から，プロセス名が同じであるプロセスは同一のものとして扱った．同様に，実行ファイルの正確な特定にはファイルコンテンツを検査する必要があるが，今回は拡張子が”.exe”であるファイルを実行ファイルとして扱った．

Web プロクシは Squid 2.7 を基に作成した．拡散活動検知装置は C#及び Microsoft SQL Server で作成した．

5 評価実験

5.1 実験概要

提案方式の評価実験について述べる．実験では，ある組織内ネットワークに設置された同一部署に属する PC30 台に対して，拡散活動を伴う標的型攻撃が仕掛けられることを想定し，提案方式の検知性能を評価した．

提案方式の評価には，通常プロファイルの作成を行う学習フェイズと，検知精度を評価する検知フェイズを実施する必要がある．我々は，PC30 台の挙動を 2 ヶ月間監視・記録し，前半の 1 ヶ月間をプロファイル作成に用い，後半の 1 ヶ月間を誤検知頻度の評価に用いた．実験期間中 PC は様々な業務のために頻繁に利用されていた．また検知率評価のためにクローズドなネットワーク環境を構築し，3 種類の標的型攻撃を模擬した．

検知率評価をクローズドネットワークで実施したのは，組織のセキュリティポリシーの理由で攻撃の模擬を業務に使用しているネットワークで行うのは好ましくないと判断したためである．また，既存手法と比較することで本方式の

優位性を検証した. 本評価は実験期間中に組織ネットワーク内では標的型攻撃は発生していないことを前提とする.

以下に, 各実験の詳細及び比較対象手法の概要, および実験結果について示す.

5.1.1 通常プロファイル作成実験

通常プロファイル作成実験は, 2014 年 12 月 10 日から 2015 年 1 月 10 日までの 1 ヶ月間, 同一組織ネットワーク内で稼働する PC30 台(以下, 評価用端末)を対象に実施した. 評価用端末は Windows Vista/7 がインストールされている汎用 PC であり, 多数のユーザからアクセスされる専用サーバは含まれない. 但し, 一部の端末はファイル共有や実験, Windows が提供するサービスの影響で他端末からアクセスされる.

多くの端末内で行われるオペレーションは書類作成などの一般業務・WEB 閲覧・ソフトウェア開発などである. 各ユーザは, 自身の使用端末の管理者権限を有している.

実験の結果導出された 4 種類の活動の通常プロファイルの件数は表 1 の通りである.

表 1 通常プロファイル件数

活動	通常プロファイル件数	
	30 台累計	1 台平均
プロセス起動	4983	166.1
ポートオープン	491	16.4
内部通信	205	6.8
Web 通信	4545	151.5

内部通信は, 通信先端末が評価用端末である通信のみを分析対象とした. 30 台の評価用端末間で行われる内部通信の送受信端末の組み合わせ数は 870(=30x29)通りである. このため, 1 ヶ月の間に可能な組み合わせの 27%(=205)が実際に発生したことになる. 通信プロトコルとしては TCP/5357, TCP/2869, TCP/137 が大半を占めた.

5.1.2 誤検知頻度評価実験

誤検知頻度評価実験では, 2015 年 1 月 11 日から 2015 年 2 月 10 日までの一か月間に発生したアラート数を, 誤検知頻度として測定した. 実験期間中に組織ネットワーク内では標的型攻撃は発生していないことを前提とするため, 実験期間中に発生したアラートは全て誤検知と判断する.

提案検知手法のパラメータのデフォルト値は表 2 の通りである.

表 2 デフォルトパラメータ

パラメータ	デフォルト値
window	3600 秒
TH_{SUS}	4
gap1	windows の 1/2 の値
gap2	window と同値
TH_{GRAPH}	2

不審端末の分析期間である window のデフォルト値は, 速効型の標的型攻撃の持続時間(攻撃開始から目的達成までにかかる時間)が最長で数時間以内であるという知見[13]を基に設定した. gap1, gap2 は window 値に依存するものとした. 不審端末の判定閾値である TH_{SUS} のデフォルト値は, 攻撃者は, 端末侵入後にプロセス起動などの活動を数回以上行うという知見[10]を基に設定した.

不審活動グラフの判定閾値である TH_{GRAPH} のデフォルト値は, 拡散活動により起因するグラフスコアの最小値を設定した. また, SUS(x,t)における各活動の重み付けは全て 1 とした.

尚, 数日・数週間など, window 時間より長い期間をかけて行う攻撃であっても, 各端末で, 各 window 期間に TH_{SUS} 件以上の不審活動を伴うものであれば, 原理的に検知が可能である.

5.1.3 検知率評価実験

検知率評価実験では, 3 台の端末(PC1,PC2,PC3)を対象とした標的型攻撃をクローズドネットワークで模擬した. クローズドネットワークは, 組織ネットワークに相当するサブネット及び C&C サーバが存在するサブネットから構成される. そして, 誤検知頻度評価実験で得られた不審活動ログを基に, 当該攻撃が 30 台の評価用端末の一部に対して行われた場合の検知率を評価した.

検知率評価では, 通常オペレーション時に発生する不審活動の検知率への影響を測るために, 以下の 2 種類の検知率を評価した.

- ベースライン検知率・・・攻撃を受ける端末内で, 攻撃に関係しない活動が一切行われていない場合を想定したときの検知率
- 実質検知率・・・端末内で, 攻撃に関係しない通常オペレーションに関する活動が行われている場合を想定したときの検知率. 誤検知頻度評価実験で用いたログを元に測定する

実験では, 文献[10][11][12]に基づき, 3 種類のシナリオを実施した.

シナリオ 1 では, PC1 を乗っ取った攻撃者が PC2 に拡散し, PC2 内にある機密ファイルを, PC1 を経由して C&C サーバ(C2)に送信する. 表 3 に詳細を示す.

シナリオ 2 はシナリオ 1 と同様だが，PC1 を介さず，PC2 から機密ファイルを，直接，異なる C&C サーバに送信する．このシナリオでは，攻撃者が，攻撃に関与する端末やサーバを分散させ，個々の端末・サーバに着目した検知方法を回避しようとするというケースを想定している．

シナリオ 3 では，攻撃者は PC1,PC2 を介して PC3 にある機密ファイルを窃取する．このシナリオでは，機密ファイルが存在する PC3 に対して PC1 から直接アクセスできない場合に，PC2 を経由するというケースを想定している．

シナリオ1は表 3 の 7 ステップから構成される． シナリオ 2･3 は，これに独自のステップを加えたものとなる．

表 3 シナリオ１・攻撃ステップ

1	PC1 が不正なショートカットファイルを実行し，RAT ツールを D.L.サーバからダウンロード・実行する
2	RAT ツールが C2 サーバに SSL で接続し，遠隔操作が開始される．攻撃者は攻撃に使うツールをダウンロードする．また，トンネリングに用いるポートを開く
3	PC1 に侵入した攻撃者は，Windows 標準コマンドを実行し，PC1 の端末情報を調査する．この端末情報の送信量は累計で 10KB を超えることを想定する．また，アドレススキャナを実行し，PC1 の近傍にいる PC2 を発見する．
4	攻撃者は Paexec を実行して，PC1 から PC2 に拡散する．また必要な攻撃ツールをコピーする
5	攻撃者は PC2 の端末情報を調査する
6	攻撃者は PC2 内のファイルを圧縮し，Windows ファイル共有機能を介して PC1 にコピーする
7	攻撃者は圧縮ファイルを C2 サーバにアップロードする．また PC1,PC2 に残されたファイルを削除し，攻撃を完了する

攻撃に用いた RAT ツール，アドレススキャナは，実験のために作成した独自なものであり，一般的なアンチウイルスソフトで検知されないことを確認している．

本シナリオの実行に際して，攻撃開始から完了までにかかる時間は一時間程度である．また， PC2，PC3 は Paexec（Psexec の互換ツール）による遠隔操作が可能であると仮定する．クライアント型 Windows 端末を Paexec で遠隔操作するには，事前のレジストリ設定が必要である．実験中にこの設定がされている評価対象端末は半数以下であったが，各端末で本設定が有効であり，拡散活動が可能であることを仮定して検知性能を見積もった．本方式は遠隔管理ツールや脆弱性攻撃など何らかの手段を用いて拡散活動を行う攻撃を想定しており，特定の遠隔操作方法やプロトコルには依存しない．このため，本シナリオでは実施が容易な Paexec を用いたが，他のツールや脆弱性攻撃を用いた場合も検知結果の一般性は失われない．

シナリオ 1･2 では 2 台の PC，シナリオ 3 では 3 台の PC が攻撃に含まれる．また，評価用端末は 30 台である．このため，シナリオ 1･2 に対しては 870（＝30x29）通り，シナリオ 3 に対しては 24360(=30x29x28)通りの PC の組み合わせに対する攻撃を実施し，検知率を求めた．

5.1.4 比較対象手法

本方式の比較対象手法として，文献[8]での提案を基に，プロセスホワイトリスト手法（PWL）を用いる．PWL では，学習フェイズで 30 台の端末内で起動したプロセスを，端末の区別無くホワイトリストに格納する．そして，検知フェイズでホワイトリストに載っていないプロセスが起動するとアラートを上げる．

5.2 実験結果

表 4 に誤検知頻度実験において発生した不審活動数を示す．プロセス起動に関する件数が最も多く，次いで内部通信，実行ファイル作成となっている．

表 4 不審活動数

活動種類	不審活動件数
プロセス起動	8004
ポートオープン	548
内部通信	3156
Web 通信	137
実行ファイル作成	1250
ICMP Echo Request	63

図 4 に検知率の比較を示す．提案方式のベースライン検知率及び PWL の検知率は全てのシナリオに関して 100%となる．一方，提案方式の実質検知率は，シナリオ 1･2 において 3%程度低下する．検知ミスは，PC2 が攻撃と無関係のオペレーションにより，PC1 から内部通信を受信する前に不審端末に転換する場合に発生する．この場合，PC1 と PC2 は同一のグラフに含まれなくなる．シナリオ 3 では，PC1→PC2 のグラフに加え，PC2→PC3 のグラフも検知に利用できるため，実質検知率の低下は 0.1%程度になる．また，提案方式では，攻撃に含まれる C&C サーバが単一であること前提とする既存技術[9]では対応が難しい，複数の C&C サーバを用いたシナリオ 2 が検知可能である．

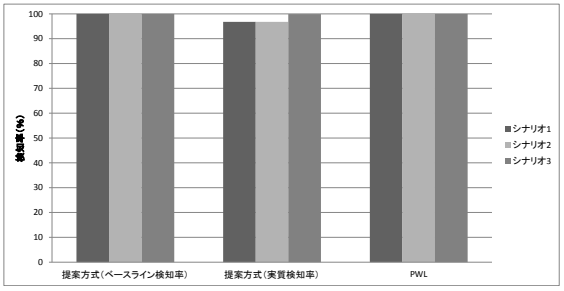

図 4 検知率評価

一方で，測定の結果，PWL の誤検知頻度が 250 件／月

であるのに対し，提案方式はその約 1/10 の 26 件となった．これは，提案方式では複数の活動を基に検知を行うため，通常オペレーション時に新しいプロセスが起動したのみではアラートを上げないためである．

次に，図 5 に，攻撃者が高度なスキルを持ち，プロセスやファイルシステムに攻撃の痕跡が残らず，通信関係のみ監視可能な場合の検知率を示す．提案方式では SUSNET に基づき不審端末を検知する．SUSNET の閾値は 1 とする．

PWL ではプロセスに痕跡が残らない攻撃の検知が出来ないため検知率は原理的に 0%となる．一方提案方式では，通信関連の監視のみでも検知率は 70%を超える．また，シナリオ 2 ではシナリオ 1,シナリオ 3 よりも検知率が高くなる．これは，シナリオ 2 では，他のシナリオと異なり，拡散先の端末(PC2)もC2サーバと通信を行うため，検知に利用できる通信関連の活動が多くなるためである．またシナリオ3ではPC2→PC3の拡散活動を検知に利用できるため，シナリオ 1 より検知率が高くなる．

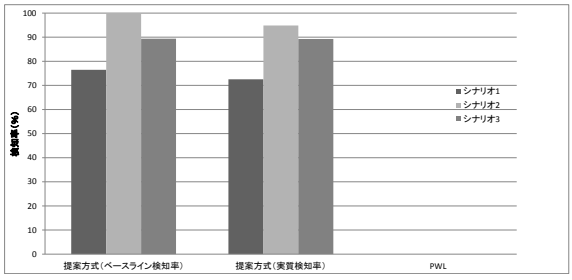

図 5 検知率評価（通信関係のみ監視）

一方，SUSNET の閾値を 1 に設定した場合の誤検知頻度は 52 に増加するが，PWL の 250 に比べると低い値となる．

図 6 に THSUS 値の検知性能への影響を示す．THSUS が大きくなるにつれ誤検知頻度は低下する．一方，THSUS が 8 を越えると検知率が低下する．このため今回の実験では THSUS は 6～8 程度の値であるとき，検知・誤検知の両面で最適な結果が得られることがわかる．

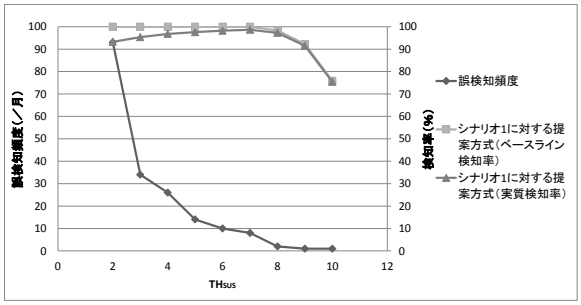

図 6 THSUS 値の検知性能への影響

6 おわりに

本稿では，拡散活動を不審活動グラフとして抽出することで標的型攻撃を検知する方式を提案した．性能評価実験を通じ，提案方式は，各パラメータが最適に設定された場合，(i)拡散活動を伴い(ii)各端末で数時間に数件程度の不審活動を行う標的型攻撃を検知でき，誤検知頻度を1ヶ月で一桁台まで抑えることができることを示した．提案方式は検知性能及び高度な攻撃に対する有用性の点で既存方式より優れていることを確認した．今後は，監視対象とする不審活動の拡充，様々なネットワーク環境・攻撃シナリオを用いた性能評価を通じて，提案方式の有用性をさらに検証していく．

参考文献:

[1] Jennifer Bielski.: Looking Ahead: The State of Incident Detection and Response in 2015, https://www.mandiant.com/blog/state-incident-detection-response-2015/

[2] SECURITY WEEK.: Dropbox Abused in Targeted Attacks Using PlugX RAT With Time Bomb, http://www.securityweek.com/dropbox-abused-targeted-attacks-using-plugx-rat-time-bomb

[3] Trend Micro.: LATERAL MOVEMENT :How Do Threat Actors Move Deeper Into Your Network? (2013)

[4] LAC, co.ltd., Cyber GRID View Vol.1 (2014)

[5] Stefano Ortolani et al.: KILMAX: Profiling Memory Write Patterns to Detect Keystroke-Harvesting Malware, In Proc. of RAID 2011 (2011)

[6] C.Willems, et al.: Toward Automated Dynamic Malware Analysis Using CWSandbox, IEEE Security and Privacy Magazine, Vol.5, Issue 2, 2007 (2007)

[7] D.Inoue, et al.: Automated Malware Analysis System and its Sandbox for Revealing Mal-ware's Internal and External Activities, IEICE Trans. Information and Systems, Vol.E92-D, No.5, 2009 (2009)

[8] 中里ら.:ホスト型 IDS を用いた不審プロセスの特定, SCIS2015 予稿集 (2015)

[9] 山田ら.:組織内ネットワークにおける標的型攻撃の振る舞い検知に向けた複数センサの連携手法, SCIS2015 予稿集 (2015)

[10] 寺田ら.:研究用データセット「動的活動観測 2014」の検討,CSS2014 予稿集 (2014)

[11] Mandiant.: M-Trends Report 2010 (2010)

[12] Eric Hut Chins, et.al.: Intelligence-Driven Computer Network Defense Informed by Anal-ysis of Adversary Campaigns and Intrusion Kill Chains, In Proc. of CIW2011 (2011)

[13] トレンドマイクロ.:"国内標的型サイバー攻撃分析レポート 2015 年度版"(2015)

[14] 独立行政法人情報処理推進機構(IPA).:「標的型メール攻撃」対策に向けたシステム設計ガイド,(2013)