Vol.2020-CSEC-91 No.9
Vol.2020-SPT-40 No.9
Vol.2020-EIP-90 No.9
2020/11/25

ログ可視化手法を用いた脅威ハンティングにおける詳細分析支援の検討

山岸 伶[1]　片山 貴大[1]　川口 信隆[1]　重本 倫宏[1]　野澤 篤史[2]　山口 環[2]

概要：サイバー攻撃の高度化に伴い，既存のセキュリティ機器にのみ頼った脅威検知の見逃しが指摘されている．このような背景から，機器に依存せず，既に組織内に侵入している脅威の発見を目的とした脅威ハンティングが注目されている．脅威ハンティングでは，分析者がネットワークやエンドポイントといったログを分析し，痕跡を収集することで，脅威発見に至る．一方で，高度な脅威を発見しうる分析者不足が脅威ハンティング実施の障壁となっており，分析知見を十分に持たない分析者への支援が重要となっている．既存技術ではログ内の情報を抽出し，可視化することで概要把握や不審点の発見を支援するが，ログエントリ単位での詳細分析支援の観点に欠ける．そこで本研究では，脅威ハンティングにおける詳細分析の支援技術を検討する。まず、詳細分析時のログエントリの内容理解や関連するログエントリの発見を支援することに加え，関連手法とは異なった観点で不審点に関する気づきを与えることを要件としてあげた．次に，要件を満たしうるキーアイデアとして，ログエントリのアイコン化，ログの値間の共通性可視化，ログエントリ間の関連性の可視化をあげ，本アイデアを実現しうるシステムを提案した．

キーワード：脅威ハンティング，ログ分析，可視化

A study of Log Analysis Support for Threat Hunting by Log Visualization

REI YAMAGISHI[1]　TAKAHIRO KATAYAMA[1]　TOMOHIRO SHIGEMOTO[1]
NOBUTAKA KAWAGUCHI[1]　ATSUSHI NOZAWA[2]
TAMAKI YAMAGUCHI[2]

Abstract: With the sophistication of cyber attacks, it is pointed out that countermeasures relying only on security devices may miss significant threats. From this background, threat hunting, a technique to deeply analyze logs to discover evidences of already penetrating threats, has been attracting attention. In threat hunting, analysts analyze logs such as networks and endpoints to collect traces of advanced threats. On the other hand, the shortage of analysts capable of discovering them has become a barrier to threat hunting and it is important to support analysts who do not have sufficient analytical knowledge. The existing technologies extract the information from the logs and visualized them, but they usually focus on grasping the whole picture of the logs in total and often lack the support to the step-by-step analysis of each log entry. In this study, we discuss the requirements to support detailed analyzes, and consider the idea of the key visualization. Then, we propose a system that realizes a key idea and demonstrate the dashboard prototype.

Keywords: Threat Hunting, Log Analysis, Visualization

1. はじめに

サイバー攻撃は年々増加し，その手口も高度化している．特に，APT(Advanced Persistent Threat)攻撃は，攻撃者が特定の組織を標的にし，戦略的かつ執拗に高度な活動をする．この APT 攻撃では，攻撃者が対象となる組織を調査し効果的な脅威を選択するため，組織が対策を講じるのがより困難になってきている．

攻撃の高度化に伴い，組織が実施してきた従来のサイバー攻撃対策プロセスだけでは，脅威を見逃してしまうことが指摘されている[1]．従来の対策プロセスでは，セキュリティ機器を導入・監視し，セキュリティ機器が発するアラートの内容を調査･分析することで,攻撃に対応してきた．しかし，攻撃者は対象組織を熟知しており，こうしたセキュリティ機器の検知を逃れ，組織内に侵入している．

脅威の見逃しの対処として，セキュリティ機器の検知率を向上させて見逃し自体を軽減するアプローチと，見逃した脅威が重大になる前に早期発見するアプローチが考えうる．後者のアプローチとして，アラートでの検知後に分析を始める従来のプロセスに加えて，分析者自らが脅威を探し出す脅威ハンティングが着目されている．脅威ハンティングにおいて，分析者は脅威がすでに侵入していると仮定し，セキュリティ機器やネットワーク機器，メール，端末といったあらゆる機器から出力されるログを分析し，脅威の痕跡を発見する．

このように多種かつ膨大なログを分析し，高度な脅威の

1 株式会社 日立製作所
Hitachi Ltd.

2 株式会社 日立情報通信エンジニアリング
Hitachi Information & Telecommunication Engineering, Ltd.

Vol.2020-CSEC-91 No.9
Vol.2020-SPT-40 No.9
Vol.2020-EIP-90 No.9
2020/11/25

痕跡を発見する際には，セキュリティに関する専門的な知識が分析員に求められる．しかし，セキュリティ人材不足が問題視される現状[2]にともない，分析員として高い専門的な知識を有した人材を確保することが困難となっている．分析員やセキュリティ管理者を主対象としたアンケート調査によると,「組織の脅威ハンティング実施を妨げる要因が専門性を有した人員不足」と回答した人は 72.1%であった[3].

こうした人材確保の問題を背景に，専門的な知識を十分に持たない分析者への支援が重要になってくると考える．支援の一手法として，ログの可視化が挙げられる．ログ可視化の既存技術は，ログ内の情報を抽出し，可視化することでログの概要の理解や分析の起点となる不審点の発見に利点がある一方で，ログエントリ単位での詳細分析に対する支援の観点に欠ける．

そこで本研究では，詳細分析に必要な支援の要件をあげ，キーとなる可視化のアイデアを検討する．また，キーアイデアを実現しうるシステムを提案する．以降本稿では，2 章で本研究の背景や関連技術について述べ，3 章で提案手法について述べる．また，4 章で今後の課題について述べ，5 章で本研究についてまとめる．

2. 研究背景

本章では，研究背景として，脅威ハンティングのプロセスについて述べる．また，脅威ハンティング実施中に活用されるインテリジェンスとその種類について述べる．最後に，本研究の関連技術および関連研究について説明する．

2.1 脅威ハンティングのプロセス

1 章で述べたように，脅威ハンティングでは，脅威がすでに侵入していると仮定し，ログを分析することにより，高度な脅威の痕跡を発見する．分析対象となるログは，エンドポイントにおけるプロセス実行ログ, レジストリログ, ファイルデータ，ネットワークデータ，ネットワークにおけるセッションログ，ネットフロー，プロキシログ，DNS ログ，ファイアウォールログ，スイッチやルータログ，関連するインテリジェンス，過去のアラート情報，チケット情報，メールログなど多岐に渡る[4][5]．以下では，如何にしてこのログ分析を実施し，脅威の痕跡の発見に至るかのプロセスについて説明する．

脅威ハンティングのプロセスに関して，様々な企業・組織がモデルを提案している．例えば, Sqrrl 社による Hunting Loop[6]，Carbon Black 社による The Carbon Black Hunt Chain[7]が挙げられる．各モデルの詳細は異なるが，ある根拠をもとに仮説を立て，ログ分析により仮説を検証し，検証結果を自組織内/外で共有するといったおおよその流れは共通している．

本節では代表して，The Carbon Black Hunt Chain について述べる．The Carbon Black Hunt Chain は．図 1 が示すように, 脅威ハンティング開始から継続的な防御策の向上まで, 黒字の 8 段階に分かれている．以下では，本研究の対象となる脅威の発見に至るまでの 4 段階について詳細を述べる．

(1) 脅威ハンティングの開始(Starting the Hunt)

脅威ハンティングでは，はじめに関連するデータを収集し，明確な目的と範囲を定義する．これは，目的なく実施すると十分な効果を得られない可能性が高いためである．

(2) 正規な活動の除外(Filter out legitimate activity)

分析者は，疑わしい活動発見の際に分析する範囲を狭めるため，分析対象となるログから正規の活動を除外する．正規の活動の除外の際に，分析者は対象組織の環境やアプリケーションに関する知識を活用する．

(3) 疑わしい活動の発見(Find suspicious activity)

(2)の除外により残された活動は，すべて疑わしい可能性がある．これらの活動から，疑わしい活動を見つけ出す．なお，脅威ハンティングでは，侵入に関する活動だけを見つけるわけではない．

(4) 詳細分析(Deeper investigation)

(3)で絞り込んだ活動の詳細分析を実施する．この際, OS, アプリケーション，ネットワークデータフロー，過去の事例，異常な活動に関して専門家に支援を求めることも必要となる．詳細分析により，既知ではない攻撃の痕跡（アーティファクト）を発見することがある．

図 1 が示すように、分析者は(2) 正規な活動の除外(Filter out legitimate activity)，(3) 疑わしい活動の発見(Find suspicious activity)，(4) 詳細分析(Deeper investigation を繰り返すことで，脅威発見に至る．

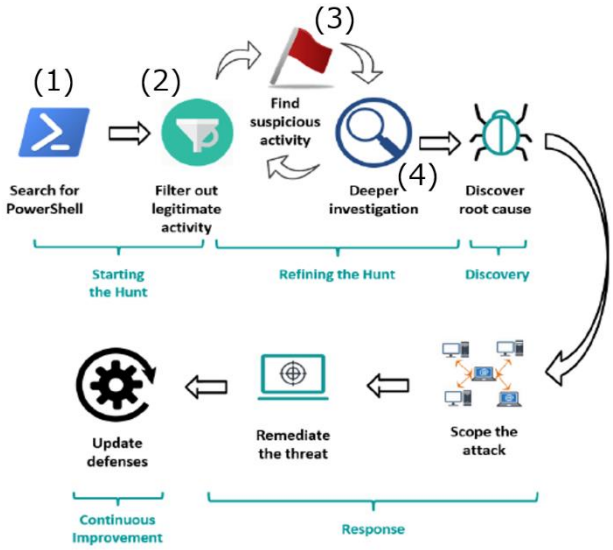

図 1. The Carbon Black Hunt Chain のプロセス[7]
（図中の(1)~(4)は著者が追記）
Figure 1. The Procedure of Carbon Black Hunt Chain [7].

Vol.2020-CSEC-91 No.9
Vol.2020-SPT-40 No.9
Vol.2020-EIP-90 No.9
2020/11/25

2.2 脅威インテリジェンス

脅威インテリジェンスは攻撃者に関する知識であり，脅威ハンティングでは分析員が脅威インテリジェンスを活用する． 以下では，脅威インテリジェンスの分類の一つである Pyramid of Pain[8]について述べる． David Bianco は Pyramid of Pain として，図 2 が示すように脅威インテリジェンスを 6 種類に分類した．Pyramid of Pain は上層ほど，攻撃者にとって変更するのが困難な情報である．以下では，6 種類の脅威インテリジェンスのうちを下層から順に説明していく．

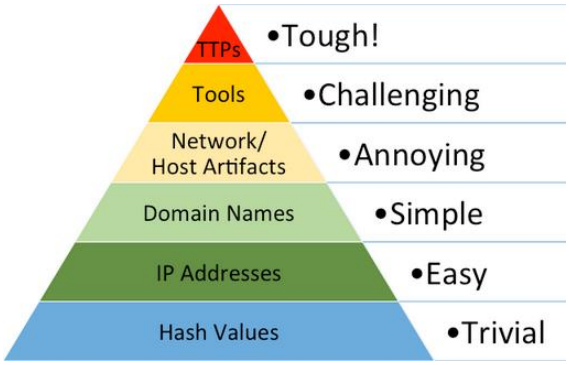

図 2. Pyramid of Pain[8]
Figure 2. Pyramid of Pain [8].

ハッシュ値

ハッシュ値は SHA1 や MD5 などのハッシュによって計算されたファイルごとの値である．ファイル内の文字等が一文字でも異なるとハッシュ値は変わるため，攻撃者もハッシュ値を容易に変更可能である．

IP アドレス

IP アドレスは，攻撃者がマルウェアの配布，通信する際に活用するサーバのアドレスを示す．攻撃者は Tor といった匿名プロキシサービス，クラウドサービスを活用により IP アドレスの検知を回避可能であるため，ファイルのハッシュ値と比較して変更は困難であるが，全体的にみると変更が容易である．

ドメイン名

ドメイン名は IP アドレスと同様に，攻撃者が活用するサーバのドメインを示す． IP アドレスと比較すると，ドメイン取得の際に DNS プロバイダーのサービス上で手続きをする必要がある点や，ドメイン取得からインターネット上で範囲されるまで時間を要する場合がある点から，変更が困難になる．

ネットワーク・ホストのアーティファクト

ネットワーク・ホストのアーティファクトは，攻撃者のふるまいに関する情報である．例えば，マルウェアがファイルをドロップする場所，攻撃者サーバにおける URL やパラメータのパターン，マルウェアが操作するレジストリキーなどが該当する．上記の 3 点と比較し，ネットワーク・ホストのアーティファクトはプログラムレベルで攻撃者が修正することが想定されるため，変更が困難になる．

ツール

攻撃者は，攻撃活動中に，被害環境内に存在するツールでなく，自身のツールを活用する．ツールを変更するのは攻撃者にとって，使い慣れたソフトウェアを変更することであるため，比較的困難である．

TTPs

TTPs は攻撃者の戦術，戦法，行動に関する情報であり，攻撃者の目的や知識にも関連する情報である．例えば，トロイの木馬の PDF ファイルによるスピアフィッシングやスピアフィッシング自体が TTPs に該当する．先述したように，攻撃者の目的や知識に関連するため，攻撃者が TTPs を変更するには新しい知識を学ぶなど時間を要する．

この TTPs は, ナレッジベースのフレームワークとして，MITRE ATT&CK[9]がまとめている．

ATT&CK は米国の非営利団体 MITRE によって開発された TTPs のフレームワークである. 当該フレームワークは，4 種類存在するが，以降では企業向けの攻撃を対象とした Enterprise のモデル(Matrices)を説明する． Matrix は縦軸と横軸で形成される表となっている．その横軸は tactics と呼ばれる攻撃者の達成する目的を手順に沿って表している．左側の Tactics は攻撃の初期段階での目的を表しており，表の右に行くほど最終的な目的となっており，計 12 段階の手順で構成されている．Matrix の縦軸は，Techniques と呼ばれる tactics を実現するための技術や手法を表しており，具体的な TTPs に関連付けられる．例えば，tactics の Execution の下にある AppleScript といったセルは Execution を実現する Technique である．各 Techniques に付随する情報として，当該技術を利用する groups（攻撃者グループ）とそのツールを示す Software，検知手法を示す Detection，緩和策を示す Mitigation が記述されている．これらの情報を活用することで, TTPs をより具体的にすることが可能となる．

2.3 関連手法

本節では，関連手法として，脅威ハンティングで活用される可視化ツールやテーブル形式で表示されたログへの可視化手法について述べる．最後に，関連技術・研究の課題について整理する．

なお, 以降ではログの中の IP アドレスの「192.168.0.1」などの一つ一つを値と呼び，「送信元 IP アドレス」といった値の種別をログのフィールドと呼ぶ．また，ログの値で構成される一行一行をログエントリと呼び，ログのエントリで構成される全体をログと呼ぶ．

ネットワークグラフ

脅威ハンティングに活用される可視化手法として，ログ

Vol.2020-CSEC-91 No.9
Vol.2020-SPT-40 No.9
Vol.2020-EIP-90 No.9
2020/11/25

内の値をノードとした図 3 のように，ネットワークグラフがある[5] [10]．ツールによりノードの種別は異なるが，IP アドレス，ドメイン名，端末名，ユーザ名，ファイル名などが該当する．また，エッジの色や形状により，内包や参照，実行といった関係性を表す．また，ノードをクリックすることで，図 3 の下部のように，当該ノードを含むログエントリをテーブル形式で表示し，詳細分析することが可能となる．分析者はノードとその関係性を俯瞰しながら不審な点を探し出し，一覧表示から詳細分析することで脅威を発見する．

kibana

kibana[11]はダッシュボードにおいて、様々なグラフとしてデータを可視化するオープンソースシステムである．グラフは，時系列グラフ，ヒートマップをはじめとして様々な種類が提供されている．脅威ハンティングでは，こうしたグラフからログの不審点を発見し，不審点に関連する検索クエリを駆使することで脅威が発見できる．

テーブル形式で表示されたログへの可視化手法

上記の手法は，ログの情報を抽出して可視化するが，図 3 下部に示すようなログのテーブル形式の画面に対し可視化を施す手法も存在する．見えログ [12]は高田らによって提案されたログ可視化のブラウザである．ログの分析に必要な機能を出現頻度，周期性としてあげ，ログのテーブル形式の画面に対し，可視化処理を施した．可視化処理として，出現頻度把握を目的とした出現頻度ごとの単語の背景色変更処理や，周期性把握を目的としたアウトライン表示処理を提供する．Emoji-nized Log Browser [13]は，高田らによって提案されたログ可視化ブラウザである．ログ情報の直感的な理解を目的に，事前に定義された絵文字と文字の対応に従い，ログ中の文字を絵文字に置換するログ可視化ブラウザである．

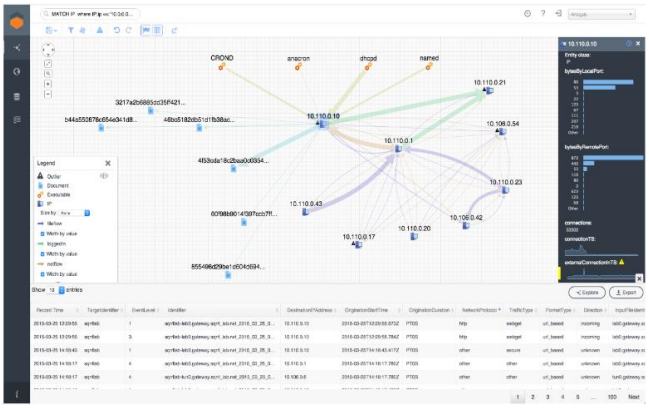

図 3. Sqrrl 社のシステムの画面例[5]
Figure 3. One of Sqrrl Enterprise UIs [5].

関連手法における課題

ネットワークグラフや kibana のグラフは，ログの情報を抽出して可視化することで，ログの概要や性質の理解を支援し，2.2 節で述べた(2) 正規な活動の除外や(3) 疑わしい活動の発見に貢献する．一方で，2.2 節の(4) 詳細分析におけるログのテーブル形式表示に対する可視化は対象としていない．例えば，(3) 疑わしい活動の発見においてネットワークグラフで不審な IP アドレスのノードと当該アドレスとエッジでつながるファイルやドメインを発見したとする．詳細分析で，分析員はファイルやドメインを含むログエントリをテーブル形式で表示し，ファイルが実行したコマンドなどのより詳細なファイルの挙動を分析する．このようにテーブル形式でログエントリを閲覧し，分析することがあるが，この詳細分析を対象とした可視化による支援がない．また，見えログや Emoji-nized Log Browser はテーブル形式の画面に可視化処理を施しているが，多種ログを同時に分析するといった脅威ハンティングの観点で支援を検討していない．

3. 設計方針

2.3 節で述べた通り，脅威ハンティングを対象としたテーブル形式のログを可視化により支援する手法は検討されていない．したがって，本研究では，テーブル形式のログを分析において，専門的な知識を十分に持たない分析者に対し，分析支援の提案を目的とする．本提案の意図は，ネットワークグラフといった関連手法を代替することでなく，関連手法で主目的でない(4)詳細分析を支援することや，(2) 正規な活動の除外や(3)疑わしい活動の発見において異なった観点で分析員に気づきを与えることにある．

本章では，可視化手法の設計方針について述べる．まず，テーブル形式のログ分析支援に必要な要件を述べる．次に当該要件を実現しうるのキーアイデアについて説明する．

3.1 テーブル形式のログ分析支援の要件

上述した通り，本研究では，テーブル形式での可視化により，関連手法とは異なった観点で，専門的な知識を十分に持たない分析員を支援する．本節では，テーブル形式でのログ分析に必要な要件を以下にあげる．

なお，これらの要件では，詳細分析のタスクを以下の 3 つに細分化して考える．まず 2.2 節で述べた(3)疑わしい活動の発見で疑わしいと考えたログエントリの示す内容やその特徴を理解する．次に，当該ログエントリと他のエントリを比較し，同じ/類似した値があるか確認する．ここでいう類似した値の一例は，同一ネットワーク環境にある IP アドレス，類似プロセス，同一機器の IP アドレス，DNS と IP，ユーザ名と所有端末，端末と IP アドレス，親子関係のプロセスであり，これらは分析時に必要な類似性をもつ値である．最後に，次に確認するログエントリを決定し，上記のタスクを繰り返し分析すると考える．これらの各タス

クに関する詳細は，以下の要件の中で述べる．

（要件 1）ログエントリの特徴理解を支援すること

ログエントリの特徴理解のタスクは，専門的な知識を十分に持たない分析者にとって困難であると考える．2 章で述べたように，ログエントリを TTPs の観点から分析することは重要となるが，ログエントリの内容を理解できた分析員でもこの観点をもつには，より多くの知識量が求められる．例えば，ATT&CK の technique の概念を知る分析員でも，各 technique を覚えることはより多くの知識を要求され，ログエントリの特徴と technique を対応付けるにはさらに多くの知識を要求される．したがって，このようにログエントリに対し対応する特徴を示し，理解を支援することが求められる．

（要件 2）ログエントリ間の比較を支援すること

ログエントリ間を比較し，同じ/類似した値があるか確認することは，ログエントリ数が多い点や類似する値は見えにくい点から，困難であると考える． 詳細分析で着目しているログエントリの値と同じ値が他ログエントリに含まれるか把握することは，当該ログエントリの特異性把握に必要であり，類似した値の把握も当該ログエントリの傾向把握に貢献する．例えば，同一ネットワーク環境にある IP アドレスがどの程度ログの中に含まれるか確認することで，感染拡大の有無の判断に貢献しうる．一方で，テーブル形式のログから一目で同じ/類似した値を把握することは困難であるため，支援が必要となる．

（要件 3）次に確認するログエントリ判断を支援すること

着目していたログエントリから次に確認するログを判断する際には上記の要件であげたログエントリの特徴や，同じ/類似した値が活用されると考える．例えば，同じ groups の特徴の持つログを分析するといった観点で分析が進行する．一方で，大量の他ログエントリに対し，同じ特徴等をもつか判定し，次に確認するログエントリを判断することは困難となるため，支援が必要となる．

（要件 4）詳細まで示すテーブル形式のログを俯瞰的に見ることによる疑わしい活動の発見を支援すること

ネットワークグラフや kibana では，情報を抽出し外観を可視化することで，疑わしい活動の発見を支援していたが，重要な情報を逃してしまう懸念がある．例えば，ネットワークグラフでは，時系列情報の把握が困難となる．一方で，テーブル形式のログはすべてのログエントリを表示するため，情報を切り捨てることはない．したがって，(2) 正規な活動の除外や(3) 疑わしい活動の発見において，本形式を俯瞰的にみることで，他手法とは異なる気づきを与えうると考える．例えば，特定の攻撃者グループに関連する活動を示唆するログエントリの発見は誤検知の可能性もあるが，全体俯瞰して見たときに，同攻撃者グループに関連する活動が多い場合，当該エントリを疑わしいと思うには十分な証拠になると考える．しかし，こうした観点でログエントリの傾向一つ一つを把握することは困難となるため，支援が必要となる．

3.2 キーアイデア

本節では，3.1 節であげた要件を実現しうるキーアイデアについて述べる．キーとなるアイデアは，ログエントリのアイコン化，ログの値間の共通性可視化，ログエントリ間の関連性の可視化であり，以下で詳細を述べる．

ログエントリの特徴のアイコン化

(要件 1)，(要件 4)を満たすことを目的として，図 4 上部が示すように，ログエントリごとに特徴定義セットからパターンマッチングを実施し，当該ログエントリの特徴をアイコンとして可視化する．したがって，アイコンはログエントリの持つ特徴を表しており，アイコンを閲覧することで，ログエントリの特徴を把握することができる（要件 1）．また，アイコンを俯瞰してみることにより，特定攻撃グループを示唆する活動の多発といった気付きを与えることが可能となる(要件 4)．なお， Emoji-nized Log Browser はログの値の置換として絵文字を活用したが，本手法はログエントリの特徴の可視化に利用したため，対象が異なる．

本研究における特徴は，一例として図 4 下部にあげている．種別として，ATT&CK の technique の特徴，tactics の特徴，groups の特徴や，異常値を含む特徴，アーティファクト(ホスト，ネットワーク)の特徴である．ATT&CK に関する特徴は，detection や Procedure Examples の項目から，異常値やアーティファクトは既存の異常検知，ふるまい検知のシグネチャからの活用により作成される．

ログの値における類似性の可視化

(要件 2)を満たすことを目的として，あるログの値をマウスオーバした際に，当該値と同一の値や類似した値の背景色を変更し，可視化する．なお，同一の値の場合は当該値と同一背景色になり，類似した値の場合は同系統の別の背景色となる．俯瞰的にこれらの背景色をみることで，ログエントリの特異性の理解することが可能となり，（要件 2）を満たしうる．

ログエントリ間の関連性の可視化

(要件 3)を満たすことを目的として，あるログエントリと他のログエントリとの関連性を線の色や太さとして可視化する．分析員によりあるログが指定された際，当該ログエントリから他のログエントリに対し，可視化した線が引かれ，次に見るべきログエントリの判断の参考に活用される．

本研究において，ある２つのログエントリの関連性は，当該ログエントリの関連度を表示する全ログエントリの関連度の最大値で割った数として計算される．なお，２つのログエントリの関連度は，共通する特徴アイコン(tactics を除く)の数，tactics の特徴値，同じ値の数，類似した値の数

Vol.2020-CSEC-91 No.9
Vol.2020-SPT-40 No.9
Vol.2020-EIP-90 No.9
2020/11/25

×0.5 の和で計算される．なお，tactics の特徴値は，tactics 上での距離(同じ場合は 1 とする)の逆数で表され，例えば Lateral Movement と Command and Control の差は 2 なので 1/2 となる．この関連性が線の太さの値となり，閾値で色を変化させる．

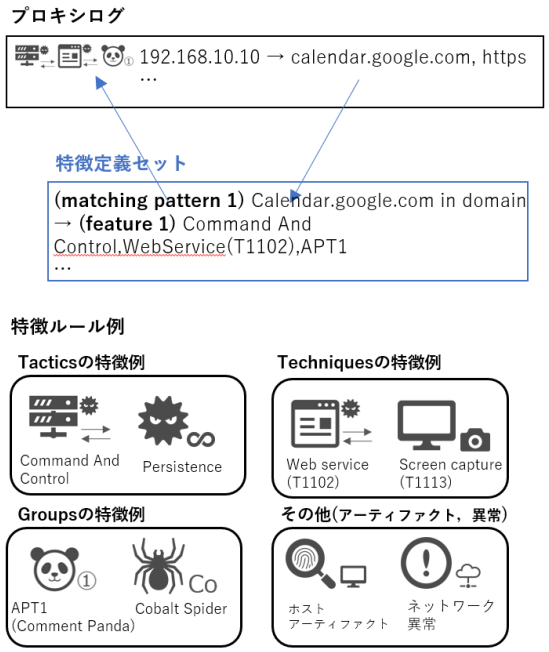

図 4. ログエントリの特徴のアイコン化詳細

Figure4. The Details of iconization of features in log entries .

4. 提案システム

本章では，3.2 節で述べたアイデアを実現しうるシステムを提案する．提案手法の構成図を図 5 に示す．以下では，図 5 の各部に分けて処理を説明する．

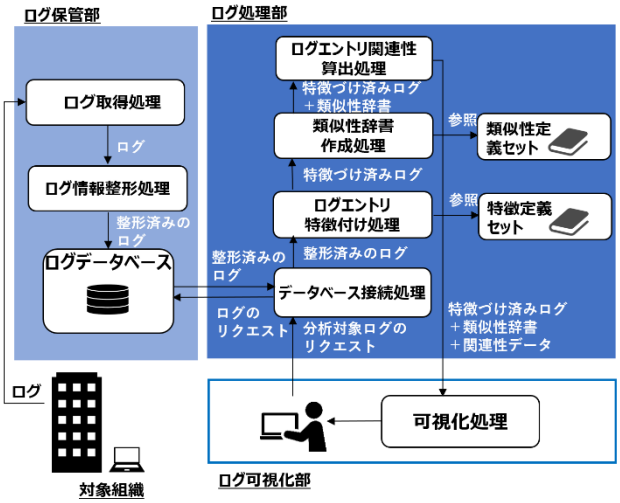

図 5. 提案手法構成図

Figure5. The Configuration of Our Proposal Method.

対象組織

対象組織は，脅威ハンティングの対象となる組織内の IT 環境を指す．本組織内では，PC やセキュリティ機器，ネットワーク機器等が存在しており，これらの機器はログを生成する．後述するログ保管部より，ログ要求の通信を受け取り，これらのログを当該部に送信する．

ログ保管部

ログ保管部は，対象組織よりログを取得し，保管する役割を持つ．具体的には，ログ保管部は，ログ取得処理，ログ情報整形処理，ログデータベースから構成される．ログ取得機能は，定期的に対象組織環境へログのリクエストを送信し，ログを取得する．次に，ログ情報整形処理は，可用性向上を目的として，取得したログを整形し，ログフォーマットを統一する．例えば，あるログでは「ip アドレス=192.168.0.1」，別のログでは「ip->192.168.0.1」と表記され，そのままデータベースで保管されると以降の操作可用性は低い．したがって，形式の異なるログを整形し，統一フォーマットとする．最後に，ログデータベースでは，ログ情報整形処理で，整形されたログを保管する．

ログ処理部

ログ処理部は，分析員のログのリクエストを起点とし，ログ保管部からログを取得し，ログに処理を施して，画面にログを表示する．この際ログは複数同時に指定可能である．以降で，具体的なログの処理を説明する．リクエストに応じて，データベース接続処理は，ログデータベースに要求を送信しログを取得する．次に，ログエントリ特徴付け処理は，ログエントリの特徴のアイコン化の実現を目的とし，事前に定義したログ特徴セットを参照し，パターンマッチ形式でログエントリの特徴付けを行う．ここで出力される特徴付け済みログは，ログエントリごとに，その特徴のフィールドを加えたログである．次に，類似性辞書作成処理はログの値間の共通性可視化の実現を目的とし，ログ中の値ごとに，当該値をキーとした場合に類似する値を一覧化した辞書を作成する．最後に，ログエントリ関連性算出処理は，ログエントリ間の関連性の可視化の実現を目的とし，ログエントリごとの特徴をもとに，関連性を算出する．ここで出力される関連性辞書は，ログエントリをキーとして，他のログエントリとの関連性(可視化する線の色と太さ)を一覧としているデータである．上記機能で処理された特徴づけられたログ，類似性辞書，関連性辞書はログ可視化部に送信される．

ログ可視化部

ログ可視化部は，分析員が脅威ハンティングを実施するインターフェースである．当該インターフェースで指定したログは，ログ処理部で処理され，特徴づけ済みログ，類似性辞書，関連性辞書の組として受け取る．当該部では，これらのログや辞書をもとに可視化処理を施し，図 6 のよ

Vol.2020-CSEC-91 No.9
Vol.2020-SPT-40 No.9
Vol.2020-EIP-90 No.9
2020/11/25

うに分析員に可視化・提示する．ログ可視化部は，上述したログを指定する部分等を有するが，本稿では可視化に関連する 1)ログ表示部と 2)関連性表示部を説明する．

1) ログ表示部

ログ表示部は指定されたログを表示する部分である．なお，ログ表示部は，相関分析を目的として，ログの種別ごとのコンポーネントを複数同時に表示可能である．図 6 では左側にプロキシログが，右側に EDR ログの例が表示されている．

ログ表示部では，それぞれ，1 行目にログの各フィールド名が存在し，それ以降の行はログエントリを表示する．

ログ表示部の 1 列目は，(a)のように，ログ特徴付け処理で算出されたログの特徴を表すアイコンが表示される．また，最終列には，関連性ボタンが表示される．当該ボタンをクリックすることで，後述の 2)に，クリックしたエントリと他のエントリの関連性が表示される．

なお，(b)で示すように，ログの値にマウスオーバーした際には，類似性辞書をもとに，背景色が変更され，他エントリの同値は同背景色に，他エントリの類似性のある値は関連する色の背景色に変更される．図 6 では，「https://mal.com」にマウスをあわせた際に，同一値は同じ青色に，対応する IP アドレスの背景色が薄い青色に変更されている．

2) 関連性表示部

関連性表示部では，(c)で示すように，ログエントリ間の関係性を，線形式で表示する．これらの線は，関連性辞書をもとに色や太さが決められており，該当する種別の数ごとに線の色は異なり，関連性が高いほど，線は太くなる．また，線は一度にすべて描画されず，上述したボタンをクリックすることで，クリックしたエントリと他のエントリの関連性が表示される．

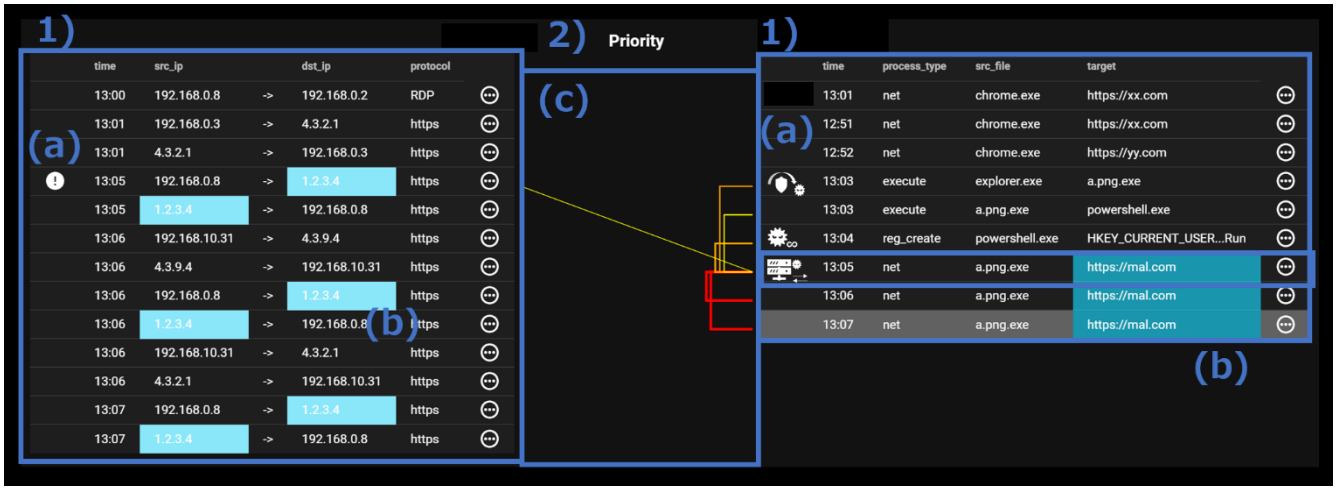

図 6. 提案手法構成図

Figure 6. One fo Visualization Screens with Our Proposal Method

5. 今後の課題

本節では，今後の課題として，実際のログによる実用性の評価と特徴の再考，ユーザビリティの評価と可視化手法の再考，特徴ルールセット作成の支援，関連性付けの高度化の 4 点をあげる．以下で詳細について述べる．

実際のログによる実用性の評価と特徴の再考

本研究では，ログエントリの特徴付けによる可視化を提案した．一方で，検討してきた特徴の種別で十分か，逆に，特徴付けが過剰に行われ本来見るべき特徴が埋もれないか，といった懸念点が上がる．実際のログを用いて，提案手法がこれらの懸念点に耐えうるか検証が必要である．

ユーザビリティの評価と可視化手法の再考

本研究では，専門的な知識を十分に持たない分析員の支援を目的とし，必要な可視化による支援を検討してきた．一方で，十分な支援ができているかユーザビリティの観点での評価が必要である．また，ユーザビリティ評価に基づき，より分析員を支援しうる可視化手法を再考することが今後の課題となる．

特徴定義セット作成の支援

特徴定義セットは，事前に定義するが，一つ一つの定義セット作成の負担が大きいことが想定される．人手による

作成は困難であるため支援が必要であると考える．

関連性付けの高度化

関連性付けの高度化が今後の課題としてあがる．関連性は, 単純に同一の特徴を持つかといった観点で算出したが, 組み合わせによってより関連性が高いと判定できるケースもある．例えば，感染拡大に関する脅威特徴かつ同一ネットワーク環境内の端末の類似性が重複する場合，関連性が高くなる．こうした特徴の組を考慮して，関連性を算出することが課題として残る．

6. おわりに

既存のセキュリティ機器に依存せず，既に組織内に侵入している脅威の発見を目的とした脅威ハンティングが注目されている．脅威ハンティングでは，分析者がネットワークやエンドポイントといったログを分析し，痕跡を収集することで，脅威発見に至る．一方で，高度な脅威を発見しうる分析者不足が脅威ハンティング実施の障壁となっており，分析知見を十分に持たない分析者への支援が重要となっている．既存技術ではログ内の情報を抽出し，可視化することで概要把握を支援するが，ログエントリ単位でのテーブル形式のログの詳細分析支援の観点に欠ける．そこで本研究では，テーブル形式でのログ分析に必要な要件を検討した．また，要件を実現しうるキーアイデアとしてログエントリの特徴のアイコン化, ログの値間の共通性可視化, ログエントリ間の関連性の可視化をあげ，これらのキーアイデアに基づいたシステムを提案した．今後は，実際のログによる実用性の評価と特徴の再考，ユーザビリティの評価と可視化手法の再考，特徴ルールセット作成の支援，関連性付けの高度化に取り組む．

参考文献

[1] DomainTools：2019 Threat Hunting Report, 入手先< https://www.domaintools.com/content/2019-Threat-Hunting-Report.pdf>(2020 年 9 月 14 日確認).
[2] 我が国のサイバーセキュリティ人材の現状について, 総務省，入手先 <https://www.soumu.go.jp/main_content/000591470.pdf>(2020 年 9 月 14 日確認).
[3] Mathias Fuchs：Is Your Threat Hunting Working? A New SANS Survey for 2020, 入手先<https://www.sans.org/reading-room/whitepapers/analyst/membership/39600> (2020 年 9 月 14 日確認).
[4] Sqrrl：Hunt Evil Your Practical Guide to Threat Hunting, 入手先 <https://www.threathunting.net/files/hunt-evil-practical-guide-threat-hunting.pdf> (2020 年 9 月 14 日確認).
[5] Sqrrl：Hunt Pedia, 入手先 <https://www.threathunting.net/files/huntpedia.pdf > (2020 年 9 月 14 日確認).
[6] Sqrrl：The Threat Hunting Reference Model Part 2: The Hunting Loop, 入手先 <https://www.threathunting.net/files/The%20Threat%20Hunting%20Reference%20Model%20Part%202_%20The%20Hunting%20Loop%20_%20Sqrrl.pdf> (2020 年 9 月 14 日確認).
[7] Carbon Black：Threat Hunting: The Thrill of the Hunt , 入手先 <https://www.carbonblack.com/blog/threat-hunting-thrill-hunt/> (2020 年 9 月 14 日確認).
[8] David J Bianco： The Pyramid of Pain, 入手先< http://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html/> (2020 年 9 月 14 日確認).
[9] MITRE：ATT&CK, 入手先<https://attack.mitre.org/> (2020 年 9 月 14 日確認).
[10] Siren: Cyber threat hunting and discovery with Siren link analysis: The tale of GoScanSSH, GandCrab, 入手先 <https://siren.io/cyber-threat-hunting-with-siren-link-analysis/> (2020 年 9 月 14 日確認).
[11] elastic：Kibana, 入手先< https://www.elastic.co/jp/kibana>(2020 年 9 月 14 日確認).
[12] 高田哲司, 小池英樹:見えログ: 情報視覚化とテキストマイニングを用いたログ情報ブラウザ, 情報処理学会論文誌, Vol.41, No.12, pp.3265-3275 (2000)
[13] T. Takada, T. Abe：Emoji-nized log browser: visualization of server-logs by emoji for system administrators, In Proceedings of the 2018 International Conference on Advanced Visual Interfaces, pp. 1-3.(2018)