サイバー攻撃に対する 能動的観測による収集データのモデル化と正規化手法

金谷 延幸† 津田 侑† 遠峰 隆史† 高野 祐輝† 井上 大介†

概要：標的型攻撃のような組織内に潜伏し持続的に発生するタイプのサイバー攻撃の観測では，攻撃者による標的組織内での活動状況を把握することが本質となる．そのためには，攻撃者を誘引できる解析環境を準備し，解析者のノウハウを活かして試行錯誤しながら攻撃者の継続的な活動を促す必要がある．我々はこのような観測を「能動的観測」と呼ぶ．能動的観測で収集されるデータには，観測事例毎の試行錯誤で生じる偏り，特に解析環境の構成による差異が含まれ，複数の観測事例で収集されたデータに対する環境横断的な解析は困難を伴う．そこで本稿では，観測事例毎の偏りを排除しつつ，本来のデータの意味を損なわずに正規化する手法を提案する．能動的観測における環境構成とそのデータをモデル化し，観測時の環境構成を定義することで収集データに生じる偏りの正規化を実現する．

キーワード：サイバー攻撃対策，モデル化，正規化，能動的観測，データセット

Data Modeling and Normalization for Active Monitoring of Cyber Attacks

Nobuyuki Kanaya† Yu Tsuda† Takashi Tomine† Yuuki Takano† Daisuke Inoue†

Abstract: Understanding the activities in the victim's environment of an adversary is the essence for the monitoring of the targeted attacks and the cyber attacks which persistently intrude into a victim organization. To understand the activities, cybersecurity analysts prepare analysis environments which lure such adversaries and encourage the adversaries to be continuously active through trial and error. We call the monitoring the active monitoring. However, the data of the active monitoring is biased because of the differences in the environments of trials, which are built based on analysts' knowledge, and cross environments analysis is thus non-straightforward. Therefore, this paper formally defines a standard data model for the analysis environments and, then, introduces a normalization method which removes the biases, non-essential information for analysis.

Keywords: cyber attack, modeling, normalization, active monitoring, dataset

1. はじめに

サイバー攻撃の対策には，攻撃者や攻撃手段に関する知見を蓄積する必要がある．これには，実際のサイバー攻撃の観測によるデータ収集が不可欠である．収集したデータは，データセットとして広く公開することで，さらなる研究開発の発展に寄与できる．これまでも，MWS データセットに代表されるサイバー攻撃対策技術開発向けのデータセットが開発，提供されてきた．例えば，情報通信研究機構（NICT）が提供する「NICTER Dataset」では，NICT が所有するダークネットセンサで収集したパケットデータを提供している[1]．このような，特定のセンサを設置して攻撃を待ち受けてデータを収集する観測方法を受動的観測と呼ぶ．インターネットにおける無差別型攻撃を主な対象とする受動的観測では，長期に渡って大規模観測網で大量のデータを収集することで，データのゆらぎが排除され均一かつ定常的に良質なデータセットを提供することができる．

一方で，標的型攻撃のような組織内に潜伏し持続的に発生するタイプのサイバー攻撃の観測は，攻撃者の標的とする組織内での活動の状況を把握することが本質となる[2]．そこで，解析者のノウハウを活かした試行錯誤により，攻撃を誘引し観測を行う**能動的観測**が必要となる．例えば，攻撃者を解析環境に誘引するために，攻撃者のプロファイルやマルウェアの静的解析結果に応じたドメイン名の変更，OS の変更といった環境構成の変更により，攻撃の標的となる組織を模倣する[3]．これは，有効なデータを収集するには非常に重要であるが，得られる観測データは解析者のノウハウを多分に含み，データにはこの試行錯誤の結果として偏りが含まれ，データ利用に不都合が生じる．特に，データの偏りを生じさせる主要因の一つである環境構成による差異は，環境横断的な解析を困難にし，データの有効性・汎用性を著しく低下させる．このため，能動的観測によって得られたデータから偏りを排除した観測データを作成する手法が必要となる．

そこで，我々は環境構成による差異を排除する正規化手法を提案する．この正規化手法は，環境構築時に使われた構成の定義を使いデータを変換することで正確な正規化を

†国立研究開発法人 情報通信研究機構
National Institute of Information and Communications Technology

実現する．本論文では，まず 2 章にて能動的観測により得られるデータにおける課題と提案手法の概要について述べ，次に 3 章にて標的型攻撃が発生する環境のモデルとそこで得られるデータについて考察し，データの形式化について議論する．次に 4 章にて，提案する観測データの正規化手法について述べ，5 章にて実装と応用事例について述べ考察を行う．最後に，6 章にて関連研究について考察し，7 章にてまとめと今後の課題について述べる

2. 能動的観測によるデータ収集の実現

ここでは，組織内で発生する標的型攻撃対策研究のために行われる能動的観測について議論し，収集されるデータの課題を明らかにするとともに，提案手法の概要について説明する．

2.1 能動的観測によるデータ収集に対する課題

標的型サイバー攻撃を対象として得られるデータには，以下の課題があるため，そのデータ収集が難しく，利便性・有効性が高いデータセットを作成することは難しい．

1. 解析者による能動的観測が必要なため，均一な環境での多量の事例蓄積が困難
2. 解析者毎に異なる環境でデータ収集されるため，単純比較などによる環境横断的な解析が困難
3. データに環境の詳細が含まれることによる環境情報漏洩の懸念

能動的観測では，攻撃者という人間を相手に長期に誘引することで解析結果が得られるため，均一な環境で多くの事例を蓄積することが難しいという課題がある．攻撃活動のデータを得るためには，誘引実験の成功に向けて実在の組織を模倣するといった解析者の試行錯誤が必要となり解析のコストが大きい．また，攻撃者という人間と対峙している以上，マルウェア単体の動的解析とは異なりその成功率は低くなる．このように，サイバー攻撃対策研究に活かせるようなデータセットを作成するには，多数の解析者が協力し，多様な環境で得られたデータを組み合わせることが必要となってくる．

次に，多様な環境で得られたデータを組み合わせて作られたデータセットには，その得られた環境毎の違いが，データの機械的な解析を難しくするという課題がある．例えば，解析者がある実験 A において攻撃行動「攻撃者は，乗っ取ったクライアント PC を使い，管理者に成りすまし，AD サーバに対し認証要求を送信した」というイベントを発見したとする．この事例によって得られた通信データは，その実験 A が行われたネットワーク環境によって，IP アドレスやドメイン名，管理者のユーザ名などが異なる．他の異なる環境で実施された実験 B にて同じようなサーバに対する攻撃行動が発生したとしても，サーバのアドレスが異なれば異種の通信と認識する可能性がある．実験 B の環境において，あるマシンが AD サーバであり，そのマシンの IP アドレスに対する通信である場合，この通信データは，実験 A のイベントと同等であると解析できる．つまり，異なる環境で得られたデータの解析には，環境の構成と関連付けて解析する必要が生じ，環境横断的な解析が困難となり，データの利便性・有効性を低下させてしまう．

さらに，誘引実験によって得られたデータを公開することによって，実験環境の詳細が漏洩してしまうという懸念がある．どのような実験環境を用意するかは，攻撃者誘引に関する貴重なノウハウである．また，誘引実験は，実在組織を標的とした検体を用いることもあり，実験データを公開することは標的となった組織に関する情報の漏洩に相当するとみなされる場合もある．これは，誘引実験だけではなく，現実に発生したインシデントのデータを利用する際にも同じ問題が発生し，標的型攻撃に関するデータの公開を困難にしている大きな要因でもある．一般的なインシデントレポート等では，被害組織の情報を隠すために IP アドレスをマスクするなどの処理が行われるが，データセットに対して不適切なマスク処理を施すと，データの意味が失われてしまう．

このように能動的観測に基づくデータセットを生成する際には，以上のような課題に対応する必要がある．そのために多様な環境構成でデータを収集し，環境横断的な解析を可能とするために実験間の差異を吸収し，場合によってはデータの意味を保持しつつ環境情報を隠蔽できる変換を施す必要がある．

2.2 提案手法

我々は，標準的な ICT 環境を表すモデルとして**標準モデル**を用意し，収集データをこの標準モデル上に射影することでデータを正規化する．この標準モデルに射影したデータを蓄積することで実験間の差異を吸収し，データ解析が容易となり，収集元となった環境情報を隠蔽したデータセットを提供することができる．この標準モデルへの射影は，標準モデルに基づきデータ収集先となった環境構成を定義し，この定義に従いデータを変換することで実現する．

データの形式化にあたり，攻撃者の振る舞いを想定して標準的な ICT 環境を設計する．そして，そこで発生するデータに対する分析によって攻撃活動の意味付けや関連性を明らかにする．異なる複数の環境で得られたデータに対し共通の意味付けをするために，本研究では標準モデルとデータ入手先の環境構成との関係を用いる．

解析者は自信のノウハウに基づき観測対象となる模擬環境の構成を変更するため，観測毎に環境構成に差異が生じる．この差異を**環境特徴情報**として定義する．我々はこれまでに標準モデルと環境特徴情報を用いて模擬環境をカス

タマイズする環境構築手法を提案してきた[4]．この手法では，まず典型的な ICT 環境の構成を事前に定義する．そして，能動的観測を行う際には，実験に応じて変更する部分を環境特徴情報として定義することで，解析者が望む模擬環境を構築する．すなわち，サイバー攻撃の誘引実験を行う模擬環境について変化しない要素と変化する要素を分離して考えたときに，変化しない要素が標準モデルであり，実験毎に変化する要素が環境特徴情報として捉えることができる．逆にデータの正規化では，環境特徴情報によって生じた差異を変換することにより環境毎の差異を吸収することができる．

標準モデルは，汎用的な ICT 環境のモデルである必要はなく，サイバー攻撃のデータ収集に最低限必要な単純なモデルであることが必要である．このためには，攻撃者によるサイバー攻撃を想定し,その行動に適合した構成をもち，さらに想像上の理想環境ではなく，実際の攻撃者を誘引しその活動を観測できる現実的な環境である必要がある．

次章では，まず標準モデルと環境特徴情報を明確化するため，観測対象となるサイバー攻撃を明らかにし，これに応じた標準モデルについて議論し，さらに可変となる要素について議論する．次に，この標準モデルに従い収集可能なデータについて議論することでデータの形式化を行う．

3. 模擬環境とデータのモデル化

ここでは，観測対象となるサイバー攻撃とその環境について考察し，さらにそこで得られるデータについて検討することでモデル化を試みる．

3.1 観測対象となるサイバー攻撃

本研究では，標的型攻撃のような組織内ネットワークで行われ，長期間に渡り潜伏しながら重要なデータを窃取するような攻撃の観測を対象とする．このようなサイバー攻撃では，以下のような手順にて攻撃されると考えられている[5].

Phase 1. 諜報: 標的に関する情報を収集
Phase 2. 侵攻: 標的組織内部の PC への攻撃
Phase 3. 潜伏: 標的組織内部の PC への侵入
Phase 4. 橋頭堡確保: 標的組織内部の PC を支配
Phase 5. 索敵: ネットワーク情報の探索
Phase 6. 浸透: ネットワーク内の踏み台を増殖
Phase 7. 占領: 目的となるサーバを支配
Phase 8. 収奪: 目的情報の入手と搬出
Phase 9. 撤収: 攻撃の痕跡の消去

このような標的型攻撃に対する観測は，これまでマルウェア等の解析に用いられてきたサンドボックスやハニーポットによる観測と比較し,Phase3 以降の観測が本質となる．

```
標準モデル = 外部ネットワーク，模擬環境;
外部ネットワーク ＝ セグメント;
模擬環境 = {セグメント}，ドメイン情報;
セグメント = {サーバマシン}，{クライアントマシン}，セグメント情報;
サーバマシン = "サービス"，{"サービス"}，マシン情報;
クライアントマシン = マシン情報;
マシン情報 = ネットワーク情報 ， プログラム情報，
                ファイル情報，マシン固有情報;
ネットワーク情報 = "IP アドレス"，"ホスト名"，"MAC アドレス";
プログラム情報 = {"プログラム"};
ファイル情報 =  {"ファイル"};
マシン固有情報 = "OS" ，{"ユーザ"};
ドメイン情報 = セグメント情報 ，"ドメイン名";
セグメント情報 = "ネットワークアドレス" ，"ネットマスク"，
                "ゲートウェイアドレス";
```

図 1　標準モデルの形式表現

このために，侵攻に成功し，橋頭堡と化した PC から，索敵・浸透先となる複数のマシンを用意した模擬環境を用意する．

サイバー攻撃のデータセットを作る上では，攻撃活動の対象となる環境とそこでの行動を正確に把握し，データを収集する必要がある．次に，これらの攻撃行動が発生する模擬環境を想定し，環境構成を表現するための標準モデルと攻撃行動についての関係について議論する．

3.2 環境構成の標準モデル

我々は，攻撃者行動の分析により環境の特性を明らかにし，その環境を単純化することで，観測対象の環境を標準モデルとしてモデル化した．この標準モデルの拡張バッカス・ナウア記法による形式表現を図 1 に示す．この標準モデルでは， 前節の攻撃行動に現れるマシンを列挙し，攻撃活動の対象となる環境は以下から構成されるとした.

1. 外部ネットワーク（インターネット等）のマシン
2. サービスを提供するサーバマシン
3. 次の浸透先となるクライアントマシン

また，これらのマシンは攻撃が発生する環境に応じた以下の状態を含む.

1. ネットワーク
   例： IP アドレス，ドメイン名，ホスト名など.
2. プログラム
   例： 事務処理用の文書作成ソフト，Apache WEB サーバなどのサービス用のプログラムなど.
3. ファイル
   例： 文書ファイル，設定ファイルなど．共有フォルダ上のファイルをアクセスする場合もある.

4. マシンの固有の情報
例： OS，ログインユーザなど.

この形式表現に従い模擬環境の例を表す．ある模擬環境はサーバセグメントと事務所を想定したネットワーク（事務所セグメント）の 2 のセグメントで構成される．サーバセグメントはサービスを提供するサーバマシンを複数含み，事務所セグメントは次の浸透先となるクライアントマシンを複数含む．サーバマシンは，固定の IP アドレスとホスト名を割り当て（ネットワーク情報），提供するサービスに応じたソフトウェアがインストールされ（プログラム），設定ファイル等にアクセスし（ファイル），Linux 系の OS で動作する（マシン固有情報）．クライアントマシンは，事務所セグメントに DHCP によって IP アドレスが割り振られ(ネットワーク情報)，事務処理用のソフトウェアがインストールされ（プログラム），PC 内や共有フォルダ内の文書ファイルにアクセスし（ファイル），Windows OS 上に異なるユーザでログインする（マシン固有情報）．

先に示した標準モデルの形式表現に従い，ある模擬環境を定義するデータ構造を JSON 形式とその擬似コードで表すことができる. 例えば, ある模擬環境 A のドメイン名は，以下のように表せる．

```
A{"模擬環境"}{"ドメイン情報"}{"ドメイン名"}
```

サーバマシンはサービスで識別し，例えばサーバセグメントに配備された AD サーバを以下のように表す.

```
A{"模擬環境"}{' サーバセグメント'}{"サーバマシン"}{'AD サービス'}
```

ここでは便宜的に，標準モデルで定義された属性はダブルクォートで，インスタンスの識別子はシングルクォートで記載する. サーバマシンが複数のサービスを提供する場合，１つのサーバマシンインスタンスは複数の要素から参照される．クライアントマシンは 識別をせず同等に扱うため，配列として表す以下の表現とする．

```
A{"模擬環境"}{' 事務所セグメント'}{"クライアントマシン"}[1]
```

このルールに従い，模擬環境構成に関する情報を形式化することが可能となる．たとえば，ある模擬環境の AD サーバの IP アドレスは以下のように表現できる．

```
A{"模擬環境"}{' サーバセグメント'}{"サーバマシン"}
    {'AD サービス'}{"マシン情報"}{"ネットワーク情報"}{"IP アドレス"}
```

このように､サイバー攻撃が発生する組織の ICT 環境を、単純化された標準モデルへと射影できる．この標準モデルに基づき標的型攻撃の誘引実験を行う模擬環境を定義し、そこで収集されたデータの意味付けを行う．

3.3 標準モデルにおけるデータの形式化

能動的観測によって，環境内では攻撃行動の痕跡として多様なデータを収集できる．例えば，通信データ，ホスト内のプロセスのデータ, サービスが出力するログ, Windows のイベントログなどである．本節では，これらの収集したデータの特徴を把握した上で形式化する．データの形式化は前節の環境モデルに基づいて議論を進め，サイバー攻撃対策のためのデータセットの作成を目指す.

```
{
   "proto": "smb",
   "src": A{"模擬環境"}{' クライアントセグメント' }
            {"クライアントマシン"}[1]{"マシン情報"}
              {"ネットワーク情報"}{"IP アドレス"},
   "dst": A{"模擬環境"}{' サーバセグメント' }
            {"サーバマシン"}{'AD サービス'}
              {"マシン情報"}{"ネットワーク情報"}{"IP アドレス"},
   "type": "auth_request",
   "time": 1525100809
}
```

図 2　SMB 解析結果に対する意味付け

3.2 節における模擬環境で攻撃者が行動すると，それに伴い様々なイベントが発生する．例えば，攻撃者が RAT を使って橋頭堡と化した PC を遠隔操作すれば，C&C サーバとの通信が発生する．またあるイベントをきっかけとして別のイベントが発生するなど，データ間に関連性が生じる．解析者は，サイバー攻撃の分析をする上で，これらいくつかの特徴的なイベントに着目し，攻撃者の行動を把握する．さらに，解析者は彼らのノウハウに基づき攻撃者の行動を観測する．ここで解析者は以下の攻撃行動に伴うイベントに着目する場合を考える.

1. インターネットへの通信
2. プログラム実行
3. ファイルアクセス
4. サービス利用の通信
5. その他のイントラネット通信

上記の解析者が着目するイベントが発生した場合の、収集可能なデータとその関連について，先に示した標準モデルを使いその関係性を表す. 例えば, ある模擬環境 A にて，サービス利用の通信「あるクライアントマシンから AD サーバに認証要求」というイベントが発生した時，SMB プロトコル（Windows 認証に使われるプロトコル）を解析した結果，以下の観測データが得られたとする.

```
{"proto":"smb",
   "src":"10.1.3.30","dst":"10.1.2.8",
   "type":"auth_request", "time": 1525100809}
```

この JSON データの意味は，標準モデルとそのデータ表現を使い図 2 のように表せる.

```
データ = ″プロトコル識別子″，送信元アドレス，送信先アドレス，″種別″，″時刻″;
送信元アドレス = A{″模擬環境″}{*}{″サーバマシン″}{*}{″マシン情報″}{″ネットワーク情報″}{″IP アドレス″} |
        A{″模擬環境″}{*}{″クライアントマシン″}[*]{″マシン情報″}{″ネットワーク情報″}{″IP アドレス″};
送信先アドレス = A{″模擬環境″}{*}{″サーバマシン″}{'AD サービス'}{″マシン情報″}{″ネットワーク情報″}{″IP アドレス″}
```

（注：簡便化のためここでは，ある要素に含まれるすべてのインスタンスを列挙するために，″*″という表現をつかう）

図 3　SMB 解析結果の形式表現

さらに，データ上の各要素と標準モデル上の要素を対応させることで，SMB プロトコルの解析結果は，図 3 のように形式化できる．

以上のように，環境の標準モデルを使い収集されたデータの意味と形式を,模擬環境の構成と対応させ定義できる．次の章では，標準モデルによる環境の定義と正規化手法について述べる．

4. 標準モデルとその差異に基づく正規化手法

これまでに，我々は標準モデルと環境特徴情報を定義することで，観測対象となる模擬環境の自動構築を可能にする手法を提案してきた．本章ではこの手法を応用し，能動的観測で得られたデータの意味付けや正規化したデータの生成について述べる．

4.1 標準モデルによるデータの意味付け

先に示した標準モデルとデータの形式化により，模擬環境の構成を定義し，複数の環境にて得られたデータに対する共通の意味を与えることができる．例えば，ある実験 A では，模擬環境として AD サーバに IP アドレス 10.1.2.8 を与えサーバマシンを構築したとする．また，別の実験 B では，AD サーバに IP アドレス 192.168.13.6 を与えサーバマシンを構築したとする．この場合，実験 A における IP アドレス 10.1.2.8 も，実験 B における 192.168.13.6 に対する通信も共に，標準モデル上の

```
A{"模擬環境"}{' サーバセグメント'}{"サーバマシン"}
    {'AD サービス'}{"マシン情報"}{"ネットワーク情報"}{"IP アドレス"}
```

で示される同一の要素に対応するデータを発生させると言える．

また，データ間の関係も，標準モデルと対応させることで明らかになる．例えば，実験 A では，標準モデルの Windows クライアントから IP アドレス 10.1.3.30 を与えたマシンαを構築した場合，マシンαから 10.1.2.8 のマシンに AD サービスに対する認証要求の通信が発生すれば，同期して 10.1.2.8 のマシン内に，10.1.3.30 からの認証結果のログデータが生成されるはずである．これは，標準モデルに示されるサービス

```
A{"模擬環境"}{' サーバセグメント'}{"サーバマシン"}{'AD サービス'}
```

が出力するログの関係により得られ，またこの関係はどの模擬環境においても同じ関係性が維持される．

このように，まず標準モデル上にてデータの意味と関係性を定義しておき，実際の攻撃が行われた環境におけるデータが，標準モデル上のどの要素に対応し，標準モデルにおけるどのベントに対応するかで，実際の攻撃データの意味と関係性を明らかにすることができる．

4.2 環境特徴情報を用いたデータの正規化

環境特徴情報を用いた定義手法では，環境特徴情報は標準モデルに従った模擬環境の差分を表すが，これは標準モデルに従った環境間の対応関係を表しているとも言える．この，環境特徴情報の定義から得られる模擬環境の差異を利用し，逆に模擬環境で得られたデータを，標準モデル上の対応関係に従い変換することで，データの正規化を実現できる．

例えば，ある攻撃観測事例 A では，IP アドレス 10.1.2.8 をもつ AD サービスが動作するサーバに対して通信が発生したとする．また，別の攻撃事例 B では，IP アドレス 192.168.13.6 をもつ AD サーバに対して通信が発生したとする．これら 2 つのデータの通信先は標準モデルの同じ AD サービスを示しており，これらを同じアドレスに置き換えればどちらも同じ AD サーバに対する通信として同等に取り扱う事ができ，環境横断的な分析が可能なデータセットを実現できる．異なる環境で収集されたデータであっても，標準モデル上の同じ要素に対応するデータを同じ値に変換する限り他データにおける要素との関連性は維持される．例えば，ある実験 A で得られた AD サービスに対する認証サービス通信に含まれる送信元 IP アドレスと，AD サーバのログデータに含まれるクライアント IP のアドレスは，標準モデルの同じ要素に対応し，両者が同じく置換されればその関係性を維持できる．

我々の自動構築手法では環境特徴情報を正確に反映した模擬環境が構築されるため，そこで得られるデータも環境特徴情報を正確に反映していると言える．したがって，データを環境特徴情報に従い標準モデル上のデータとして変換することで，実験毎に異なる模擬環境で生成されたデータの差異を正確に吸収したデータの正規化を実現できる．

5. 実装とケーススタディ

我々は，実際の攻撃者を誘い込み，そこで行われた攻撃から標的型攻撃に対するインテリジェンスや解析ノウハウを得るために，サイバー攻撃誘引基盤「STARDUST」を開発してきた[3]. STARDUST は解析者に提供され，各々の解析者による誘引実験の結果，多様なデータが蓄積されてきた.

本章では，STARDUST でのデータ収集の仕組みを説明し，データ正規化の実装とそのベンチマーク結果について述べ，さらに正規化手法とデータセットの適用事例について説明し，最後にこれらに対し考察する.

5.1 STARDUST でのデータの収集

STARDUST は解析で利用する模擬環境を自動構築する機能がある．この機能で構築された模擬環境やそれを構成する仮想スイッチは攻撃者に気づかれずにデータを収集するセンサを有する. 現在の STARDUST では，以下のデータを収集している.

1. ネットワークトラフィック
   模擬環境内の仮想スイッチからミラーした，模擬環境内を流れるすべてのパケットデータ
2. ホスト内情報
   Windows クライアント内のエージェントで収集され送信される，ホストの OS 情報や実行中プロセスの情報

さらに，ネットワークトラフィックを入力として，以下のプロトコルに関する通信を解析し蓄積する.

1. HTTP
2. DNS
3. ICMP
4. Syslog
5. SMB（Windows ファイル共有，リモートコマンド実行）

これらのプロトコル解析には，Takano らによる L7 レベルでのネットワークトラフィック解析が可能な「SF-TAP」 [6]（プロトコル１-４）と海野らによる Windows のファイル共有プロトコルを対象としたネットワークフォレンジック手法[7]（プロトコル５）を利用している．リアルタイムで解析された結果は，JSON 形式のファイルとして保存し（図 4），リアルタイムで解析者に参照される.

なお，STARDUST では，解析者が誘引実験を行う模擬環境を自動構築ツール「Alfons」を使って構築する[8]．構築された模擬環境に攻撃者を誘引し，観測によって得られたデータは模擬環境毎に分離され蓄積される．次に，この収集データからデータセットへの正規化に関する実装について述べる.


```
[
    {
        "a": null,
        "dst": "1.1.1.1",
        "name": "SOUMU02.abc.com",
        "src": "1.1.3.12",
        "time": 1525100809,
        "type": "query"
    },
    {
        "a": "1.1.3.11",
        "dst": "1.1.3.12",
        "name": "SOUMU02.abc.com",
        "src": "1.1.1.1",
        "time": 1525100809,
        "type": "answer"
    },
    {
        "a": null,
        "dst": "1.1.1.1",
        "name": "dns.msftncsi.com",
        "src": "1.1.3.12",
        "time": 1525100851,
        "type": "query"
    },
    {
        "a": null,
        "dst": "1.1.1.1",
        "name": "ns1.msft.net",
        "src": "192.12.94.30",
        "time": 1525100851,
        "type": "answer"
    }
]
```


図 4　収集データの例

5.2 データ正規化の実装とベンチマーク

本節では，STARDUST にて収集されたデータに対する本正規化手法を用いた実装について説明する.

正規化は構築時に定義された環境特徴情報に対応する要素から，データセットの環境定義に対応する要素へ変換することで実現する．まず，生成するデータセットに対応する模擬環境の構成を標準モデルに基づき定義する. 例えば，「AD サービスを提供するマシンの IP アドレスは 10.1.2.8」とした場合，この IP アドレスが変換先となる．変換元の要素には，観測対象となる模擬環境を構築したときに用いた定義を適用する．例えば，自動構築システムに環境特徴情報として「AD サービスを提供するマシンの IP アドレスは 192.168.13.6」と指定したとすれば，192.168.13.6 を 10.1.2.8 に変換する.

ある時点の，ある模擬環境で得られたデータを正規化する場合，以下のようにして正規化データを変換する.

1. 自動構築ツールから，模擬環境構築の際に定義した「環境特徴情報」を取り出す.
2. 環境特徴情報から，変換に必要な特徴を取り出す.
3. 各特徴がどの JSON 要素に対応し，その要素をどう置き換えるかを表す，変換テーブルを生成する.
4. 変換テーブルから，正規表現で置換を行う，変換スクリプトを自動生成する.
5. 自動生成された変換スクリプトを使い，蓄積された JSON ファイル中のデータを置換する.

表 1　ベンチマーク対象のデータのサイズ

プロトコル	ファイル数	データ数	ファイルサイズ
HTTP	14,894	3,017,176	21,935,721K
ICMP	14,897	1,502,787	826,719K
Syslog	17,415	148,044,741	42,348,632K
ホスト内情報	14,905	1,951,114	3,344,833K

環境特徴情報を反映した変換スクリプトを自動生成することで，高速なデータ正規化を実現している．さらに，正規表現による文字列置換にて実装することで，JSON データを読み込み，処理し，JSON 形式で書き出す実装と比べ高速な処理を実現する．この変換スクリプトは，定期的に実行されており，常に収集されたデータから正規化されたデータを生成している．

次に，実際に収集されたデータを使いベンチマークを行う．ベンチマークには，もっとも長期間（900 日弱）利用されている模擬環境から収集された，表 1 に示すサイズのデータを利用する．これらの全データに対して正規化処理を施した結果，297 分で処理が完了した．

5.3 データセットの適用

本節では，我々の正規化手法によって生成されたデータセットの適用事例として海野らの研究[7]について述べる．

海野らの研究では，STARDUST で取得したデータを利用し，提案手法の評価を実施している．文献[3]では，誘引実験によって様々な RAT と C&C サーバ間の HTTP 通信を取得した．この通信の中から攻撃者による浸透活動の目的と考えられるコマンド列を抽出した（図 5）．

次に，この誘引実験で得られたデータから正規化されたデータセットを生成し，海野らの解析手法を適用した．その結果，AD サーバに対する以下の通信が発生したことを確認した（表 2，参考文献[7]より引用）．

図 5 と表 2 に示す結果を比較すると，*net group* の実行が一致している（項番 1-5）．この結果から，誘引実験を実施した解析者による観測時に行われたインターネット通信に対する分析結果と，正規化したデータセットを利用した

```
netview
netgroup /domain
netview /doamin
netview /domain
arp -a
dir "c:¥program files"
dir c:¥
dir "c:¥Program Files (x86)"
net group /domain
net group "domain admins" /domain
```

図 5　インターネット通信に対する解析結果

AD サービス利用の通信に対する分析結果の対応関係が確認できた．さらに，解析時の観測では認識されていないファイルサーバに対する参照を抽出している（項番 6，7）．

5.4 考察

まず，5.3 節に述べたデータセットの適用事例が示すとおり，我々の正規化手法により生成されたデータセットは，観測時に得られたデータに対し正しく関係性を維持することがわかり，妥当性があると言える．さらに，正規化したデータセットを利用して得られた未認識攻撃イベントの抽出という事実は，このデータセットの有効性を証明すると言える．

また 5.2 節で示したベンチマークの結果からは，1 模擬環境の 1 日分のデータを 20 秒ほどで処理しており，複数の模擬環境に変換処理を実行しても十分な時間で処理が終わることがわかる．つまり，模擬環境が更新されると同時に新たな変換スクリプトを生成し，定期的に実行することで，常に最新のデータに対する最新のデータセットを維持することができる．つまり，我々の実装は，実用に十分な速度を実現していると言える．

表 2　AD サービス通信の解析結果（他の SMB プロトコル通信を含む）

項番	操作元	操作先	操作内容
1	192.168.10.11	192.168.20.31	net group /domain adserver.jp
2	192.168.10.11	192.168.20.31	net group /domain adserver.jp
3	192.168.10.11	192.168.20.31	net group /domain adserver.jp
4	192.168.10.11	192.168.20.31	net group /domain adserver.jp
5	192.168.10.11	192.168.20.31	net group 'domain admins' /domain
6	192.168.10.11	192.168.20.35	net use ¥¥FileServer¥総務部
7	192.168.10.11	192.168.20.35	READ 公募 2016¥**28 年度**再エネ**導入事業¥1*2*3*.docx

6. 関連研究

MITER 社は，サイバー攻撃を段階的に表現することで攻撃行動のモデルと情報蓄積のためのフレームワークを定義し，これに沿った攻撃・防御に関する様々な情報を蓄積している[9]．我々はこれらの知識に基づき観測環境を構築するとともに攻撃対象となる環境や攻撃行動を形式化し，環境とデータをモデル化した．我々の蓄積したデータはこれらのモデルとフレームワークに還元可能なものだと考える．

Gao らは，監視ツールによる攻撃調査を効率化するため，攻撃調査に必要な情報に特化したデータのモデル化を実施し，それに合わせたデータベースと問い合わせ言語，検索エンジンを提案している[10]．これはサイバー攻撃の観測から得たデータのモデル化という点で，我々のアプローチに類似している．ただし，彼らは攻撃調査に必要な情報を最小限に絞ることで効率化を実現しているが，我々は観測によって得られるすべてのデータを対象とする点が異なる．

データセットに関する議論として，Ahmed らは，ネットワーク異常検知技術を分類・調査している[11]．この中では，ネットワークトラフィック解析技術の評価に適用できるデータセットはプライバシー上の理由により限定されていると述べられている．さらにデータセットが作成時の環境（OS 等）に深く依存しているという問題意識が述べられている．Sharafaldin らは，IDS/IPS のベンチマークに適用可能な十数のデータセットについて，彼らの指標による有用性の評価を行っている[12]．データセット評価の動機について，データセットがプライバシー上の理由によりデータへ過度な変更が加えられ，現実データとの乖離が生じる可能性を述べている．これらは，サイバー攻撃研究のためのデータに関し，我々の研究と同じ問題意識を有している．

7. まとめと今後の課題

我々は，環境特徴情報により環境構成の差異を定義し，標準モデルを使い射影することで，能動的観測によるデータの偏りを正規化する手法を提案し，利便性・有効性が高いデータセットを実現した．これまでに我々が提案した標準モデルによって構築された模擬環境では，環境特徴情報が反映されたデータを収集できる．逆に，収集データを環境特徴情報で表される写像により変換することで，実験毎に異なる模擬環境で収集したデータの差異を吸収した正規化処理を実現する．本稿では能動的観測が可能な STARDUST に本手法を適用し，正規化されたデータセットを作成した．このデータセットは実際の研究開発で活用され始めている．

本正規化手法による変換は，サイバー攻撃に関連するデータの意味や関係性を維持しつつ，任意の環境構成に射影することを可能とする．これによりサイバーレンジを利用した演習シナリオの作成に適用することが考えられる．我々が提案した攻撃シナリオ定義が可能な標的型攻撃再現環境[13]に対して本研究で得られたデータセットを活用することで，より実際の攻撃に近い演習環境を提供できる．

今後の課題は，サイバー攻撃と得られるデータに対するより汎用的なモデル化である．本論文におけるデータ正規化のアプローチをさらに発展させ，標準モデルを利用したデータ表現の形式化を目指す．これにより，攻撃手法のモデル化や，サイバー脅威インテリジェンスの蓄積，解析ノウハウの形式化などが実現できると考える．

参考文献

[1] D. Inoue, M. Eto, K. Yoshioka, S. Baba, K.Suzuki, J. Nakazato, K. Ohtaka, and K. Nakao, "nicter: An Inciden Analysis System Toward Binding Network Monitoring with Malware Analysis," In WOMBAT Workshop on Information Security Threats Data Collection and Sharing, pp.58--66, 2008.

[2] Mandiant, "APT1: Exposing One of China's Cyber Espionage Units," 2013. https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf（2017 年 12 月 18 日参照）.

[3] 津田侑, 遠峰隆史, 金谷延幸, 牧田大佑, 丑丸逸人, 神宮真人, 高野祐輝, 安田真悟, 三浦良介, 太田悟史, 宮地利幸, 神薗雅紀, 衛藤将史, 井上大介, 中尾康二, "サイバー攻撃誘引基盤 STARDUST," マルウェア対策研究人材育成ワークショップ 2017 （MWS2017）, 2017.

[4] 金谷延幸, 津田侑, 遠峰隆史, 安田真悟, 井上大介, "環境特徴情報による模擬環境自動構築効率化手法の提案と実装," 2018 年暗号と情報セキュリティシンポジウム（SCIS2018）, 2018.

[5] 特定非営利活動法人日本セキュリティ監査協会 APT による攻撃対策と情報セキュリティ監査研究会, "APT 対策入門 新型サイバー攻撃の検知と対応, " 2012.

[6] Yuuki Takano, Ryosuke Miura, Shingo Yasuda, Kunio Akashi, and Tomoya Inoue, "SF-TAP: Scalable and Flexible Traffic Analysis Platform Running on Commodity Hardware, " In Proceedings of the LISA '15, pp. 25–36, 2015.

[7] 海野由紀, 森永正信, 及川孝徳, 古川和快, 金谷延幸, 津田侑, 遠峰隆史, 井上大介, 鳥居悟, 伊豆哲也, "標的型攻撃の被害範囲を迅速に分析するネットワークフォレンジック手法の提案, " 2018 年暗号と情報セキュリティシンポジウム（SCIS2018）, 2018.

[8] 安田真悟, 三浦良介, 太田悟史, 高野祐輝, 宮地利幸, "ビルディングブロック型模擬環境構築システム," インターネットコンファレンス 2015 （IC2015）, 2015.

[9] MITER, "Adversarial Tactics, Techniques & Common Knowledge". https://attack.mitre.org/.

[10] Peng Gao, Xusheng Xiao, Zhichun Li, Fengyuan Xu, Sanjeev R. Kulkarni, and Prateek Mittal, "AIQL: Enabling Efficient Attack Investigation from System Monitoring Data, " In Proceedings of the 2018 USENIX Annual Technical Conference (USENIX ATC 18), pp. 113-126, 2018.

[11] Mohiuddin Ahmed, Abdun Naser Mahmood, and Jiankun Hu, "A survey of network anomaly detection techniques," Journal of Network and Computer Applications vol. 60, pp. 19-31, 2016.

[12] Iman Sharafaldin, Amirhossein Gharib, Arash Habibi Lashkari, and Ali Ghorbani, "Towards a Reliable Intrusion Detection Benchmark Dataset," Software Networking, pp.177-200, 2017.

[13] 津田侑，神薗雅紀，遠峰隆史，安田真悟，三浦良介，宮地利幸，衛藤将史，井上大介，中尾康二, "標的型攻撃再現のための攻撃シナリオ定義インタフェースの実装, " コンピュータセキュリティシンポジウム 2014（CSS2014），2014.