セキュリティ・オペレーションにおける 秘匿データ分析システムの提案

西嶋克哉[1] 川口信隆[1] 重本倫宏[1] 近藤賢郎[2] 中村修[3]

概要：突発的に起こるサイバー攻撃からネットワークやシステムを守ることは，セキュリティ・オペレーションにとって不可欠の要件である．しかし，大規模化し巧妙になる攻撃を，自組織だけで守り抜くのは困難である．対策の一つとして，複数組織で攻撃関連情報等の情報共有を行い，得られた情報を元に事前に対策を行うことが考えられる．しかし，プライバシ・機密を含む生データの開示リスクや，安全に大量の生データを共有することのコストが，円滑な情報共有を阻害している．ここで我々は，セキュリティ情報の共有には，必ずしも生データ自体を必要としているわけではないことに着目する．例えば，自組織と同じ攻撃を受信した組織の発見や，コミュニティ内で同攻撃を受信した組織数を調査する場合には，共通の情報を持っているかどうかという情報のみが必要とされる．そこで本稿では，データ開示を伴わない分析を支援する，秘匿データ分析システムを提案する．本システムでは，データ自体は共有せず，データを分析する機能を配付・実行し，データ分析のために必要十分な情報のみを共有する．これにより，機密データの開示や大量データの共有を抑えることができ，情報共有を促進することが期待できる．

Proposal of confidential data analysis system for security operation

KATSUYA NISHIJIMA[1] NOBUTAKA KAWAGUCHI[1] TOMOHIRO SHIGEMOTO[1]
TAKAO KONDO[2] OSAMU NAKAMURA[3]

1. はじめに

突発的に起こるサイバー攻撃からネットワークやシステムを守ることは，セキュリティ・オペレーションにとって不可欠の要件である．しかし，大規模化し巧妙になる攻撃を，自組織だけで守り抜くのは困難である．

対策の一つとして，複数組織で攻撃関連情報等の情報共有を行い，得られた情報を元に事前に対策を行うことが考えられる．実際に国や各業界団体により，様々な情報共有の取組みが行われている[1][2][3].

一方で，前述の情報共有の取組みでは先ず，参加会員組織は取りまとめを行うオペレーションセンターに情報提供を行う．次いで，オペレーションセンターが情報の集約，分析を行い,会員に情報を発信するという形をとっている．このため，会員はオペレーションセンターの情報発信を待つことになり,リアルタイムに情報を得ることができない．また，分析された情報を受け取るということは，整理された情報が得られるというメリットがあるが，反対に，必要な情報が削減されてしまうデメリットがある．

これに対し，日立製作所では，慶應義塾大学と協力して，複数組織の SOC（Security Operation Center）を跨ったセキュリティ・オペレーション連携により，サイバー攻撃への集団防御を実現する,分散 SOC アーキテクチャを提案している[4]．本アーキテクチャは，組織間で生データを共有・分析することにより，リアルタイムで他組織の情報を活用することを目的としている．

一方で，共有するデータによっては，プライバシ・機密を含む生データの開示リスクや，安全に大量の生データを共有するコストにより，情報共有が阻害されてしまう．

ここで我々は，セキュリティ情報の共有では，必ずしも生データ自体を必要としているわけではないことに着目する．例えば，自組織と同じ攻撃を受信した組織を発見することや，コミュニティ内で同攻撃を受信した他の組織数を調査する場合には，攻撃に関するメールやログといった生データ自体は必要ではなく，共通の情報を持っているかどうかという情報のみが必要とされる．

そこで本稿では，生データの開示を伴わない分析を支援する，秘匿データ分析システムを提案する．本システムでは，データ自体は共有せずに，データを分析する機能を配付・実行し，データ分析のために必要十分な情報のみを共有する．これにより，機密データの開示や大量データの共有を抑えることができ，情報共有を促進することが期待できる．

本稿の構成は以下の通りである．まず 2 章で関連研究を説明し，3 章で提案システムの要件定義について説明する．4 章では提案手法の設計について述べ,5 章でまとめと今後の課題を述べる．

2. 関連研究

機微情報を扱ったセキュリティ情報共有に関する文献はいくつか存在する．

1 株式会社日立製作所 研究開発グループ
Hitachi, Ltd. R&D Group
2 慶應義塾インフォメーションテクノロジーセンター本部
Headquarters of Information Technology Center, Keio University
3 慶應義塾大学環境情報学部
Faculty of Environment and Information Study, Keio University

東野[5]は，標的型攻撃メールか否かの判断が困難化しているという課題に対し，分散型の情報セキュリティ教育システムを提案している．同システムでは先ず，ある組織が受け取った標的型攻撃メールの無害化・匿名化を行う．次いで，複数の組織が持つサーバ間で同メールを共有することで，共有した情報を元に標的型メール攻撃対応訓練を行うことを可能とする．

また，齊藤ら[6]は，標的型攻撃対策のためにマルウェア検体情報の共有が必要であるが，マルウェア検体は機密情報等を含むファイルである場合があり外部に提供しづらい，という課題に対し，検体に影響を与えることなく，個人，組織に関連する情報を消去する技術を提案している．同技術は、標的型攻撃で利用されることが多い PDF 形式の検体に関して，その挙動に影響を与えるテキストやメタデータなどのコンテンツ部分を特定し，当該箇所以外を削除または固定値に置換する墨塗り手法である．

これらの文献では，機微情報の共有に，匿名化を用いている．その他，機微情報を保護しながら情報分析を行う方法として，データを隠したまま任意の関数を計算する秘密計算という技術が存在する[7]．この秘密計算をセキュリティ情報共有に活用した文献は，以下がある．

Melis ら[8]は，複数組織で得られたログやアラートから将来の攻撃元を予測する，Collaborative predictive blacklisting (CPB)を提案している．同稿では，semi-trusted authority (STA)と呼ばれる機能が，各組織の暗号化された攻撃元 IP アドレス情報を用いて，生ログへのアクセス無しに各組織のクラスタリングを行う．そして，各組織は同一クラスタ内で生ログの共有を行う．これにより，各組織はデータ開示を抑えながら，データ共有の効果を得る事が可能となる．

上記の参考文献のように，機微情報のセキュリティ情報共有では，共有される情報が標的型攻撃メール，マルウェア付き文書ファイル，攻撃元 IP アドレスと多岐にわたる．また，分析の方法も，匿名化や暗号化状態のままクラスタリングを行うなど，複数存在する．そこで本稿では，このような多種多様な分析要件に対して柔軟に対応するプラットフォームとなるシステムを提案する．

3. 要件定義

本章では，組織間のセキュリティ情報共有の現状において，情報共有を阻害している要因について説明し，当該要因を解決するセキュリティ情報共有のユースケースを示す．また，それらのユースケースを実現するために提案システムが満たすべき主な要件について整理する．

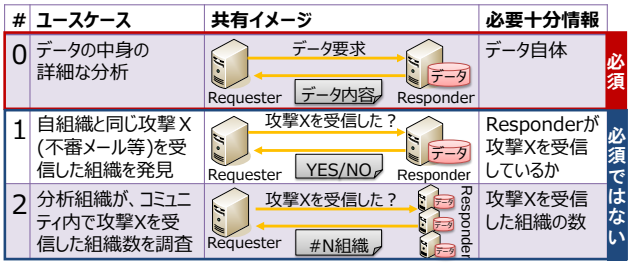

図 1 セキュリティ情報共有のユースケース

3.1 組織間のセキュリティ情報共有の現状

前述のとおり，複数組織が協力して防御を行うためには情報共有が不可欠であるが，機微な生データの共有は困難である．これは，主に以下の 2 つの要因から成る．

- 要因 1 : プライバシ・機密を含むデータの開示リスク
- 要因 2 : 安全に大量データを共有するための高いコスト

要因 1 の具体例としては，顧客との契約，General Data Protection Regulation (GDPR)といった法の違反や，共有相手組織からの情報漏洩があげられる．また，要因 2 は具体的には，データ共有にかかるネットワーク帯域やストレージ，暗号化処理などのコストが，データ共有のメリットを上回ることである．

これらの要因により，情報共有の取組等で同一コミュニティを形成していても，機微な生データ共有へのハードルは依然として高い状態である．

3.2 セキュリティ情報共有のユースケース

セキュリティ情報共有のユースケースイメージを示す．（図 1）．

図 1 に示す通り，セキュリティ情報共有のユースケースには，以下の 3 つが考えられる．

- ユースケース 0 : 共有データの中身を利用した詳細な分析
- ユースケース 1 : 自組織と同じ攻撃 X を受信した組織の発見
- ユースケース 2 : 分析組織による，コミュニティ内で攻撃 X を受信した組織数の調査

ユースケース 0 では，データの中身を必要とするため，データに機微情報が含まれている場合には，実現が困難となる．一方でセキュリティ情報共有では，ユースケースの 1 や 2 のように，データ自体の共有が必須ではないユースケースも考えられる．これらの場合，データ自体を共有せずに分析を行うことにより，3.1 項の要因 2 が解決する．また，連携組織が情報分析に必要十分な情報の開示を許容し，必要十分情報を超える開示が無いことを技術で保証することができれば，要因 1 が解決される．

従って，ユースケース 1 や 2 のような分析を行うことを可能とするプラットフォームシステムがあれば，情報共有

表 1 主要要件

#	ユースケース	必要十分情報	開示要件	
			共通	個別（ユースケース毎）
1	自組織と同じ攻撃Xを受信した組織の発見	Responder が攻撃 X を受信しているか	【開示制御性】 ・データ自体を開示しない（要因 2） ・分析対象の開示データを Responder が制御できる（要因 1）	【秘匿検索性】 攻撃 X を受信していない Responder には，Requester の問合せ対象が攻撃Xだと判らない（要因 1）
2	分析組織による，コミュニティ内で攻撃 X を受信した組織数の調査	攻撃Xを受信した組織の数	同上	【匿名性】 Requester は，どの組織が攻撃を受けているかは判らない（要因 1）

が促進されることが期待できる．

3.3 実現すべき主要要件

3.2 項のユースケース 1，2 を実現するために提案システムが満たすべき主要な要件について示す（表 1）．

表 1 に示す通り，提案システムの主要要件には，以下の 2 つが考えられる．

- 開示要件：「データ自体の非共有，必要十分を超える情報の非開示」を保証
- 分析要件：ニーズに応じた多様な分析の実現

開示要件は，ユースケース毎に異なる個別のものと，ユースケース共通のものが存在する．個別のものとしては，ユースケース 1 では，秘匿検索性が必要となる．これは，検索を行う側(Requester)が何を検索したのかを，検索をされる側(Responder)に対して秘匿化する性質である．

一方，ユースケース 2 では，匿名性を満たす必要がある．これは，Requester が，Responder のどの組織が攻撃を受けているかを知りえないという性質である．また，ユースケース共通の開示要件としては，開示制御性がある．これは，データ自体を開示しないことや，Responder が開示する結果を制御できる性質であり，これにより Responder の不本意な情報開示を抑制することが可能となる．Requester, Responder が必要十分情報の開示に合意し，かつ情報共有手段が開示要件を満たすことが保証されれば，要因 1 及び 2 は解決されると考える．

分析要件は，ユースケース毎に異なるニーズ（ユースケース 1 では自組織と同じ攻撃を受けた組織の発見，ユースケース 2 ではコミュニティ内で自組織と同じ攻撃を受けた組織数の調査）に対して，必要な分析を行うために，実施する分析を柔軟に変更可能とすることである．

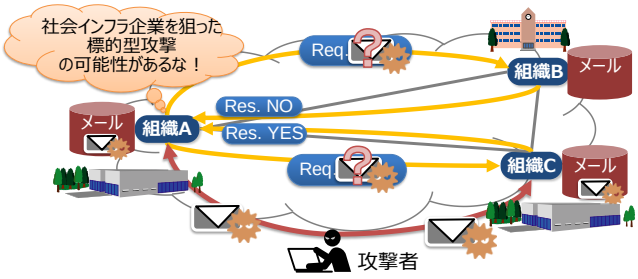

図 2 具体的なユースケースイメージ

3.4 具体的なユースケース

3.2 項のユースケース 1 の具体的なユースケースとして，不審メールの情報共有により，攻撃の目的・規模を推定するケースを示す（図 2）．

上記ユースケースの具体的な処理は以下である．

(1) 社会インフラ系組織である組織 A（Requester）は，受信した不審メールと類似したメールを受信したかどうかを，学術系組織である組織 B（Responder），および，社会インフラ系組織である組織 C（Responder）に問い合わせを行う．

(2) 組織 B，組織 C は，組織 A の問い合わせ内容を知ることなく，また，自身のメール内容を開示することなく，類似メールの有無を組織 A に返答する．

(3) 組織 A は，組織 B には類似メールが届いておらず，組織 C には類似メールが届いているという事実から，攻撃者は社会インフラ系を対象として攻撃を行っていると推定する．

本ユースケースを用いて，問い合わせの返答ごとに，各組織が得られる情報について説明する．

表 2 共有するデータ例

#	データ	用途
1	メール	他組織が受信した不審メールの情報により，攻撃の目的・規模を推定する
2	ドメイン名･IP アドレス	他組織の IP アドレスやドメイン名のブラックリストにより，自組織のブラックリストの更新を行う
3	添付ファイル	他組織が受信した添付ファイルの情報により，攻撃の目的・規模を推定する

- 回答が NO であった場合

  Requester は，Responder が不審メールを受信していないか，あるいは，Responder によって開示が拒否されたと判断する．一方 Responder は，Requester から問い合わせがあったことはわかるが，Requester が問い合わせたメールの文面は得られない．
- 回答が YES であった場合

  Requester は，Responder が不審メールを受信している可能性が高いと判断する．一方 Responder は，Requester から問い合わせがあったことはわかるが，Requester が問い合わせたメールの文面は得られない．Requester と Responder が得られる情報に公平性を保つ場合は，Requester が検索したメールを Responder に開示する等の工夫が必要となる．

また，上記では，類似メールの検索を例として述べたが，セキュリティ情報の共有で共有されるデータは，メール以外にも考えられる（**表 2**）．また，共有するデータによっては，例えば IP アドレスなどでは，類似検索ではなく，完全一致による検索が考えられる．

4. 設計

4.1 機能設計

3.3 項の主要要件を満たすために提案システムが具備すべき機能について示す（**表 3**）．

機能 1-1，1-2 は分析要件を満たすために必要な機能である．Responder 側に分析用プラットフォームを設置し（機能 1-1），分析要求に応じたロジックをコンテナ化・配信・実行（機能 1-2）することで，Responder 側で分析処理を行いその結果だけを共有することが可能となる．また，Requester は様々な分析要件を柔軟に選択することが可能となる．

機能 2 は，ユースケースに依らない共通の開示要件を満たすために必要な機能である．Responder 側プラットフォームにアクセス制御の仕組みを導入することで，分析対象のデータを制御することが可能となる．

表 3 機能設計

#	分析・開示要件	機能
1	[分析要件] 多様な分析要求に柔軟に対応	1-1. Responder 側に分析用プラットフォームを設置 1-2. 分析要求に応じたロジックをコンテナ化・配信実行
2	[共通開示要件] 分析ロジックのデータアクセスを制御	2 Responder 側プラットフォームにアクセス制御の仕組みを導入
3	[個別開示要件] 分析ロジックがユースケースの要件を満たすことを保証	3-1. 分析ロジックの健全性を検証するエンティティを導入 3-2. 数理技術で分析ロジックの動作を数学的に保証

機能 3-1，3-2 は，ユースケース毎の開示要件を満たすために必要な機能である．分析ロジックの健全性を検証するエンティティを導入し（機能 3-1），数理技術で分析ロジックの動作を数学的に保証する（機能 3-2）ことで，秘匿検索性や匿名性等，ユースケース毎に求められる開示要件を満たすことが可能となる．

4.2 アーキテクチャ

4.1 項の機能要件を満たすことを目的として設計した提案システムのアーキテクチャを示す（図 3）．

以下に，各役割について説明する．

- Requester

  情報を検索・分析するシステム．分析を行う相手（Responder）の選択，分析のために利用するロジックの選択，分析ロジックを利用した分析を行う．
- Responder

  情報を検索され，分析に必要な情報を Requester に提供するシステム．Requester に指定された分析ロジックを，後述する Distributor から取得し，健全性を保証し実行する．
- Distributor

  分析に利用する分析ロジックの保管･配付を行うシステム．分析ロジックは事前に動作健全性の検証を行っておく．

また，以下に，各要素について説明する．

- プラットフォーム I/F（インターフェース）

  Responder の選択，分析のために利用する分析ロジックの選択操作を行うための I/F．

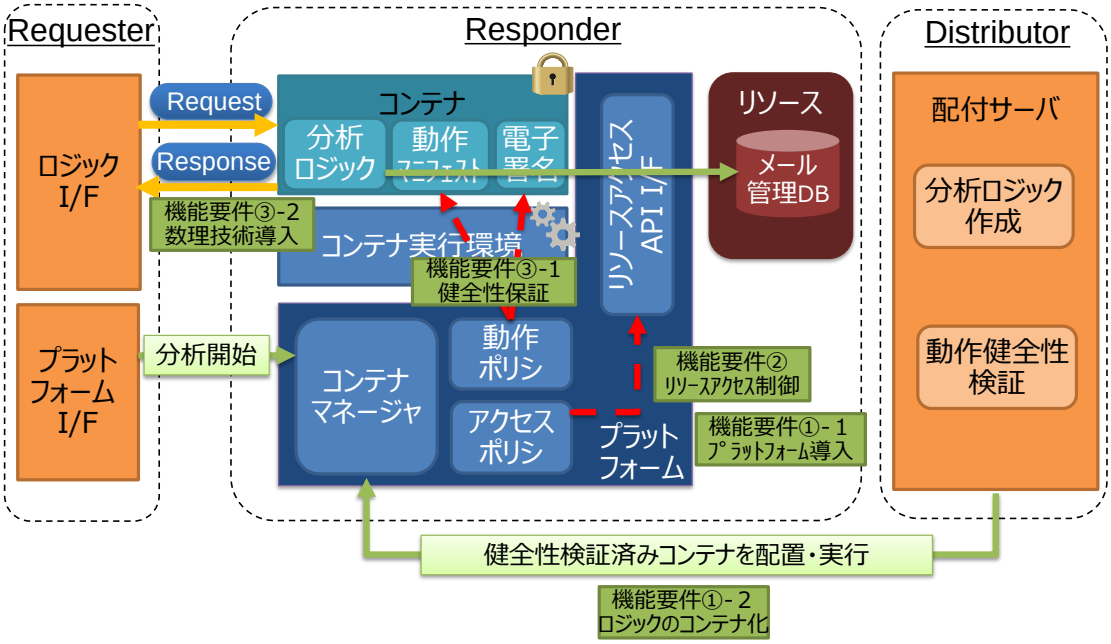

図 3 アーキテクチャ

- ロジック I/F
  分析ロジックを用いた分析操作を行うための I/F.
- コンテナマネージャ
  Distributor からコンテナイメージの取得，コンテナイメージの実行，停止，削除を行う.
- 動作ポリシ
  Responder が許可する動作についてのポリシ.
- アクセスポリシ
  分析ロジックの，Responder のリソースへのアクセス制御についてのポリシ.
- リソースアクセス API I/F
  分析ロジックがリソースにアクセスする際の I/F. アクセスポリシを参照し，リソースへのアクセス可否を判断する.
- 分析ロジック
  暗号化した状態で一致した値を検索するなど，個別の分析を行うためのロジック.
- 動作マニフェスト
  分析ロジックが実行する内容を宣言したもの.
- 電子署名
  分析ロジックに付与された署名.

また，提案システムの処理フローを以下に記述する.

(1) Requester のプラットフォーム I/F が分析要求を Responder のコンテナマネージャに送る.
(2) Responder のコンテナマネージャが Distributor から，分析したいロジック（コンテナ）を取得し，実行する．尚，当該コンテナは予め Distributor にて作成，動作の健全性検証を行っておく.
(3) Responder にて，予め用意しておいた動作ポリシと，コンテナに含まれる動作マニフェスト，電子署名を用いて，コンテナの健全性を確認する.
(4) Requester のロジック I/F に，検索要求を行う.
(5) Responder の分析ロジックが，リソースにアクセスし，検索結果を返答する．この際，予め用意したアクセスポリシを参照することで，Responder 側のデータ開示を制御する.

5. まとめと今後の課題

本稿では，データ開示を伴わない分析を支援する，秘匿データ分析システムを設計し，提案した.

データ分析機能を相手組織に共有することにより，柔軟な分析要件に対応することが可能となる．また，分析に秘密計算を活用すれば，互いの情報を暗号化したまま分析が行える.

本システムにより，セキュリティ情報共有を阻害している要因である，(1)機微情報の開示リスクや，(2)情報共有にかかるコスト，を低減することができ，情報共有が促進されることが期待できる.

今後は，提案システムの実装，評価を行っていく.

参考文献

[1] サイバーセキュリティ協議会 https://www.nisc.go.jp/conference/cs/kyogikai/index.html (Last Visit:2020/04/17)
[2] サイバー情報共有イニシアティブ（J-CSIP（ジェイシップ））https://www.ipa.go.jp/security/J-CSIP/index.html (Last Visit:2020/04/17)
[3] 一般社団法人 ICT-ISAC https://www.ict-isac.jp/ (Last Visit:2020/04/17)
[4] 近藤 賢郎: 分散型 SOC アーキテクチャに基づいた複数組織

間におけるセキュリティ・オペレーションの連携，マルチメディア，分散協調とモバイルシンポジウム 2018 論文集，pp.872-878 (2018)
[5] 東野 正幸: 組織間協働可能な標的型メール攻撃対応訓練システムの設計，研究報告モバイルコンピューティングとパーベイシブシステム（MBL），No.20，pp.1-5 (2019)
[6] 齊藤 真吾 他: 標的型攻撃情報共有のための文書型マルウェアの墨塗り手法，Computer Security Symposium 2013，pp.9-16 (2013)
[7] 菊池 亮 他: 秘密計算の発展，電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review，2018，12 巻，1 号，p. 12-20 (2018)
[8] On collaborative predictive blacklisting: L. Melis, A. Pyrgelis, E. De Cristofaro, ACM SIGCOMM Computer Communication Review, Volume 48 Issue 5, (2018)