様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　元 年　 ６ 月 １０ 日現在

機関番号：１２６０８
研究種目：基盤研究(C)（一般）
研究期間：2015～2018
課題番号：１５Ｋ００１１５

研究課題名（和文）高度・多様化するセキュリティ機器を考慮した大規模分散運用システム

研究課題名（英文）A large-scale distributed log operation system considering advanced and diversified security equipments

研究代表者
松浦　知史（MATSUURA, Satoshi）

東京工業大学・学術国際情報センター・准教授

研究者番号：００５３３８４５

交付決定額（研究期間全体）：（直接経費）　　3,500,000円

研究成果の概要（和文）：サイバー攻撃やインシデント対応の種類によって、セキュリティログに対する検索要求は大幅に異なり、特定の静的な論理空間では検索パターンを吸収しきれない。そこで時間軸に基づく分散環境の構築が妥当だと結論づけ、その設計思想に合うシステムを実際にSOCに設計/導入し検証を重ねた。時間軸の制約を持った分散環境では、他の属性(例えはIPアドレス)のみによる検索を行うと全件調査となり効率が極端に下がるが、サイバー攻撃の種類によって要求は異なるものの、時間制約を課した検索は多くの場合問題にならず、幅広いケースで利用でき実際の運用環境においても有用である事が示された。

研究成果の学術的意義や社会的意義
サイバー攻撃が高度多様化する中で多層防御が一般的となり、各組織で対策が進んでいる。一方でそれらのセキュリティ機器を扱うためには多くの計算機資源と高度な専門知識を有する技術者が必要である。本研究では時間軸に着目し、実際の運用現場で利用可能なログ分析基盤の構築を行い検証を重ねた。またその提案システムをさらに活用するために技術者の知見蓄積および再利用にも取り組み、セキュリティの現場において成果を得る事が出来た。このような具体的で実践的な取り組みは他組織のセキュリティ現場においても利活用出来る成果であり、一定の社会的意義を持つ研究であると考えられる。

研究成果の概要（英文）：Depending on the type of cyber attack or incident response, search requests for security logs may differ significantly, and search patterns may not be matched in a particular static logical space. Therefore, we concluded that the construction of a distributed environment system based on time axis is appropriate, and we have actually designed / constructed and verified the proposed system on our SOC. In a distributed environment with time constraint, searching by other attributes (for example, IP address) will result in a complete survey and the efficiency will be extremely low. The search imposed was not a problem in many cases, and was shown to be usable in a wide range of cases and also useful in the actual operation environment.

研究分野：情報セキュリティ運用技術

キーワード：サービス構築基盤技術

１．研究開始当初の背景

多層防御の体制を整える事が情報セキュリティにおいて重要である。これは攻撃の高度化および多様化が主な原因であり、例えばアンチウイルスソフトを最新の状態で運用したとしても半数の攻撃は防げないと報告されている。そこでこれらの攻撃を防御するために、より詳細な監視を可能とした様々な次世代型のセキュリティ機器が運用されている。一方、高度多様化するセキュリティ機器を運用するには技術力や人員の面でもコストが高く、SIEM(Security Information and Event Management)と呼ばれる統合運用システムが利用される場合がある。しかし、SIEM は多くの機器をまとめ上げるために負荷の一極集中が避けられず、性能を上げるためには非常に高価な製品を導入せざる得ない状況であった。加えて、過去のセキュリティ事案発生時の調査を行う場合には、ログの検索性能に問題があり実用に耐えないケースも多く見られた。そのため、様々なセキュリティ機器から発生する多様かつ大量のログを収集および分析し、リアルタイムに監視結果をレポート可能である規模拡張性に優れた統合運用システムが求められていた。

２．研究の目的

本研究で想定している機器は一般的な価格帯のサーバである。SIEM と比べて 2 桁程度価格が低い。このような安価なサーバを実際の運用環境(大学内)で複数台利用し、実利用可能であるかを明らかにする事が研究の目的である。具体的にはログの種類や発生頻度に対して提案システムの処理性能を計測する。ここで重要な事はログの種類や発生頻度の増加に対して、提案システムが規模性を持つかどうかである。SIEM のように高いマシン性能をもつ機器に置き換えて処理性能を上げるアプローチでは最上位機種以上の性能は期待出来ない。サーバを追加していくことで規模性を確保し、提案システムが大規模なネットワークを抱える大学等でも利用可能な性能を備えるかどうかを明らかにし、また収集・分析されたログ情報の検索性能も計測し、過去ログも含めて実用的な検索性能を持つかどうかを明らかにすることも研究の目的である。

３．研究の方法

グローバル規模のセンサデータを共有・管理するために、地理位置に基づくオーバレイネットワークやデータ処理機構を持つ Publish/Subscribe システムを開発してきた。本研究では、研究開始当初これらの研究成果を利用し、多様なセキュリティ機器から発生するデータをリアルタイムに収集・分析可能なシステムを構築する事を想定していた。これまでの研究では実空間(2 または３次元空間)という共通の軸があり、この軸を基本とした論理空間の分割処理を基本としてデータ管理の分散化やリアルタイムデータ処理の分散化を実現してきた。しかし今回対象としているセキュリティ機器の情報は決定的な共通の軸が存在せず、複数存在するデータ収集・分析過程がそれぞれ違ったパラメータを利用している状況である。多次元データの高速検索を行うためにはインデックス化が基本であるが、分散化するためには一つ共通の次元を決定する必要があり、分散環境における多次元検索は根本的に困難な問題である。単純な共通の軸が無い事から、これまでの研究成果を単に当てはめるだけで問題が解決するわけでは無い。現状のセキュリティ機器から発生する情報を効率的に管理可能な論理空間を構築する事が本研究の核心部分である。この論理空間の設計は処理能力や規模性にも決定的な影響を与える。そ

のため高い性能を発揮できる論理空間を発見できた場合はセキュリティ機器マネージメントに対して一極集中だけで無く、分散化のアプローチを提示でき、大規模ネットワークにおけるセキュリティレベルの向上等に貢献できると期待できる。様々な検索要求に対して規模拡張性を考慮しながらオーバーレイネットワークの設計および評価を重ねた。しかし、多種多様なログつまり性質の異なる高次元の空間を対象とした論理空間の設計は根本的に困難であり、様々な検索要求に応えられる ID 体系を見出せない可能性もある。研究期間の中盤以降は ID 体系の設計過程で得られた知見なども利用しながら機械学習や自然言語処理技術などを利用してログに加工処理を施すことで目的とする規模拡張性のあるログ分析システムの構築を目指し、特徴ベクトル抽出のためのインシデント対応フローにおけるデータの正規化やセキュリティ機器のログ分析に耐えうる拡張性を備えた仮想化基盤の設計および運用、その環境上でのログ分析基盤の構築や運用などに取り組んだ。加えて、位置情報、論理空間、仮想化技術といった研究過程で得られた知見を利用し、本研究とも関連性の深いセキュアなデータ交換や共有に関する研究に関しても取り組んだ。

４．研究成果

大量かつ多種多様なセキュリティログを効率的に処理し続けるための分散環境を実現するために論理的な ID 空間の設計に取り組み、安定性向上など一定の成果を得た。しかし、一つの論理空間上で多種多様なログを扱いながら様々なクエリーに一定の検索速度を維持して返答することは極めて困難であり、もう一つの研究方法として示したデータの正規化や機械学習手法の応用などに取り組み、システムの構築や評価を行った。また、論理空間の設計やログ分析基盤の構築過程において得られた知見を利用して、本研究とも関連性の深い、セキュアなデータ管理に関してプロトタイプ実装や評価を行い、一定の成果を得た。以下に詳細を記す。

最初に論理空間に関する取り組みを記す。オーバーレイネットワークを構成するノードが安定している場合には、高い分散効率を保ちながらシステム全体が良好に動作する。一方でネットワークの一時的な切断や機器の再起動などノード間のコミュニケーションが安定しない場合は論理空間の再構成などシステム全体に大きな負荷が発生する。システムを安定化させるために確認メッセージを高頻度でやり取りする事でこの問題は改善されるが、システム全体に高い負荷をかけ続ける事になってしまう。そこで軽量かつ効率の高い確認メッセージのプロトコルを提案し、評価を行った。論理空間上での検索時を中心にクエリーが転送される過程で確認を行い、該当ノードの近隣との確認を重視する事で余分なコミュニケーションを抑えつつ、高い安定性を確保する事が可能となっている。オーバーレイネットワークの効率の良い安定化技術を用いることで、本研究の対象としている多種・大量のセキュリティ機器のログ処理を分散環境で安定して実行する事が期待出来る。

次に実践的なログ分析基盤の構築に関する取り組みについて記す。大量かつ多種多様なセキュリティログを効率的に処理し続けるための分散環境を実現するために論理的な ID 空間の研究を進めてきたが、一方でサイバー攻撃の種類 やインシデント対応の種類によって、セキュリティログに対する検索要求は大幅に異なり、特定の静的な論理空間では検索パターンを吸収しきれない。言い換えると特定の検索パターンには非常に高速に返答出来るが、他のパターンでは極端に時間がかかったりまたシステム全体に過剰な負荷を与えかねない状況が発生する。検

討を重ねた結果、時間軸に基づく分散環境の構築が妥当だと結論づけ、その様な設計思想に合うシステムを実際に SOC に設計/導入し、検証を重ねた。時間軸の制約を持った分散環境では、他の属性(例えば IP アドレス)のみによる検索を行うと全件調査となり効率が極端に下がるが、サイバー攻撃の種類によって要求は異なるものの、数時間から数ヶ月の範囲で時間制約を課しながら検索する事は一般的に問題にならず、幅広いケースで利用でき実際の運用の現場でも利用できる事が示された。実際に一般的な IA サーバ群を用いて、大量のセキュリティ機器を扱うために拡張性を考慮した仮想化基盤を構築し、その基板上で多数のアプリケーションをコンテナで実行する方法を確立した。具体的には外部からのアクセスおよび内部での認証の処理を一元化することによって、多くのアプリケーションを低い運用コストで扱う方法を提案した。これにより、大量のセキュリティログを扱う実験基盤が整うと同時に、それらを使ったアプリケーションを実行する検証環境が構築された。また、上記仮想化基盤環境を利用して、時間軸に基づく分散環境の構築といった方針の下に、複数の仮想マシンによるクラスタ環境を構築し、多種多様なログを一元管理し、セキュリティの現場において十分に利用可能な性能を備えることが示された。加えて、インシデント対応時における知見の蓄積に関しても研究を行った。実際のインシデント対応時には被害の最小化が最優先されるために、将来のための情報蓄積は見過ごされがちである。一方で、事案発生時に担当者がいないと判断が先延ばしになったり、過去の同様の事案であっても一から対策を講じたりと、知見の蓄積がされていないために現場のコストが一層高まっているのが一般的な状態である。どのような粒度および箇所で情報を蓄積すると良いか、現場の負担を考慮しながら蓄積すべきデータの正規化を行った。上記の様に分散環境システムそのものの研究およびシステムを効率的に行うためのデータの正規化という 2 つの方向性を持って研究開発を進めた。このデータの正規化はインシデント対応の自動化や機械学習等における特徴ベクトルを意識して設計されており、設計で無く OSS の GitLab をベースに知見蓄積をサポートするシステムを試作し、セキュリティの現場でも運用評価した。インシデント対応フローに関連する知見蓄積はログ分析基盤システムの性能を直接的に向上させるものでは無いが、実際には事案解決に向けた適切なクエリーが発行されそれに対してログ分析基盤が回答する事が望ましい。知見蓄積の枠組みは適切なクエリーを発行をサポートするためのものであり、その観点からは本研究の分散環境におけるログ分析基盤システムの効率的な利用に直結する成果と考えられる。

最後に研究過程で得られた知見を利用した派生の研究成果について述べる。オーバーレイネットワークを構築するには論理的な ID 空間が必要であり、その ID に基づいてシステム全体が分散管理される。これまで地理位置に基づく ID 空間の構築手法に取り組んでおり、その応用として地理情報に基づいたセキュアなファイルシステムを提案した。現在、計算機を使った日常業務は据え置き型の PC だけで無く、ラップトップやスマートフォンなどのモバイルデバイスが多く利用されている。多くのデータが処理、蓄積される中で情報漏洩はインパクトの大きなリスクとして認識されている。一方で個人情報の漏洩などの事件は後を絶たない。提案した地理位置に基づくファイルシステムはデータの機密性と位置情報の組み合わせを考慮し、ユーザのコンテキストに合わせてデータへのアクセス機能を提供している。正当なユーザは追加の認証手順を意識すること無く計算機を利用できる一方で、紛失や窃盗など不正に持ち出されたデバイスに関してはアクセス制限がかかり、情報漏洩を抑止する効果が期待できる。またこれまで大量のセキュリティ機器のログデータを扱うために拡張性を考慮した仮想化基盤を構築してきた。その中で Docker に代表されるコンテナ技術も積極的に採用しており、その知見を生かし安

全なデータバック方法を提案した。現在、ランサムウェアによる被害が公的機関、企業、個人と非常に広い領域を対象に多発しており、ランサムウェアに感染した場合であってもどのようにデータを守るかが重要な課題となっている。提案するバックアップ機構ではアクセスを厳しく制限されたコンテナを通してリモートにバックアップを作成している。この事によりランサムウェア感染時においても本体の計算機からはリモートバックアップに対するアクセス手段が存在せず、バックアップデータを保護できるシステムを開発し評価した。

５．主な発表論文等

〔雑誌論文〕（計　2　件）

[1] Yong JIN, Masahiko TOMOISHI, Satoshi MATSUURA, Yoshiaki KITAGUCHI.
A Secure In-Depth File System Concealed by GPS-Based Mounting Authentication for Mobile Devices. IEICE TRANSACTIONS on Information and Systems, Vol.E101-D, No.11, pp.2612-2621, 2018.査読あり

[2] Kimihiro Mizutani, Takeru Inoue, Toru Mano, Osamu Akashi, Satoshi Matsuura, and Kazutoshi Fujikawa, "Living Will for Resilient Structured Overlay Networks," IEICE Trans. Commun., vol.E99-B, pp.830-840, April 2016.査読あり

〔学会発表〕（計　8　件）

[1] 森 健人, 石井 将大, 松浦 知史, 金 勇, 北口 善明, 友石 正彦
セキュリティ事案における知見の蓄積・活用を可能とする対応フローの提案と実装
情報処理学会 研究報告インターネットと運用技術（IOT）, Vol. 2019-IOT-46, Jun. 2019.

[2] 石井 将大, 森 健人, 松浦 知史, 金 勇, 北口 善明, 友石 正彦
東工大 CERT におけるインシデント対応の分析とその自動化に関する考察
情報処理学会 研究報告インターネットと運用技術（IOT）, Vol. 2018-IOT-43, No. 2, pp. 1-8, Sep. 2018.

[3] Yong Jin, Masahiko Tomoishi, Satoshi Matsuura, Yoshiaki Kitaguchi, "A Secure Container-based Backup Mechanism to Survive Destructive Ransomware Attacks", IEEE International Conference on Computing, Networking and Communications (ICNC2018), Maui, Hawaii, USA, March 5-8, 2018.

[4] Yong Jin, Masahiko Tomoishi, Satoshi Matsuura, "An In-Depth Concealed File System with GPS Authentication Adaptable for Multiple Locations", 2017 IEEE 41th Annual Computer Software and Applications Conference (COMPSAC), Torino, Italy, July 4-8, 2017.

[5] 森 健人, 松浦 知史, 金 勇, 友石 正彦. "オンプレミスで実現する業務効率化のための OSS 基盤環境構築", 研究報告インターネットと運用技術（IOT）, Vol. 2016-IOT-35, No. 10, pp. 1-8, Sep. 2016.

[6] Y. Jin, M. Tomoishi and S. Matsuura, “Enhancement of VPN Authentication Using GPS Information with Geo-Privacy Protection,” 2016 25th International Conference on Computer Communication and Networks (ICCCN), Waikoloa, HI, August 2016.

[7] Y. Jin, M. Tomoishi and S. Matsuura, “Design of a Concealed File System Adapted for Mobile Devices Based on GPS Information,” 2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC), Atlanta, GA, June 2016.

[8] 松浦 知史, 森 健人, 金 勇, 友石 正彦. "拡張性を考慮した小規模仮想化基盤の構築", 研究報告インターネットと運用技術（IOT）, Vol. 2016-IOT-32, No.27, pp.1-8, 3rd-4th Mar. 2016.

６．研究組織

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。