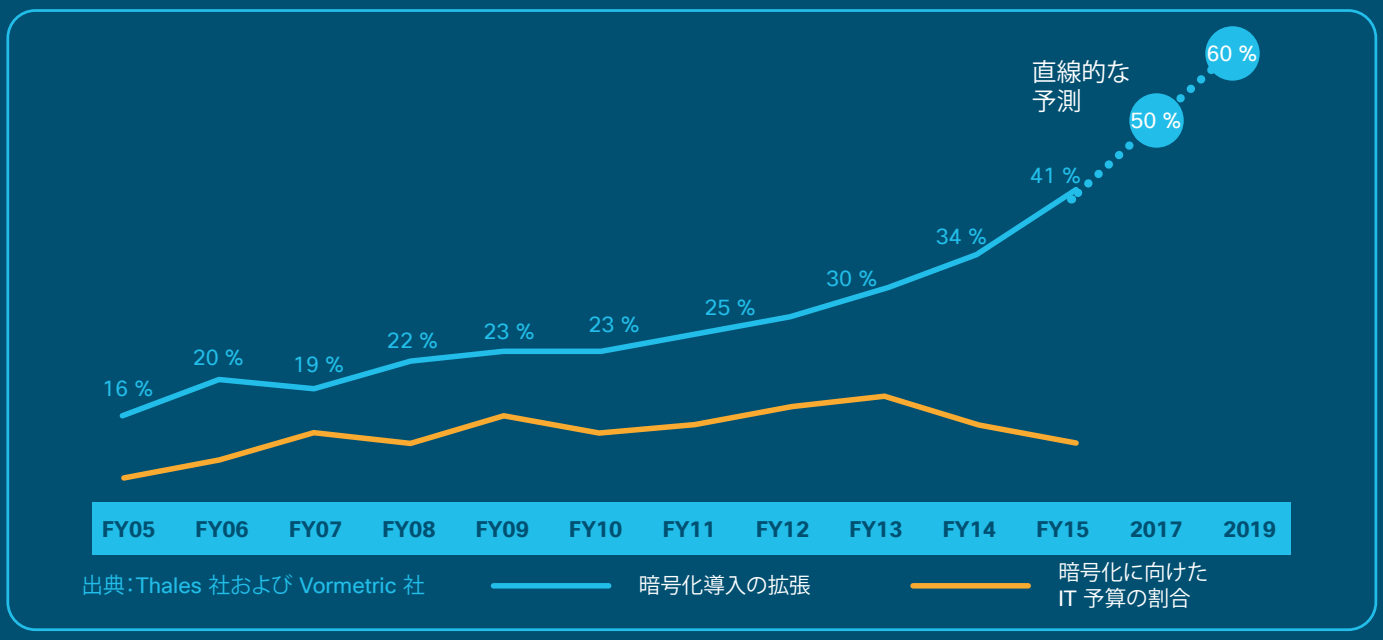
暗号化トラフィック分析

はじめに

暗号化トラフィックの急激な増加により、脅威の状況が変化しています。デジタル化を進める企業が増えるにつれ、情報保護の主要な手段として暗号化が多くのサービスとアプリケーションで使用されています。具体例として、暗号化トラフィックは対前年比で 90 % 以上の増加を続けており、2015 年に 21 % だった Web サイトの暗号化トラフィックは 2016 年には 40 % 以上を占めています。Gartner は、2019 年には 80 % の Web トラフィックが暗号化されると予測しています。

ビジネスにおいてオンラインでの通信と処理にインターネットを利用する企業では、暗号化技術がプライバシーとセキュリティの実現に貢献してきました。モバイル、クラウド、Web アプリケーションは、キーと証明書を使って適切に実装された暗号化メカニズムによりセキュリティと信頼性を確保しています。しかし、暗号化の恩恵を受けるのは企業だけではありません。攻撃者はこの恩恵を、検出の回避と悪意のあるアクティビティの保護に活用したのです。
図 1 に、このような攻撃の経済的影響を示します。

暗号化で変わる脅威の状況



目次

ネットワーク全体の可視化はますます難しくなり、従来の検出方法ではデータを検査対象にできないことが想定されます。デジタル ビジネスが暗号化によってどれほど保護されているか、または保護されていないかを同時に評価できるようにし、悪意のあるトラフィックと無害なトラフィックを分別するための評価も行う必要があります。

Gartner は、2019 年に発生するマルウェアの活動の半数が何らかのタイプの暗号化を使ってマルウェアの配信、コマンド アンド コントロール、またはデータ漏えいを隠すと考えています。

図 1.　悪意のある攻撃による経済的影響

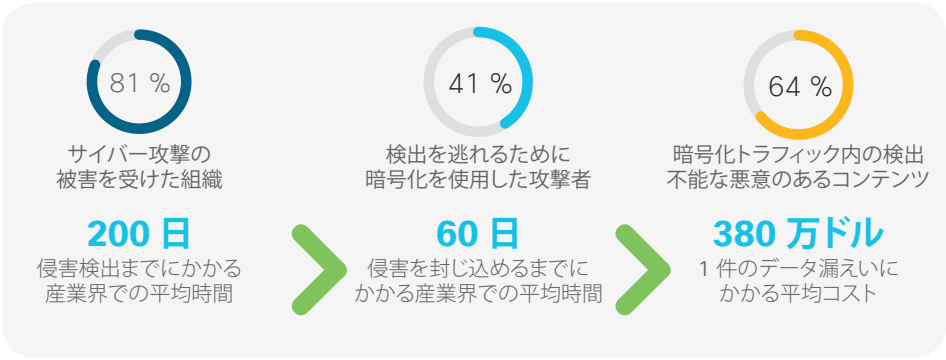


表 1 に、暗号化トラフィックの特性に基づく新しい脅威ベクトルを示します。

表 1.　暗号化トラフィックの特性に基づく新しい脅威ベクトル

検査されない暗号化トラフィック	脅威
HTTPS を介した従業員の Web 閲覧	• マルウェア感染 • コマンド アンド コントロール サーバとのコバート チャネル • データ漏えい
ネットワーク エッジ（DMZ）サーバに安全に接続する内部ネットワークの従業員	感染したホストから横への被害拡大
暗号化されたプロトコルを使用して企業の公開サーバに接続するインターネット ユーザ	受信トラフィックを検査する技術が 1 つしかないことによる、多層防御機能の低下

暗号化トラフィックのセキュリティに関する課題

今ある組織の大多数が、暗号化トラフィック内の悪意あるコンテンツを検出するソリューションを持っていません。また、ネットワーク インフラストラクチャ全体に展開できるソリューションをネットワーク速度の低下なしに実装するツールやリソースが不足しています。

脅威の検査時に復号、分析、再暗号化を一括して行う従来の方法は、パフォーマンスやリソースを考えると、常に実践または実現できるとは限りません。ただし多くの場合、悪意のあるフローを特定する高度な分析技術を使って、復号技術による詳細な調査を実施できます。

どのようなときも、デジタル化されたビジネスがどれほど暗号化されているのか、またはされないままなのかを明確にはできません。トラフィックが暗号化されている場合、特定のセキュリティ ポリシーへの対応を求めるコンプライアンス要件を満たすように暗号化されるのが一般的です。

暗号化トラフィック分析の概要

従来のフロー モニタリングでは、フローのアドレス、ポート、バイト数とパケット数のレポートによってネットワーク通信の概要を表示できます。それに加え、内部フロー メタデータまたはフローの内部で発生したイベントにまつわる情報は、フロー監視フレームワーク内で収集や保存、分析できます。パケット詳細調査が今後実行可能ではないため、このデータはトラフィックが暗号化されているときには特に重要となります。暗号化トラフィック分析と呼ばれるこの内部フローのメタデータは、フロー内のメッセージの長さや到着時間など、プロトコル詳細とは独立した新しいタイプのデータ要素またはテレメトリを使って導き出されます。これらのデータ要素には、暗号化および非暗号化フローの両方に等しく利用できるという便利な性質があります。

暗号化トラフィック分析は、これらのデータ要素または内部フロー テレメトリを使って暗号化トラフィック内のマルウェア通信を特定する手段であり、一括復号を行わなくても暗号化フローの健全性を維持できます（図 2）。表 2 に、暗号化トラフィック分析の利点を示します。

図 2.　暗号化トラフィック分析：テクニカル ソリューションの概要

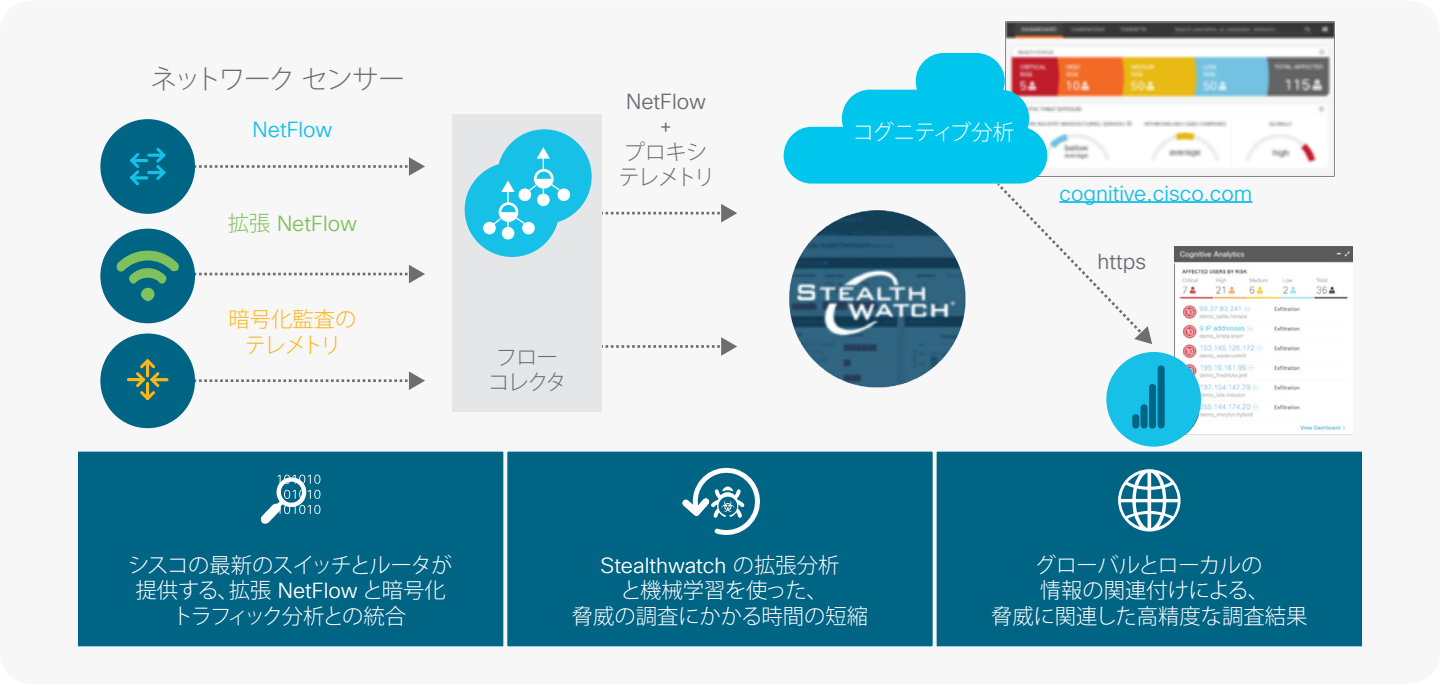


表 2. 暗号化トラフィック分析の利点

利点
• セキュリティの可視性：ネットワークの分析により、暗号化トラフィック内の脅威に関する情報を把握できます。リアルタイム分析により、ユーザとデバイスの情報に相互に関連するコンテキスト上の脅威インテリジェンスが得られます。 • 暗号化アセスメント：企業が暗号化プルトコルに準拠していること、およびネットワーク内の暗号化データと非暗号化データに関する可視性と知識を確保できます。 • 応答時間の短縮：感染したデバイスとユーザをすばやく抑制できます。 • 時間とコストの削減：セキュリティ ポスチャの基盤としてネットワークを使用し、ネットワーク セキュリティに対する投資を十分に活用します。

暗号化トラフィック分析：暗号化トラフィックの新しいデータ要素

暗号化トラフィック分析は、パッシブ モニタリングや関連データ要素の抽出、クラウドベースのグローバルな可視性を備える管理された機械学習を通じて、暗号化トラフィック内に潜むマルウェア通信を特定することに重点を置いています。

Transport Layer Security（TLS）はアプリケーションにプライバシーを提供する暗号化プロトコルです。TLS は通常、Web 閲覧で使用する HTTP や電子メールで使用する Simple Mail Transfer Protocol（SMTP）のような共通プロトコルの最上位に実装されます。HTTPS は HTTP 通信に TLS を使う方法です。これは、Web サーバとクライアントとの間の通信をセキュアにする最も一般的な方法で、ほとんどの主要な Web サーバでサポートされています。

暗号化トラフィック分析は、4 つの主要なデータ要素を抽出します：パケット長とパケット時間のシーケンス、バイト分布、TLS 固有の特長、初期データ パケットです。シスコ独自の特定用途集積回路（ASIC）アーキテクチャは、データ ネットワークの速度を低下させずにこれらのデータ要素を抽出することができます。

- パケット長とパケット時間のシーケンス（SPLT）：SPLT はパケット間の着信時間の間隔に加えて、各パケットのアプリケーション ペイロードの長さ（バイト数）を伝えます。SPLT は、一連のパケット サイズ（単位：バイト）を、1 つ前のパケットがモニタされてからの一連の時間（単位：ミリセカンド）とともに表すことができます。
- バイト分布：バイト分布は、フロー内のパケットのペイロードに特定のバイト値が現れる可能性を表します。フロー内のバイト分布は一連のカウンターを使用して計算できます。バイト分布に関連付けられた主なデータ タイプは、フル バイト分布、バイトの欠如、バイトの平均または標準偏差です。たとえば、バイト値ごとに 1 つのカウンターを使って HTTP GET リクエストである「HTTP/1.1.」を計算できます。その場合、「H」に対するカウンターを 1 つ増やし、続く 2 つの「T」に対して別のカウンターを 2 つ増やし、同様に処理を続けます。バイト分布は、一連のカウンターによって維持されますが、合計バイト数で正常化することにより、簡単に適切な分布に変えることができます。
- 初期データ パケット（IDP）：IDP は、フローの最初のパケットからパケット データを取得するために使用します。IDP により、HTTP、URL、DNS ホスト名または IP アドレスなどのデータ要素といった分析に役立つデータを抽出できます。TLS ハンドシェイクは、分析に利用できる非暗号化メタデータを持つ複数のメッセージで構成されます。メタデータは、暗号化スイート、TLS のバージョン、クライアントの公開鍵の鍵長などの抽出に使用します。

付録 A に、新しいデータ要素の詳細を示します。

暗号化トラフィック分析を利用したセンサーとしての拡張ネットワーク：コンポーネント

拡張 NetFlow

NetFlow アーキテクチャでは、レコード形式のデータが、エクスポータからコレクタに送信されます。データ セット内の各レコードは同じフォーマットをもち、フォーマットはテンプレートを使って指定します。データ レコードは、NetFlow の一連の情報要素または「フィールド」で構成され、特定の ID 値が各フィールドに割り当てられています。情報要素の ID 値は、Internet Assigned Numbers Authority（IANA）がグローバルに定義およびアーカイブ、または企業固有のもので個々の組織が定義する場合があります。

NetFlow のテンプレートは、IANA が管理する、グローバルに定義された複数の要素を使用します。グローバル要素の一部、たとえば IP アドレスやレイヤ 4 のポート番号などは、よく知られた 5 タプルとなり、一意のフロー識別子（フロー キー）として使用されます。追加の要素を使って基本的なパケットまたはオクテット統計、およびタイムスタンプをレポートできます。

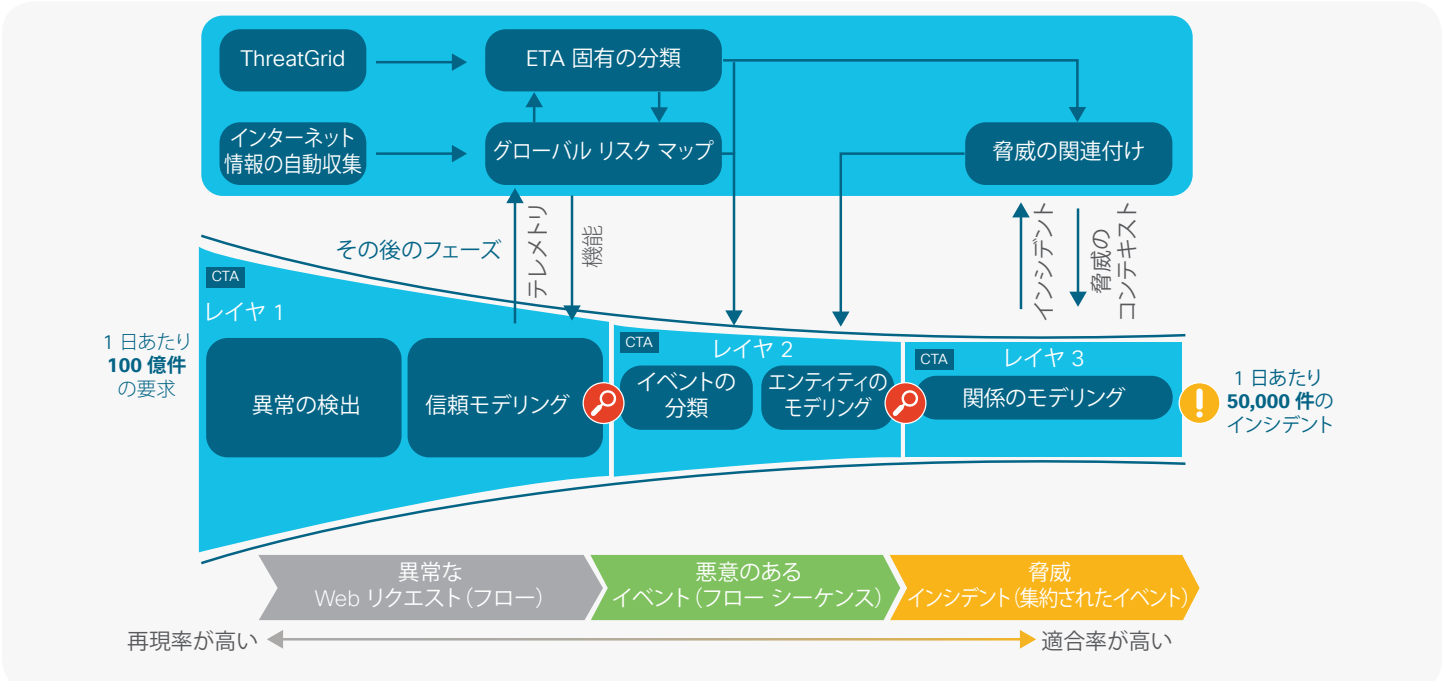
これらのグローバル定義済み要素は、前述と付録 A で示したベンダー固有の（シスコ ベンダー ID）データ要素を使用して拡張します。ベンダー固有のデータ要素を Cisco Stealthwatch® で分析することにより、暗号化トラフィック内の脅威と脆弱性に関するインサイトが得られます。

コグニティブ分析を統合した Stealthwatch

Cisco Stealthwatch は、NetFlow、プロキシ サーバ、エンドポイント テレメトリ、ポリシーとアクセス エンジン、トラフィック セグメンテーションなどを使用し、企業全体にわたるホストとユーザの「正常」なふるまいを示すベースラインを確立します。コグニティブ分析と統合することにより、クラウドベースの分析エンジンである Stealthwatch はトラフィックとグローバルな脅威との相互関係を示すことができ、感染したホスト、コマンド アンド コントロールの通信、および疑わしいトラフィックを自動的に特定します。

コグニティブ分析は、グローバル リスク マップ（インターネット上のサーバに関する、ふるまいの広範なプロファイル）を保持し、攻撃と関係があるか悪用されている可能性のある、または将来攻撃の一部になりうるサーバを特定します（図 3）。これはブラックリストではなく、セキュリティに関する予測を基にした全体像です。コグニティブ分析は機械学習と統計モデルを活用し、拡張 NetFlow から得られる暗号化トラフィックの新しいデータ要素を分析します。コグニティブ分析エンジンにより、グローバル リスク マップと、暗号化トラフィック分析のデータ要素はお互いの機能を強化します。コグニティブ分析を利用する Stealthwatch はトラフィックの復号ではなく機械学習アルゴリズムを使って悪意のあるパターンを暗号化トラフィックから検出し、脅威の特定とインシデント対応の改善を支援します。

図 3.　コグニティブ分析エンジン



Stealthwatch Management Console のセキュリティ インサイト ダッシュボードを使うと、リスク タイプ別のコグニティブ分析により、影響を受けたユーザを特定し、表示できます。コグニティブ分析のダッシュボードを展開すると、上位リスクへの深刻化と、脅威にさらされる相対的な度合いについての詳細情報を表示できます。表 3 に、暗号化されたコマンド アンド コントロールの通信を使う、危険度の高い脅威の例を示します。

図 4. コグニティブ分析を統合した Stealthwatch のセキュリティ インサイト ダッシュボード

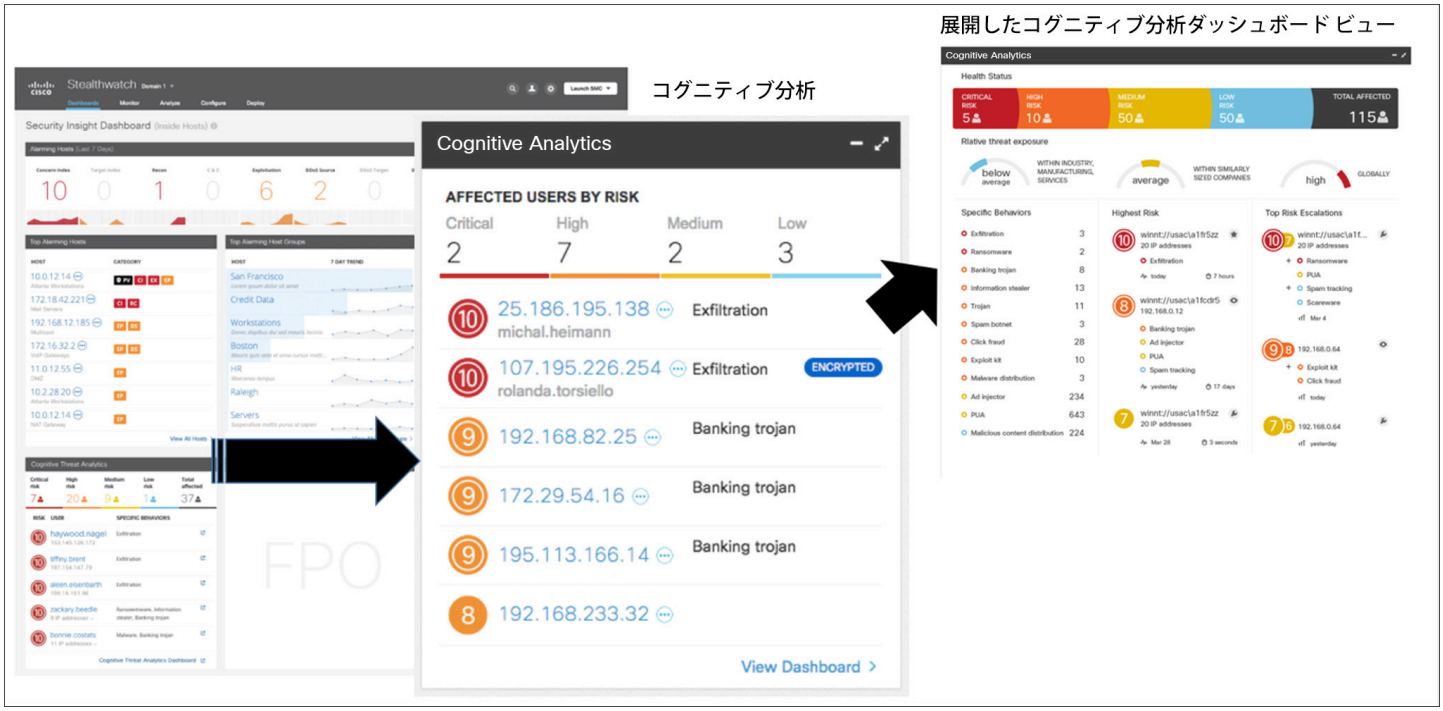


表 3. 暗号化されたコマンド アンド コントロールの通信を使う、危険度の高い脅威の例

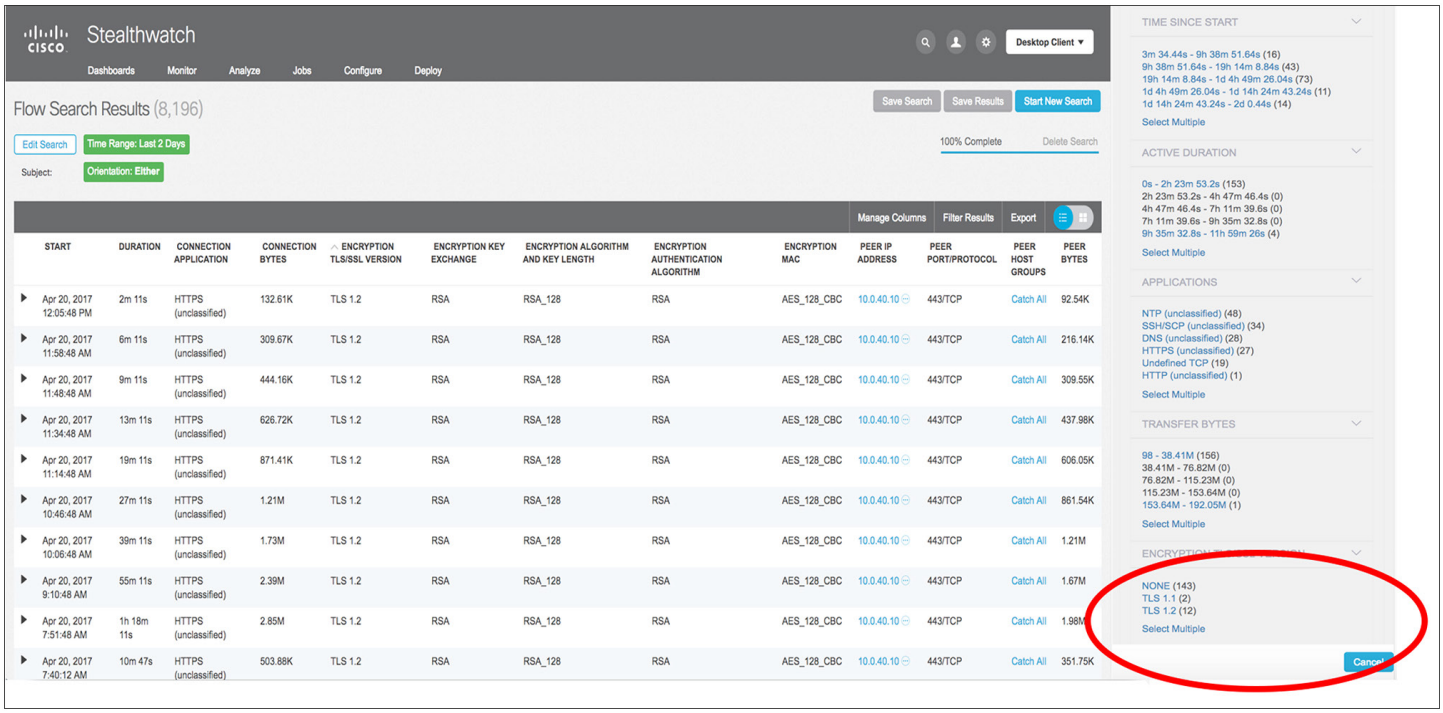
名前	タイプ
Gamarue および Andromeda	モジュール型ボットネット
Sality	ファイル感染型、モジュール型ボットネット
Necurs	情報窃盗、バックドア型、ボットネット
Rerdom	クリック詐欺、ボットネット

暗号化された悪意のあるフローが検出されると、Stealthwatch でブロックまたは隔離できます。Cisco TrustSec® と Software-Defined Access（SD-Access）を統合した Cisco Identity Services Engine（ISE）を使用し、pxGrid を経由するポリシー駆動の修復動作により、シンプルで迅速なネットワーク セキュリティ運用が可能になります。

暗号化のアセスメント

暗号化トラフィック分析では、企業が暗号化プロトコルに準拠できるように、あらゆるネットワーク通信の暗号化の品質をすばやく特定して可視性を提供します。これによってネットワーク上で暗号化されているデータや暗号化されていないデータの情報を提供でき、デジタル ビジネスが確実に保護されていることを主張できます。この暗号化のアセスメントは Stealthwatch で表示され、API 経由でサードパーティ製のツールにエクスポートすることで暗号化コンプライアンスの監視と監査を実施することもできます（図 5）。

図 5.　暗号化のアセスメント



機能のサポート

シスコの最新のネットワーキング機器は、Cisco IOS XE 16.6 以降で、暗号化トラフィック分析機能を使用する拡張 NetFlow をサポートします。

1. 暗号化トラフィック分析機能を統合した拡張 NetFlow をサポートする、互換性のあるシスコ機器：
   - スイッチ：Cisco Catalyst® 9300 シリーズ（Cisco IOS XE リリース 16.6 以降）、および 9400 シリーズと 9500 シリーズ（Cisco IOS XE リリース 16.8.1 以降）
   - ルータ：ASR 1001-X、ASR 1002-X、ASR 1001-HX、ASR 1002-HX、ASR 1004、ASR 1006-X、ASR 1009-X、4221 ISR、4321 ISR、4331 ISR、4351 ISR、4431 ISR、4451-X ISR、サービス統合型仮想ルータ（ISRv）（5400 エンタープライズ ネットワーク コンピューティング システム、クラウド サービス ルータ（CSR）1000V（Cisco IOS XE リリース 16.7 以降）を含む）
2. Stealthwatch（リリース 6.9.2）では、機械学習と統計モデリング機能を追加できます。これにより、暗号化トラフィック分析と統合した拡張 NetFlow のデータを分析することができます。
3. ルータの Stealthwatch 学習ネットワーク ライセンス（v2.0）では、暗号化トラフィックのふるまいに関するプロファイルを作成する機能が利用でき、暗号化トラフィックで検出した異常にフラグを付けることができます。

効果およびシスコの調査結果

実データに基づく実験で、99 % 以上の精度とともに 0.01 % の誤検出（10,000 TLS 接続ごとに 1 件のみの誤検出）を達成できています。これは、シスコの調査結果に記載しているように、実際の HTTPS セッションの膨大なサンプルに基づいています。

まとめ

要約すると、いまやネットワークは、暗号化トラフィック内の脅威を検出できる、高性能のセキュリティ センサーです。デジタル ネットワーク アーキテクチャ対応のインフラストラクチャによって、ネットワークは、セキュリティに対して今後現れる高度な脅威を検出して封じ込める、エンドツーエンドのセンサーとエンフォーサに変わります。

付録 A

暗号化トラフィック分析から抽出できるデータ要素

データ要素名	説明
パケット長とパケット時のシーケンス（SPLT）	LENGTH 値を表す数値列で、その後に INTERARRIVAL TIME 値が続きます。LENGTH の数値列は、アプリケーション ペイロードを含むフローの最初の N 個のパケットを表します。各 LENGTH は 16 ビットの整数としてエンコードされ、20 バイトの数値列になります。この直後に、各 INTERARRIVAL TIME が続き、16 ビットの整数としてエンコードされ、別の 20 バイトの数値列になります。
バイト分布	フローに含まれるアプリケーション ペイロードの最初の N バイトの、各バイト値（または値の範囲）の発生頻度を表すヒストグラムです。各「発生頻度」は、16 ビットの整数で表されます。
初期データ パケット（IDP）	実際のペイロード データを含むフローの最初のパケットの内容で、IP ヘッダーの先頭から開始されます。
TLS レコード	LENGTH の数値列を表します。この値の後に INTERARRIVAL TIME 値、CONTENT TYPE 値、HANDSHAKE TYPE 値が、それぞれ続きます。この配列には TLS フローの最初の N 件のレコードが表されます。
TLS レコード長	TLS フローで、最初の N 件のレコード全体にわたる長さを表すシーケンスです。
TLS レコード時間	TLS フローで、最初の N 件のレコード全体にわたる到着時間間隔を表すシーケンスです。
TLS コンテンツ タイプ	TLS フローで、最初の N 件のレコード全体にわたる ContentType 値のシーケンスです。
TLS ハンドシェイク タイプ	TLS フローで、最初の N 件のレコード全体にわたる HandshakeType 値のシーケンスです。
TLS 暗号スイート	TLS フローで、クライアントから提供される、またはサーバが選択する N 件の暗号スイートのリストです。

データ要素名	説明
TLS 拡張	LENGTH 値の配列の後に EXTENSION TYPE 値の配列が続きます。この配列は、TLS フローの Hello メッセージ内にある TLS 拡張を表します。
TLS 拡張の長さ	TLS フローの Hello メッセージ内にある最初の N 件の TLS 拡張全体にわたる、拡張の長さのリストです。
TLS 拡張のタイプ	TLS フローの Hello メッセージ内にある最初の N 件の TLS 拡張全体にわたる、拡張のタイプのリストです。
TLS バージョン	TLS フローの Hello メッセージ内にある TLS バージョン番号。
TLS キー長	TLS ClientKeyExchange メッセージ内にあるクライアント キーの長さ。
TLS セッション ID	フローに TLS Hello メッセージがある場合に使用されるセッション ID 値。
TLS ランダム	フロー内の TLS Hello メッセージにあるランダム値。

参考資料

1. Gartner：セキュリティ リーダーは増加する SSL トラフィックが生む脅威に立ち向かうべき [英語]
2. Ponemon Institute：暗号化トラフィックに潜む脅威の解明、2016年 [英語]
3. NSS Labs：TLS/SSL の現状 [英語]暗号化 Web パート 1：増加の傾向 [英語]
4. コンテキスト フロー データによる、暗号化されたマルウェア トラフィックの特定、Blake Anderson と David McGrew による共著、AISEC 2016 年 [英語]