挙動に基づく分散型サイバー攻撃のオンライン検知システムとその性能評価
A Behavior-based Detection Engine for Distributed Cyber Attacks and its Performance Evaluation

フォン　ヤオカイ *　　堀　良彰 **　　櫻井　幸一 *
(*九州大学大学院システム情報科学研究院 **佐賀大学全学教育機構)

1　はじめに

分散型サイバー攻撃とは多くの攻撃者が協力しながら被害者を攻撃するものである。分散型攻撃の典型的な例は、サービス拒否（DDoS 攻撃）攻撃である。関連の検知技術に関する研究は、サイバーセキュリティコミュニティで最も重要なトピックの一つとなっている。多くの検出方法が提案されているが、様々な問題が残ってある [1, 2]。近年、挙動に基づく方法は、多くの研究者や開発者から大きな注目を集めている。挙動に基づくアプローチでは、歴史のトラフィック・データから通常の動作モードを抽出して、それを異常検出に利用する。本稿では、分散型サイバー攻撃を検出するための挙動に基づく検知エンジンを実装する方法について説明した上で、実装した検知システムの検知結果も報告する。

2　挙動に基づく方法の長所

挙動ベースのサイバー攻撃検知は、監視ネットワークの歴史トラフィックから通常モードを抽出し、それを用いる検出方法である。次の長所がある。

- トラフィックから異常を区別するためのしきい値を自動的に歴史的なトラフィックから抽出される。閾値を事前に決定する必要がない。
- 抽出された通常モードは、特定の監視対象ネットワークの特徴を反映することができる。異なる組織でのネットワークトラフィックが互いに大きく異なっているので、これは重要である。
- 新種攻撃および既知攻撃の変種も対応できる。
- 検出は高速である。検出プロセスがリアルタイムトラフィックの統計量と事前に抽出した通常モードの単純な比較である。

3　挙動に基づく分散型攻撃の検知エンジンの実装

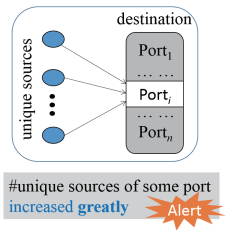

図 1: All the destination ports are monitored one by one.

我々の検出エンジンでは、すべての宛先ポートは個別に監視されている。あるポートにアクセスしたユニークなソースホストの数が突然大幅に増えると警告アラートを出す。本検知エンジンでは、通常モードが監視対象のネットワークの前月 1 日から 28 日までのトラフィックデータから抽出される。それを利用してオンラインモードで異常検知を行う。ユーザーは事前に定義された閾値を与える必要はない。また、通常モードは不変でないので、新しい状況を反映するために通常モードは自動的に更新される。例えば、私たちは 2016 年 1 月のトラフィックを検出するために始めたときに、2015 年 12 月 1 日から 12 月 28 日まで 4 週間のトラフィックデータを利用して通常モードを抽出する。その抽出された通常モードは 2016 年 1 月の終わりまで使用される。すなわち、新しい月の開始時に学習アルゴリズムが呼び出され、通常モードが更新される。図 2 で示される。

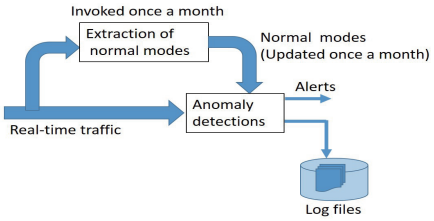

図 2: General idea.

本検知エンジンには次の 4 つのプロセスがある。1) データの収集；　2) 度数分布の作成；　3) 通常モードの抽出；3) 異常検知。通常モードの抽出アルゴリズムは検知性能を大きく影響する。論文 [3, 4] で関連アルゴリズムを詳しく議論している。

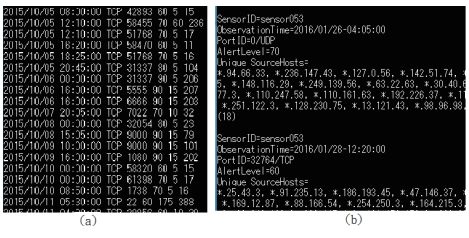

図 3: An example of detection result.

参考文献

[1] Varun Chandola, Arindam Banerjee and Vipin Kumar, Anomaly Detection: A Suryvey, *ACM Computer Survey*, Vol. 41, No.3, pp. 1-72, 2009.

[2] Feng Y., Hori Y., Sakurai K. and Takeuchi J.: A Behavior-Based Method for Detecting Distributed Scan Attacks in Darknets, Journal of Information Processing (JIP), Vol.21, No.3, pp. 527-538, 2013.07.

[3] 王サン, フォン　ヤオカイ, 川本 淳平, 堀　彰良, 櫻井 幸一, 挙動に基づくポートスキャン検知の自動化に向けた学習アルゴリズムの提案とその性能評価, 情報処理学会論文誌, 56, 9, 1770, 1781, 2015.09

[4] Feng Y., J., Hori Y., Sakurai K., A Proposal for Detecting Distributed Cyber-Attacks Using Automatic Thresholding, *Proc. 10th Asia Conference on information security (AsiaJCIS2015)*, 2015.