

TCP ヘッダの特徴による攻撃分類†

長田和樹†* 沖野浩二††

†* 富山大学大学院理工学教育部 †† 富山大学総合情報基盤センター

1 はじめに

インターネットを経由したサイバー攻撃による被害が多発している。サイバー攻撃に対する対策はいろいろとあるが、その代表的なものとしては、FW や AntiVirus ソフトウェアが挙げられる。インターネットの通信は IP と呼ばれるプロトコルで通信が行われる。IP における通信では、通信パケットは、通信の制御情報を記すヘッダ部分と実際の内容とするデータ部分に分かれる。従来の攻撃検出方法は、これらの通信に対してパターンマッチなどを行うことにより攻撃を検出していた。だが、現在の通信では、盗聴対策などのセキュリティ向上を目的として、SSL などの暗号化が施されている。このため、従来の方法では検出が難しくなっている。そこで、本研究では、宛先に通信を届けるため必要である、暗号化されていないヘッダ部分を利用し、実際の攻撃において、ヘッダ部の情報が通常の通信と異なる点に着目し、これらの特徴量による攻撃の分類ができないかを検証する。

2 TCP 通信について

2.1 3-wayhandshake について

IP 通信は、大きく分けてコネクションを保証する TCP 通信と保証しない UDP 通信がある。TCP 通信においてはコネクションを保証するため実際のデータのやり取りの前にコネクションの確立を行う。

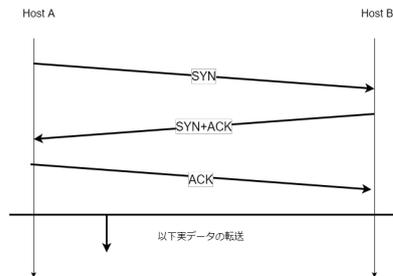


図1 基本的な 3-wayhandshake

まず HostA は通信したい HostB に SYN パケットを送信する。すると、その SYN パケットを受け取った HostB は SYN+ACK パケットを返信する。そしてその SYN+ACK パケットを受信した HostA はその SYN+ACK を受信した旨を HostB に通知してから通信が確立できたとして、実データの転送を始める。

攻撃者は攻撃対象の稼働状況の確認や OS の特定のために 3-wayhandshake プロトコルに細工を施してポートスキャンと呼ばれる調査を行っていることが知られている。代表的な

ポートスキャンである fullscan(tcpscan) と呼ばれるものと halfscan(stalthscan) と呼ばれる手法を今回取り上げる。

2.2 halfscan

halfscan は 3-wayhandshake を成立させず、攻撃者は調査対象に SYN パケットのみを送信して、調査対象から返信されるパケットがある場合でも何も応答しない、という調査手法である。このとき、調査対象の標的ポートが稼働状態であれば SYN+ACK パケットが返信される。非稼働状態であれば、何も返信がないまたは RST+ACK パケットが返される。この非稼働ポートに対する返答方法は OS の IP スタックの実装や設定に依存する。

2.3 fullscan

fullscan は halfscan とは異なり、クライアント側から 3way 目の通信が帰ってくるものを指す。fullscan は通常のコネクション確立を行うものや、ACK に代わり FIN や RESET などを利用し、通信の終了を通知するものがある。

3 提案手法

提案手法は、取得された通信データのうち、コネクションを保証する TCP 通信のヘッダ部情報を利用して、攻撃分類を行う。TCP は利用する理由は、UDP と比較して、接続の偽装が困難であり、取得できる情報が多いからである。攻撃の種類を halfscan,fullscan, 実攻撃 (attack) の 3 種類とし、以下の方法で分類する。

まずパケット群の内 TCP ヘッダの Control flag が SYN のみ立っているものを抽出する。これを syn として分類の基準とする。

次に、この syn と送信元 IP, 送信先 IP, 送信元ポート番号, 送信先ポート番号が同一であることに加え Control flag が SYN1 つのみではないパケット群を same_stream_candidate とする。次に same_stream_candidate の内、syn から 1 分以内に送信されてきているパケットを related_syn と定義する。

この related_syn 内のパケット数やオプション値によって以下のように分類する。

- halfscan
 - related_syn に属するパケット数が 0 である場合
- fullscan
 - related_syn の中にパケットが 1 つのみ存在
 - その 1 つのパケットのシーケンス番号が syn のシーケンス番号に 1 を足したもの
- attack
 - related_syn の中に 2 つ以上パケットが存在する

加えて、調査や攻撃に使われているユニーク IP アドレス数が調査・攻撃数と共にどのように変異するかも合わせて実験する。

† Attack Classification by TCP Header Features

†* Kazuki Nagata, Graduate School, University of Toyama

†† Koji Okino, Information Technology Center, University of Toyama.

これらの条件で得られたデータを AS ごとに集計し、調査・攻撃の時系列的な傾向を観察する。

4 実験データ取得環境

本実験の取得環境を表 1 にまとめる。

表 1 検証環境

データの期間	2020.02/01 - 02/29
分類値の合計幅	1 日
AS 情報	GeoLite 2019/12/27
honeypot	Ubuntu 18.04 LTS (kernel ver 4.15.0-128-generic)
Pcapfile の総容量	約 30GB
変換後 csv の総容量	約 45GB

尚、本実験で用いたハニーポットは FW の設定で特定ポートを解放してあるだけでその他は通常の web サーバと同じである。

5 実験結果

とある AS に属する IP アドレス群からの攻撃傾向を可視化する。

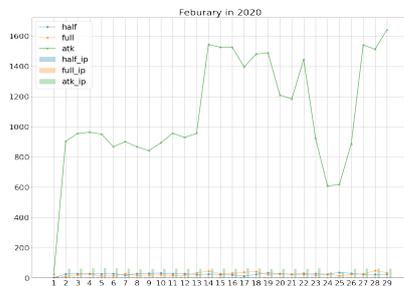


図 2 AS:A の調査・攻撃パターン

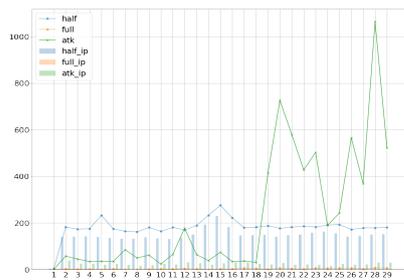


図 3 AS:B の調査・攻撃パターン

6 考察

まず、図 2, 3 に関して基本的に halfscan 数や fullscan 数とその scan に用いられたユニーク IP 数は平行している。図 2 においては halfscan と攻撃が同じように推移しており、図 3 では絶えず毎日同量の halfscan が観測できる。

図 3 の 5 日と 13 - 17 日の halfscan 数は他と比べて特徴的である。この両者では用いられているユニーク IP アドレスの偏移に差があることに着目して、図 3 を参照すると先ほどの 2 件の特徴的な halfscan の後に攻撃数が増加している傾向がある。よってこの 2 つの特徴的な halfscan は攻撃するための予備調査であり、その他の halfscan は常設的な情報収集のためのものではないかと思われる。

5 日の halfscan の後、7-10 日と 11-14 日に 2 回攻撃数の増加が確認できる。しかし攻撃に用いられたユニーク ip 数があまり変化しないことから、常設されていた攻撃サーバーから 5 日の調査を経て 2 度攻撃してきたのではと考えられる。

また 13 - 17 日の halfscan 数は ip と共に増減している。この特徴的な halfscan の後、攻撃数が急激に上昇していることからこの halfscan を皮切りに攻撃の手法が変化すると推察できる。またこの攻撃手法の変化の直前のユニーク ip 数の増減は、この後攻撃の手法の大幅な変化と合わせると FW の IP フィルタに検知されない範囲を検索し、その後その調査に用いた IP 群の中から新しく構築する攻撃ネットワークの IP アドレスの最適化を行うために実行された scan だと考えられる。

7 まとめ

本実験では既に収集されたパケット群から tcp プロトコルに沿って攻撃の種類を判別する方法を提案した。この手法では攻撃変化の予兆や、常設されたサーバによる攻撃と局所的に発生した攻撃などを分類する手がかりの 1 つとなるのではないかとと思われる。また本実験においては tcp ヘッダの Control flag やシーケンス番号などを参考にしたため、その他のオプション値が詳細な分類に有効であるかは追実験が必要である。

参考文献

- [1] RFC793 TRANSMISSION CONTROL PROTOCOL
- [2] 有本純, 曾根直人, 森井昌克, "ダークネット観測によるスキャンパケットの傾向と分析" Computer Security Symposium (2019, 10)
- [3] 牧田大佑, 島村集平, 久保正樹, 井上大介, "全ポート待受型の簡易 TLS ハニーポットにより観測されたサイバー攻撃の分類, 電子情報通信学会, (2020,03)"