研究指導 中澤 真 教授

Slow HTTP DoS 攻撃に対する攻撃検知のための一考察

宮﨑 真琴

1. はじめに

急速なインターネット普及に伴い，サイバー犯罪が増加しており，中でも不正アクセス等のインターネット上に対する攻撃が最も大きく増加している．2021年度のサイバー犯罪の検挙件数の調査によると，2019年から2021年にかけて9,875件が12,209件にまで増加している[1].

インターネット上に対する攻撃の代表例としてDoS攻撃がある．DoS攻撃とは，攻撃者がサーバに対して短時間に大量のパケットを送り続け，サーバのサービス停止を図る攻撃である．DoS攻撃の判別方法の現状として，サーバに向けて特定のクライアントから一定時間内に送信されるパケット量が多いか少ないかを比較し，閾値を超えたものを攻撃者と判別する手法が用いられてきた[2]．しかし，既存の判別方法では対応できない攻撃が登場した．それがSlow HTTP DoS攻撃である．

Slow HTTP DoS攻撃とは，Webサーバに対し大量のHTTPリクエストを長時間維持することで，Webサーバの処理中リクエストを飽和させ，サービス不能状態にする攻撃である．この攻撃では，少ないパケット量で攻撃が可能なため，これまでのDoS攻撃の判別手法ではパケット量を比べても正常な通信のクライアントと攻撃者の判別がつかなくなってしまう[3][4].

この問題に対して，サーバのタイムアウトの設定やクライアント別の同時リクエスト数などに閾値を設定する対策方法がある[2][3][4]．特に平川ら[3]は1つのHTTPリクエストを1つの確立したTCPコネクションとみなし，重複したTCPコネクションの閾値に基づいてSlow HTTP DoS攻撃の判別をしている．しかしこの手法では，多数の攻撃元を用いる分散型の攻撃の場合，正常な通信のクライアントを誤って攻撃者として判別してしまうことがある．

本研究では，重複したTCPコネクションの閾値のみを使用した場合に，多数の攻撃元を用いる分散型のSlow HTTP DoS攻撃を，正常な通信のクライアントと攻撃者を適切に判別することができないという問題を解決に導くために，重複したTCPコネクション数での判別が可能か否か，実際の正常な通信での重複コネクション数の分析をし，この結果について考察する．

[1] パケットの通信を行うためにネットワーク上に確立される仮想的な通信路のこと[5]

2. Slow HTTP DoS 攻撃

2.1 Slow HTTP DoS 攻撃の登場

インターネットの普及に伴い登場したセキュリティの脅威の1つがDoS攻撃である．最近では従来の攻撃検知手法が通用しないSlow HTTP DoS攻撃が登場した．

Slow HTTP DoS攻撃とは，Webサーバに対し大量のHTTPリクエスト（以下リクエストとする）を長時間処理中の状態に維持し，処理中リクエストを飽和させる攻撃である．そして処理中のリクエスト数がサーバに設定された上限値に達すると，新たなリクエストを受け付けることができなくなり，異常終了，無反応といったサービス不能状態に陥る．リクエスト数の上限値の設定は主記憶容量によって制約を受ける．一方でこの攻撃では長時間にわたって少ないパケットで大量のリクエストを維持し続けることで，これまでのDoS攻撃の判別手法ではパケット量を比べても正常な通信のクライアントと攻撃者の判別がつかなくなる．そのため，別の対策が必要になる[3][4].

2.2 重複した状態のコネクション

平川らは1つのTCPコネクション1（以下コネクションとする）を1つのリクエストとみなして，同時処理リクエストのひっ迫を判断している．リクエストを追跡するには，アプリケーション層での解析が必要となるため，1つのリクエストを1つのコネクションとみなすことで，トランスポート層での解析が可能になる．

コネクションはクライアントからサーバに開始を意味するSYNパケットが送信されることにより確立し，クライアントとサーバのいずれかが終了を意味するFINパケットを送信することで終了する．基本的には前のコネクションが終了すると次のコネクションが確立されるが，FINパケットを送信してコネクションが終了する前に次のコネクションのSYNパケットが送信されてしまう図 1のような重複している状態（以下重複コネクションとする）になる場合がある．コネクション終了前に，サーバに対する新たなコネクションが次々に確立して，サーバが多くのコネクションを維持する必要に迫られると，同時処理がひっ迫してしまう可能性があり，場合によってはサービス不能状態に陥ってしまう可能性もある．

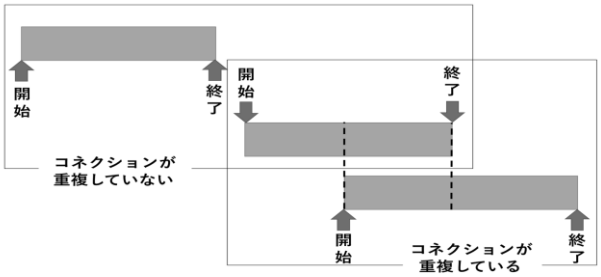

図 1 コネクションが重複している状態の例

2.3 Slow HTTP DoS 攻撃の検知手法と問題

平川らの手法[2]では，重複コネクション数の閾値を設定し，新たなコネクションを確立した後，その閾値に達したクライアントの重複コネクション数を比較する．この数が最も多いクライアントを攻撃者と判別してコネクションを切断する．この時，閾値をサーバが同時処理できるコネクション数の上限よりも低く設定することによりサービス不能状態に陥ることを防ぐ．

しかし，正常な通信でも重複コネクション状態になるケースもあり，その数が必ずしも少なくなるとは限らない．また，複数箇所から分散して攻撃されると，1クライアントあたりの重複コネクション数を小さくすることができるため，正常な通信の重複コネクション数が攻撃者の重複コネクション数を上回ってしまう場合もあり，結果として正常な通信をしているクライアントを誤って攻撃者と判別し，コネクションを切断してしまう可能性がある．そこで，正常な通信の重複コネクション数を分析し，閾値を見直す必要がある．

本研究では，正常な通信の重複コネクション数を分析し，Slow HTTP DoS攻撃の検知手法について考察する．

3. 正常な通信の重複コネクション数

3.1 トラフィックデータからパケットの抽出

正常な通信の重複コネクション数を分析するために，トラフィックデータと呼ばれる実際にネットワーク上を流れるパケットの通信状況が記されたデータを用いた．使用するトラフィックデータは2022年7月20日14時から14時15分までの15分間のデータである．このデータは，トラフィックの計測および解析を目的とした「WIDEプロジェクト2」の「MAWIワーキンググループ」のデータであり，一定時間，一定場所のデータを毎日計測し，ネットワーク上に公開しているものである．さらに，このトラフィックデータでは15分間で約800万件のパケット通信が行われており，常に通信のクライアントのパケットのみで構成されていると仮定する．

本研究では，トラフィックデータから特定の条件に合致したデータを抽出するために図 2に示した「WireShark」でパケット解析をする[6][7].

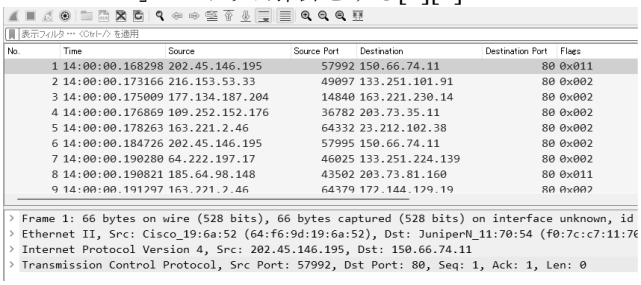

図 2 Wireshark の画面例

分析の手順についての概要を図 3に示す．まずSlow HTTP DoS攻撃はWebサーバに向けての攻撃であるため，トラフィックデータからHTTPの通信と，セキュリティが施されているHTTPSのパケットデータを抽出する(①)．次にコネクションの開始時間と終了時間を把握するために①で抽出したデータからSYNパケットとFINパケットのみを抽出する(②)．最後にコネクションを構成する1対のSYNパケットとFINパケットに対し，同一のクライアントとサーバ間でのコネクションが5つ以上となっているパケットのみを抽出する(③)．理由として，クライアントとサーバ間でのコネクションが多くなると重複コネクションが発生する可能性が高くなるためである．

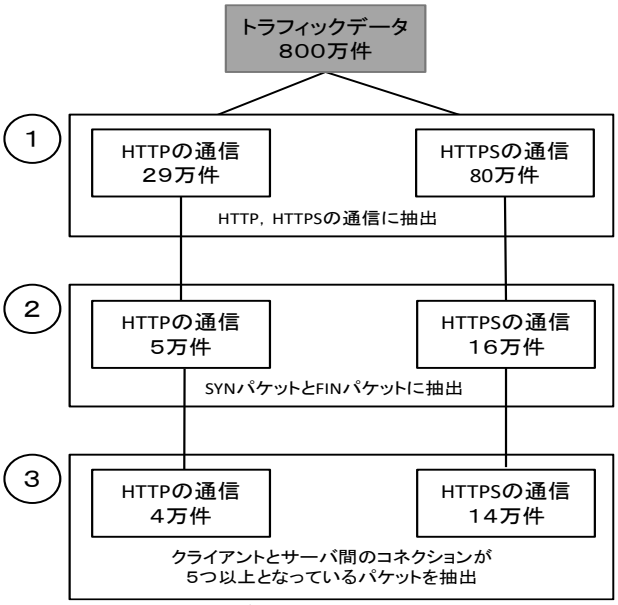

図 3 分析データの抽出処理の手順

最終的に抽出されたパケットデータ群から，クライアントIPアドレス，サーバIPアドレス，クライアントポート番号が一致しているSYNパケットとFINパケットの対を1つずつ抽出することで，コネクションの一覧を生成する．この時，SYNパケットのタイムスタンプがコネクションの開始時刻となり，FINパケットのタイムスタンプが終了時刻となる．その結果の例を 表 1に示す．

[2] https://mawi.wide.ad.jp/mawi/

なお，同一のクライアントとサーバ間の通信であることが明示できるように，それぞれのペアにA，B，C…という識別名を付与してある．

表 1 特定のペアのコネクションデータ

ペア	クライアントのIPアドレス	サーバのIPアドレス	コネクション開始時刻	コネクション終了時刻
A	202.76.96.204	13.135.88.44	14:00:00.1	14:00:00.2
A	202.76.96.204	13.135.88.44	14:00:00.2	14:00:00.5
B	14.248.97.94	150.66.73.139	14:00:00.2	14:00:00.6
B	14.248.97.94	150.66.73.139	14:00:00.5	14:00:00.8
B	14.248.97.94	150.66.73.139	14:00:00.6	14:00:01.0
C	202.76.96.204	14.248.97.94	14:00:01.0	14:00:01.5

3.2 パケットデータからの重複コネクションの抽出

表 1のデータから正常な通信の重複コネクション数を抽出するために，0.1秒単位の時間区分ごとに各コネクションが継続した状態か否か示した表 2を作成した．この表では各行が1つのコネクションを表し，そのコネクションが接続した状態である時間区分のセルを1としている．なお，「14:00:00.1」は14時0分0.1秒という示し方であり，Aの1つ目のコネクションは14時0分0.1秒から14時0分0.2秒で継続しているため，「14:00:00.1」，「14:00:00.2」の2つのセルが1となる．この表の作成にあたっては，分析対象データ数がHTTPで約2万件，HTTPSで約7万件と大量であり，Excelの関数処理では膨大な時間がかかってしまう．これを避けるため，本研究ではVBA[8][9]でプログラムを作成することにより，分析処理時間の大幅な短縮を図っている．

表 2 コネクションの時間区分ごとの接続状態

ペア	コネクション開始時刻	コネクション終了時刻	14:00:00.1	14:00:00.2	14:00:00.3	.....	14:15:00.0
A	14:00:00.1	14:00:00.3	1	1	1		
A	14:00:00.2	14:00:00.5		1	1	.....	1
B	14:00:00.2	14:00:00.6		1	1		
B	14:00:00.5	14:00:00.8					1
B	14:00:00.6	14:00:01.0					

次に重複コネクション数をカウントするために，特定の二者間通信，すなわち特定のペアのコネクションに対し，各時間区分の1の数を集計する．これにより，各ペアの時間区分別の重複コネクション数を求めることができる．なお，この分析もVBAを用いたプログラムにより行っている．例として，表 3のような時間区分ごとの接続状態であれば，各ペアの時間区分ごとの重複コネクション数は表 4のような結果となる．この分析をHTTP，HTTPSのプロトコルそれぞれに対して行う．

表 3 時間区分ごとの接続状態の例

ペア	14:00:00.1	14:00:00.2	14:00:00.3	14:00:00.4	14:00:00.5	14:00:00.6	16:00:00.7	14:00:00.8
A	1	1						
A		1	1	1	1			
B		1	1	1	1	1		
B					1	1	1	1
B						1	1	1

表 4 ペア・時間区分別の重複コネクション数集計表の例

ペア	14:00:00.1	14:00:00.2	14:00:00.3	14:00:00.4	14:00:00.5	14:00:00.6	14:00:00.7	14:00:00.8
A	1	2	1	1	1			
B		1	1	1	2	3	2	2

4. 重複コネクション数の比較

表 4のデータをHTTP，HTTPSの通信で分けて2つのデータを作成し，IPアドレスのペアごとに重複コネクション数の最大値を抽出した．縦軸に最大値がすべてのコネクションに占める割合を算出して2つのデータを比較する．その結果を図 4，図 5に示す．

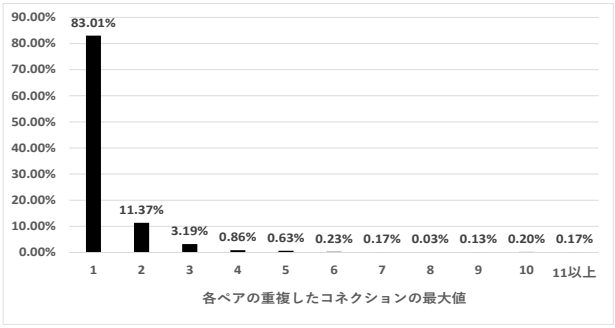

図 4 HTTP の重複したコネクション数の割合

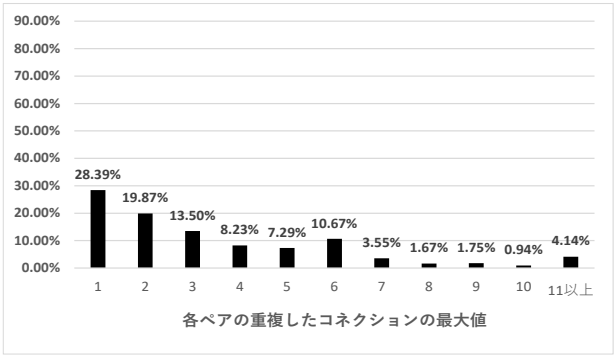

図 5 HTTPS の重複したコネクション数の割合

図 4では重複がない状態を意味する最大値が1の二者間通信の割合が全体の83%を占めているため，HTTPでは重複がない二者間通信よりも，重複が発生する二者間通信の割合が少ないことが明らかになった．

それに対し，図 5では重複がない二者間通信の割合が全体の28％しか占めていない．そのため，HTTPSでは重複がない二者間通信よりも，重複が発生する二者間通信の割合が多いことがわかる．

この結果から，HTTPSのプロトコルでは正常な通信のクライアントでもコネクションの重複は頻繁に発生することがあるため，従来の重複コネクションの閾値による手法では正常な通信のクライアントを誤って攻撃者として判別することがある．このため，判別の基準が重複コネクション数の閾値だけでなく，ほかの判別要素も検討する必要がある．

5. 重複コネクション数とコネクションの継続時間の関係性

Slow HTTP DoS攻撃の新たな判別要素としてコネクションの継続時間について検討する. Slow HTTP DoS攻撃では大量のコネクションが長時間にわたって維持されているという特徴がある. それゆえ, 重複コネクションが存在すると, コネクションの継続時間は重複コネクションが存在しない場合よりも一般的に長くなる傾向になり, 重複コネクション数とコネクションの継続時間の間には相関関係があることも考えられる.

もし相関関係があれば, 2つの値を用いた判別では多重共線性[10]が生じ, 重複コネクション数で判別できなければコネクションの継続時間を説明変数として追加しても判別制度が向上しない可能性が高くなる. 一方, 相関関係がなければ, 重複コネクション数で攻撃を判別できない場合でも, 独立性のある情報としてコネクションの継続時間の判別に寄与できる可能性がある.

この相関関係を明らかにするために, ペアごとの重複コネクションの最大値と, ペアごとのコネクション継続時間の平均値(以下コネクションの平均継続時間とする)に対する散布図と相関係数を分析した. HTTP, HTTPSそれぞれの散布図を図 6, 図 7に示す.

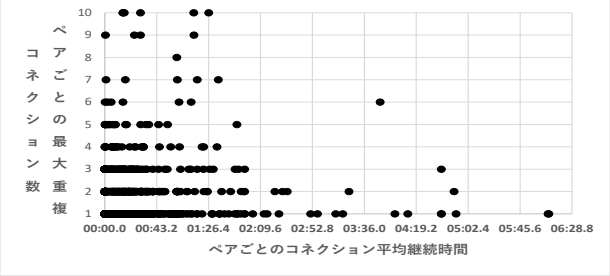

図 6 HTTP の通信の重複コネクションとコネクション平均継続時間の関係を示した散布図

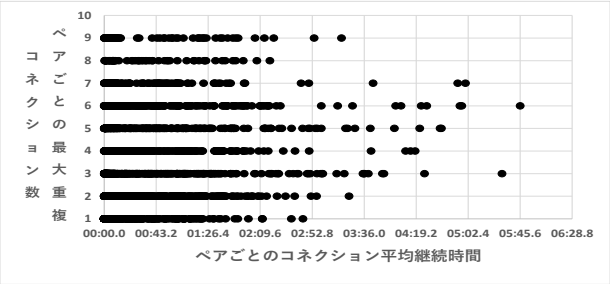

図 7 HTTPS の通信の重複コネクションとコネクション平均継続時間の関係を示した散布図

それぞれの相関係数は, HTTPの通信は0.20, HTTPSの通信は0.26となり, 重複コネクション数とコネクション継続時間には相関関係が見られなかったため, コネクションの継続時間はSlow HTTP DoS攻撃の新しい判別要素として利用できる可能性がある.

6. まとめ

本研究は, 正常な通信の重複コネクション数を分析し, HTTP, HTTPSの通信ではコネクションが重複する頻度が異なるため, 判別の基準が重複コネクション数の閾値のみでは正常な通信のクライアントを誤って攻撃者と判断する可能性があることを明らかにした. さらに, 重複コネクション数とコネクションの継続時間に相関関係がないことを示し, コネクションの継続時間をSlow HTTP DoS攻撃検知のための新たな判別要素として利用できる可能性も示した.

今後, 攻撃者のデータを用いて, コネクションの継続時間を判別要素として加えた場合の実証実験を行う必要がある.

参考文献

[1] 警察庁, 令和 3 年におけるサイバー攻撃をめぐる脅威の情報等について, https://www.npa.go.jp/publications/statistics/cybersecurity/data/R03_cyber_jousei.pdf, (参照 2022-4-7).
[2] 平川哲也, "攻撃コストの観点による Slow HTTP DoS 攻撃に対する防御手法の研究", 岩手県立大学機関リポジトリ, 2020.
[3] 平川哲也ほか, "TCP コネクション数と継続時間に基づく Slow HTTP DoS 攻撃に対する防御手法", 情報処理学会論文誌, Vol,61, No.3, pp.581-590, 2020.
[4] 川橋裕ほか, "Slow HTTP DoS 攻撃の手法と防御方法について", 学術情報処理研究, No.20, pp.128-136, 2016.
[5] ネットワークマガジン編集部, すっきりわかった！TCP/IP, 2005.
[6] 久米原栄, 上田浩, Wireshark パケット解析リファレンス, ソフトバンククリエイティブ, 2009.
[7] 竹下恵, パケットキャプチャ実践技術 Wireshark によるパケット解析応用編, 株式会社リックテレコム, 2009.
[8] 大村あつし, Excel VBA をはじめるまえに絶対知っておきたい「マクロ」の本, 株式会社技術評論社, 2014.
[9] 立山秀利, Excel VBA のプログラムのツボとコツがゼッタイにわかる本, 秀利システム, 2014.
[10] 菅民郎, Excel で学ぶ多変量解析入門 Excel2013/Excel2010 対応版, オーム社, 2013