CDN のキャッシュサーバを騙った DDoS 攻撃の二段階検知法の最適閾値設定法

Optimum Threshold Design of Two Stage Detection Method of DDoS Attack Mimicking CDN Caches

谷口 和也　　上山 憲昭
Kazuya Taniguchi　　Noriaki Kamiyama
立命館大学　情報理工学部　情報理工学科
Faulty of Information Science and Engineering, Ritsumeikan University

1. はじめに

インターネット上ではコンテンツの多くが, 遅延時間とトラヒック量を削減するため, CDN（Content Delivery Network）を用いて配信されている. 一方で近年, ネットワーク上に広く存在するボットから大量のパケットをターゲットホストに送信することで、ターゲットサーバを機能不全とする DDoS（Distributed Denial of Service）攻撃が頻繁に発生している. そのような攻撃に対して, ボットが OS の IP アドレス宛てに直接パケットを送信した場合は，OS に対し CS から要求が届くことから，CS 以外からの配信要求をファイアウォールで棄却することで対処できる [1][2]. しかし, ボットがキャッシュサーバ (CS) の IP アドレスを発アドレスとして偽り, 標的オリジンサーバ（OS）へパケットを送信した場合に, ファイヤーウォールで検知ができない問題がある. それに対して著者らは，DDoS 攻撃が短い時間間隔で多数の CS から発生することに着目し, 要求発生の違いを利用した, CDN を騙った DDoS 攻撃の二段階検知法を提案した [3]. 本稿では本方式の, 最適な閾値の設定法を提案する.

2. DDoS 攻撃の二段階検知法

CS からの正常な問い合わせ時には, コンテンツ提供者の DNS サーバに名前解決のログが残るのに対し, ボットからの IP アドレスを直接用いた要求は DNS の名前解決を用いないため名前解決のログが残らない. そのため DNS のログを調べれば, CS からの正常な配信要求か, ボットからの DDoS パケットか判別が可能である. しかし, OS には大量の配信要求が到着することからすべての配信要求に対して DNS のログを調べると OS の処理負荷の増大が懸念される. そのため, 先行研究では問い合わせ間隔に閾値 T を設定し, ホストごとに要求の到着間隔の計測を行い, T に対しての要求の到着間隔の長さにより DNS サーバの問い合わせの有無を確認する必要のある要求を絞り込んでいる. 具体的な判別の方法としては, T より長い間隔の要求は CS からの正常な要求と見做し, T より短い間隔の要求はボットからの DDoS 攻撃の可能性を考え問い合わせの有無を確認する. さらに問い合わせが有れば CS からの正常な要求と見做し, 問い合わせがなければ DDoS 攻撃と見做しアクセスを棄却する. 本方式により DNS のログの検索負荷を減らし, DDoS 攻撃を効率よく検知することができる。

3. 最適閾値設定法

閾値 T を大きく設定しすぎると DNS サーバでの問い合わせの有無を確認する回数が増え OS の負荷が増大する. 一方閾値 T を小さく設定しすぎると DDoS 攻撃を検知できない可能性があり, DDoS 攻撃に対しての強度が低下する. そのため, 閾値 T は DNS の検査レートの許容上限を満たす範囲で最大化することが望ましい. このように閾値を設定することで, DNS サーバの処理能力上限を考慮しながら, DDoS 攻撃を最大限, 検知することができる. コンテンツ数を M, 閾値 T に対してのコンテンツ m の DNS の検査レートを $r_m(T)$ とすると, T に対する DNS の総検査レート R_n は

$$R_n(T) = \sum_{m=1}^{M} r_m(T) \qquad (1)$$

で求まる. DNS の検査レートの上限値を U_n とするとき, $R_n(T) = U_n$ となる T の最大値を T に設定する.

4. 性能評価

提案方式の有効性を計算機シュミレーションにて評価する. キャッシュ置換方式は LRU 方式とし, コンテンツ数を N=100, シュミレーション時間を 10,000 秒とする. また, ユーザのコンテンツ選択確率はパラメタ θ の Zipf 分布に従う. 図 1 は CS や Zipf 分布の θ を変化させたときのコンテンツ毎のキャッシュミスの発生間隔の平均値を示している. コンテンツ ID は人気順を示していて, 低くなるほど人気度が高くなる. コンテンツ ID が大きくなるほどキャッシュミスの発生回数が少なくなるためキャッシュミスの発生回数の平均値が大きくなるのは予想できる. しかし一方で, 特に Zipf 分布の θ が大きく, 高人気コンテンツの要求比率が高い場合は, 高人気コンテンツのキャッシュミスが減少するため，キャッシュミスの発生間隔が増加する. そのため平均キャッシュミス発生間隔は下に凸な曲線となる. 図 2 は CS や Zipf 分布の θ を変化させたときの上限値を満たす最適な閾値 T を示している. これは, 閾値を仮に設定した際のコンテンツ全てに対しての DNS の総検査レートのグラフより閾値を設定した際のグラフとの交点から求めることができる. CS や Zipf 分布の θ を変化させても大きな変化はなく, いずれの場合も上限値の増加に伴い T の最適値は増加するが, 上限値がある値に近づくと急激に T の最適値は増加する.

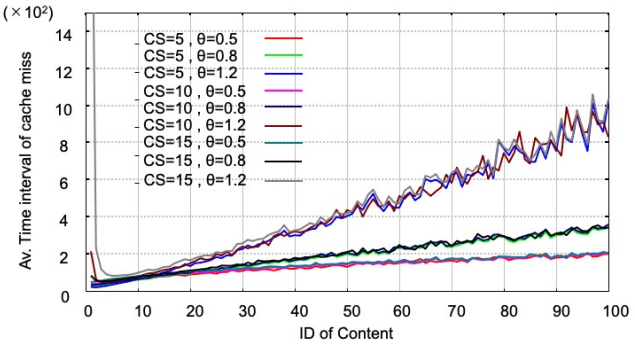

図 1　キャッシュミスの平均発生間隔

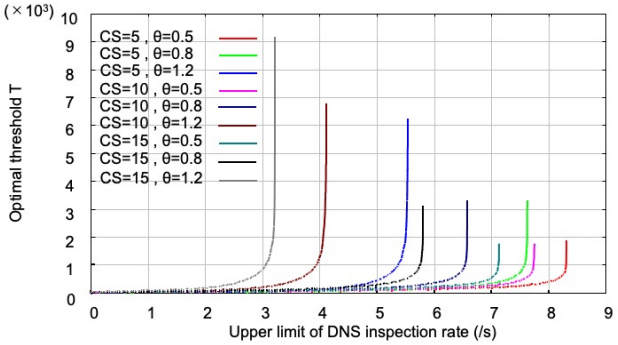

図 2　上限値に対する最適な閾値 T

謝辞本研究成果は JSPS 科研費 18K11283 および 21H03437 の助成を受けたものである. ここに記して謝意を表す.

参考文献

[1]T. Vissers, et al., Maneuvering Around Clouds: Bypassing Cloud-based Security Providers, ACM CCS 2015
[2]D. Gillman, et al., Protecting Websites from Attack with Secure Delivery Networks, Comp. Mag., 2015
[3] 宮崎椋平, 上山憲昭, CDN のキャッシュサーバを騙った DDoS 攻撃の防御方式, 信学会 2021 年総合大会, B-11-10

2022/9/6 ～ 9　オンライン開催