4X-03

トラヒックの安定性に着目したダークネット観測データの解析

○金井　登威†，　角田　裕‡，　キニ グレン マンスフィールド§
†東北工業大学大学院工学研究科，‡東北工業大学工学部情報通信工学科，
§株式会社サイバー・ソリューションズ

1. はじめに

ダークネット[1]は到達可能だが特定のホストに割り当てられていないアドレス空間であり，ワームによるスキャン活動やDoS攻撃によるバックスキャッタなどのパケットを効率的に観測できる．そのため，ダークネット観測データを分析し，これらの不正な活動の傾向を把握する研究が広く行われている．

本研究では，ネットワークトラヒックの特性として知られている安定性[2]という概念が，ダークネットのトラヒックに対しても適用できることを示し，ダークネット観測データからの安定性に基づいたイベント抽出について検討した．また，安定性が失われた時間帯の観測データを詳細に調査し，安定性を失わせた要因の分析結果を報告する．

2. トラヒックの安定性

トラヒックの安定性とは，ネットワークトラヒックの大部分は常にある一定数の種類のもので占められているという考え方である．文献[2]では，トラヒックをタイムスロット毎に分割し，各スロットにおいて支配的なパケットの種類数Top$N$を求め，それを安定性の判別指標とする方式を提案している．

2.1. Top$N$の定義

文献[2]ではスロットあたりのパケットをトラヒック情報項目毎にグループ化し，各グループが占める割合に基づいてTop$N$を定義している．Top$N$の求め方の概要を図 1 に示す．

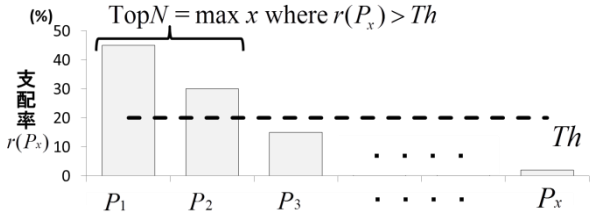

図 1 Top$N$の概要

ここで，$n(P_i)$はグループP_iのパケット数であり，$n(P_i) > n(P_{i+1})$が成り立っているとする．あるグループP_xの全トラヒックに対する支配率$r(P_x)$を数式(1)で表すとき，Top$N$を数式(2)で定義する．つまり，Top$N$は閾値Th以上の支配率を持つグループの数を示している．

$$r(P_x) = \frac{n(P_x)}{\sum n(P_i)} \times 100\,(\%) \qquad (1)$$

$$\mathrm{Top}N = \max x \text{ where } r(P_x) > Th \qquad (2)$$

Analyzing the darknet traffic focusing on stability of traffic
†Toui KANAI・Tohoku Institute of Technology
‡Hiroshi TSUNODA・Tohoku Institute of Technology
§Glenn Mansfield Keeni・Cyber Solutions Inc.

2.2. 安定性の有無の判別方法

安定性がある状態とはスロット毎のTop$N$の変動幅が小さい状態を指す．安定性が失われた状態とはTop$N$が過去と比較して大きく変化した場合であり，原因となる何らかのイベントが発生したと考えられる．例えば宛先ポート番号毎にパケットをグループ化しているとき，数種類の特定のポートへのパケットが支配的な状態が続いていたとする．このときDoS攻撃などにより，特定の単一の宛先ポート番号へ大量にパケットが到着すれば，Top$N$は大きく変化し安定性が失われることになる．

本研究では，図 2 に示すように当該スロットと過去TスロットのTop$N$の値に基づいて各スロットの安定性を判別する．

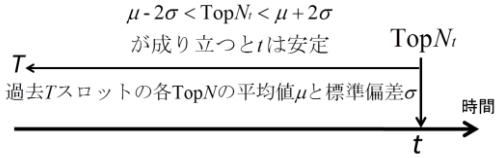

図 2 tでの安定性の有無の判別

判別対象スロットtのTop$N$をTopN_tとするとき，tを基準として過去TスロットのTop$N$ (TopN_{t-1},…,TopN_{t-T})の平均値と標準偏差をそれぞれμとσとし，数式(3)によってスロットtの安定性の有無を判別する．

$$\mathrm{t}\text{ は}\begin{cases}\text{安定} & (\mu - 2\sigma < \mathrm{Top}N_t < \mu + 2\sigma)\\ \text{安定ではない} & (\text{otherwiseif})\end{cases} \qquad (3)$$

3. ダークネットトラヒックにおける安定性

3.1. 対象データ

本研究では，一般社団法人 JPCERT コーディネーションセンターの定点観測システム[3]が 2010/1/1～2012/12/31 に観測したトラヒックを分析した．これは IX の近傍や xDSL のエッジ等，ユーザが利用するネットワークと同等の環境に分散配置されたセンサによる観測結果である．

このトラヒックを 1 日幅のタイムスロットに分割し，各スロットにおける TCP および UDP パケットを宛先ポート番号毎にグループ化した．宛先ポート番号に着目した理由は，ダークネットで多数観測される DoS 攻撃やポートスキャンが影響する情報項目と考えたためである．

本節では各スロットで求めたTop$N$の傾向とそれに基づいたダークネットトラヒックの安定性の調査結果を説明する．

3.2. Top$N$の値の傾向

図 3 に，$Th = 1.0\%$で算出したTop$N$の値の出現頻度を示す．

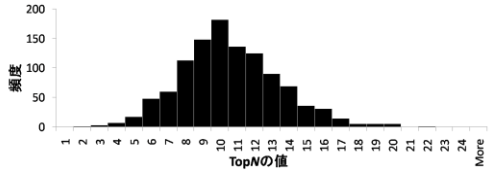

図 3 日毎のTop$N$の度数分布

図 3 から，Top$N$の値は 10 を中心に 2〜22 の範囲をとり，ダークネットにおいても，タイムスロット毎のTop$N$は一定の範囲に収まり，極端に大きいまたは小さい値を取ることは少ないと言える．即ち，Top$N$は概ね安定しており，値が大きく変化した時には何らかの通常では見られないイベントが発生していると考えることができる．

3.3. 安定性が失われたスロット数

$T = 2, \cdots, 10$および$Th = 0.5\%, 0.75\%, 1.0\%$とし，3 年間の全 1,096 スロットの安定性を調査した．安定性が無いとされたスロット数を図 4 に示す．

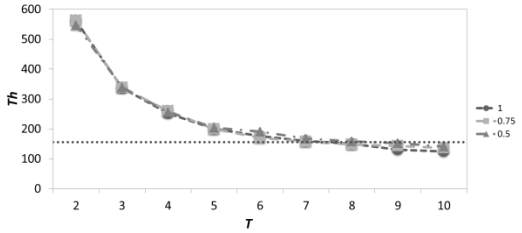

図 4 安定性がないスロット数

図 4 から，安定性の無いスロット数はThにはほとんど左右されないが，Tを小さくすることで大きく増加することが確認できる．

4. 安定性が失われたスロットの調査

本節では，安定性が無いとされた日のトラヒックを詳細に調査し，発生していた事象について考察する．ここでは，週 1 日(3 年間で計 156 日)程度の頻度で生じるイベントの検知を目的とし，図 4 において安定性が失われたスロットが最も 156 日に近く検知される$T = 7$を採用した．

この時，$0.5\%, 0.75\%, 1.0\%$の全てのThで共通して安定性がないとされたスロットは 52 日分あった．各日のトラヒックを調査した結果，表 1 に示すように 32 日分でポートスキャン，バックスキャッタ，DoS 攻撃とみられる事象が観測されていた．

表 1 発生していると考えられる事象

ポート スキャン	バック スキャッタ	DoS 攻撃	不明
20	15	7	20

これらのうち実際の裏付けが得られた 2010 年 4 月 23 日と 2012 年 4 月 4 日の事象について説明する．

4.1. 2010 年 4 月 23 日

宛先ポート毎のパケット数を確認したところ，66 種類のポートに 10 件ずつのパケットが到着していた．そのようなポートは過去 1 週間には 1 日当たり 0〜5 種類しか存在しておらず，何らかの原因により発生した事象であると考えられる．

66 種類のポートに向かうパケットを抽出したところ，中国のあるホストの TCP6532 番ポートから送られた計 658 個の SYN/ACK パケットを確認した．文献[4]によれば，2010 年 4 月下旬に TCP 6532 番ポートを送信元とする中国からのバックスキャッタが多数観測されており，本研究で確認した SYN/ACK パケットもその一部であると考えられる．すなわち，このバックスキャッタの存在が安定性を失わせる要因になったと考えられる．

4.2. 2012 年 4 月 4 日

頻繁に攻撃手段や攻撃対象となる TCP80 番ポートを送信元または宛先とするパケットを調査した結果，ある米国のホストから対象ポートを用いて送信された 1,1191 個の SYN/ACK パケットを確認した．このとき宛先ポートの種類数は 1,180 種類であり，各ポートへ約 1 個ずつパケットが送信されていることがわかった．文献[5]によれば，対象の日を含む期間において TCP80 番ポートを送信元とする米国からのバックスキャッタが増加傾向にあったことが報告されており，これらのパケットはその一部であると考えられる．

以上から，2012 年 4 月 4 日に安定性を失わせた要因は米国のサーバから届いたバックスキャッタであると考えられる．

5. まとめと今後の課題

本研究では安定性という考え方をダークネットでのトラヒックに適用し，トラヒックに含まれる事象を発見できることを示した．本稿では宛先ポートによりパケットをグループ化し安定性の有無を調査したが，今後は送信元ポートなど他の要素でも同様に安定性の有無の判別による分析を進めていく．加えて，様々な期間や閾値での調査も今後の課題となる．

謝辞

定点観測システムのデータを提供していただいた一般社団法人 JPCERT コーディネーションセンターに感謝いたします．

参考文献

[1] M. Bailey, E. Cooke, F. Jahanian, A. Myrick, and S. Sinha, "Practical Darknet Measurement" in Proceedings of the 40th Annual Conference on Information Sciences and Systems (CISS '06), pp. 1496-1501, Princeton, New Jersey, USA, March 2006

[2] IPA 情報セキュリティ対策研究開発評価等事業　高トラヒック観測・分析法に関する技術調査 http://www.ipa.go.jp/security/fy15/reports/traffic_mon/documents/traffic_mon.pdf

[3] TSUBAME(インターネット定点観測システム),http://www.jpcert.or.jp/tsubame/

[4] @police インターネット観測結果等(平成 22 年度第 1/四半期(4 月〜6 月)), http://www.npa.go.jp/cyberpolice/detect/pdf/20100730.pdf

[5] @police インターネット観測結果等(平成 24 年度第 1/四半期(4 月〜6 月))http://www.npa.go.jp/cyberpolice/detect/pdf/20120831.pdf