パケットカウントの分布に基づく Darknet トラフィックの解析

杉本　周 †　福田 健介 §　廣津 登志夫 ¶　菅原 俊治 †
† 早稲田大学 基幹理工学研究科　§ 国立情報学研究所/科学技術振興機構
¶ 法政大学 情報科学部

1 序論

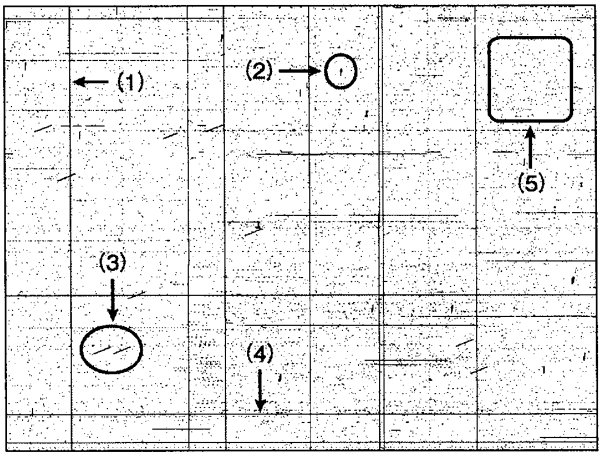
インターネットは我々の日常生活において電気、水道、ガス、電話に続く第 5 の生活インフラとなりつつあり、安心して利用できる安全なネットワークを提供することがインターネットに関する研究の主要なテーマの一つになっている。その一例として Darknet を利用したトラフィック解析 [1] が行われている。Darknet とは経路広報されているが実際には使用していないアドレス空間を指し、そこでは通常の通信によって発生するトラフィック以外の異常なパケットを効率よく収集できる。

我々は Darknet を利用した広域分散協調可能な攻撃性トラフィック監視アーキテクチャ [3,5–7] を提案している。これは複数の拠点に分散配置した Darknet を監視し、各拠点での観測データをもとに監視外アドレス空間への攻撃性トラフィックを予測するものである。したがって、このアーキテクチャを実現する課題の一つとして Darknet トラフィックデータから精度の高い攻撃予測を行う手法を確立することが挙げられる。そこで本研究では Darknet トラフィックデータのパケットカウントの分布に基づく解析を行い、これを利用して攻撃予測の精度向上に必要不可欠なアドレススキャン型攻撃を抽出する手法を提案する。

エントロピーや画像処理を利用する従来の異常トラフィック検知手法 [2,4] は主にバックボーントラフィックを解析対象としており、Darknet トラフィックデータにおける攻撃検出には必ずしも適しているとは言えない。特に従来手法では、時系列データにおける特徴量の変化に対する影響力の弱い攻撃は検出不可能である。これに対して提案手法ではトラフィックデータをブロック分割して解析することで Darknet トラフィックデータの中に埋もれている小規模かつ伝播速度の遅い攻撃をも抽出できることを示す。

2 パケットカウントの分布

本研究では日本国内に設置された /18 の観測アドレス空間を持つ Darknet において収集されたパケットトレースを使用した。図 1 は 2008 年 7 月 14 日の 24 時間分のパケットトレースのうちトラフィッククラスが tcp-syn のパケット[*1]について、(a) パケット到着パターン[*2]、および (b) 到着パケット数の推移を図示したものである。これによると (a) において観測アドレス空間全域を縦断するアドレススキャン型攻撃 (1) が確認できる時間帯でパケット数が急激に増加しているのがわかる。対して対象範囲が /24 程度の小規模なアドレススキャン (2) や伝播速度の遅いもの (3) は単純なパケット数の時系列データに対する影響力が弱いことがわかる。またパケット到着パターンに見られる横方向の線 (4) は特定の宛先ホストに対して集中的にパケットを送信するポートスキャン型攻撃が表れたものであり、上記以外のアドレス空間全域に散らばる点 (5) は背景雑音に分類される。提案する監視アーキテクチャの実現には、アドレススキャン型攻撃をポートスキャン型攻撃や背景雑音と区別し、効率的に抽出する手法が必要である。



(a) パケット到着パターン（86,400sec × 16,384ip）

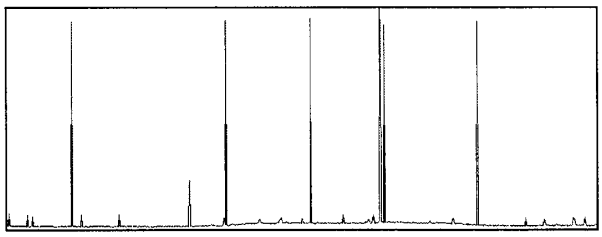

(b) 到着パケット数の推移（every 60sec）

図 1　Darknet トラフィックデータ @ 2008-07-14

本研究ではまずトラフィックデータをパケット到着時刻と宛先 IP アドレスについて適当な大きさを持つブロックで格子状に区切り、各ブロック中に含まれるパケット数 (PiB) を算出し、この PiB を階級とするブロック数の度数分布を調べた。図 2 は図 1 (a) のトラフィックデータのブロックサイズ 60sec × 256ip における度数分布である。これによると PiB の小さい部分に多数のブロックが集中しており、PiB = 100 でブロック数の累積相対度数は全体の 94% に達する。一方でパケット数の累積相対度数は 38% にとどまる。

Extracting Address-Scan Attacks from Darknet Traffic
Shu SUGIMOTO†, Kensuke FUKUDA§, Toshio HIROTSU¶, Toshiharu SUGAWARA†
†Waseda University, Graduate School of Fundamental Science and Engineering
§National Institute of Informatics/Japan Science and Technology Agency
¶Hosei University, Faculty of Computer and Information Sciences

[*1] TCP のフラグフィールドにおいて SYN ビットのみが 1 にセットされた TCP パケット

[*2] x 軸にパケットの到着時刻、y 軸に宛先 IP アドレスをとってプロットした散布図

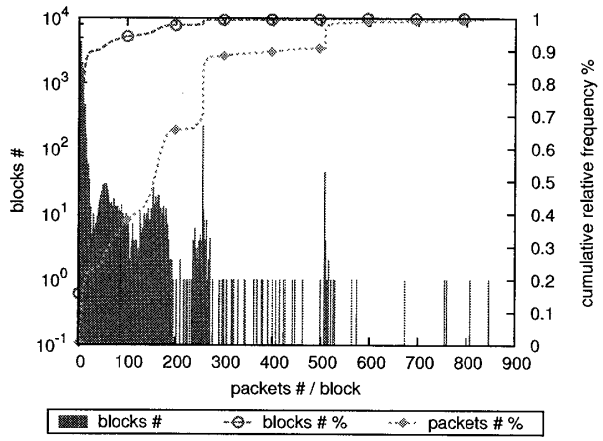


図 2　度数分布（60sec × 256ip）

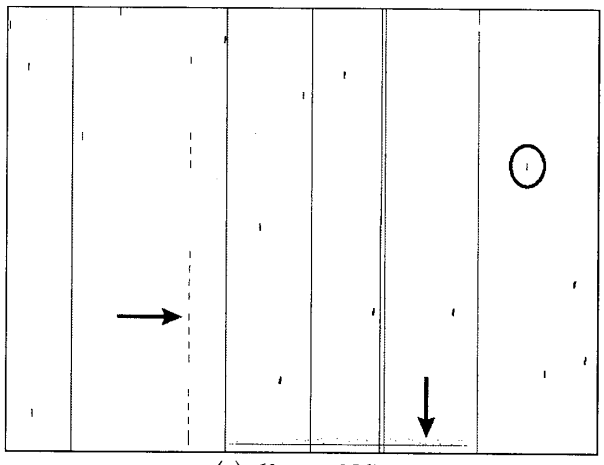
3　アドレススキャン型攻撃の抽出

図 1 より大規模なアドレススキャン型攻撃のパケットを含むブロックの PiB は大きな値を取り、またブロック分割を適用したことで小規模なアドレススキャンを含むブロックの PiB も同様に大きな値を取ると推測される。したがって図 2 より PiB の上位から数パーセントのブロックを取り出せば小規模なものも含め、アドレススキャンが抽出できると考えられる。そこで次にブロック数の累積相対度数が 95% を超える点を境界とし、これより下位のブロックを分離した。なお、この時の PiB は 134 packets/block、パケット数の累積相対度数は 42% である。分離後のトラフィックデータのパケット到着パターンを図 3 (a) に示す。分離前の図 1 (a) と比較するとアドレススキャンを示す 7 本の縦方向の線のうち左から 2 番目の線に部分的な欠落が見られるが、それ以外は大部分が抽出できており、また小規模なアドレススキャンも抽出できているのがわかる。下部にポートスキャン型攻撃を表す横方向の線が 1 本残っているが、観測アドレス空間全域に散らばる背景雑音は殆ど全て除去されているといえる。

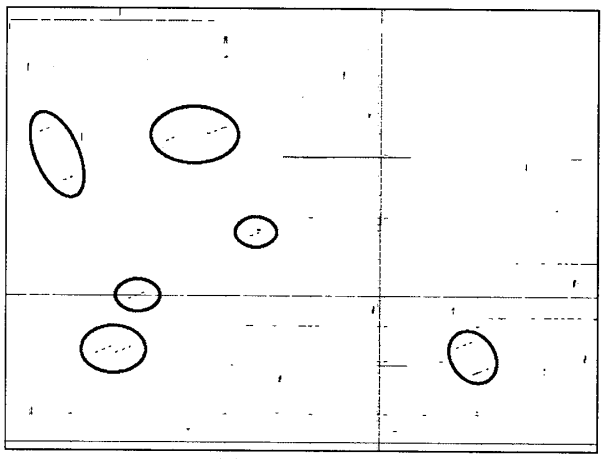
同様にブロックサイズ 480sec × 32ip で抽出した結果を図 3 (b) に示す。完全ではないが 60sec × 256ip で抽出できなかった伝播速度の遅い小規模な攻撃を抽出できているのがわかる。また縦方向の線が減り、横方向の線が増えていることもわかる。このようにブロックサイズを変化させると抽出できる攻撃に違いが出るのは、ブロックサイズの時間/空間比によって各々の攻撃の PiB に対する影響力が変化するためである。時間/空間比の値が小さければ伝播速度の速いアドレススキャン型攻撃の影響力が強まり、反対に大きければ伝播速度の遅いアドレススキャン型攻撃やポートスキャン型攻撃の影響力が強くなる。

4　結論

本研究では 86,400sec × 16,384ip の Darknet トラフィックデータをブロックに分割し、各ブロック中に含まれるパケット数（PiB）についての度数分布を調査した。その結果、ブロックサイズを 60sec × 256ip とすると大半のブロックが PiB の小さい部分に集中することがわかった。これに基づき PiB の上位からブロックを取り出したところ、小規模なものも含めてアドレススキャン型攻撃を精度良く抽出できた。さらにブロックサイズの時間/空間比を変化させることで、より抽出が困難な伝播速度の遅い攻撃も抽出できることがわかった。



(a) 60sec × 256ip



(b) 480sec × 32ip

図 3　分離後のパケット到着パターン

References

[1] Michael Bailey, Evan Cooke, Farnam Jahanian, Jose Nazario, and David Watson. The Internet Motion Sensor: A distributed blackhole monitoring system. In *NDSS '05*, 2005.

[2] Romain Fontugne, Toshio Hirotsu, and Kensuke Fukuda. An image processing approach to traffic anomaly detection. In *AINTEC '08*, pp. 17–26. ACM, 2008.

[3] Kensuke Fukuda, Toshio Hirotsu, Osamu Akashi, and Toshiharu Sugawara. Correlation among piecewise unwanted traffic time series. In *GLOBECOM*, pp. 1616–1620, 2008.

[4] Anukool Lakhina, Mark Crovella, and Christophe Diot. Mining anomalies using traffic feature distributions. In *SIGCOMM '05*, pp. 217–228. ACM, 2005.

[5] 今間俊介, 福田健介, 廣津登志夫, 菅原俊治. 断片ダークネットのためのパケット観測用ブリッジの提案. インターネットテクノロジーワークショップ（WIT）, 2008.

[6] 杉本周, 福田健介, 廣津登志夫, 明石修, 菅原俊治. 特定のアドレス空間を基準とした遅延相関解析によるインターネット上の攻撃予測の可能性. インターネットテクノロジーワークショップ（WIT）, 2009.

[7] 廣津登志夫, 福田健介, 栗原聡, 明石修, 菅原俊治. 断片アドレスを用いた分散協調インターネット監視に関する一考察. 情報処理学会 OS 研究会研究報告（83）, pp. 39–45, 2007.