ダークネットUDP通信の可視化画像とトラフィック統計量を用いたDDoSバックスキャッタ判定の改善

野村 竜也[1]　小澤 誠一[1,2,a)]　班 涛[3]　島村 隼平[4]

概要：本研究では，ダークネットで観測された UDP 通信によるトラフィック情報から DDoS 攻撃の跳ね返りパケットであるバックスキャッタか否かを判定する手法を提案する．パケットデータの振る舞いを可視化した画像を畳み込みニューラルネットワークで学習した埋め込み特徴と，UDP 通信に関連した 17 個の統計情報を組み合わせ，DDoS バックスキャッタの判定を行う．実験データは NICT のダークネットセンサから得られたパケットデータを使用する．実験の結果，0.98 を超える F 値が得られ，バックスキャッタの判定を行うのに十分な精度が得られることを示す．

キーワード：機械学習，サイバーセキュリティ，DDoS，ダークネット，深層学習

Enhancement in DDoS Backscatter Detection Using Visualization Images of Darknet UDP Communication and Traffic Statistics

Tatsuya Nomura[1]　Seiichi Ozawa[1,2,a)]　Tao Ban[3]　Jumpei Shimamura[4]

Abstract:
In this study, we propose a method that detects backscatter events—rebounding packets from a DDoS-attacked host—from malicious network traffic observed in the darknet, with a focus on UDP communication. The classification of DDoS-backscatter events is implemented with a convolutional neural network that leverages two types of features: 17 statistical features that characterize the communication from a host to the darknet and image features obtained from a graph that visualizes the attack behavior during the event. When the proposed method is evaluated on packet data collected at NICT's darknet, it shows a F-measure value above 0.98—an accuracy that is considered promising for practical security operation in a real-world scenario.

Keywords: machine learning, cyber security, DDoS, darknet, deep learning

1. はじめに

対象サーバの処理不能を目的として一斉にパケットを送り付ける DDoS (Distributed Denial of Service) 攻撃による被害が拡大している．DDoS 攻撃には大きくボットネットを利用した攻撃，リフレクター攻撃の 2 つに分けられる．ボットネットを利用した攻撃に関しては攻撃者が多数のインターネット機器をマルウェアに感染させるなどして操り，一斉に攻撃命令をかけ攻撃を行う．近年では脆弱性が指摘されている IoT の普及やそれらを狙ったマルウェアの出現に伴い，DDoS 攻撃の踏み台となる機器が増え規模が大きくなり脅威となっている [1]．一方，リフレクター攻撃は攻撃者が攻撃対象の IP アドレスになりすまして DNS サーバなどにパケットを送り，応答パケットを攻撃対象に

[1] 神戸大学 大学院工学研究科
Graduate School of Engineering, Kobe University
[2] 神戸大学 数理・データサイエンスセンター
Center for Mathematical and Data Sciences, Kobe University
[3] 国立研究開発法人 情報通信研究機構
National Institute of Information and Communications Technology
[4] 株式会社 クルウィット
clwit inc.
[a)] ozawasei@kobe-u.ac.jp

送らせる攻撃である．また，こういったDDoS攻撃の代行サービスを行う業者も存在し手軽に攻撃を行えることもあり，攻撃回数は年々増えている．

DDoS攻撃は攻撃対象に短時間で大量のパケットを送り付けて処理不能に陥らせるため，DDoS攻撃を早期に検知し対策を行うことが重要である．しかし，攻撃対象からすると送られてきたパケットが正常なものなのか，DDoS攻撃によるものなのかの区別は容易ではない．また，攻撃者は自身の身元が特定されるのを防ぐために送信元IPアドレスを詐称して攻撃を行うため，その特定も困難である．DDoS攻撃の検知について，ダークネット観測に関する研究が行われている．ダークネットとは，インターネット上で到達可能なIPアドレスのうち特定のホストやサーバが割り当てられていないIPアドレス空間である．ダークネットは未使用なアドレス空間であるにもかかわらず，相当数のトラフィックが観測される．それらのほとんどが脆弱性探索を目的とするスキャンや，DDoSバックスキャッタである．DDoS攻撃を受けたサーバは送られてきた全てのパケットに対し応答パケットを返すが，そのパケットは詐称されたIPアドレスに対して送られる．詐称されたIPアドレスが未割当のIPアドレスの場合ダークネットで観測される．このDDoS攻撃による応答パケットがバックスキャッタである．ダークネットではリフレクター攻撃のような応答パケットの存在しないDDoS攻撃は検知できない．しかし応答パケットが存在する攻撃であれば可能である．また，ダークネットに届くトラフィックの多くは不正な振る舞いなため，ライブネットでDDoSを検知するよりも容易である．更に，ダークネットに届くパケットデータを短時間で分析しバックスキャッタであると判断できれば，DDoS攻撃によってサーバが処理不能に陥る前に対策を講じることができると考える．ダークネット観測を行うもう一つメリットとして，広域な観測ができることである．ダークネットは世界中からパケットが届くため，DDoS攻撃のグローバルなモニタリングシステムとしても利用できる．

古谷ら [2] や宇川ら [3] は，ダークネットで観測されたパケットデータのうちそれぞれTCP通信，UDP通信によるデータを短時間のホスト毎に分割し，それらの統計量を特徴としDDoSバックスキャッタを判定する手法を提案している．それによってTCP通信によるバックスキャッタの判定精度は年間を通して0.98を超えるF値が得られたが，UDP通信によるバックスキャッタの判定精度は0.89に留まっている．UDP通信によるバックスキャッタの数がTCPと比べても少なく，少し誤判別しただけで精度が大きく変わってしまうこともあるが，既存の特徴だけではUDP通信によるバックスキャッタを判定するのに十分でないと考える．そこで著者らは，UDP通信によるバックスキャッタの判定精度を上げるための手法を提案する．既存手法において，機械が外れ値とみなしたり，判別の信頼度が低かったり，誤判別した際などに，監視者が目視でバックスキャッタか否かを判別するためにトラフィックの振る舞いを可視化した画像の作成を行っていた．本研究では，その可視化画像を畳み込みニューラルネットワーク(CNN)[4] に学習させ得られた埋め込み特徴と，既存研究で用いられていた17次元の特徴ベクトルを組み合わせてバックスキャッタの判定を行う．この手法によってより高精度でDDoSバックスキャッタの判定が行えることを示す．

第2章で既存手法の説明，第3章で提案手法の説明，第4章で性能評価実験と実験結果の考察を行う．第5章で結論を述べる．

2. トラフィック特徴量に基づいたDDoSバックスキャッタ判定

宇川らが行っていたDDoSバックスキャッタ判定手法について説明する．宇川らはダークネットトラフィックから特徴ベクトルを作成し，それをSVM[5] やone-class SVM[6] などの分類器に学習させ判定を行った．図2に示すように，特徴ベクトルは，ホストごとに20パケット以上のパケットが送信されたときに作成され，短期から長期にわたって多数のパケットが送信される様々な攻撃に対応するため，観測期間を30秒から30秒間隔で3600秒まで変化させる方法を提案している．3600秒以内に20個以上のパケットが観測されなかった場合，DDoSバックスキャッタではないと判断し，そのホストでは特徴ベクトルは作成されない．また，特徴ベクトルが作成されたホストでも，1時間パケットが観測されずその後新たにパケットが観測された場合，新たな攻撃が行われていると考えパケットの抽出を開始する．本研究では特定のポートを用いる攻撃だけでなく，未知の攻撃にも対応できることが目標なため，ポート番号を直接特徴には用いない．特徴には送信先や送信元のポート，送信先IPの統計情報やペイロードに関する情報を用いる．また特徴の値による偏りを無くすためにすべての特徴を最大値1，最小値0となるように正規化を行う．これらの工程で特徴ベクトルを作成する．特徴ベクトルは17個の特徴量で構成されている．それらを以下に示す．

パケット総数，パケットの観測時間間隔

- パケット総数
- パケットの観測時間間隔の差分の平均・分散

プロトコルの種類，送信元のポートに関する統計量

- プロトコルの種類の総数
- 送信元のポート番号の総数
- 送信元のポート番号ごとのパケット数の平均・分散

送信先のIPアドレス，送信先のポートに関する統計量

- 送信先のIPアドレスの総数
- 送信先のIPアドレスごとのパケット数の平均・分散

特徴ベクトルの作成

Input：ダークネットトラフィック
1：トラフィックをホスト毎に分割
2：function：make_vector(ホスト毎のトラフィック)
3：for 30 秒から 3600 秒まで 30 秒ずつ
4：　if パケット数 ＞20
5：　　return：特徴ベクトル作成
6：　　break
7：　else continue
8：function end
9：
10：for i = 1：ユニークなホスト数
11：　make_vector(i 番目のホストのトラフィック)
12：
13：作成された特徴ベクトルのホストを収集
14：(特徴ベクトルとして抽出されたパケットは削除)
15：for j = 1：作成された特徴ベクトル数
16：　if 1 時間パケットが観測されなかった
17：　　make_vector(j 番目のホストのトラフィック)

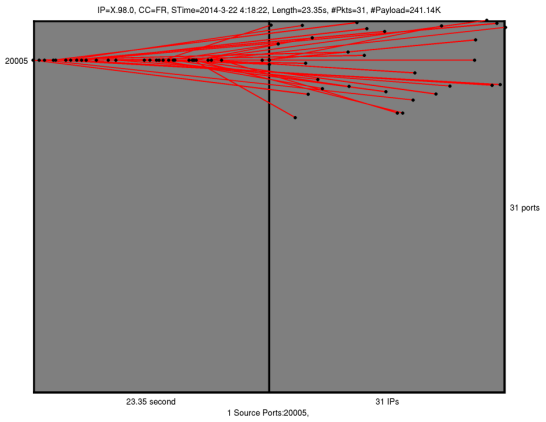

図 1　トラフィックの振舞い可視化画像

表 1　評価データの内訳

	バックスキャッタ	非バックスキャッタ	合計
訓練データ	5,269	5,924	11193
評価データ	1,975	4,5037	47012

- 送信先の IP アドレスの差分の平均・分散
- 送信先のポート番号の総数
- 送信先のポート番号ごとのパケット数の平均・分散

ペイロードの統計量

- ペイロードサイズの平均・分散

3. 提案手法

3.1 トラフィックの振る舞い可視化画像

本研究では，既存研究で定義された 17 個の特徴ベクトルだけでなく，CNN を用いた特徴量抽出も行い，これらを組み合わせて DDoS バックスキャッタの判定を行う．CNN には，ダークネットトラフィックの特徴を可視化した画像 [7] を入力として用いる．

可視化例を図 1 に示す．左半分の縦軸が送信元のポート番号，横軸が観測時間，右半分の縦軸が送信先のポート番号，横軸が送信先の IP アドレスを表す．この可視化画像は特徴ベクトル 1 つごとに作られる．左半面の点から右半面の点を結ぶ線分は，送信元から送信先への 1 パケットの送信を表しており，一定期間に特定の送信元から送信されたパケットの振る舞い (トラフィックパターン) を画像で確認できる．CNN に入力する際は周りの情報はトリミングを行って省き，画像部分のみを入力としている．

既存研究では，この可視化画像をシステムが外れ値と判定したものや判定信頼度が低いものに対し，監視者が直接目視で判定するために使われていた．CNN は生物の視覚野の構造からヒントを得た画像認識のための人工ニューラルネットモデルであり，セキュリティ監視者がトラフィック可視化画像から得ている DDoS バックスキャッタの特徴を捉えらえる可能性があり，本研究では，これを特徴量として追加する．

3.2 ネットワーク構成

本研究で提案する手法のネットワーク構成を図 2 に示す．まず，可視化画像を入力として 16 層の CNN による畳み込みを行う．畳み込みを終え全結合層で 126 次元の特徴ベクトルに変換する．その後，17 次元の特徴ベクトルを加えた 143 次元特徴ベクトルを 3 層の全結合層へ入力し，出力層でバックスキャッタか非バックスキャッタかの判定を行う．

4. 評価実験

4.1 評価データ

実験では NICT のダークネットセンサで観測された 2016 年 4 月 1 日から 6 月 7 日までの期間のパケットデータを用いる．これらのパケットデータから特徴ベクトル，トラフィックの振舞い可視化画像を作成し，実験を行う．訓練データ数，評価データ数を表 1 に示す．今回用いているデータはバックスキャッタの数が非バックスキャッタの数と比べて非常に少ない不均衡なデータである．よって訓練データにおいて，同じ日数のデータを用いるとバックスキャッタの数が非常に少なくなるため，バックスキャッタのデータだけ 4 月 1 日から 5 月 31 日までのデータを使用し，非バックスキャッタは 6 月 1 日のデータのみを使用している．また，すべての評価データに対して，どの程度正しい判定ができたかを示す識別率では正確に評価できない (すべてのデータを非バックスキャッタと判定しても識別率は 95.8%であるが，本来の目的に沿わない)．よって本研究では，Precision，Recall，F 値を評価指標として用いる．

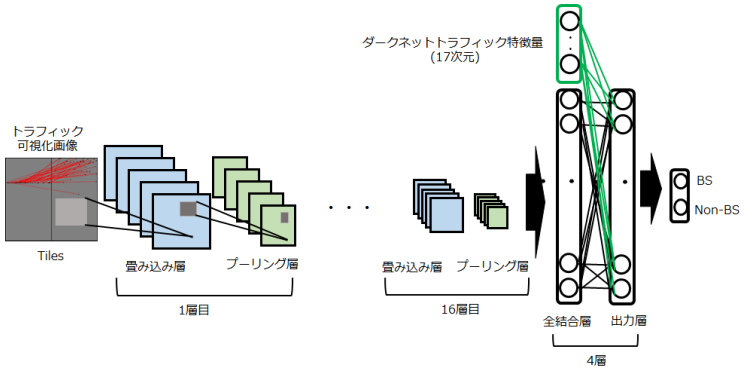

図 2　提案モデルの概略図

表 2　PR 曲線の AUC

	Precision	Recall	F 値
既存手法	0.916	0.877	0.892
CNN のみ	0.844	0.998	0.903
提案手法	0.986	0.983	0.984

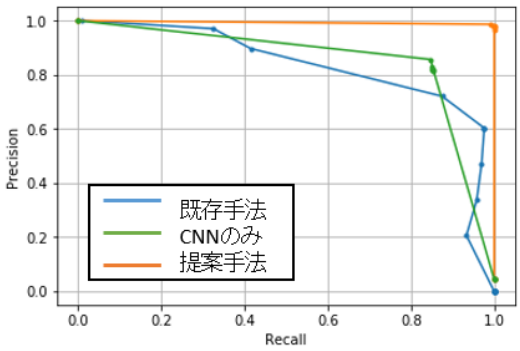

図 3　PR 曲線

4.2　実験結果

表 2 に，6 月 2 日から 6 月 7 日までの期間で実験を行った結果を示す．比較として既存手法，トラフィックの可視化画像を CNN で学習させただけの手法，提案手法での結果を示す．ただし，既存手法は宇川らが実験を行った 1 年間の平均精度である．また，図 3 に 3 つの手法それぞれの PR (Precision-Recall) 曲線を示す．表 2 より，提案手法は既存手法と比べて全ての項目で精度が高いことがわかる．CNN のみの手法と比べて，少し Recall が下がっているが，Precision が高く，F 値も高いことがわかる．表 2 からも提案手法における AUC が一番大きく UDP のバックスキャッタを判定するのに提案手法が有効であるといえる．

4.3　考察

表 2 より，提案した CNN モデルにおいて，可視化画像のみを入力とした CNN や 17 個のダークネットトラフィック特徴量を用いた既存手法に比べて，高い精度が得られた．既存手法の Recall は低く，DDoS バックスキャッタを多く見逃している．誤判定した例として，可視化画像を見ると同じような振る舞いをしていても，パケット送信の間隔やペイロードサイズ等が違うと間違って判定していた (図 4)．これに対し，画像情報をつけ加えることでバックスキャッタと正しく判定できたと考える．また，可視化画像を CNN に学習させて判定を行っただけの結果よりも提案手法のほうが精度がよい．CNN のみで誤判定を起こしていた例を図 5，図 6 に示す．これらは，両方とも既存手法においては正しく判定できていた．図 5 では送信元ポート，送信先ポートともに複数になっている画像に対して誤判定を起こしていた．このようなデータは 1 つ 1 つ振る舞いが大きく異なるため誤判定したと考えられる．また，図 5 では，複数ポートから送信されており，バックスキャッタではない例をバックスキャッタと判定している　．これは可視化画像の送信先ポートの軸が正規化されてプロットされているためにポートが近いと同じ点にプロットされてしまい，送信先のポート数の情報が欠如してしまっているからだと考える．その他にも，可視化画像は一度に複数のパケットが送られてきたときに同じ場所にプロットするためパケット数の情報が欠如したり，画像情報だけではペイロード等の情報も欠如しており誤判定の原因となっていると考える．よって提案手法では，17 個の特徴，可視化画像それぞれに不足している特徴を互いに補完することで精度が大きく向上したといえる．しかし，提案手法において図 6 のような CNN のみのときに多く誤判定していたデータが大幅に削減できたが，依然としていくつか誤判定している．送信元ポート数が比較的少ないものは画像の特徴に大きく影響されているからだと考える．よって今後は送信元ポートが近い値だとしても視覚的に複数のポートがあると判定できるように軸の取り方を変える必要があると考える．

5.　結論

本研究では，ダークネットで観測された UDP 通信によるトラフィックから DDoS バックスキャッタを判定する改善手法を提案した．既存手法で用いられていた 17 個の統計的特徴量だけでなく，監視者がバックスキャッタかどうかの判断を下すために用いていたトラフィックの可視化画

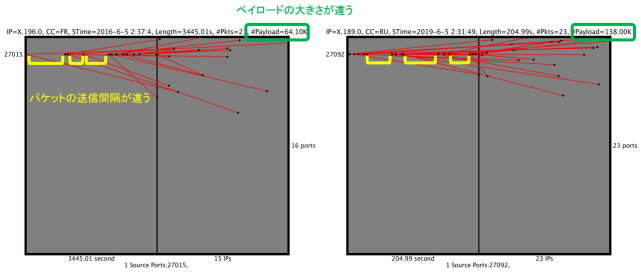

図 4　既存手法による誤判定例．左図は非バックスキャッタと誤判定した．

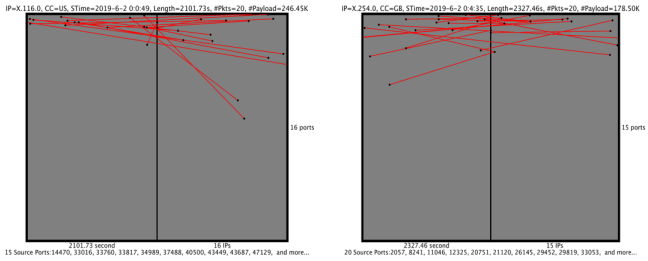

図 5　CNN のみのモデルで誤判定した例

像を入力とした CNN による埋め込み特徴を組み合わせて判定を行った．この提案により 17 個の特徴量，CNN による特徴だけでは不足する情報を互いが補完しあうことにより精度が大きく向上した．

評価実験では NICT のダークネットセンサで観測されたパケットデータのうち 2016 年 4 月 1 日から 2016 年 6 月 7 日までの期間のデータを用いて評価を行った．実験の結果提案手法では 0.98 を超える F 値が得られ，比較手法と比べても精度が大きく向上した．また，今回の実験で得られた F 値は古谷らが TCP において評価実験で示した 0.98 と変わらない結果が得られた．よって提案手法は UDP 通信における DDoS バックスキャッタの判定手法として有効であると考えられる．

今回の実験では送信先のポートが近い画像に対して誤判定を起こしていた．よって送信ポートが近いパケットの振る舞いだとしても視覚的に判断できるように軸の取り方を変える必要がある．また，今回は約 1 万 1 千個のデータを用いて学習を行ったが，より多くのデータを学習させたり可視化画像を上下反転させるなどして，より精度を上げたり，新たな攻撃にも対応できるようにしたい．さらに今回の評価は 6 月 2 日から 6 月 7 日までの期間でしか行っていないため，今後はさらに評価期間を延ばして実験を行い有効性を示していきたい．

参考文献

[1] Seiichi Ozawa,Tao Ban, Naoki Hashimoto, Junji Nakazato, Jumpei Shimamura, “A study of IoT malware activities using association rule learning for darknet sensor data,” *International Journal of Information Security*, pp. 1-10, 2019.

[2] Nobuaki Furutani, Jun Kitazono, Seiichi Ozawa, Tao Ban, Junji Nakazato, Jumpei Shimamura, “Adaptive DDoS-Event Detection from Big Darknet Traffic Data,” *Neural Information Processing*, vol. 9492, LNCS, Springer, pp 376-383, 2015.

[3] 宇川雄樹，北園　淳，小澤誠一，班　涛，中里純二，島村隼平，“ダークネットトラフィック解析による学習型 DDoS バックスキャッタ検出システム，” 信学技報, vol. 115, no. 488, ICSS2015-67, pp. 123-128, 2016.

[4] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *NIPS’12 Proceedings of the 25th International Conference on Neural Information Processing Systems*, vol. 1, pp. 1097-1105, 2012.

[5] Shigeo Abe, *Support Vector Machines for Pattern Classi cation*, Springer-Verlag, London, UK, 2010.

[6] L. M. Manevitz, M. Yousef, “One-class SVMs for document classification,” *Journal of Machine Learning Research*, vol. 9, pp. 2579-2605, 2008.

[7] 井上大介，“サイバーセキュリティの可視化技術，”　可視化情報, vol.36, No.141, 2016.

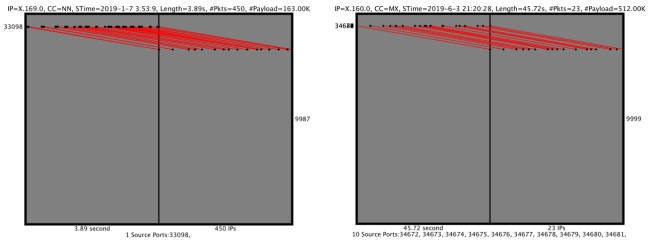

図 6 画像だけで判定が難しい例．左図がバックスキャッタで右図が非バックスキャッタ．