80/TCP ポートへの攻撃の時間的変化

沖野　浩二†　　片山　昌樹‡　　占部 優希§

†富山大学
総合情報基盤センター
okino@itc.u-toyama.ac.jp

‡ナビプラス株式会社
masaki.katayama@naviplus.co.jp

§有限会社マギシステム
urabe@forensic.jp

あらまし　筆者らは、DNS 登録や Web からのリンクがないサーバをハニーポットとして運用している。ハニーポットに蓄積される情報は、その観測環境（応答や履歴）に影響を受けることが知られており、このサーバへのパケットを取得することで実際の攻撃パケットの収集を実施している。

本論文では、このハニーポットに対する 80/TCP ポートへの攻撃に着目し、2014 年 09 月に公表された GNU bash の脆弱性 (CVE-2014-6271 等) の攻撃がどのように行われたか分析する。さらに、攻撃元 IP アドレスの国情報、AS 情報等を用い、それらの攻撃のパターンの分析を行う。

Temporal change in the attack on 80/TCP port

Koji Okino†　　Masaki Katayama‡　　Yuki Urabe§

†Infomation Technology Center, University of Toyama.
3190 Gofuku,Toyama,Toayama, Japan.
okino@itc.u-toyama.ac.jp

‡NaviPlus Co., Ltd.
Ebisu-nishi 2-20-3 daikanyama-CA-building, Shibuya-ku, Tokyo, Japan.
masaki.katayama@naviplus.co.jp

§Magisystem Co., Ltd.
1-13-1 Nihonbashi-Muromachi, Chuo-ku, Tokyo, Japan.
urabe@forensic.jp

Abstract　We have been running servers which have no URL link and no DNS registration. And those servers are a honeypot. It is known that stored information in the honeypot is affected by the observation environment (response or history) . Actually we operate a server as a honeypot, and we can obtain a real attack packet.

In this paper, we analyze 80/TCP port packet capturing by this honeypot, focusing on vulnerabilities GNU bash on Sept. 2014 (CVE-2014-6271, etc.). In addition, countries and BGP AS information, etc. from the attackers source IP address, we analyze of the pattern of those attacks.

1　はじめに

サイバー攻撃の高度化に伴い、攻撃元の詳細な分析が求められている。特に攻撃の動機や技術の多様化によって、すべての攻撃を事前に防ぐことは難しく、実際のインシデント対応の点から考えると、攻撃を的確に把握する技術や攻撃の前兆を捉えて防御方法を細かく変更する必要がある。

そのためには、攻撃者が行う攻撃にどのような特徴があるのかを分析するだけでなく、その攻撃の前後に行われた通信も含めて解析を行うことにより、事前情報によりその攻撃パターンにどのような差が生じるか、また、攻撃後はどのような行動を行うかを分析する。加えて、攻撃元の IP 情報を解析し、国や AS 単位で攻撃の特徴を抽出するための手法を提案する。

従来、攻撃手法や攻撃元の個別ホストについての詳細な分析は多く行われてきたが、攻撃主体がどのような情報に基づき攻撃を行っているのかを、国や AS の属性に基づいてグループ化し、それらのグループから特徴抽出などを行う研究は少ない。

現況ではサイバー戦争の脅威の顕在化などが議論されるにつれ、元来は不明な点が多かった攻撃元の集団の解析の重要が増している。本論文では、攻撃を含めた 8 か月間に取得できたハニーポットのデータを分析することにより、攻撃の前後でどのようなパターンが生じるかを検討する。

さらに、攻撃元集団の解析を行うため、観測された PCAP データに対して、国と BGP4 の AS の属性を付与し、ポート番号などの頻度別集計可能な項目でデータ処理を行い、PCAP 分析に新たな観点を提供する。

2　データ取得環境

2.1　ハニーポット

ハニーポットとしてデータを取得するために、Ubuntu14.04Desktop を準備し以下のサービスを起動した。このサーバへの通信を PCAP による取得し、分析データとする。

- Apache(HTTP)
- Samba
- SSH
- BIND

ハニーポットが応答を行うと、その応答内容により攻撃パターンが変更することが知られている [1]。本実験環境では、実際のサーバをおとりとして利用することにより現実の攻撃を観測することできる。また、観測環境には、FW や NAT の設置はなく、ポート変換やポート遮断はないので、攻撃者がこのサーバがハニーポットであることを確認するコストは、限りなく高くなっている。

2.2　比較データ

処理対象データについては、表 1 の 2 つのハニーポットを利用した。

表 1: ハニーポット一覧

ホスト名	目的
FTP_O	過去 FTP 利用 IP
New_A	別セグメント A 設置

FTP_O は過去の FTP サーバとして利用していたアドレスであり、New_A は別セグメントに設置したアドレスとなっている。これにより、攻撃条件に差が生じた場合、攻撃者が意図を持って行っている可能性を示していると考えられる。

2.3　取得データ

2014 年 8 月から 2015 年 3 月までの 8 カ月で取得できたフレームの変化は、次の図 1 の通りである。

8 カ月の間に観測されたフレーム数は、FTP_O が 28,632,216、New_A が 20,774,273 であり、一日平均およそ 11 万フレームと 8.5 万フレームとなっている。

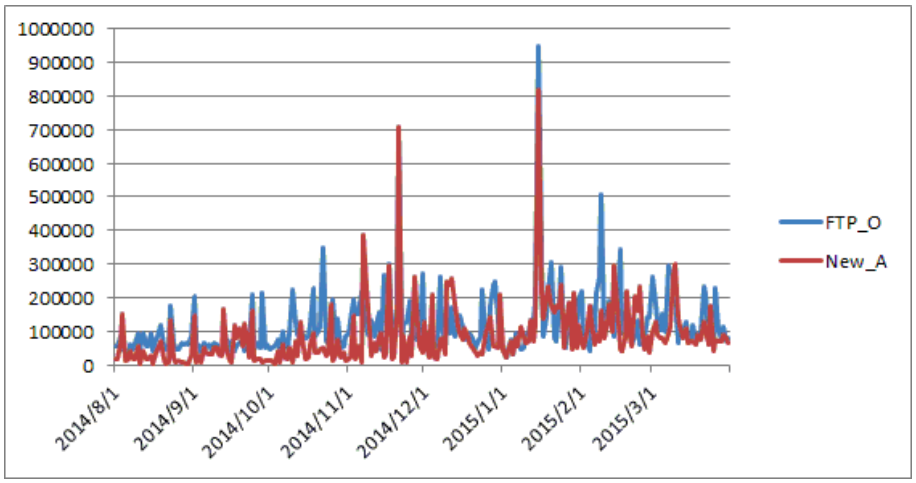

図 1: 取得フレーム数の変化

フレーム数は、FTP_O は、New_A の 1.3 倍のアクセスとなっており、過去にサービスを提供している IP への通信は多くなっている。

表 2: 主な攻撃先ポートフレーム数

	FTP_O	New_A
ICMP	708,018	21,571
21/TCP	286,672	626
22/TCP	22,113,950	20,156,414
23/TCP	9,416	8,966
53/UDP	2,846	2,796
80/TCP	5,012,561	33,100
137/UDP	308,542	414,869
445/TCP	45,987	63,559
1433/TCP	4,065	2,488
3389/TCP	2,986	2,617

表 2 に、主なポートへの攻撃フレーム数を示す。それぞれの攻撃の状況は、22（SSH) ポートに関しては、どちらのホストもほぼ同じ攻撃回数である。FTP_O に関するアクセスでは、80(HTTP), 21(FTP) は、過去のサービスに応じた攻撃となる。SSH への攻撃を除けば、攻撃先ポートに対しては、サービスを行っていたポートに関してのアクセスが頻発していることが見受けられている。

攻撃をしてきた IP の時間的変化を図 5 と 6 に示す。

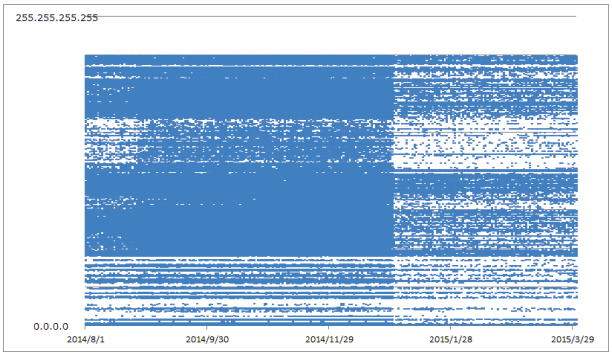

図 2: FTP_O への攻撃 IP の時間分布

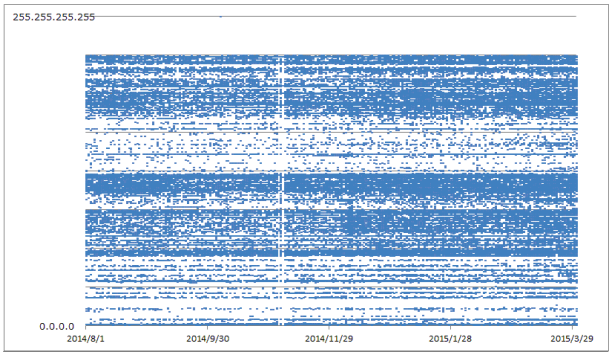

図 3: New_A への攻撃 IP の時間分布

攻撃は、絶えず行われているものと、集中的におこなわれるもの、散発的なものに分類される。

3 Shellshock 攻撃分析

2 つのハニーポットに対する Shellshock 攻撃の分析を行う。

3.1 Shellshock 攻撃

Shellshock 攻撃 [2] とは、日本時間 2014 年 9 月 24 日の夜に公開された UNIX 系 OS で利用されている shell の一種である Bash で発見された深刻な脆弱性である。

Shellshock の脆弱性は、全部で 6 種類 (CVE-2014-6271,7169,7186,7187,6277,6278) あるが、そのうち 4 種類が任意のコードが実行可能とされている。これらの脆弱性すべてに Patch があり、これを適応することで、対応することが可能である。

この脆弱性は、Bash の環境変数の取扱いによって引き起こされ、Bash を利用しているサービス、例えば、Apache における mod_cgi など shell 環境を実行している場合や ssh や telnet などリモートアクセスを公開している環境でユーザが実行できるコマンドを制限している場合に制限を回避し、任意のコードを実行することが可能となる。

特に、Web サーバ（80/TCP) に対する攻撃は、公開後すぐに具体的な方法が示され、容易に実行可能であったために、インターネット上で広く観測されることとなった。本論文では、この攻撃データを検討するために、80/TCP ポートへの攻撃を中心に解析を行う。

3.2 80/TCP ポートへの攻撃

図 4 は折れ線が 80/TCP へのフレーム数を示し、棒グラフが Shellshock 攻撃が行われた回数を示している。Shellshock の攻撃は、それぞれ、FTP_O に 1009 回 New_A に 928 回行われていた。本研究のハニーポットへの攻撃は、日本時間　2014 年 9 月 25 日から散発的に観測さ

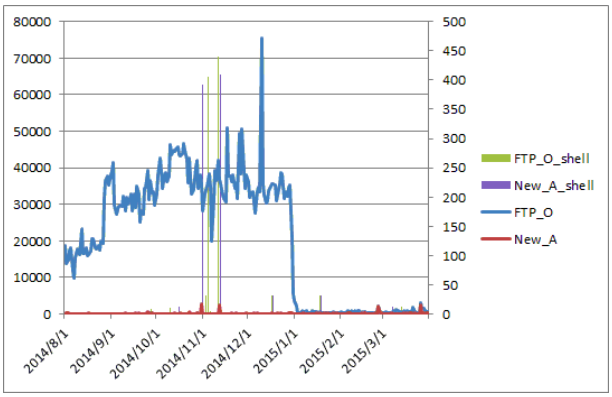

図 4: 80/TCP ポートへのアクセスと Shellshock 攻撃数

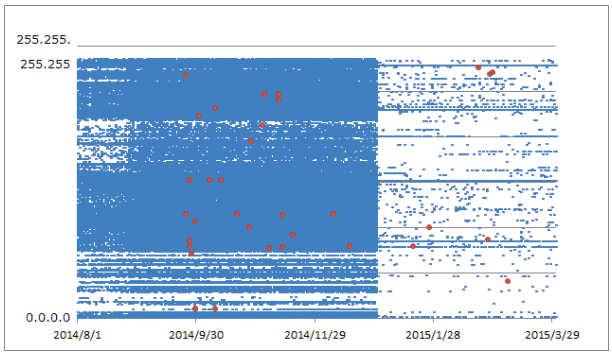

図 5: FTP_O 80/TCP への攻撃 IP の時間分布

れ、11 月前後に最大の攻撃を受けていることが判明した。しかし、11 月前後の最大回数の攻撃の中身を調査すると、cgi-bin 以下のファイルに対して、多量に繰り返してアクセスしており、Apache は 404 を返していた。これは、攻撃を成功させるためにサーバが有していると思われる CGI を実行させようとしたことにより回数が多くなっていることが判明した。

80/TCP へ攻撃をしてきた IP の時間的変化を図 5 と 6 に示す。青い点が 80/TCP へのアクセスした IP を示し、赤い点は Shellshock の攻撃を行った IP を示す。

図からだけでは、攻撃を行ってきた IP アドレスは、分散しており、その規則性を見つけ出すことは難しい。

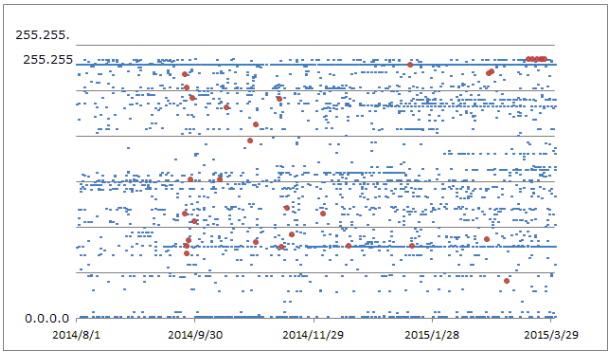

図 6: New_A 80/TCP への攻撃 IP の時間分布

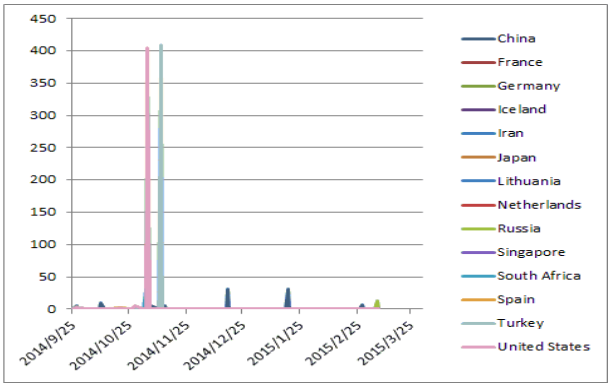

図 7: FTP_O 国別の Shellshock 時間分布

3.3 Shellshock 攻撃の変化

Shellshock 攻撃分析では、攻撃の開始が脆弱性の公開後、すぐに行われたことが判明した。加えて、攻撃は、散発的に行われ、本論文のデータ期間の 3 月末までも継続が確認できた。攻撃に関して、IP アドレスからの規則性を見つけることは難しく、どのような集団が行っているかを検討することは難しいと考えられる。

Payload を分析すると、実際の Shellshock の攻撃コードは、当初は、ping などを利用し、脆弱性の存在を確認するものだったが、時間が進むにつれて、wget 等を利用し、実際にプログラムを download し、実行する攻撃に変化していったことが確認できた。

4 個別解析

4.1 既存解析の問題点

既存の解析では、IP 毎に解析を行っている。しかし、これは攻撃側から考えると、同一 IP で絶えず攻撃を行うことはいろいろな意味で現実的ではなく、攻撃者の IP アドレスは変化している可能性は高いと言える。そこで、攻撃 IP アドレスを国または AS 情報に変更し、分析を行う。これにより組織的な攻撃なのかなど、攻撃体の分類を行うことが可能なのかを検討する。

4.2 国による差

今回のハニーポットの 80/TCP のアクセスを調べると FTP_O には 226 国から、New_A には 175 国からアクセスがあったが、図 3 に示すようにそれぞれ 14 国、10 国が Shellshock 攻撃を行った国であった。国別に考えた場合、実際に攻撃コードを送信する国は少ないことがわかる。

表 3: 国別攻撃数

	FTP_O	NEW_A
China	95	95
France	31	1
Germany	1	1
Iceland	1	1
Iran	1	0
Japan	1	1
Lithuania	1	0
Netherlands	2	3
Russia	13	13
Singapore	1	1
South Africa	31	0
Spain	2	0
Turkey	409	409
United States	421	404
ToTal	1010	929

また、国別の攻撃の時間分布を図 7,8 に示す。

攻撃を繰り返し行っている国と一時期に集中的に攻撃を行っている国に分けられることが分かる。

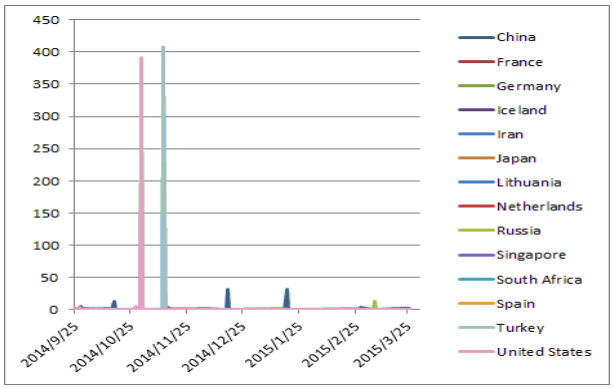

図 8: New_A 国別の Shellshock 時間分布

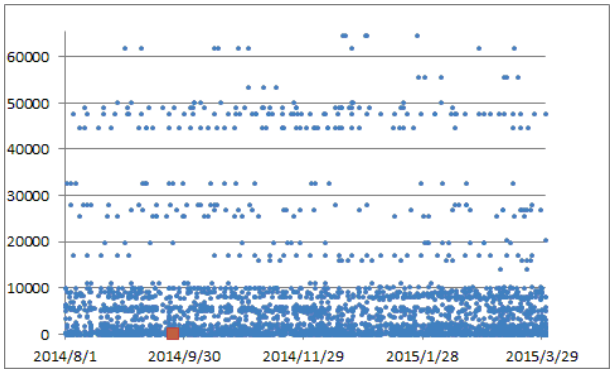

図 9: FTP_O への 1st AS からの Port/時間分布

4.3 AS による分析

データを集計し、AS 毎の攻撃ポートの時間変化を確認する。

まずは、本ハニーポットに最初 (図 9) および 2 回目 (図 10) に到着した AS のデータを表示する。横軸は時間で、縦軸は攻撃をしている Port（ハニーポット側サービス）を示す。

どちらの AS からの攻撃も、FTP_O と NEW_A とほぼ同様な形跡であったため、FTP_O のデータを提示する。初回に到着した AS はその前後でも大量な攻撃を行っており、組織的にデータを収集している可能性があると考えられる。また 2 回目に到着した AS は、その前後で少しの通信が存在していることが確認できる。

また、多量の攻撃をしてきたある AS からの攻撃を図 11 に示す。この AS も両方のハニーポットに同様な攻撃であったために、FTP_O への攻撃分布を示す。この AS からは、ほぼ毎日一定ポートへのアクセスを観測している。

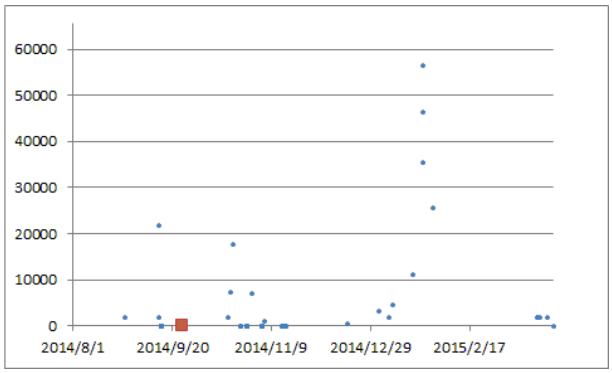

図 10: FTP_O への 2nd AS からの Port/時間分布

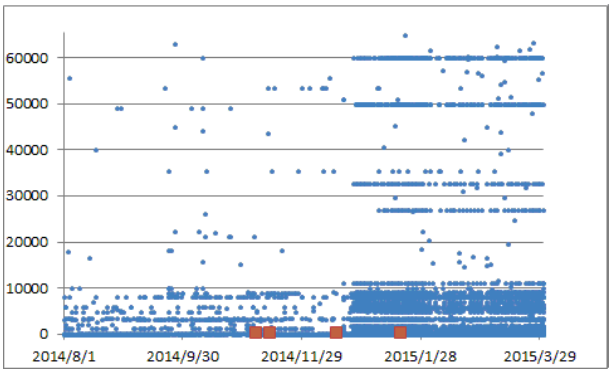

図 11: 多量の攻撃があった AS からの Port/時間分布

4.4 Shellshock の攻撃の差

パケットの Payload 部分を解析して文字列を抽出して程度判定を実施し、green、yellow、red の 3 つにカテゴライズした。

分類する基準は以下の通りである。

geen echo だけの実行と判断したもの

yellow ping を実行していると判断したもの（echo との複合含む）

red リモートから制御を取得することを目的とした挙動であると判断したもの

これにより、攻撃者の目的が、調査目的なのか、実際の攻撃なのかを判別する。

攻撃の分布を AS 毎に整理し、その攻撃の時間変化を図 12 に示す。

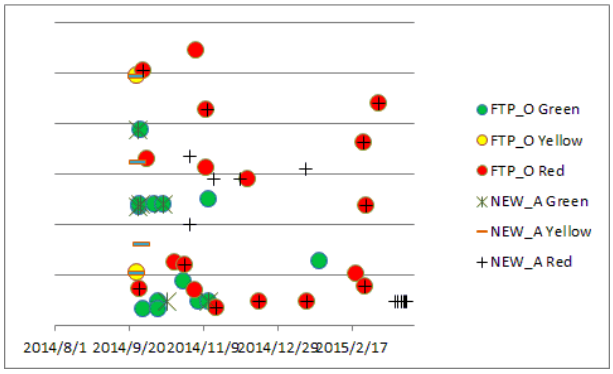

図 12: 攻撃 AS からの目的判定分布

多くの攻撃は、両方のハニーポットに対して同時期に行われていることが判断できる。解析の結果、今回の攻撃は、ランダムな IP アドレスへの攻撃ではなく、多くの場合には、総当たりによる攻撃であることが示されている。

4.5 個別解析のまとめ

既存解析と比較し、国別では、FTP_O のみに攻撃を行っているものがあることが分かり、より詳細な解析ができる AS 毎の解析では、その通信パターンに特徴があることが示されている。

個別解析の結果、実際の攻撃を行ってくる AS の多くに、事前および事後に調査を行っており、また、その調査対象は幅広い Port に対して行っていることが判別した。

5 関連研究

Honeypot をはじめとする deception system の研究の歴史は長く、システム仮想化を用いたものからネットワークトラフィックの大規模な処理まで多岐にわたる。攻撃トラフィックや関連情報の分類や特徴抽出は [3][4] で行われている。[5] は、ハニーポットの運用から一歩進めて、抽出した特徴をもとに攻撃のシグニチャを新たに生成する手法を提案している。インターネットの Darknet や Background radiation、マルウェアの可視化の研究には [6][7][8] がある。攻撃トラフィックを大規模データをとして捉え分析を行ったものに [9] がある。

攻撃データの IP アドレスを地理情報に変化した上で、SOM を適応し、攻撃者を分析したものに [10] や、地理情報だけでなく AS 情報と組み合わせ、加えて、条件の異なる複数のハニーポットと比較することで、攻撃者の分析 [11] を行っているがある。

また、最近盛んに研究されているネットワーク観測項目として、DNS がある。[13] は、DNS への悪意のある行為への早期発見と対策のための観測手法を提案している。

このようにパケットを蓄積し、分析することで、攻撃者の動向を把握し、対策を行う研究が進んでいる現状がある。

6 まとめと今後の課題

サイバー攻撃の高度化に伴い、攻撃元の詳細な分析が求められている。特に攻撃の動機や技術の多様化によって、防御方法を細かく変更する必要がある。本論文では、80/TCP へ Shellshock 攻撃に着目し、 2 台のハニーポットにおいて観測されたデータを調査することにより、攻撃者がどのような攻撃を行ってきたかを、観測データの攻撃元 IP アドレスではなく、攻撃元の AS 番号に変換することにより、管理組織ごとの攻撃パターンの特徴として個別解析するための手法を検証した。

今回の個別解析では、AS によりパターンを分析する手法を適応した結果、一部の攻撃者が定常的に Port 調査および脆弱性確認を行っていることが確認できた。

従来、攻撃手法や攻撃元の個別ホストについての詳細な分析は多く行われてきたが、実際に行われた攻撃を起点として、その前後の通信に着目し、攻撃者が事前、事後に情報をどのくらい利用しているのかを検討している研究は少ない。

元来は不明な点が多かった攻撃元解析の重要性が増しているなか、本論文では、実際に攻撃が行われたハニーポットデータを利用することで、攻撃者がどのような攻撃を行っているかを、取得された PCAP データに AS の属性を付与し、ポート番号などの頻度別集計可能な項目でデータ処理を行い、攻撃分析に新たな観点を提

供した。

今後の課題としては、観測点を増やすとともに、長期的な観測データから、複数の解析結果を行い、時系列的な推移から含意のある結論を引き出すことである。また、今回は80/TCPでの集計という比較的単純な処理をおこなったが、観測データに対して他ポートへの攻撃解析やグループ間のクラスタリング解析などを行うことにより、詳細な分析を行うことを検討している。

攻撃データの解析や調査を行うことで、攻撃元のグループの解析に別観点からの知見を得ることなどができると想定される。また、関連研究で議論したダークネットの観測結果とあわせることで、早期対策や、防御側のフィルタリングや動的構成の粒度を高度化するための情報が得られる可能性がある。

参考文献

[1] 横田凌一, 大久保諒, 曽根直人, 森井昌克,” ダークネット観測に対してハニーポットが与える影響 (その 2)”, 信学技報 113(43), 97-100, 2013-05-16, 電子情報通信学会 ,2013.

[2] 上田隆一,”Shellshoch の顛末書”, 情報処理 Vol.55 No.12, p1320-1323, 情報処理学会 ,2014.

[3] An internet protocol address clustering algorithm, Robert Beverly, Karen Sollins, in Proc. SysML'08 Proceedings of the Third conference on Tackling computer systems problems with machine learning techniques, 2008.

[4] Honeycomb - Creating Intrusion Detection Signatures Using Honeypots, Christian Kreibich, and Jon Crowcroft. Proceedings of the Second Workshop on Hot Topics in Networks Hotnets II,2007.

[5] J. M. Agosta, Carlos Diuk, Jaideep Chandrashekar and Carl Livadas, An Adaptive Anomaly Detector For Worm Detection, Workshop on Tackling Computer Systems Problems with Machine Learning Techniques (sysML-07) 2007

[6] Characteristics of Internet Background Radiation, Ruoming Pang, Vinod Yegneswaran, Paul Barford, Vern Paxson, and Larry Peterson Appeared in IMC 2004, Taormina, Sicily, Italy, October 2004

[7] nicter: An Incident Analysis System Toward Binding Network Monitoring with Malware Analysis, Inoue, D. Eto, M. ; Yoshioka, K. ; Baba, S. ; Suzuki, K. ; Nakazato, J. ; Ohtaka, K. ; Nakao, K, WISTDCS '08 Proceedings of the 2008 WOMBAT Workshop on Information Security Threats Data Collection and Sharing

[8] Wei Zhuo, Yacin Nadji ”MalwareVis: Entity-based Visualization of Malware Network Traces” Symposium on Visualization for Cyber Security (VizSec) 2012

[9] Guofei Gu, Roberto Perdisci, Junjie Zhang, Wenke Lee. ”BotMiner: Clustering Analysis of Network Traffic for Protocol- and Structure-Independent Botnet Detection”. USENIX Security Symposium 2008

[10] 沖野 浩二 , 安藤 類央 , 片山 昌樹,” 自己組織化マップを用いたハニーポット送信元地理情報の特徴抽出と分類”,CSS2013 論文集,2013(4),716-722, 情報処理学会,2013.

[11] 沖野 浩二 , 片山 昌樹, 占部 優希”IP アドレスの履歴が攻撃に与える影響に関する考察”,CSS2014 論文集,2014(2),56-63, 情報処理学会,2014.

[12] Shuang Hao, Nick Feamster and Ramakant Pandrangi. Monitoring the Initial DNS Behavior of Malicious Domains. ACM SIGCOMM Internet Measurement Conference. Berlin, Germany. November 2011.