Vol.2017-IOT-39 No.4
Vol.2017-SPT-25 No.4
2017/9/29

UTM トラフィックログに自己組織化マップを用い高度な攻撃を検出する試み

宇井 哲也 † **長谷川 理** ‡ **鈴木 彦文** ‡

Tetsuya Ui Osamu Hasegawa Hikofumi Suzuki

1. 研究背景

近年インターネットの普及に伴い、インターネット上の攻撃は増加、多様化する傾向にある。この巧妙で多様化する攻撃の中で最も多い手口の一つが DoS(Denial of Service)/DDoS(Dstributed DoS) 攻撃である。通常、DoS/DDoS 攻撃パケットは、クライアントへのサービスを提供する HTTP, DNS, SMTP といった well-known ポートと呼ばれるポートに集中しており、攻撃のみならず、サーバを破壊するためのセキュリティホールや踏み台となるような設定手段を探すための手段としても用いられる [1]。

これを防御する手段の一つとして、通信ログを解析し攻撃を検知するといった手法が考えられる。しかし、この手法は明確に攻撃性が認められる通信の取得、大規模な通信ログを取得する環境、個人情報保護などの観点から未発達な部分が多い。そこで、本研究では信州大学のネットワークに対し仮想的な攻撃を行い、信州大学で運用されている UTM(統合脅威管理:Unified Threat Management　以下、UTM) 装置に記録された通信ログに対し自己組織化マップを用いデータマイニングを行った。本論文では対象ネットワークの概略、攻撃、データマイニングの手法、及び結果について述べる。

2. 目的と概要

通常のセキュリティの手段として、ポリシーに基づいた通信の制御を行う FireWall や IPS(侵入防止システム:Intrusion Prevention System)/IDS(侵入検知システム:Intrusion Detection System)、UTM 装置のような装置が運用されている。これらの装置は様々な通信に対して多くの機能を提供しているが、多くの場合、閾値やシグネチャーパターンに基づいた制御である。しかしこの防御方法では、シグネチャーパ ターンの開発や高帯域、高速通信においてはボトルネックになることが非常に多い。また近年これらの装置をすり抜ける攻撃も増加している。

そこで我々の研究では、このような問題を解消するようなセキュリティエンジン、機器の研究開発を行っている。このために、トラフィクログの解析時に教師データとして使用するための攻撃の通信が必要になってくる。そこで、模擬的な攻撃環境のを開発し DNS に対する DDoS 攻撃を仕掛け、それを含むトロフィックログに対し自己組織化マップを用いたデータマイニングを行った。

†信州大学大学院総合理工学研究科工学専攻電子情報システム工学分野
‡信州大学総合情報センター

3. 理論

本章では前知識として、研究環境、攻撃の手法、データマイニングの手法の一つの自己組織化マップの説明を行う。

3.1. 研究環境

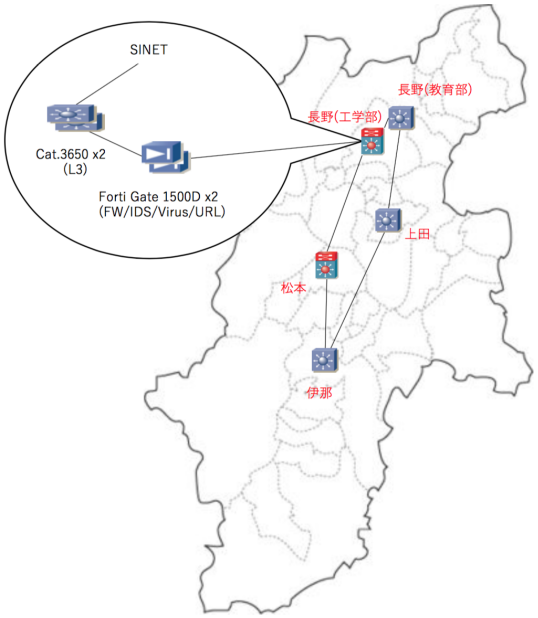

図 1: 信州大学のネットワークの形態

本研究を行うためには大規模なトラフィックログデータを入手する必要がある。しかし現実に運用されている大規模なネットワークのトラフィックログデータの入手は難しい。また通信事業者のような組織は個人情報保護や通信事業法の観点からこのようなトラフィックログデータを解析することは困難である。そこで今回は信州大学全域における通信ログを利用することにした。

信州大学は総合大学であり図 1 からもわかるように長野 (工学部)、長野 (教育学部)、松本、上田、伊那と５つのキャンパスが長野県内の様々な場所に点在しており各キャンパスの通信を接続するため大規模なものとなっている。このため、各キャンパスの通信が長野 (工学部) に集まり、長野 (工学部) にある総合情報センターからインターネットへと接続する構成になっている。ネットワーク利用者の規模は学生、教職員を含め 14,437 人である [1]。これは地方自治体と同等の規模と言える。通信量は 1 日に 2 億セッション以上である。

3.2. 攻撃の手法

本研究にはトラフィックログの解析を行う際に教師データとなる攻撃の通信が必要となる。しかし攻撃と

[1]信州大学：広報・刊行物 大学概要 2017

Vol.2017-IOT-39 No.4
Vol.2017-SPT-25 No.4
2017/9/29

わかっているデータを含む通信を定期的に入手することは困難である。そこで教師データとして利用するための攻撃環境を開発することとした。

[6] によると、近年では 2016 年の 9 月に Mirai と呼ばれるボットネットによる DDoS 攻撃が話題になった。表 1 に示すのは Mirai の攻撃の種類である。今回この中から、DNS に対するリゾルバフラッド攻撃環境の開発を行った。

表 1: Mirai の攻撃の種類 [6]

種類	特性
UDP フラッド	UDP パケットを大量に送り付ける攻撃
VSE フラッド	ゲームエンジン Source Engine」に対する UDP フラッド攻撃
DNS リゾルバフラッド	DNS に存在しないドメインの名前解決要求を送り付ける攻撃
SYN フラッド	SYN パケットを大量に送り付ける攻撃
ACK フラッド	ACK パケットを大量に送り付ける攻撃
TCP STOMP フラッド	PSH パケットと ACK パケットを大量に送り付ける攻撃
GRE IP フラッド	GRE プロトコルでカプセル化されたパケットを大量に送り付ける攻撃
GRE イーサネットフラッド	イーサネットと GRE プロトコルでカプセル化されたパケットを大量に送り付ける攻撃
プレーン UDP フラッド	ヘッダなどを省略して高速化した UDP フラッド攻撃
HTTP フラッド	HTTP リクエストを大量に送り付ける攻撃

攻撃環境を構築する前にまずローカルな環境で DNS サーバの設定と攻撃が正常に行われているのかを確認した。DNS サーバの設定では意図しない外部との通信が行われていないか、特に権威サーバに対して問い合わせをしていないか、また DNS サーバとしての機能が正常であるかの確認をした。図 2 は工学部にある総合情報センター内に設置した攻撃環境の概略である。

図 2 にあるように攻撃の PC は信州大学外部の IP アドレスに設定した。信州大学内部に攻撃 PC を設置することも可能であったが、攻撃は信州大学外部からあるものと仮定し環境の構築を行った。攻撃 PC から送られた通信は、ルーターを通り、SINET ルーターを経由し、信州大学内のルーターを通り UTM を通り信州大学内部に設置された DNS サーバへと送られる。この時 UTM で記録されたすべてのトラフィックログを解析に用いる。

攻撃は DNS に対する DDoS 攻撃を再現した。その中の特に DNS リゾルバフラッド攻撃を再現した。単一 IP アドレスからの攻撃であると UTM の機能で直ぐに攻撃が判別できてしまうため検証としては好ましくない。しかし、DDoS 攻撃に十分と言えるほどの大量の PC を用意することは難しい。そこで UDP パケットの送信元 IP アドレスの部分を偽装するプログラムにすることで一台の PC で模擬的な DDoS 攻撃を再現した。

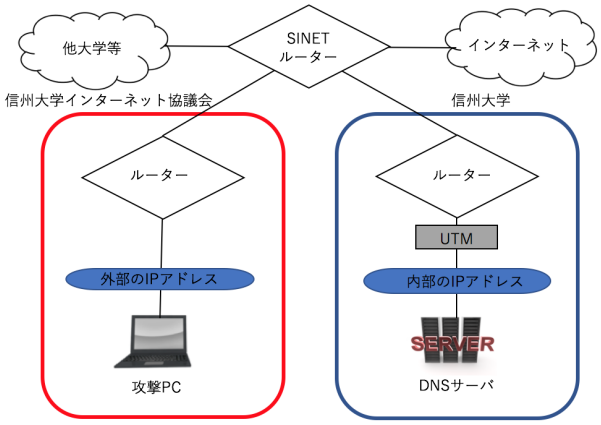

図 2: 信州大学総合情報センターに設置した攻撃環境の概略

3.3. 自己組織化マップ

トラフィクログの解析には自己組織化マップ (Self-Organizing Maps:以下、SOM) を利用した。[2] より SOM は以下のような特徴を示す。SOM は高次元のデータ間に存在する非線形な統計学的関係を、簡単な幾何学的関係を持つ像に変換する。それらは通常は 2 次元のニューロンの格子状に表示されるため、高次元空間の可視化に用いることが可能である。また、これらの多次元のデータを予備知識なしでクラスタリングすることもできるため、高次元の情報を視覚化することができ、データ同士の関係が人間にとって直感的に理解しやすくなる。SOM のこのような特徴を利用しトラフィックログの情報をパラメータとして与え解析を行った。

次に [3] より som の理論を数学的に説明する。図 2 は SOM の構造イメージである。

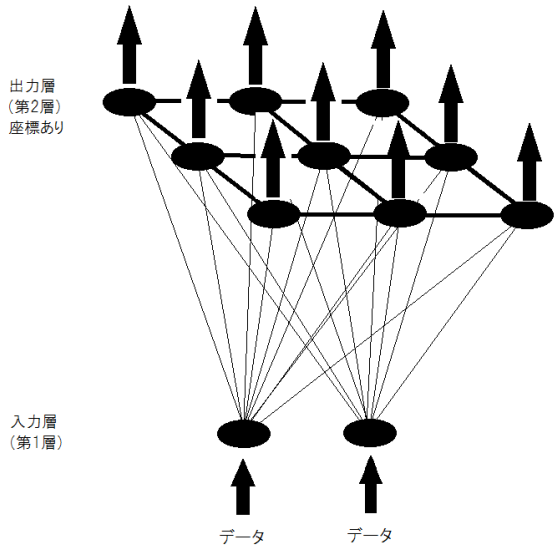

図 3: SOM の構造イメージ

I 次元の多変量データが N 個あるとする。そこから t 番目に取り出した観測変数ベクトルを

$$x(t) = (x_1(t)x_2(t)...x_i(t)...x_I(t))^{'} \quad (1)$$

と表現する、各観測変数は平均 0、分散 1 に標準化されることが多い。ここで i は観測変数の添字であり、1 から I まで動く。i は第 1 層のユニットの添字である。I は観測変数の数である。I は第 1 層のユニットの数でもある。

t はオブザベーションの添字である。しかし t は 1 から N まで動くわけではない。N 個のデータは収束するまで、何度も繰り返し使用される。したがって、通常、t は N よりもずっと大きい値まで動く。このため t は時間の添字とも呼ばれる。

時間 t における第 2 層の j 番目のユニット状態は

$$m_j(t) = (m_1j(t)m_2j(t)...m_ij(t)...m_Ij(t))^{'} \quad (2)$$

と表現される。ここで j は第 2 層のユニットの添字であり、ユニット数は J 個である。$m_ij(t)$ は第 1 層の i 番目のユニットに関する第 2 層の j 番目のユニットの重みである。したがって第 2 層のユニットは、すべての第 1 層のユニットと結合していることになる。

ここで $x(t)$ と J 個の $m_j(t)$ の差のノルムを次々に比較して、その最小値

$$||x(t) - m_c(t)|| = \min_j ||x(t) - m_j(t)|| \quad (3)$$

を求める。ただし最小値を与える j を c と呼ぶ。つまり c 番目のユニットの重みが $x(t)$ に一番似ているということになる。

ユニット c に着目し、時期 t から $t+1$ に向けて、第 2 層ユニットの重みを変更する。ただしこの時ユニット c が最も $x(t)$ の影響を受けるようにしたい。また、その近傍ユニットはその近さに応じて順次 $x(t)$ の影響を受けるようにしたい。そして、c から遠いユニットは $x(t)$ の影響を受けないようにしたい。

この要請を実現するためにユニットの重みに

$$m_j(t+1) = m_j(t) + \alpha(t) * h_cj(t) * (x(t) - m_j(t)) \quad (4)$$

という変更を施す。ここで $\alpha(t)$ と $h_cj(t)$ は 0 以上 1 以下の重みである。$\alpha(t) = 1$,$h_cj(t) = 1$ と置くと、この式は $m_j(t+1) = x(t)$ となり、$m_j(t+1)$ が $x(t)$ の影響を完全に受けることになる。したがって $\alpha(t)$ と $m_j(t+1)$ に対する $x(t)$ の影響の強さを表現した係数である。

$h_j(t)$ はユニット c とユニット j との近さを表現した関数 (遠ければ遠いほど値が 0 に近づく関数) であり

$$h_cj(t) = exp(-||r_c - r_j||^2/2\sigma^2(t)) \quad (5)$$

が用いられる。ここで r_c と r_j は、それぞれ c と j の位置ベクトルである。$r_c = r_j$ のとき、この関数は最大になり、値は 1 となる。

$\sigma^2(t)$ は時間とともに減少する関数である。この値が大きいときには、ユニット c の変化が周囲に波及しやすく、値が小さいときには変化が周囲に波及しにくい。

最初から $\sigma^2(t)$ の値を小さく設定するとマップはすぐに収束する。しかし隣り合ったユニットの類似性がなくなり、全体として無秩序なマップになってしまう。最初は大きめに設定し、時間とともにだんだん小さくすると良い結果が得られることが多い。

(3) の式の $\alpha(t)$ も、$\sigma^2(t)$ と同様に、時間とともに減少する関数である。たとえば

$$\alpha(t) = max(1 - 1/T, 0) \quad (6)$$

が使われる。T は変化が持続して欲しい時間である。1000 回まで重みを変更したい場合は $T = 1000$ とする。

4. 実験

本章では、実験で行った攻撃、攻撃時のネットワークの動作、データマイニングによる解析結果、不審な攻撃の判定について説明する。

4.1. 攻撃内容

攻撃は 2017 年 6/19 から 6/30 にかけて行った。6/19 から 6/23 にかけては攻撃の IP アドレスを 4 つに指定し、徐々に攻撃回数を増やしながら実施した。6/26 から 6/30 にかけては IP アドレスをランダムにし同じく徐々に攻撃回数を増やしながら実施した。表 2 は固定した 4 つの IP アドレス、表 3 は攻撃日時と回数の詳細である。

表 2: 設定した 4 つの詐称 IP

128.64.32.16
168.86.66.91
77.111.222.99
10.2.77.140

表 3: 試験攻撃を行った日時と回数

日付	時間	攻撃回数	IP
2017/06/19	19:28:17	100,000	4
2017/06/20	19:13:38	250,000	4
2017/06/21	19:16:37	500,000	4
2017/06/22	19:09:24	750,000	4
2017/06/23	14:12:05	1000,000	4
2017/06/26	19:09:29	100,000	ランダム
2017/06/27	19:33:31	250,000	ランダム
2017/06/28	19:23:23	500,000	ランダム
2017/06/29	19:19:57	750,000	ランダム
2017/06/30	17:54:20	1000,000	ランダム

4.2. 攻撃結果

6/19 から 6/31 日にかけて攻撃を行った際、IP アドレスを 4 つに設定して 1,000,000 回攻撃を行った 6/23 の結果と、IP アドレスをランダムに設定して攻撃を行った 6/30 の結果を掲載する。

- 6/23 の結果 (IP アドレス 4 つ)

  図 4 は攻撃実験を行った 6/23 の UTM の CPU の変化、図 5 は攻撃実験を行った 6/23 の UTM

の totalsession の変化である。6/23 について攻撃は 14:12:05 に行ったが図 4 を見るとその前後で cpu 使用率が増加していることがわかる。図 5 の totalsession については UTM の性質上 4 つ攻撃 IP がそれぞれ 1 つのセッションとしてまとめられてしまうため上昇はしていない。

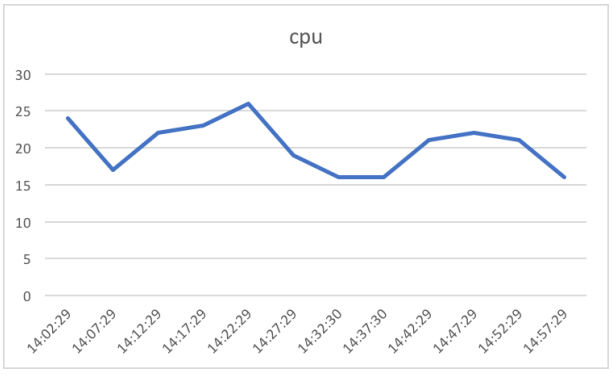

図 4: 6/23 の UTM の CPU の変化

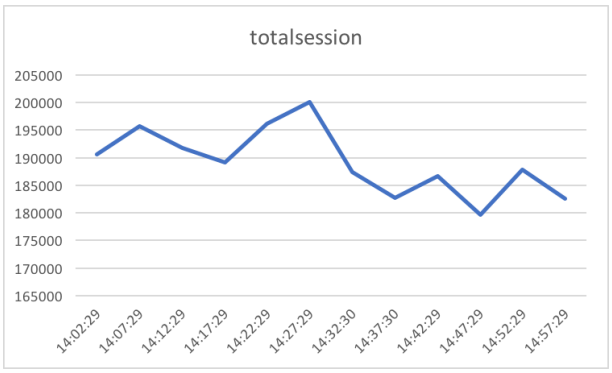

図 5: 6/23 の UTM の totalsession の変化

- 6/30 の結果 (IP ランダム)

  図 6 は攻撃実験を行った 6/30 の UTM の CPU の変化、図 7 は攻撃実験を行った 6/30 の UTM の totalsession の変化である。6/30 について攻撃は 17:54:20 に行ったが 図 6 を見るとその前後で cpu 使用率が増加していることがわかる.。図 7 を見ると、この日は IP アドレスをランダムに設定し攻撃を行ったため攻撃の前後で totalsession が上昇していることがわかる。以上のことから攻撃を行った際信州大学のネットワークに対し負荷をかけることに成功している。

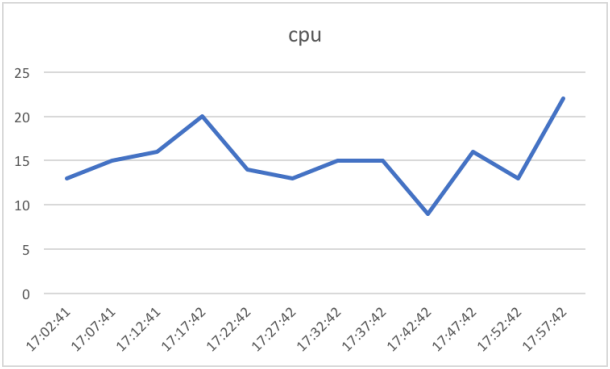

図 6: 6/30 の UTM の CPU の変化

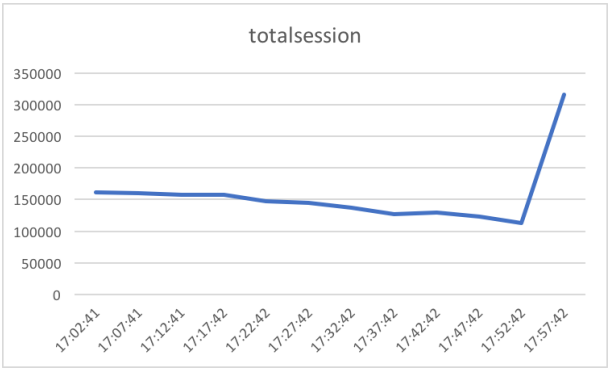

図 7: 6/30 の UTM の totalsession の変化

4.3. 自己組織化マップによる解析結果

本章では、6/19 から 6/30 にかけて攻撃を行った際に取得したトラフィクログについての解析結果を掲載する。

図 8 は、解析に用いる UTM のトラフィックログの一部である。青でハイライトされた部分が 1 セッションの情報である。表 4 に示した sentbyte, rcvdbyte, srcport, dstport の 4 つの情報を SOM のパラメータとし解析を行った 。図 9 から図 16 は、解析結果の一部である。青い点は 1 つのセッションを示し赤い点は自作した攻撃のセッションを示す。その中で、IP アドレスを 4 つに設定して 1,000,000 回攻撃を行った 6/23 の結果と、IP アドレスをランダムに設定して攻撃を行った 6/30 の結果を掲載する。

```
Jun 23 14:00:00 fw01 date=2017-06-23,time=13: 59:59,devname=slipfw01,devid=FG1K5
D3I14803098,logid=0000000013,type=traffic,subtype=forward,level=notice,vd=root,s
rcip=[redacted] srcport=5502,srcintf="port33.1",dstip=[redacted],dstport=5
087,dstintf="LAG.850",poluuid=e0fafb94-4bbb-51e5-dae6-d0c823e6e93b,sessionid=375
1890585,proto=17,action=deny,policyid=673,dstcountry="Japan",srccountry="France"
,trandisp=noop,service="udp/5087",duration=0,sentbyte=0,rcvdbyte=0,sentpkt=0,crs
core=30,craction=131072,crlevel=high
Jun 23 14:00:00 fw01 date=2017-06-23,time=13: 59:59,devname=slipfw01,devid=FG1K5
D3I14803098,logid=0000000013,type=traffic,subtype=forward,level=notice,vd=root,s
rcip=[redacted] srcport=52450,srcintf="port33.1",dstip=[redacted]stport=
23,dstintf="LAG.850",poluuid=e0fafb94-4bbb-51e5-dae6-d0c823e6e93b,sessionid=3751
890592,proto=6,action=deny,policyid=673,dstcountry="Japan",srccountry="India",tr
andisp=noop,service="TELNET",duration=0,sentbyte=0,rcvdbyte=0,sentpkt=0,crscore=
30,craction=131072,crlevel=high
Jun 23 14:00:00 fw01 date=2017-06-23,time=13: 59:59,devname=slipfw01,devid=FG1K5
D3I14803098,logid=0000000013,type=traffic,subtype=forward,level=notice,vd=VDOM4,
srcip=[redacted] srcport=54789,srcintf="LAG.501",dstip[redacted]dstport=35
44,dstintf="port33.69",poluuid=901ffade-4bbc-51e5-3ff9-2fb50b50369f,sessionid=37
51890614,proto=17,action=deny,policyid=363,dstcountry="United States",srccountry
="Reserved",trandisp=noop,service="udp/3544",duration=0,sentbyte=0,rcvdbyte=0,se
ntpkt=0,crscore=30,craction=131072,crlevel=high
```

図 8: UTM のトラフィックログの一部

表 4: パラメータとして使用したトラフィックログの４つの要素

sentbyte	送信バイト数
recvdbyte	受信バイト数
srcport	送信元のポート番号
dstport	送信先のポート番号

- 6/23、10 分間の解析結果 (IP アドレス 4 つ)

  図 9 は 6/23 の 14:10:00 から 14:19:59 の 10 分間の SOM 出力結果、図 10 は図 9 の攻撃セッションを含むクラスタの拡大である。

- 6/23、1 時間の解析結果 (IP アドレス 4 つ)

  図 11 は 6/23 の 14:00:00 から 14:59:59 の 1 時間の SOM 出力結果、図 12 は図 11 の攻撃セッションを含むクラスタの拡大である。

- 6/30、10 分間の解析結果 (IP アドレスランダム)

  図 13 は 6/30 の 17:50:00 から 17:59:59 の 10 分間の SOM 出力結果、図 14 は図 13 の攻撃セッションを含むクラスタの拡大である。

- 6/30、1 時間の解析結果 (IP アドレスランダム)

  図 15 は 6/30 の 17:00:00 から 17:59:59 の 1 時間の SOM 出力結果、図 16 は図 15 の攻撃セッションを含むクラスタの拡大である。

攻撃セッション周辺を拡大した図 10、図 12、図 14、図 16 を見てわかる通り、自作した攻撃 (赤い点) は 1 つのクラスタに分類され、自作した攻撃 (赤い点) を含むクラスタには検査対象とする通信 (青い点) も含まれていることがわかった。

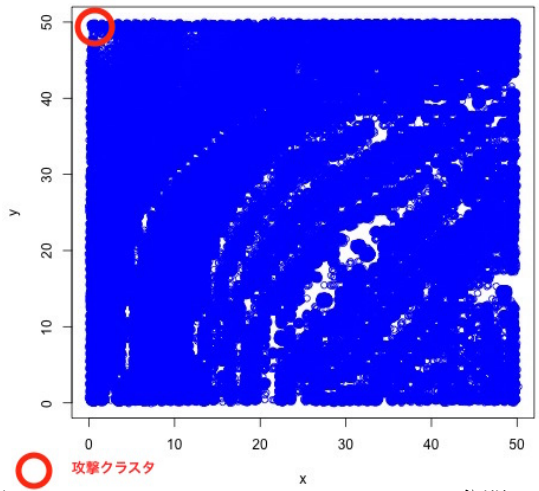

図 9: 6/23 の 14:10:00 から 14:19:59 の 10 分間の SOM 出力結果

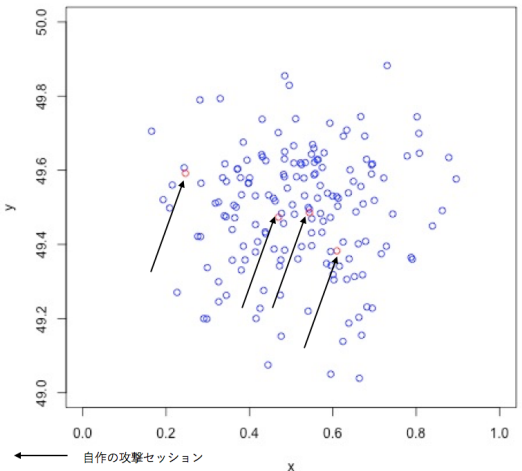

図 10: 図 9 の攻撃セッション周辺の拡大

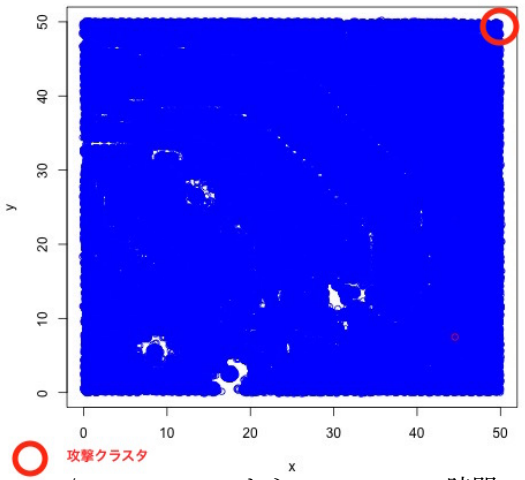

図 11: 6/23 の 14:00:00 から 14:59:59 の 1 時間の SOM 出力結果

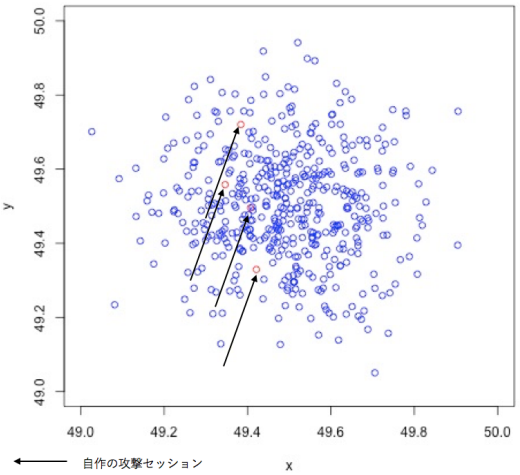

図 12: 図 11 の攻撃セッション周辺の拡大

Vol.2017-IOT-39 No.4
Vol.2017-SPT-25 No.4
2017/9/29

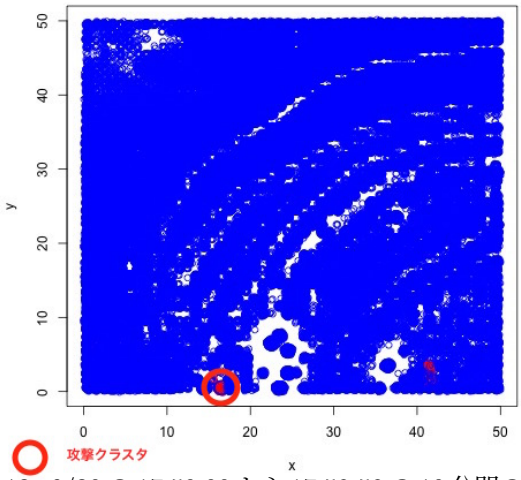

図 13: 6/30 の 17:50:00 から 17:59:59 の 10 分間の SOM 出力結果

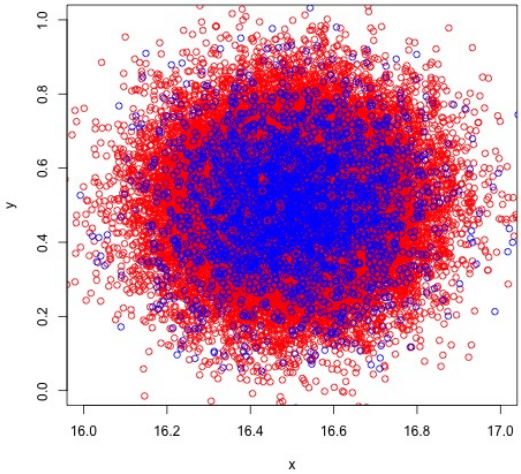

図 14: 図 13 の攻撃セッション周辺の拡大

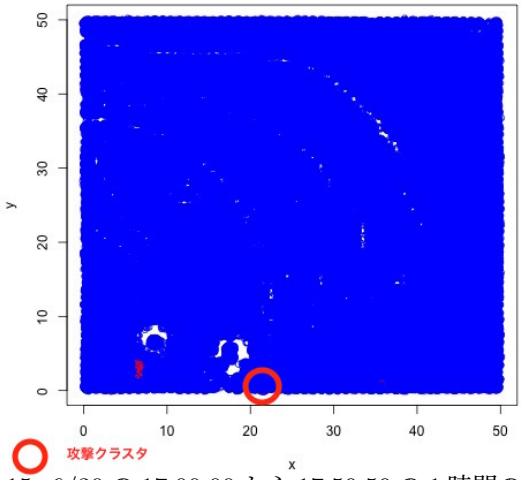

図 15: 6/30 の 17:00:00 から 17:59:59 の 1 時間の SOM 出力結果

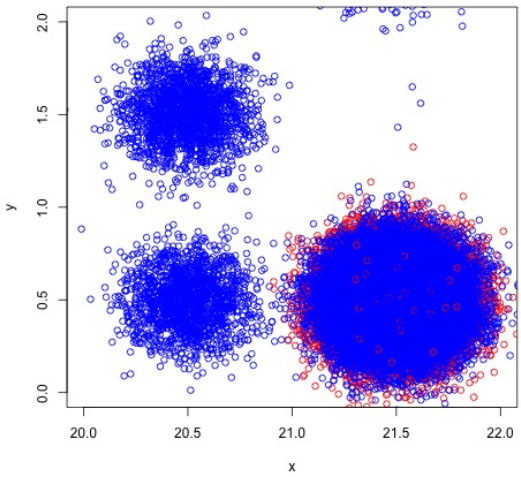

図 16: 図 15 の攻撃セッション周辺の拡大

5. 評価と考察

自作した攻撃 (赤い点) と同一クラスタに含まれる通信 (青い点) を調査した結果を以下にまとめる。

- 6/19 から 6/23(IP アドレス 4 つ)

  6/19 から 6/30 の解析結果については自作した攻撃 (赤い点) と同一クラスタに含まれる通信 (青い点) からは不審と思われる通信は発見できなかった。これは IP を 4 つに詐称した場合 UTM の性質上、通信が 1 つの大きな通信と判断されてしまい、単に通信量の非常に多い通信のクラスタに分類されてしまったためと考えられる。

- 6/26 から 6/30(IP アドレスランダム)

  6/30 の 10 分間の解析結果からは不審な通信と思われる IP アドレスが 1 件 (5 session)、1 時間の解析結果からも不審な通信と思われる同一 IP アドレスを 1 件 (2389 session) 発見することができた。6/26,27 についてもこれと同一の不審と思われる IP アドレスを発見することができた。6/28 については不審と思われる通信は発見できなかった。

このように IP アドレスをランダムにして攻撃を行いそれを解析した場合、今回は不審と思われる通信を検出することができた。また、自作した攻撃について、解析を行ったすべての日程で 1 つのクラスタに分類することに成功した。

6. まとめと今後の課題

今回は、IP アドレスをランダムに設定し攻撃を行った日の解析結果から一部、UTM では検出できない不審と思われるホストを発見することができた。第二章で述べた目的である UTM では検出できない不審と思われるホストの検出について、SOM は有用性を示すことができた。しかしながら、最終的な目的であるセキュリティエンジンへの適用は、現段階では、検出精度、データマイニングの処理時間の観点から難しい。そ

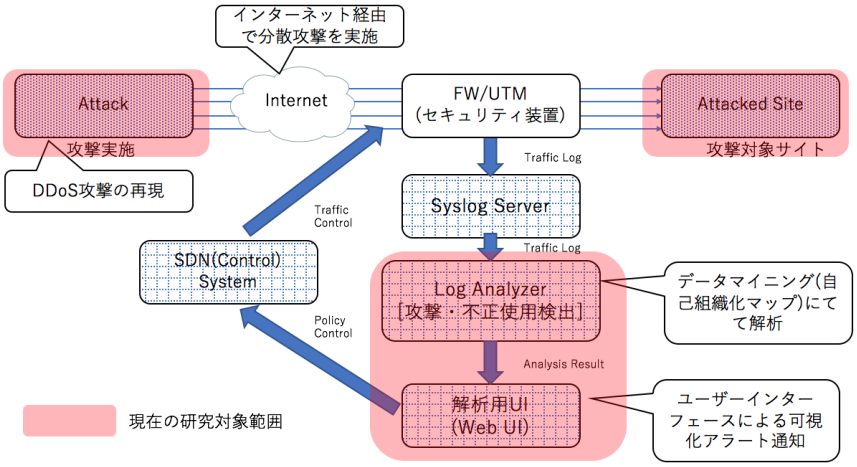

図 17: 提案する新しいネットワークの構成

のため攻撃、解析の二つの改善を行っていく必要が有る。攻撃については、今回は DNS に対するリゾルバフラッド攻撃のみ再現しているが、http に対する DDoS 攻撃など他の攻撃手法も再現していく予定である。解析については、SOM 出力の際のパラメータの種類の変更や、他のデータマイニングの手法を試していく予定である。また、研究の最終目標として、図 17 のようなネットワークの構成を考えている。現在は赤でハイライトされた部分の研究を進めているが、最終的にはデータマイニングで解析した情報をもとに SDN などの装置でトラフィックコントロールするといった構成のネットワークを実現させる予定である。

謝辞

本論文を作成するにあたって使用された UTM 装置による通信ログは、信州大学情報総合センターの協力を得て取得致しました。

参考文献

[1] 小島俊輔, 中嶋卓雄, 末吉敏則:統計的手法を用いた DoS/DDoS 検出手法とその特性; マルチメデア通信と分散処理ワークショップ集 Multimedia Communication and Distributed Processing System Workshop 2009(9),209-214,2009-09-30

[2] 椛島健, 堂園浩:SOM を用いた IP パケットトラフィックの視覚化; 日本知能情報ファジィ学会, ファジィシステムシンポジウム講演論文集 26(0), 258-258, 2010

[3] 豊田秀樹:データマイニング入門− R で学ぶ最新データ解析; 東京図書,ISBN978-4-489-02045-2,2008

[4] 宇井哲也, 成瀬愼, 湯原大二郎, 鈴木彦文:UTM では検出困難な DDoS 攻撃を統計的手法を用いて検出する研究における DDoS 攻撃環境の開発; 情報処理学会, 平成 28 年第 15 回情報科学フォーラム,FIT2016,:L-027 2016(Sep. 09)

[5] 湯原大二郎, 宇井哲也, 鈴木彦文:DNS に対する高度な攻撃を検出するためのデータマイニング; 情報処理学会, 平成 28 年第 15 回情報科学フォーラム,FIT2016,:L-028 2016(Sep. 09)

[6] 西脇春名:「Mirai」ソースコード徹底解剖とその仕組みと対策を探る (2/4); @ IT,Security & Trust 2/4