GoogleMaps と GeoIP を用いた 分散 Honeypot のログ解析と視覚化

金 子 博 一[†]　小 池 英 樹[†]

今日、ネットワークセキュリティにおいて IDS は幅広く使われている。しかし、少ない IDS では情報量が充分であるとは言えない場合がある。そこで分散 IDS のログを比較する事で 1 つの IDS ログだけでは得られない情報を得る事ができる。しかし，実際にはログのフォーマットや提供者等の問題があるため分散 IDS のログ比較は難しい。

本研究では分散 IDS のログ比較の一例として、The Honeynet Project に焦点を当てた複数拠点のログ情報の比較を行うことができる視覚化システムの提案及び実装を行った。このシステムは GoogleMaps と GeoIP を用いることで地理的視覚化を行っており、攻撃の特徴やおおまかな概要を直感的に把握できるようになり、また個々の拠点の攻撃情報の特徴を比較することが可能となった。

Visualization and Analization System for Distributed Honeypot Using Google Maps and GeoIP

Hirokazu Kaneko[†] and Hideki Koike[†]

Today, IDS is used by network security. But only one IDS data is not enough in some cases. So it is important that comparing distributed IDS data for network security. But, in fact,Making Systems is difficult for log format, log trading and so on.

By comparing distributed IDS logs, it is possible to obtain useful information which could not be obtained from one IDS log. It is, however, difficult to do such analysis since the distributed IDS logs are often recorded in their own format. This paper described a web-based visualization system which can visualize multiple IDS logs in one geographical map by focusing on the logs shared by The Honeynet Project. The system expressed by Google Maps and GeoIP enable us to understand an overview and details of attacks intuitively as well as to visually compare the attack information at different locations.

1. はじめに

近年、コンピュータネットワークの普及に伴い、インターネットが社会インフラとして重要な役割を占めるようになった。そのためインターネットに接続された計算機を標的としたサイバー攻撃の問題が深刻になっている。コンピュータワームやコンピュータウィルス、ボット等のサイバー攻撃が増加しており、これらの攻撃によってメールシステムや Web サービスといったものを停止させることがある。従ってサイバー攻撃は経済活動や公共に大きな影響を与えているといえる。

これらのサイバー攻撃は IDS(不正侵入検知システム) を通じてログとして計算機に蓄積されるが、こういった攻撃の解析には膨大なテキストデータを読む必要があり、とても一件一件を精査することはできない。そのため情報視覚化技術を用いて解析をしやすくすることが多い。

近年、インターネット上のこれらのサイバー攻撃を監視することにより、攻撃の早期発見や攻撃の予知に役立てる研究が行われている。それらの研究の一つとして、分散 IDS の比較の研究が行われている。多くの分散した IDS によるログ比較を行うことで、一つの IDS による検知ログではわからない攻撃を統合的に攻撃を解析する手法の一つである。しかし、ログの比較にはお互いのログのフォーマットや、ログの比較者に個人情報の漏洩になりうる面がある等の問題を抱えており、実現は難しい。

このような問題に対し、攻撃情報を共有することを目的とした国際的プロジェクト、The Honeynet Project がある。The Honeynet Project は、ネットワークを監

† 電気通信大学大学院　情報システム学研究科
The University of Electro-Communications

視、解析を行う行うツールの一つである Honeypot を運用している大きな団体の一つであり、世界のセキュリティ調査員はこれを用いて監視、解析を行っている。The Honeynet Project ではログのフォーマットがほぼ共通であり、交換が行いやすいといった特徴があり、上記の問題を解決できる。しかし、世界中にある拠点間で交流が少ないといった問題点があり、データのやりとりも少ない。

本研究では、The Honeynet Project の個々の拠点でのデータ比較を行うことで分散 IDS のログの比較の一例として行う。そこでログ情報を共有していることを前提に、解析結果を GoogleMaps や GeoIP を用いて地理的視覚化を用いてブラウザ上で表現する視覚化システムを考案、実装した。

2. 分 散 IDS

IDS は計算機における不正侵入と思われる攻撃をログに残すものであり、今日のネットワークセキュリティにおいて欠かせないものであるといえる。そのため多くのセキュリティ研究員にとってネットワーク上の攻撃を監視、解析し、世の攻撃を知る上で重要なツールである。しかし、一つの拠点の IDS ではその拠点に対する攻撃しかわからないという点は問題であると言える。

そこで、複数の IDS を比較することで一つの IDS ではわからない情報を得る研究というものがある。例えば、ある攻撃が広域的に行われているのか、それとも特定 IP に対して集中的に行われているのかといった情報は他の IDS のログと比較してはじめてわかることである。この場合、特定 IP に対する攻撃は広域的な攻撃よりも特徴的で危険な攻撃である可能性が高い。このような統計的な情報を得るため、複数の IDS のログを比較することは有用である。

3. The Honeynet Project

以前のネットワークセキュリティは、受動的な対処療法に過ぎなかった。ファイアーウォール、IDS や暗号化といった全てのメカニズムは、各自のリソースを守る為に受け身として用いられていた。そのため、問題が起きるまでは対処のしようがなく、常に攻撃者が主導権を握っている状況にあった。これらの状況を打破するべく、The Honeynet Project は作られたものである。

The Honeynet Project の目標は、世に広まっている攻撃に対する情報を集めることである。例えば悪意ある攻撃用のツール、攻撃手法、攻撃動機といったものが挙げられる。情報を収集するツールとして主に Honeypot を用い、現在では世界 23 拠点で活動的に攻撃の監視、解析を行っている。

3.1 Honeypot

Honeypot は脆弱な PC を模擬することで攻撃者をおびきよせ、攻撃者の行動を逐一ログにとることで監視、解析を行い、世の中の攻撃を知るセキュリティセンサの一つである。Honeypot には以下の利点を挙げられる。

- 誤検知が少ない
  通常、人が使う計算機の IDS による検知ログを見ると、正規のユーザーが行った行動でも検知されてしまう場合が多い。例えば Web アクセスによるファイルのダウンロードといったものも誤って検知してしまい、ログに記録される。Honeypot は正規のユーザーを持たず、脆弱な PC を模擬する計算機を置くというシステムである。そのためログに残っている情報は攻撃、あるいは IP アドレスを間違えてアクセスしようとしたものが対象となる。従って通常のログよりも誤検知を極端に減らす事ができる。
- 暗号化されていても活動できる
  ハニーポットは最終的な通信対象であるホストにおいて情報を捕足する。そのためネットワークなどの途中の経路で通信を監視するようなネットワークベースのファイアウォールや侵入検知システムでは対応できない SSL や IPsec といった様々な暗号化プロトコルを使用している環境や, 攻撃自体が暗号化されている状況であってもその活動を捕足することができる.
- 柔軟性
  また, ハニーポットは様々な環境に応じて柔軟にカスタマイズすることができ, 他のセキュリティアプリケーションと組み合わせて利用する事も可能である. ハニーポットはその柔軟性から, 例え組織内の犯行による情報漏洩などといった, 他のセキュリティ対策では対応不可能な問題にも適用する事ができる.

ここでは一例として Honeywall の計算機構成を図 Honeypot に記す。

3.2 Honeynet における問題

Honeypot はネットワークセキュリティ研究者において、攻撃者の攻撃動機を知る上で重要なツールである。しかし、各々の研究者が Honeypot を設置する台数は限られるため、監視、解析している Honeypot の

ログの絶対量が多いとは言えない。そのため今後個々の研究者が世の攻撃の発展に追従する事ができなくなる可能性がある。

また、The Honeynet Project 同士で互いに交流は浅いことが多いといった欠点があり、情報を提供しあったり解析情報を公開するといった、協力して物事を進めることは少ない。今日のネットワークセキュリティにおいて、攻撃側の技術も日進月歩に飛躍し続けているため、未知の攻撃に対して少数の研究者では気づけない可能性があり、これによって攻撃に対する対策が遅れる要因となりうる。

そして、ネットワークセキュリティの研究は一般に困難だと言われている。大量のログデータを見る必要があり、明らかな攻撃であっても初心者には分かりづらい。そのため新しくネットワークセキュリティの研究を始めようとする考えが生まれづらいものとなっており、攻撃を助長させる一つの要素となっている。

4. システムの基本設計

4.1 ログ情報の共有

ログ情報の公開は個人情報にもなりうる。企業や個人でも IDS を運用している場合があるが、そのログを公開することはその企業や個人の IDS のデータを利用者に公開する必要があるため全世界の IDS のユーザを対象にすることは難しい。

そこで、The Honeynet Project 内部の Honeypot ユーザならば個々の情報を公開、共有して問題ないとすることができる。ログを共有することで、研究者同士が積極的に助け合い、The Honeynet Project における問題に対処できると考えられる。特に、日進月歩のネットワークセキュリティの攻撃に対し、新しい不審な攻撃や特徴的な攻撃の早期発見につながり、周囲の攻撃ログとの比較を簡単に行う事ができる。そのため世に広まっている攻撃情報を正しく扱え、各国の特徴的な攻撃や、全世界で流行している重要な問題に対しても積極的に情報が出回るようになると考えられる。

4.2 対象とするユーザ

The Honeynet Project の研究員を対象とし、実装を行った。

インターネット定点観測システムでは、複数のセンサの情報を統合し、統計解析を行う。そのためインターネット定点観測の結果が即座に本システムが対象としているユーザにとって有益な情報であるとは限らない。そのためユーザは警告ログの解析ツールを構築する必要が有る。しかし、これらの解析ツールの構築は煩雑で、適切な運用を行う為には定期的なメンテナンスも必要であるためユーザに負担を強いることになる。

また、Honeynet における問題として、各研究者が持っているログは運用できる Honeypot の数に制限があるため少ないと言える。

本研究ではこれらの問題を解決するため、ログ情報を共有していることを前提とした、ウェブアプリケーションとしてシステムを構築した。ウェブアプリケーションの利点については後述することにする。

4.3 ウェブアプリケーションでの実装

本システムはウェブアプリケーションとして実装した。ウェブアプリケーションには以下のような利点がある。

- Web ブラウザのみで動作する
  世界共通で特別な環境に依存しないものが好ましい。ウェブアプリケーションはウェブブラウザさえあればどこからでもアクセスが可能であるため、特殊な装置や環境を必要としない。ユーザはプラットフォームの違いも計算機の違いも意識することなく、統一的な動作でアプリケーションを利用することが可能である。
- メンテナンスフリー
  ウェブアプリケーションはユーザにとってメンテナンスフリーで利用できる。システムはサービスを提供するサーバに構築されているため、ユーザは普段通りウェブブラウザからシステムにアクセスするだけでよく、特に特別な環境を必要としない。

5. システムの実装

分散 IDS のログ比較の一例として、The Honeynet Project を対象としたサイバー攻撃の地理的視覚化システムの提案及び実装を行った。

5.1 システム概要

サイバー攻撃の地理的視覚化システムをウェブアプリケーションとして実装している。

ユーザは事前に Honeypot に、本システムを利用するためログの転送システムを設置する必要が有る。その後ウェブウラウザを用いて本システムにアクセスし、解析情報にフィルタリングをかけることによって地図上に適切な解析情報を出力する。

これによりユーザは特別な環境を必要とせずウェブブラウザさえあれば手軽に解析情報を知る事ができ、また他の拠点の The Honeynet Project のデータを参照することで攻撃の特徴をとることが可能になる。

5.2 システム構成

- データ格納部
  （1） Honeypot のデータをサーバに転送する
  Honeypot にある不正検知システムのログデータを自動で送信する。今回、高対話型 Honeypot の IDS のログデータの内、Snort と Iptables を用いている。
  （2） データを格納する
  MySQL を用いて送られてきたデータを格納する。
- データ視覚化部
  （1） ブラウザを用いて HP にアクセスする
  特定の計算機、特定の環境が必要なものを避けるべく、Web アプリケーションで実装。
  （2） 解析データを視覚化する
  PHP で MySQL を用いて格納したデータを参照し、そのデータから重要拠点のデータを読み込む。読み込んだデータを GeoIP を用いて緯度経度に変換し、その情報を元に GoogleMaps を用いて地図上に描画する。また、それらのデータから攻撃に関わる IP アドレス、Port 番号、シグネチャの三種類について円グラフを作成し、上位 8 件を描画し、その他はまとめて「Etc」とした。このとき、どのデータを参照するかといったフィルタリングをすることができる。

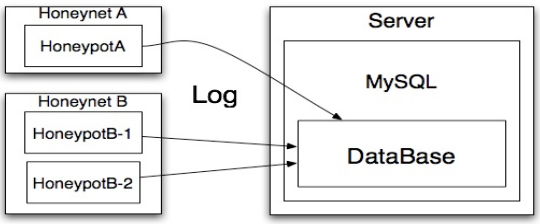

図 1 データ格納部図

今回はデータ視覚化部について実装、考察を行った。

5.3 描　　画

ここでは機種依存が激しいものや特別な設置方法が不必要な、ブラウザを用いて Web ベースの視覚化を行うことにする。

- 表示部
  操作部で指定した解析情報を基に、GeoIP を用いて攻撃元または攻撃先の IP から GeoIP を用いて地理情報に変換し、GoogleMapsAPI を用いて地図上にマーカーを描画している。

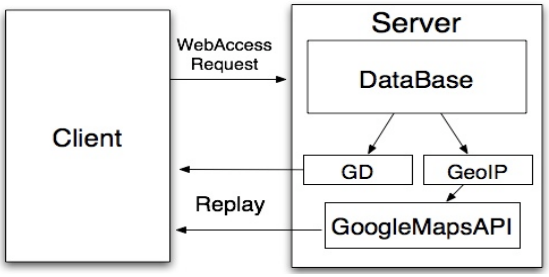

図 2 データ視覚化部図

それぞれのマーカーをクリックすることで、図 4 のようにその拠点の詳細な情報を得ることができる。IP アドレス、シグネチャ、ポート番号の三つのタグがあり、それぞれの攻撃量と分散量、攻撃先の棒グラフを得ることができる。これらの情報は攻撃量が多い順に表示されている。

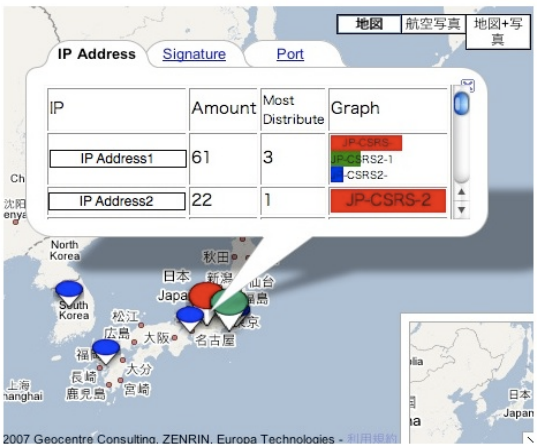

図 3 詳細表示例

そして分散量はいくつの Honeypot への攻撃があったかを知る事ができ、詳細な攻撃先は棒グラフを見ることで知る事ができる。この棒グラフは攻撃量の多い攻撃先上位三件を記し、四件目以降はまとめて Etc として表記される。Honeypot の表記は [国名]-[拠点の略称]-[Honeypot の番号] で表現している。

また、読み込んだデータから統計情報を出し、GD を用いて円グラフとした。図 5 のように全体の情報から最も量の多かった IP アドレス、シグネチャ、ポート番号をそれぞれ円グラフで表現し、上位 8 件にはその順位に対応する色をつけた。円グラフ上でマウスを乗せることでその攻撃の詳細な情報を得る事ができる。

このとき図 6 のように円グラフの色とその IP アドレスを持つ拠点のマーカーの色と同期させた。三種類の円グラフの内、どのグラフと同期させるか選択でき、色は全体として攻撃量が多かった色を優先している。また特殊な色がついているマーカーはその他のマーカーよりも大きく表現した。

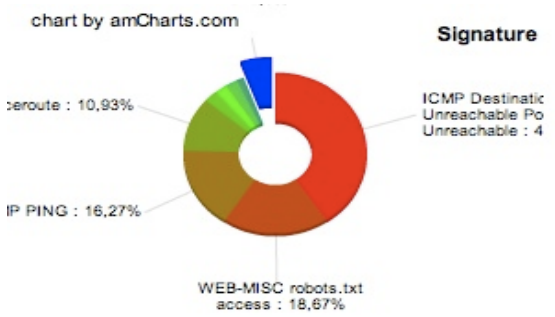

図 4 円グラフ例

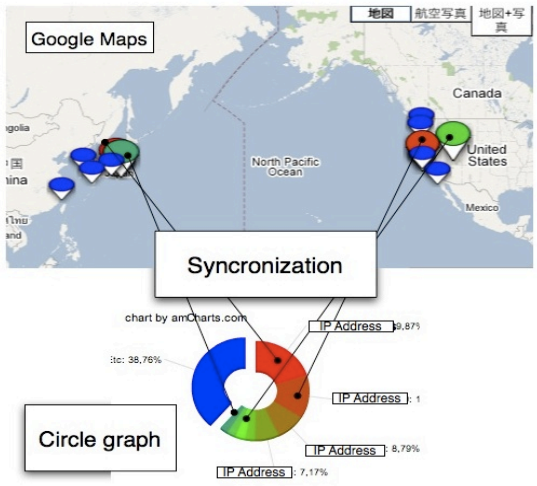

図 5 同期例

- 操作部

  検知の種類、攻撃の向き、Honeypot の指定を行い、このデータを地図表示部で反映させる。検知の種類は現在 Snort,Iptables の二種類選択可能になっている。攻撃の向きは外部から内部への IN-BOUND と内部から外部への OUTBOUND の二種類を Ajax を用いてアクティブにフィルタリングをかけれるように実装した。そのフィルタリングによりマーカーの色とどの円グラフの色を対応させるかの選択や、どの Honeypot のデータを参照するかのフィルタリングをすることができる。また特定の IP や Signature、Port のみのフィルタをかけることができるようにした。

5.4 システム動作例

システム動作例 1-1 では日本を拠点とした高対話型 Honeypot の一つ、HoneyWall の Snort ログを IN-BOUND で視覚化している。日本の 5 拠点における HoneyWall のログの内、2007/6/10-13 の 4 日間のデータを対象とし本システムを稼働させており、ここでは円グラフを IP アドレスの色を同期させている。

ここで、危険な攻撃の可能性のある、攻撃量の多い IP アドレスに着目する。システム動作例 1-2 右図は最も攻撃量が多い IP を含むマーカーの詳細情報を表示している場面である。詳細情報を確認すると、JP-CSRS-1 の Honeypot のみ攻撃している。このことからその攻撃は JP-CSRS-1 の Honeypot を狙った攻撃であり、危険度が高いといえる。次にシステム動作例 1-2 左図は、攻撃量が多い IP を含むマーカーの詳細情報を表示している。詳細情報を確認すると、JP-MORI-が最も多く、次に JP-TMR といった計 4 つの Honeypot へ攻撃していることがわかる。攻撃量が多い上位 8 件の IP にランクインする IP だが、実は多くの拠点に対して攻撃を行っており、さほど危険な攻撃ではない可能性があることがわかる。

システム動作例 2 図ではシグネチャの円グラフの色とマーカーの色を同期させた。すると、画面全体では全体の攻撃量が多い赤いマーカーは少なく、緑色のマーカーが多いことがわかる。このことから赤いシグネチャの攻撃は少数の IP アドレスからの攻撃である可能性があり、実際に調べてみたところ特定の IP からそのシグネチャがきていることがわかった。　また、システム動作例 3-1 図のように JP-MORI-の Honeypot を選択すると MySQL の攻撃量が上位にランクインしていた。そこで SQL に対する攻撃の特徴をとるためシステム動作例 3-2 図で SQL のサービスを対象としている Port 番号 1434 を対象としてフィルタリングした。この結果から SQL の攻撃は多くの国からあり、IP の円グラフを見ても上位 8 件の占める割合が少ない。そのためこれらの攻撃は、全体の攻撃量は多いが分散しているため重要でない可能性がある。

6. 考　　察

攻撃量の多い、危険な攻撃をしている可能性のある IP をとらえ、分散の視点からその攻撃の特徴をつかむことができた。特別な装置や訓練を必要とせず、解析情報を見る事ができた。

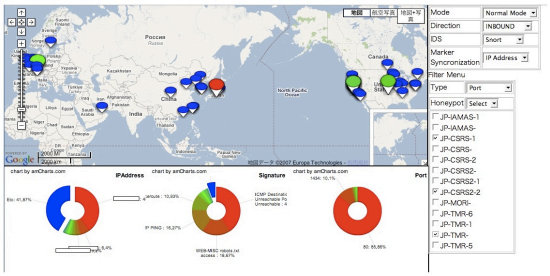

図 6　システム動作例 1-1

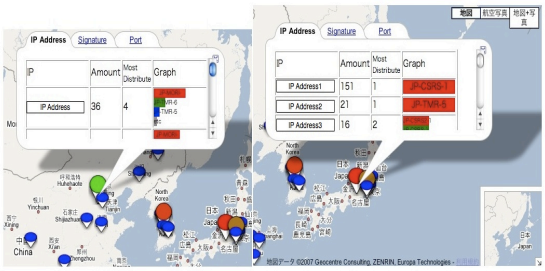

図 7　システム動作例 1-2

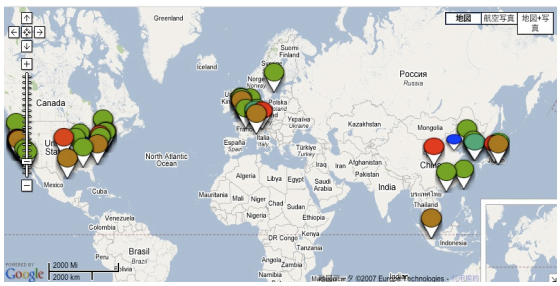

図 8　システム動作例 2

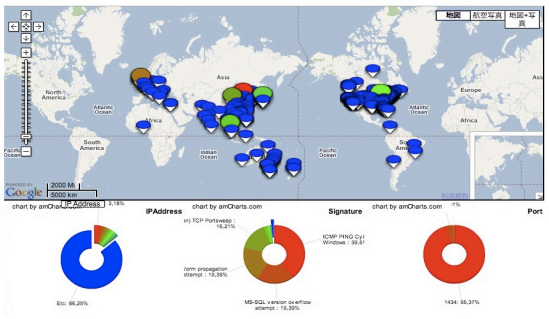

図 9　システム動作例 3-1

7. 今後の課題

本システムの課題として、以下の事項が挙げられる。

- ログの共有

  本システムで提言している IDS のログの共有

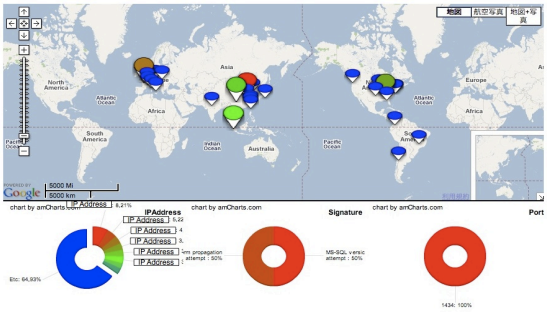

図 10　システム動作例 3-2

において、自動で解析ログをサーバーに格納するシステムが必要だと考えられる。現在のシステムは直接解析データから読み取っているため、サーバーに解析データを読み込ませさえすればそれを反映することができる。

- 時間軸

  インターネット広域監視システムにおいて、早期に発見し対策をとることが重要である。また、目的の攻撃に対して何度かステップを踏む場合がある。そのためリアルタイムであることといった時間的な要因は重要であると言える。そこで、警告ログがいつ行われたかを視覚的にわかりやすくする必要がある。

8. ま と め

本研究では分散ログを比較するためのログ情報共有視覚化システムの提案を行い、その一例として The Honeynet Project のログ共有視覚化システムの視覚化部分の実装を行った。また日本拠点で集められたデータを基に、視覚化システムの効果を示した。今後はログの共有部分を実装、システムの改良を進めシステムの完成を目指す。

参 考 文 献

1) 秀島裕介、小池英樹、「サイバー攻撃の論理的及び地理的視覚化に関する研究」、平成 18 年度電気通信大学大学院修士論文
2) The Honeynet Project, http://www.honeynet.org/index.html
3) Snort, The Open Source Network Intrusion Detection System, http://www.snort.org/
4) DShield, http://www.dshield.org/
5) Christopher P. Lee 、John A. Copeland、FlowTag: A collaborative attack-analysis, reporting, and sharing tool for security researchers