博 士 論 文

インターネット上の背景放射パケットの解明と脅威検知手法の研究

指導教官　篠田　陽一 教授

北陸先端科学技術大学院大学
情報科学研究科組込みシステムコース専攻

石黒　正揮

2009年3月

目 次

図 目 次

表 目 次

要 旨

インターネット上のワーム感染や不正侵入を意図したネットワーク攻撃は、情報通信社会の大きな脅威となっている。本研究では、インターネット上の不正なパケット (「不正パケット」と呼ぶ) を観測することにより、ワーム等による脅威を早期に検知するための手法を示す。インターネット上でネットワークサービスを提供しない IP アドレスを観測することにより、正規のネットワークサービスの利用を意図しない不正パケットを観測することができる。インターネット上でこのような IP アドレスを広域に複数地点で観測することにより、自サイトでは観測されないインターネット上の脅威を早期に検知し、自サイトが攻撃を受ける前にネットワーク防御を行うことが可能である。インターネット上では、攻撃の対象となるソフトウェアの脆弱性の対策 (ソフトウェアパッチなど) が既に取られているようなポートに対しても恒常的に不正パケットが観測される (「背景放射パケット」と呼ぶ)。ポートスキャンなどの脅威の小さい不正パケットに混在して、ネットワークサービスソフトウェアの脆弱性を攻撃する脅威の大きい不正パケットを検出することで、インターネット上の脅威を検知することが求められる。そのためには、恒常的に観測される不正パケットの統計的性質を解明するとともに、観測される不正パケットの中から脅威のレベルを評価することが必要になる。ワーム等の脅威検知においては、インターネット上に混在する多様な攻撃パターンに対応するために、多角的な脅威検知手法を同時に適用し、それらの分析結果から脅威の原因を分析することが必要になる。本研究では、不正パケットの時間的な特徴量および空間的な特徴量の変化に基づく複数の脅威検知手法について考察し、それらの特徴および限界等を評価することにより、複数の検知手法を組合わせた脅威分析手法を検討する。

第 1章

はじめに

1.1 背景

企業間の電子商取引や個人の社会生活におけるインターネットの利用が深く浸透し、インターネットは必要不可欠な社会基盤となっている。一方、インターネットの普及に伴い、ワームやボット、不正侵入などのネットワーク攻撃が増加し、情報通信社会の大きな脅威となっている [10, 36]。インターネット上の攻撃は、コンピュータや WEB サービスの動作障害を引き起こす顕在型のものやデータ消去といった初期の愉快犯的なものから、近年は、インターネット上の金銭窃盗や情報窃盗など経済犯的なものが増加している。これに伴い、ネットワーク攻撃は潜在的な方法によるものが増加し、脅威の早期検出が困難となってきている。

本研究では、インターネット上のワームや不正アクセスなどの脅威を検出するための手法として、インターネット上のパケット観測技術および観測データの分析に基づく脅威検知手法の研究開発を行った。インターネット上の攻撃は、ネットワークサービスを提供しない IP アドレスを観測し、そこで観測されるワームなどから送信される不正なパケットの変化を分析することにより検出することが可能である。ネットワークサービスを提供しない IP アドレスには、本来、外部から正規のリクエストを受ける事は無いため、外部からそのアドレスに対して送信されるリクエストは、ワームや不正アクセスなどの攻撃パケットか、ネットワーク設定の不備な機器から送信されるパケットと判断される (以下、これら両方のパケットを「不正パケット」と呼ぶ)。

ネットワークのトラフィックを観測することにより攻撃を検知する技術として、従来より IDS (Intrusion Detection System) の研究が行われてきた。IDS は、主に自サイト内のネッ

トワークトラフィックを分析することにより攻撃の有無を検知する技術である。一方、インターネット観測に基づく脅威検知技術は、インターネット上で広域的に複数の IP アドレスにおいて、不正パケットを観測することにより、新種ワームの発生や増殖などインターネット上の脅威元の集団の動的な変化を捉え、自サイトに被害が及ぶ前に早期に警戒を促すことを目的とする。インターネット上の脅威を早期に検知することができれば、特定のサービスに対する外部からのアクセスを制限したり、特定の送信元からアクセスを禁止したり、特定のサーバーのソフトウェア脆弱性を改修する（パッチの適用）などの対策に繋げることができ、被害の発生を未然に防止することに役立てることができる。近年、ワーム等による攻撃手法は多様化、巧妙化し、進化スピードが高速化している。このような状況で、単一の手法に基づく脅威検知手法では、対応が困難となっている。

本研究では、インターネット上で観測される不正パケットの IP アドレス空間上の分布の傾向について分析し、それらの分析に基づき脅威検出のための手法について検討を行い、複数の手法に基づく脅威検知を開発した。近年のインターネットにおいては、振舞いの異なる多様な攻撃が同時多発的に発生しているため、単一の特徴に基づく脅威検知手法だけでは十分ではない。本研究では、攻撃元の集団および個別の振舞いや、不正パケットの時間的な特徴や空間的な特徴に基づく、複数の脅威検知手法を開発し、それらを組合わせて利用するアプローチを提案する。

本論文の構成は以下の通りである。第 1.2 章では、研究成果の要点をまとめる。第 3 章では、関連研究についてまとめる。第 4 章では、脅威検知手法の基礎となる分析結果を示す。第 5 章では、提案する脅威検知手法について述べる。第 6 章では、脅威検知手法の使い方についてまとめる。第 7 章では、結論をまとめる。

1.2 研究成果の要点

本研究の主な成果は以下の通りである。

- ベイズ推定および時系列トレンド解析を組合せることで、不正パケット数の増加から脅威を検知する手法を開発した。検出漏れ、誤検出を総合的に判断する ROC 分析で、Az 値 0.95 を実現した。(第 5.3 章)
- ウェーブレット解析に基づき、不正パケットの周期性の変化に基づく異常検知手法を開発し、不正パケット数の増加からでは検出が困難な、脅威の予兆を検知する技術を

開発した。誤検出 25%、検出漏れ 10%を実現した。(第 5.4 章)

- 不正パケットの送信元、送信先の IP アドレスから構成されるグラフの構造を解析することにより、ワームの感染の効率性に基づく脅威を評価する手法を開発した。本手法により、不正パケット量の増加では検知できない脅威を検知できることを示した。(第 5.5 章)

- 自己相関分析に基づき、不正パケットの時系列データから、過去の周期性と異なる新しいパターンを検出する異常検知手法を開発した。(第 5.6 章)

- インターネット上の不正パケットから検出される脅威の送信元および送信先をリアルタイムに国別に地球儀上に 3 次元可視化するシステムを開発した。これにより地球規模での、脅威の様子の全体像を捕らえ易くなった。(第 5.7 章)

- インターネット上の不正パケットの送信元 IP アドレスの空間上の分布に局所性が存在することを示し、不正パケットの主な原因となるワームの振舞いとの関係について議論した。(第 4 章)

- 脅威分析手法を、攻撃元を集団あるいは個体別に分析する脅威マクロ分析・ミクロ分析として分類した。また、脅威分析手法を、不正パケットの時系列変化あるいは IP アドレス空間上の構造的変化に着目して脅威の時間特徴量分析および空間特徴量分析として分類した。これらの分類に基づき、総合的な脅威分析手法の構成アプローチについて議論した。(第 6 章)

第 2章

インターネット観測システムの概要

インターネット観測システムは、インターネット上の特定の IP アドレスでパケットを観測し、インターネット上の脅威を分析するためのシステムである。本章では、その基盤となる観測センサーに注目して観測法の分類および本研究で用いる観測システムの構成を示す。こられの基盤システムを用いて、脅威を分析するシステムの関連研究動向および本研究の成果は、第 3 章以降でまとめる。

2.1 観測法の分類

不正パケットの観測は、未使用の IP アドレスあるいはクライアントホストを設置した IP アドレスに外部から送信されるパケットを記録するセンサーを構成することにより実現することができる。

不正パケットの観測方法は，観測するパケットに対する応答の有無およびセンサーの配置方法によって分類できる．応答の有無に関しては，外部からのポートアクセスに対して一切応答を返さない受動観測 (passive monitoring) と，特定のパケットに対して応答を返し，その反応を観測する能動観測 (active monitoring) に分けられる．前者として，CAIDA telescope [16], Internet Storm Center [24], Internet Motion Sensor [27]，JPCERT/CC の ISDAS [11], WCLSCAN [39]，DShield [2] などが挙げられ．後者として，Princeton 大学の研究 [18] や, Honeynet Project の Honeypot [22] などが挙げられる．

一方，センサーの配置方法については，連続した IP アドレスを観測する CAIDA telescope [16] などのような連続アドレス型と，不連続なアドレスを観測する Internet Storm Center [24] などのような分散アドレス型に分類することができる．連続アドレス型の場合，連続

的なIPアドレスへのアクセスパターンから攻撃の種類を判別するのに有効であるが，ワームに多く見られる確率的な伝搬パターンを持つ攻撃に対しては，分散アドレス型の方が検知時間について性能が高いことがJohns Hopkins大学の研究 [23] によって示されている．

2.2 本研究における観測システムの構成

本研究では，第2.1節で分類した受動観測型で、分散アドレス配置型の定点観測を行う。

本研究におけるインターネット観測システムおよび脅威分析システムの構成を図2.1に示す．

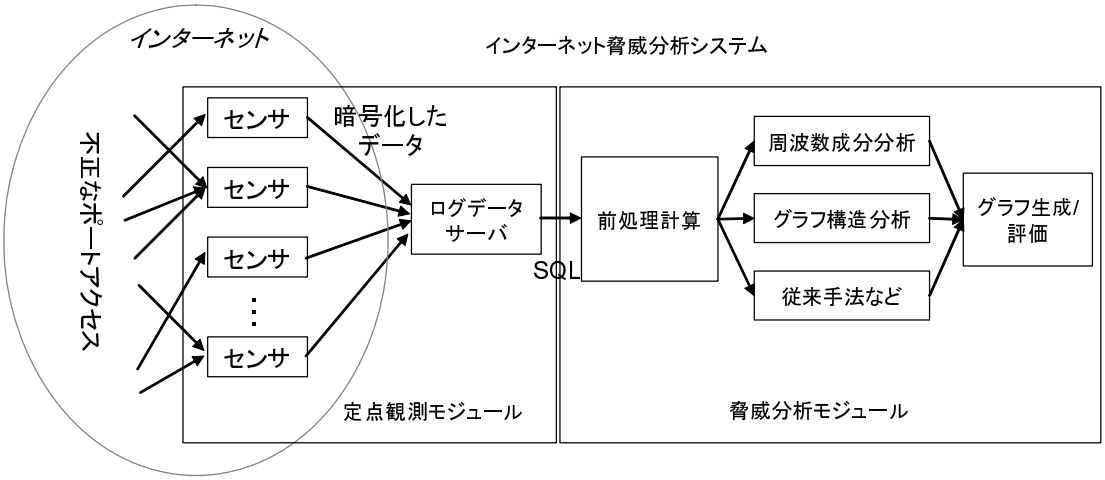

図 2.1: インターネット脅威分析システムの構成

本システムは，不正パケットの観測および観測データを管理する定点観測モジュールと観測データに対して脅威分析を行うモジュールから構成される．インターネット上に配置した複数のセンサで観測された不正パケットはログデータサーバで管理され，脅威分析モジュールで必要となるデータは，SQLを用いて定点観測モジュールから取得し，不正パケットの送受信IPアドレスの構造分析，周波数成分分析や，ベイズ推定法など複数の手法で分析した結果を出力する．

センサーは，アクセスに対しては一切応答を返さない受動観測型である。観測された各パケットについて表2.1に示すデータを記録する．インターネット上で，ネットワークサービスを提供しないIPアドレスに到達するパケットを観測することで，本来，送信されるはずの無いパケットを観測することができる．このようなパケットには，ワームやボットが

送信したパケット，送信元を改竄した DDoS 攻撃の応答パケット (Back Scatter)，ネットワーク機器の設定不備により送信されるパケット，ポートスキャンなどが含まれる．各センサーは，一定の時間間隔で，ログデータサーバに観測データを送信する．

表 2.1: 観測されるデータの属性

パケットのアクセス時刻 (年月日, 時間)
プロトコル種別 (TCP, ICMP, UDP)
送信元 IP アドレス
送信元ポート番号
送信先 IP アドレス
送信先ポート番号

定点観測システムと脅威分析システムは、標準的な SQL を用いてデータのやり取りを行い、相互に独立性の高いシステムとなっている。脅威分析システムでは、分析に必要なデータを SQL を用いて時間やセンサーの種別，パケットの送信元，送信先などの条件を指定して柔軟に取得できる環境を構成している。

第 3章

関連研究

定点観測データの分析手法は，インターネット上の不正パケットの送信元を集団として捉え，集団の特徴に対する統計的な推測を行う集団特徴分析型と，送信元別のアクセスイベントの順列パターンなどから攻撃者の振舞いパターンに注目する振舞い分析型に分類できる．また，集団特徴分析型と振舞い分析型は，それぞれ特徴量自体に時系列情報を持つか否かで，時系列特徴量分析型と空間特徴量分析型 (非時系列特徴量分析型) に分類される．

3.1 マクロ脅威分析

集団特徴分析型では，不正パケット数の時系列データから推定される特徴量と実際に観測される特徴量の統計的な偏差などを評価する．Zou 等は，ワームの拡散モデルに基づき，インターネット上のワームの感染率をカルマンフィルターを用いて推定し，感染率の推定値が一定以上の値で収束する場合を脅威と見なす手法を提案している [31]．Lakhina 等は，送信先に対するパケット数の時系列データに対して主成分分析を適用し，パケット頻度を主要成分と残差成分に分離し，残差成分 (異常成分) の増減によって異常を検知する方法を提案している [14]．また，不正パケット数の時系列データから統計的な偏差 (Z スコア) に対してベイズ推定を適用することにより危険度に対する推定値を学習する方法 [8] や，パケット頻度の自己回帰分析による推定値からの分布の偏差に対してシャノン情報量に基づき変化点の検出を行う方法 [29, 33] など統計分布に基づく分析などがある．Wagner 等は，不正パケットのアクセス先の分布をエントロピーで評価することにより，分布のランダム性や偏りに関する変化からワームの発生を検知する方法 [28] を提案している．これらの手法は，時系列データから特徴量を推定し，観測値とのずれを評価しているが，個々の特徴

量自体に時系列性を持たないため，ここでは空間特徴量分析型に分類する．

3.2 ミクロ脅威分析

振舞い分析型に関しては，送信元別に観測される送信先ポートへのアクセスパターンや，それらのクラスタリングにより，これまでに観測されていない新しいパターンの発見を行うものがある．振舞い分析型のうち，空間特徴量分析型のものには以下のような方法がある．Theriault 等は，送信元 IP 別に送信先ポートの分布に対して距離を定義することで，送信元に対するクラスタリングを行い，クラスタ構成の変化により異常を検知する方法 [26] を提案している．能動観測データを用いたものには，TCP コネクション確立の有無を観測することで，ワーム感染の尤度を求める方法 [25] や，送信元 IP アドレスごとに SYN パケットと FIN パケットの数を計測し，その SYN パケットと FIN パケットの差から攻撃を検知する方法 [13] が提案されている．一方，振舞い分析型で，時系列特徴量分析型には，送信元 IP アドレスごとに，送信元ポート番号と送信先 IP アドレスの関係を時間を追ってグラフ表示することにより，個々の送信元のパケット送信パターンを視覚的に捉える研究 [37] がある．

本研究では，集団特徴分析のうち特徴量自体に時系列情報を持つウェーブレット解析法と，空間特徴量分析で，集団特徴分析と振舞い分析の両方を性質を持つ不正パケットのグラフ構造分析に基づく手法について提案する．

第 4章

インターネット背景放射パケットの基礎分析

本章では、インターネット上の脅威分析手法の開発に必要となる背景放射パケットの分布等に関する基礎分析およびその原因について考察する。

本研究で用いる受動観測型のインターネット観測システムで観測されるパケットは、正規のネットワークサービスに対する要求パケットでは無く、ワーム等から送信される攻撃パケット、スキャンパケット、ネットワーク機器の設定不備により送信される非正規パケットなどから構成される。これらの非正規パケットの傾向を分析することにより、それらの種別の構成を推測し、効果的な脅威検知手法の開発に役立てる。

ワームの実装コードや感染動作などに関する研究 [4, 34] から、多くのワームは、自分に近い IP アドレスを優先的に確率的な探索を行うことが知られている。ワームに用いられる多くの感染探索戦略のうち、確率的な近傍優先探索法がインターネット上の環境でもっとも効率的であることが示されていること [30] からも、近傍優先探索が、ワームの感染戦略の主流であることが予想される。

第 4.1 章では、センサーと背景放射パケットの IP アドレス上の距離分布について分析し、第 4.2 章では、送信元 IP アドレスのアドレスブロック分布について分析し、第 4.3 章では、背景放射パケット数の時間周期性について分析する。第 4.4 章では、上記の分析結果およびインターネット観測システムの観測方式をもとに、観測パケットの原因について考察する。

4.1 送信元IPアドレスの距離分布に関する分析

本章では，ワームの近傍優先探索法に関する仮説について考察するために、観測される非正規パケットの送信元IPアドレスとセンサーIPアドレスの距離の分布について観測データにより検証する．

4.1.1 プロトコル別IPアドレス距離分布

図4.1は，2005年4月の1ヶ月間に4つのセンサーで観測されたパケットに関して，プロトコル別に発信元IPアドレスとセンサーIPアドレス間の距離の分布求めたものである．横軸は，センサーIPアドレスと発信元IPアドレスに関して上位ビットから連続して一致するビット数を示し，縦軸は，それら上位ビットが一致するパケットの全体比を示している．各系列は，TCP, UDP, ICMPプロトコルを対象としたもの，TCPパケットから，DDoSのバックスキャッタと考えられるパケット[1]を除いたもの，そして，IPアドレス空間全体で一様ランダムにアクセスした場合に観測されると考えられる理論値を示している．上位1ビットが一致するアドレスは，IPアドレス空間全体の半分を占めるため，横軸1の場合，一様ランダムアクセスの全体比は0.5を示し，上位一致ビットが増加するに従い，全体比は半減する曲線を描く．

TCPの曲線は，横軸8ビット，16ビットの地点で急激な変化が見られる．これは，/8ネットワーク，/16ネットワーク内の発信元アドレスからアクセスが集中していることを示している．TCPによるワーム感染の場合，接続を確立してからワーム・コードを送信する必要があるため，発信元IPアドレスをスプーフすることはないと考えられる．したがって，この結果は，ワームは効率的な感染を行うために，自分に近いIPアドレスから優先的に探索することによると考えられる[2]

ワームの局所的な探索比率を，近似モデルによって求めると以下のようになる．感染したワームは，IP空間上に一様にαの割合で存在し，すべてのワームが，/16ネットワークを一様にアクセスする比率をp, IP空間全体を一様にアクセスする比率を$1-p$とすると簡単化する．この時，センサーの属する/16ネットワークに存在するワームから送信され

[1] 発信元のポートがwell-knownサービスであるものは，発信元IPをランダムにスプーフしたパケットをネットワークサービス送信した結果SYN-ACKパケットが返ってきたものを観測したと考えることができ，DDoSのバックスキャッタパケットである可能性が高いと考えられる．

[2] ISPによるingressフィルタリングにより，センサーの属するネットワークブロック以外のパケットが破棄さることによる影響も考えられるが，ここではingressフィルタリングの影響は考慮しない．

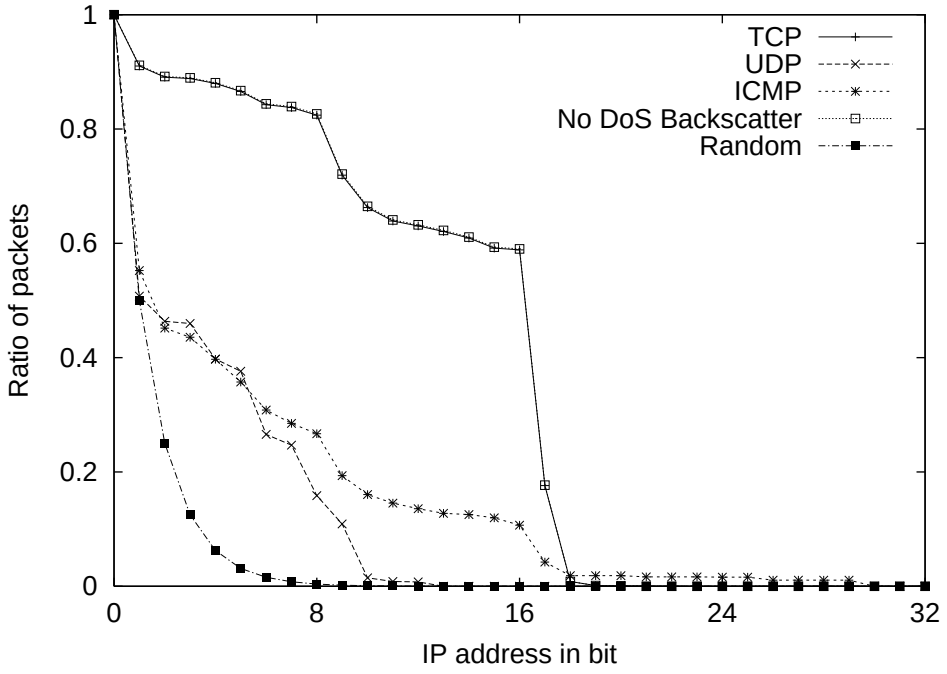

図 4.1: プロトコル別の発信元-センサー IP アドレス距離の分布

るパケットの観測数と，/16 ネットワークの外に存在するワームから送信されるパケットの観測数の比率は以下のようになる．

$$\frac{p}{2^{16}} \cdot 2^{16}\alpha : \frac{(1-p)}{2^{32}-2^{16}} \cdot (2^{32}-2^{16})\alpha = p : 1-p \tag{4.1}$$

表 4.1 は，アクセスパケットのうち，発信元アドレスとセンサーアドレスの上位ビットが一致する比率を示している．TCP, UDP, ICMP, Excluding DDoS, Random は，それぞれ TCP パケット，UDP パケット，ICMP パケット，DDoS の backscatter が中心と考えられる Well-known ポートからのアクセスパケットを除外したもの，IP 空間全体に対して一様ランダムにアクセスした場合の比率を示している．

TCP について上位 16 ビットが一致する比率は，0.589 であった．ワームが/16 ネットワーク内に優先的にアクセスする比率は 58.9% となる．TCP について上位 8 ビットが一致する比率は，0.824 である．/8 ネットワークについても上記と同様な近似式が成り立つため，ワームが/16 ネットワークの外部で/8 ネットワークの内部を優先的に探索する回数の比率は，/$82.4\% - 58.9\% = 23.5\%$ となる．この結果は，CodeRed, Nimda, Sasser 等のコード解析 [19, 1, 5] に基づく局所探索の比率の結果にほぼ一致する．

UDP, IMCP は，TCP の場合に比べて，局所的なアクセス傾向が低い．これは，UDP, ICMP は，接続を確立する必要がないため，IP スプーフィングが可能であることや，UDP, ICMP は接続確立のオーバーヘッドが小さいため，IP 空間全域のパケット送信も行いやすいことが原因と考えられる．

表 4.1: アクセスパケットのうち発信元アドレスとセンサーアドレスの上位ビットが一致する比率

	TCP	UDP	ICMP	Excluding DDoS	Random
1	0.910	0.508	0.552	0.911	0.500
2	0.891	0.464	0.452	0.892	0.250
3	0.888	0.460	0.436	0.889	0.125
4	0.880	0.398	0.397	0.881	0.063
5	0.866	0.376	0.357	0.867	0.031
6	0.843	0.266	0.308	0.844	0.016
7	0.838	0.247	0.285	0.839	0.008
8	0.824	0.159	0.267	0.827	0.004
9	0.719	0.109	0.194	0.721	0.002
10	0.663	0.015	0.160	0.665	0.001
11	0.639	0.008	0.145	0.641	0.000
12	0.631	0.008	0.136	0.632	0.000
13	0.621	0.000	0.128	0.623	0.000
14	0.609	0.000	0.125	0.611	0.000
15	0.592	0.000	0.120	0.593	0.000
16	0.589	0.000	0.107	0.590	0.000
17	0.176	0.000	0.042	0.177	0.000
18	0.008	0.000	0.019	0.008	0.000
19	0.001	0.000	0.019	0.001	0.000
20	0.001	0.000	0.019	0.001	0.000

21	0.000	0.000	0.016	0.000	0.000
22	0.000	0.000	0.016	0.000	0.000
23	0.000	0.000	0.016	0.000	0.000
24	0.000	0.000	0.016	0.000	0.000
25	0.000	0.000	0.016	0.000	0.000
26	0.000	0.000	0.011	0.000	0.000
27	0.000	0.000	0.011	0.000	0.000
28	0.000	0.000	0.011	0.000	0.000
29	0.000	0.000	0.011	0.000	0.000
30	0.000	0.000	0.000	0.000	0.000
31	0.000	0.000	0.000	0.000	0.000
32	0.000	0.000	0.000	0.000	0.000

4.1.2 センサー別 IP アドレス距離分布

図 4.2 は，2005 年 4 月の 1ヶ月間に観測された TCP パケットに関して，センサー別に発信元 IP アドレスとセンサー IP アドレス間の距離の分布求めたものである．横軸，縦軸は，図 4.1 と同じである．

センサーで観測されるパケット数は，有意な差があり，2005 年 4 月のパケット総数は，下表の通りである．センサー 3 は，アクセス数が極端に少なく，設置場所の関係から周りの IP アドレス利用率が低いことが分かっている．

センサー	パケット数
センサー 1	11185
センサー 2	62885
センサー 3	750
センサー 4	43762

図 4.2 では，センサー 3 は，局所的なアクセスが少ないことが確認される．これは，周辺の IP アドレス使用率が低く，/16 ネットワーク内にホストがあまり存在しないため，/16 ネットワーク外からのアクセスが相対的に高いためであると考えられる．

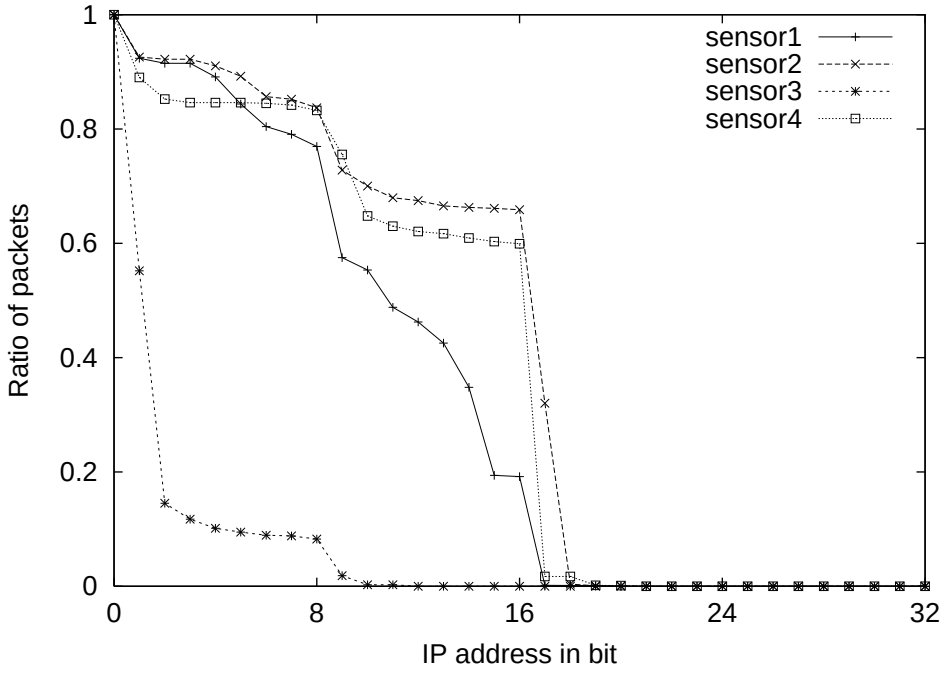

図 4.2: センサー別の発信元-センサー IP アドレス距離の分布

4.1.3 ポート別 IP アドレス距離分布

図 4.3 は，2005 年 4 月の 1ヶ月間に観測された TCP パケットに関して，アクセス頻度上位 3 つのポートごとに発信元 IP アドレスとセンサー IP アドレス間の距離の分布求めたものである．横軸，縦軸は，図 4.1 と同じである．

ポート 135, 445 は，図 4.1 で確認できる TCP の全体平均で見た場合の /16 ネットワークへの局所的なアクセスよりも強く /16 ネットワークに集中していることが確認できる．一方，ポート 1433 アクセスは，TCP 平均よりも局所的なアクセスが低いことが確認できる．

以上のような局所的なアクセスの傾向は，2004 年 11 月，2005 年 1 月のデータに関しても同様の傾向が確認できた．

4.2 送信元 IP アドレスの分布に関する分析

本章では、脅威検知のための観測センサーの最適配置などについて考察するために、非正規パケットの送信元 IP アドレス分布について分析する。

ここでは、定点観測システムで観測される不正パケットのうち、送信元 IP アドレスがあ

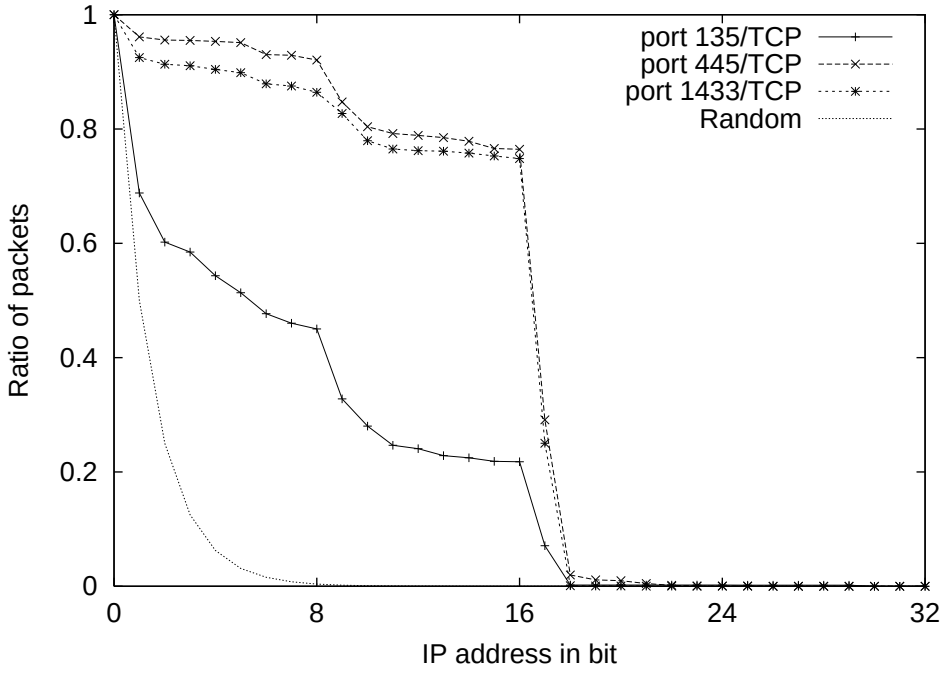

図 4.3: ポート別の発信元-センサー IP アドレス距離の分布

る程度信頼できる TCP パケットを対象として、送信元 IP アドレスの分布の分析を行う。不正パケットのうち最も多く観測されるのは、ワームの感染活動のためのパケットである。TCP 接続によるワームの感染パケットは、感染のために接続を確立する必要があるため、送信元 IP アドレスが改竄されることはないと考えられる。また、観測される TCP パケットのうち、DDoS 攻撃のランダムな反射パケット (Back scatter) と判断されるものを取り除くことで、ノイズの影響を最小限に抑える。本研究では、これらを分析対象とする不正パケットとする。

4.2.1 送信元 IP アドレスの分布

不正パケットの送信元 IP アドレスの分布を捉えるために、不正パケットの送信元 IP アドレスまたはネットワークブロック単位での分布を視覚化する。IP アドレス空間は非常に大きいため、送信元アドレス分布を可視化するために、IP アドレスを構成する 4 つのオクテットを上位から順に 2 つづつ抽出した 2 次元空間を定め、各座標に該当するアドレスあるいはネットワークブロックからの不正パケット数を濃淡で表示する。

図 4.4 から図 4.6 は、2006 年 4 月 1ヶ月間の不正パケットについて、センサーごとに送信元 IP アドレスの分布を表示したものである。図 4.4 は、第 1 第 2 オクテットをそれぞれ横軸縦軸に対応させた 2 次元平面上の各座標に位置するネットワークブロックから送信された不正パケット数の頻度を表示している。2 次元平面上の各座標は、1 つの /16 ネットワーク (2^16 個の IP アドレス) に対応している。送信元 IP アドレスごとのパケット数は大きな差があるため、視覚的に捉えやすいように頻度の対数値によって濃淡をつける。図 4.5 は、センサーと第 1 オクテットが一致するパケットに関して、第 2 第 3 オクテットをそれぞれ横軸縦軸に対応させた 2 次元平面上の各座標に対応するネットワークブロックから送信された不正パケット数の頻度を表示している。2 次元平面上の各座標は、1 つの /24 ネットワーク (2^8 個の IP アドレス) に対応している。図 4.6 は、センサーと第 1 第 2 オクテットが一致するパケットに関して、第 3 第 4 オクテットをそれぞれ横軸縦軸に対応させた 2 次元平面上の各座標に対応する IP アドレスから送信された不正パケット数の頻度を表示している。

図 4.4 から図 4.6 は、いずれも不正パケットに明瞭な偏りが見られる[3] 。図 4.4 は、インターネット全体の不正パケット分布を示している。この場合、観測したセンサーに因らず、不正パケットは、第 1 オクテットが 60 前後および 210 前後といった特定のアドレスブロックに偏っていることが確認される。これらのアドレスブロックは、APNIC, ARIN などによって一般に割当てられた IP アドレスであり、IP アドレスの利用率が高く、脆弱なホストも多いことが推測される。図 4.5 の 4 つのグラフは、それぞれインターネット全体のうちセンサーが存在する /8 ネットワークを示している。このグラフでは、センサーごとに異るパターンが確認できる。図 4.6 の 4 つのグラフは、それぞれインターネット全体のうちセンサーが存在する /16 ネットワークを示している。このグラフでは、不正パケットの送信元 IP アドレスがグラフ上から直接確認することができる。センサーごとに異るパターンを示しているが、そのパターンはセンサーの位置に因らず、周囲の感染ホストの分布によって決る。第 3 オクテットを単位として、縦方向に明瞭な筋が見られる。

4.2.2 不正パケット間の関連

同じ送信元 IP アドレスから送信されるパケットの関連性を分析するために、連続する 2 つの期間において、同一の送信元 IP アドレスから送信されるパケットの送信先ポートの種類の一致率について分析した。同一のワームから発信された相関の高いパケットであれば、

[3] 別の実験で、年間を通して季節に因らない一定の偏りを確認している。

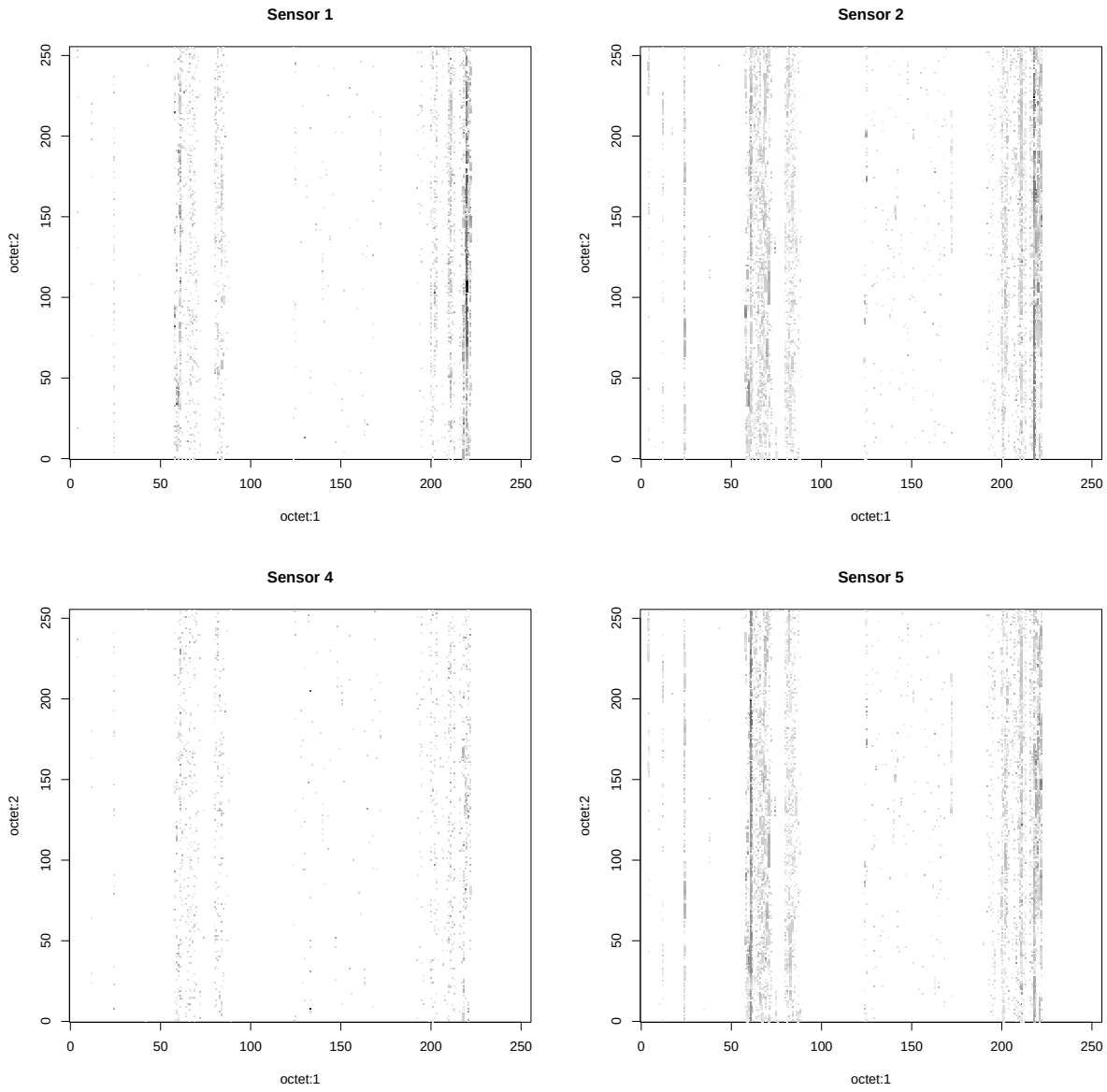

図 4.4: センサーごとの不正パケットの第 1 第 2 オクテット空間における分布

攻撃先のポート種類の一致率は高いと考えられる。

表 4.2 から表 4.4 は、2 つのオクテットによって決る送信元 IP から送信されるパケットの送信先ポートの種類の一致率を 3 通りに分け、それぞれの一致率を満たす送信元数の全体の比率を示している。この結果からは、送信元を示すネットワークブロックの単位が大きい程、一致する比率が高いことが確認される。このことから、ある程度の IP アドレスのまとまりを単位として解析することで、パケット間の相関を活用することが出来る可能性

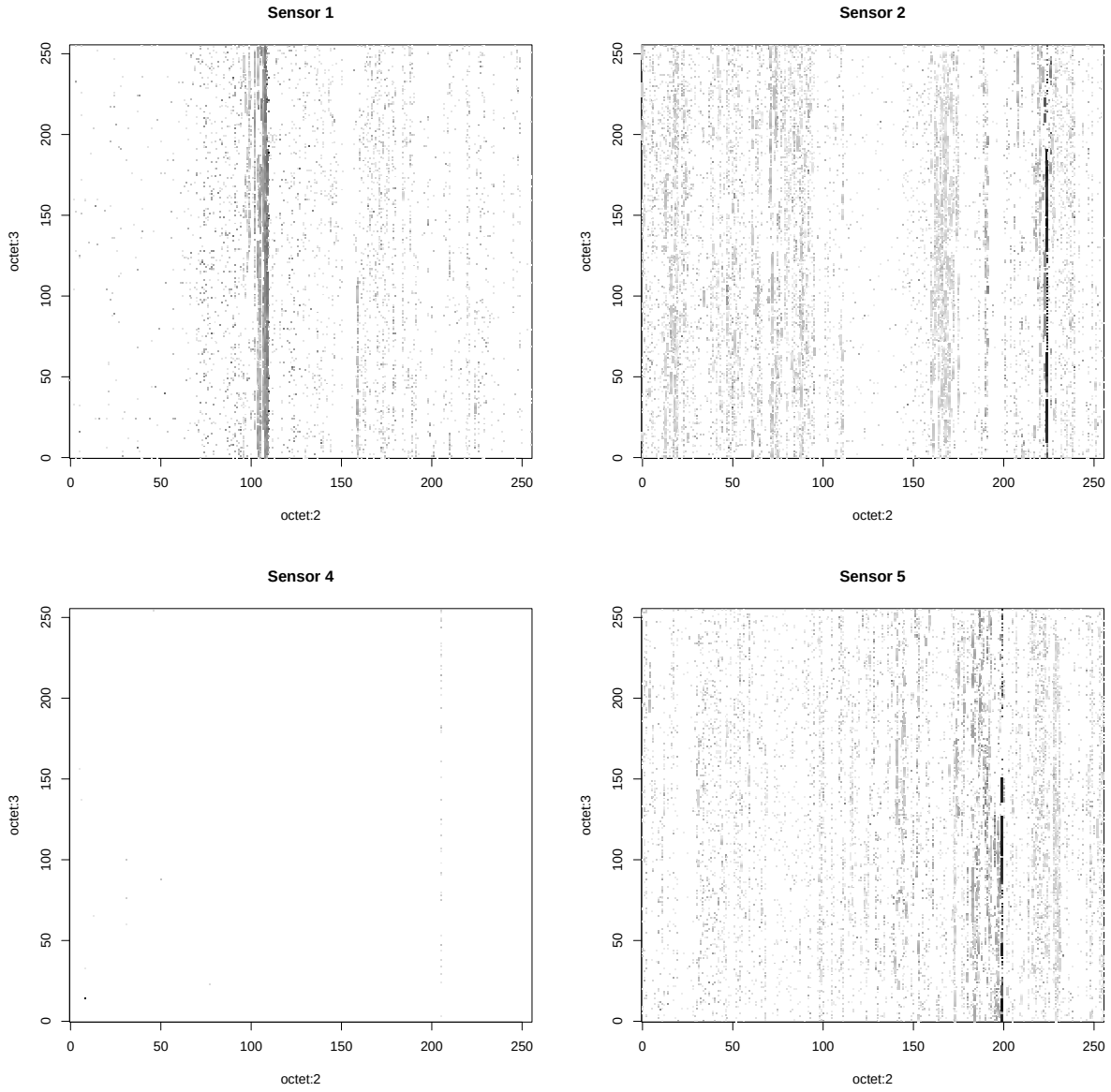

図 4.5: センサーごとの不正パケットの第 2 第 3 オクテット空間における分布

あると考えられる。

4.3 非正規パケットの周期性分析

インターネット脅威分析のために非正規パケットの時系列変化および周期性について分析する。

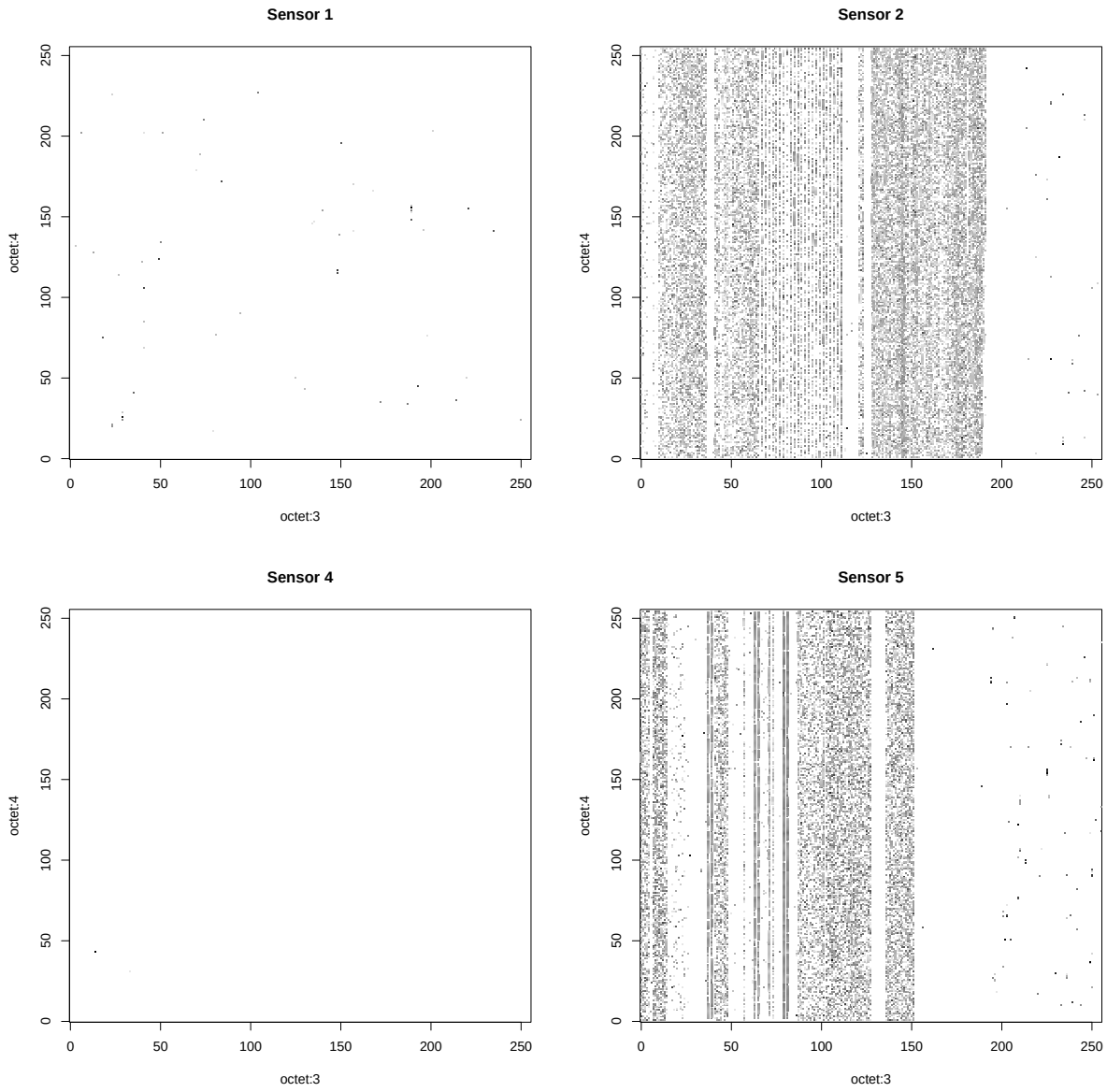

図 4.6: センサーごとの不正パケットの第 3 第 4 オクテット空間における分布

4.3.1 アクセス頻度相関

図 4.7 は，2005 年 4 月 7 日から 1 週間のアクセス頻度がもっとも多い TCP ポート 135 番に対する 1 時間単位の頻度時系列である．このグラフからは，1 日を周期とする時間周期性が確認される．

これは，インターネット IX の JPIX が公開する IX バックプレーンのトラフィック量でも同様の傾向が見られることから，インターネットを利用するコンピュータ自体の稼働率

表 4.2: 第 1 第 2 オクテットグラフ (/16 単位)

ポート種別一致基準	比率
ポート種別 100%一致	5.73%
ポート種別 70%一致	16.38%
ポート種別 1 つ以上一致	18.86%

表 4.3: 第 2 第 3 オクテットグラフ (/24 単位)

ポート種別一致基準	比率
ポート種別 100%一致	7.47%
ポート種別 70%一致	12.50%
ポート種別 1 つ以上一致	12.62%

表 4.4: 第 3 第 4 オクテットグラフ (IP 単位)

ポート種別一致基準	比率
ポート種別 100%一致	1.30%
ポート種別 70%一致	2.40%
ポート種別 1 つ以上一致	2.40%

に大きな影響を受けていると考えられる．

実際，TCP135 番の 1 時間単位の頻度系列に対する時間相関からも周期性が確認できる．図 4.8 2005 年 4 月 1ヶ月間の 1 時間間隔の頻度系列に対して，1 時間から 25 時間まで順に時間をずらした頻度系列に対して相関係数を求めたコレログラムである．横軸は，時間のずれを示し，縦軸は，相関係数を示している．このグラフから，時間のずれが 0 時間，24 時間の近辺で，相関係数が高く，12 時間前後の付近で，負の相関が高いことが分かる．

TCP135 に限らず，多くのポートにおいて平常時にはこのような時間周期性が確認される．これは，ワームの発生源であるインターネット上のパソコンの利用時間帯に周期性があることが原因と考えられる．

表 4.5 は，2005 年 4 月 1ヶ月の 1 日単位の頻度系列に関して，4 つのセンサーの全組合わ

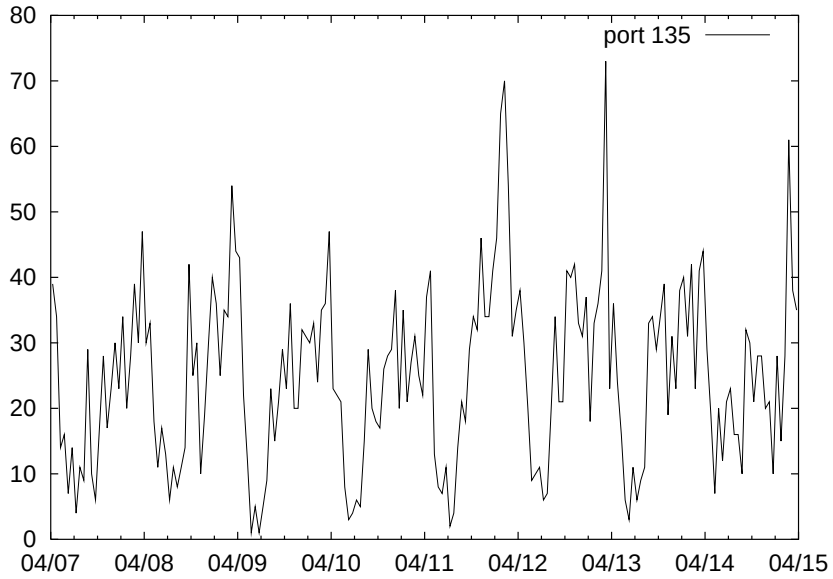

図 4.7: ポート 135/TCP のアクセス頻度の時系列変化 (1 時間単位)

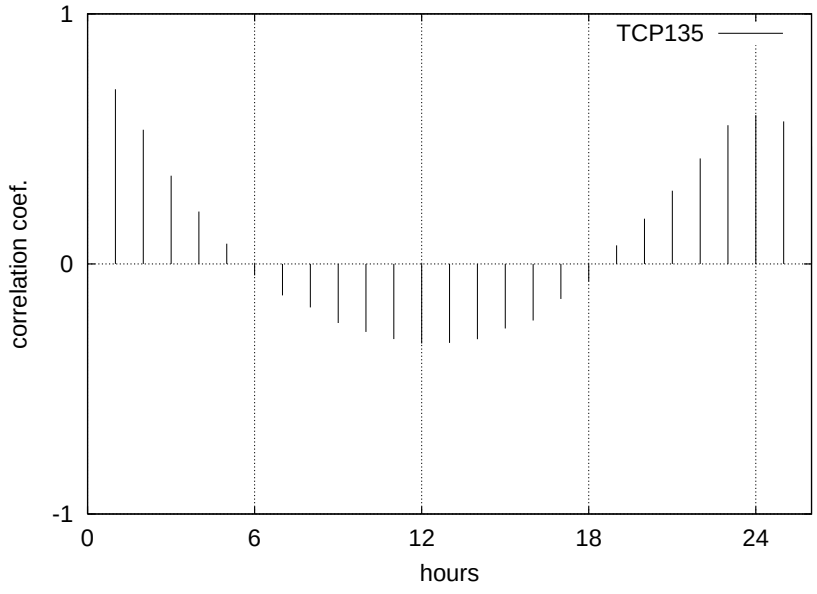

図 4.8: ポート 135/TCP のアクセス頻度に関する自己相関関数

せに対する相関係数である．

表 4.5: センサー間のアクセス頻度相関 (1 日単位)

	センサー 1	センサー 2	センサー 3
センサー 2	0.259	–	–
センサー 3	0.123	0.161	–
センサー 4	0.034	0.105	-0.073

この結果から，センサー間の相関係数は，絶対値が 0.259 以下と低いことが分かる．一方の頻度がもう一方のセンサーで説明できる割合を示す決定係数が，相関係数の 2 乗から求められることを考慮すると低いと言える．

4.3.2 ポート間相関

表 4.6 は，2005 年 4 月 1ヶ月の 1 日単位の頻度系列に関して，アクセス頻度上位の 5 つのポートに関して，相関係数を求めたものである．ポート 135, 445 の組合わせおよびポート 1025,1433 の組合わせの相関は比較的高いがそれ以外のポートの間の相関は低いことが確認できる．

表 4.6: ポート間のアクセス頻度相関 (1 日単位)

	ポート 135	ポート 445	ポート 1433	ポート 139
ポート 445	0.435	–	–	–
ポート 1433	0.378	-0.155	–	–
ポート 139	0.212	0.015	0.027	–
ポート 1025	0.211	-0.008	0.476	-0.259

4.3.3 発信元 IP 別アクセス傾向

表 4.7 は，単一ソース IP アドレスからアクセスされるポート数のヒストグラムを示している．対象データは 2004 年 11 月 11 日の TCP アクセスである．この表から，9 割り (88%) 程度の IP アドレスからの送信先が単一ポートであることが確認できる．

表 4.7: 単一ソースアドレスからの送信先ポート数のヒストグラム

送信先ポート数	件度	割合
1	1308	88.0 %
2	125	8.4 %
3	25	1.6 %
4 以上	5	1.8 %

図 4.9 は，アクセスされたポートごとに，発信元の IP アドレス数のヒストグラムを示している．対象データは 2004 年 11 月 11 日の TCP アクセスである．このグラフから，発信元 IP アドレスが 1 であるものがもっとも多いが，比較的散らばっている．アクセスを受けるポートの発信元ユニーク IP 数が 1〜15 のものは，全体の 85%であった．

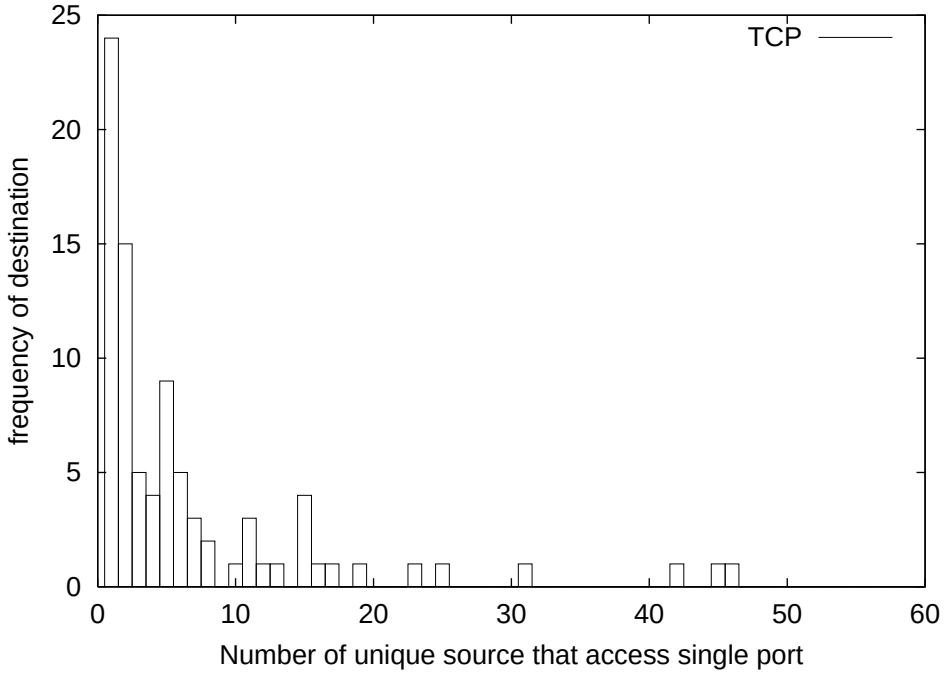

図 4.9: アクセスされたポートごとの発信元の IP アドレス数のヒストグラム

4.3.4 送信先ポート集合の時系列変化

同一ソース IP アドレスからアクセスされるポートの集合は，ソース IP のホストに感染したワームによる送信先ポートの集合を表していると考えることができる．必ずしも，単一のワームから送信されたパケットとは限らず，複数のワームに感染した PC からのアクセス集合の場合も考えられるが，ソースのワームを分類する上で有効な情報を示していると考えられる．

そこで，PC からのアクセスの単位周期と考えられる 1 日を単位として，同一ソース IP アドレスから送信された送信先ポートの集合を求め，集合が一致する件数の時系列変化について分析する．

図 4.10 は，2005 年 4 月の観測データに対して，送信先集合が 2 以上の場合について同一アクセス集合となるソース IP アドレスの件数の時系列変化を示している． 横軸は 2005 年 4 月の日付，縦軸は件数を示している．もっとも件数の多いアクセス集合は，{445/TCP, 135/TCP} である．

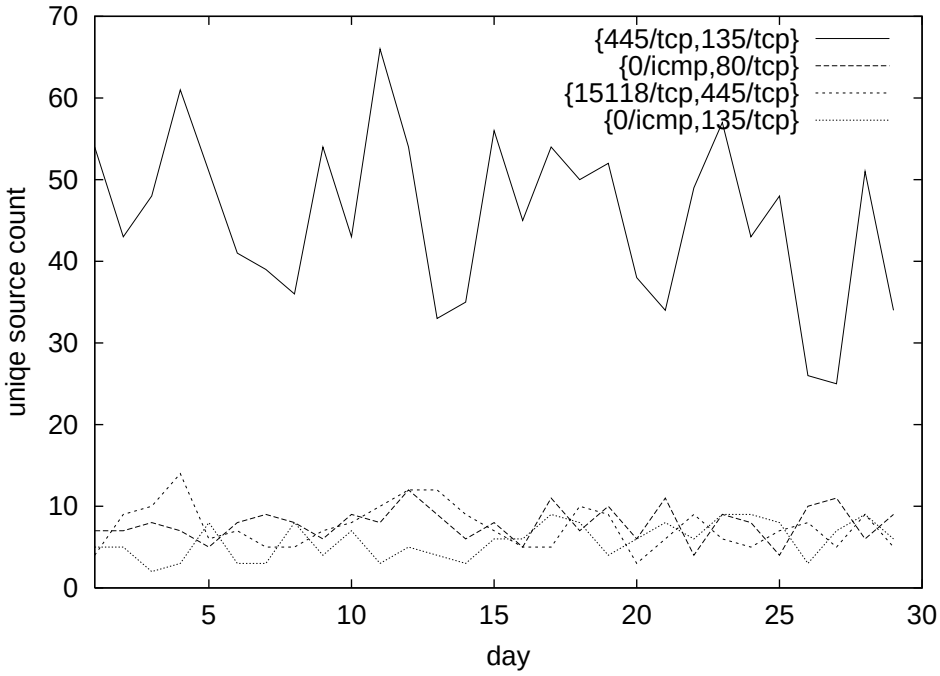

図 4.10: 同一ソース IP アドレスからの送信先ポート集合ごとの頻度時系列変化

表 4.8 は，同一ソース IP アドレスからの送信先ポート集合ごとの上位占有率を示している．135/TCP, 445/TCP など単一ポートへのアクセスを行うソースが大部分を占めること

が分かるが，複数のポートにアクセスする {445/tcp,135/tcp} が全体の 3.5%を占めている．

表 4.8: 同一ソース IP アドレスからの送信先ポート集合ごとの上位占有率

送信先ポート集合	比率 (%)
135/tcp	43.1
445/tcp	19.5
1433/tcp	7.3
445/tcp,135/tcp	3.5
4899/tcp	3.3
139/tcp	2.9
0/icmp	2.7
1025/tcp	2.3
1434/udp	2.2
137/udp	1.6
22/tcp	0.8
80/tcp	0.7
0/icmp,80/tcp	0.6
15118/tcp,445/tcp	0.5

4.4 背景放射パケットの原因分析とセンサー配置

第 4.1 章の分析により、非正規の TCP パケットは IP アドレス空間上の一様分布から極めて偏った局所的な分布を示していることが確認できた。また、局所分布の確率勾配は、ワームの実装コードや感染動作などに関する研究 [4, 34] に示されるように、ワームの近傍優先探索の確率勾配 (/16 ネットワーク内を 50%, /8 ネットワーク内を 25%、残りインターネット全体 25%) と極めて近い。また、警察庁 cyber police のインターネット観測システムによる報告 [20] における IDS のワーム検知ルールに基づく情報も合せて考えると、インターネット上の非正規パケットの大部分がワームによる感染パケットであると推定される。

第 4.1 章における UCP, IMCP パケットに関する距離分布に関しても、TCP パケット程

ではないが、インターネット上の一様分布に比較して、分布の大きな偏りが確認されるため、同様にワームの感染パケットが中心であると推定される。

一方、第 4.2 章では、非正規パケットの送信元 IP アドレスは、インターネット空間全体のうち、IP アドレスの利用率が高いネットワークブロックに極めて偏っていることが確認された。

以上のことから、ワーム等の感染パケットを早期に検知するためには、感染戦略が、確率的近傍探索法を前提とすれば、IP アドレスの 2 ビット表記の内、上位ビットができるだけことなる IP アドレス空間で、かつ、IP アドレスの割り当て率の高いネットワークブロックに優先的にセンサーを設置することが有効である。

第 5章

インターネット脅威検知手法

本章では、インターネット上の非正規パケットの観測に基づく脅威検知に関して本研究で開発した手法について述べる。

5.1 研究課題の抽出および問題解決の流れ

本研究における課題の抽出および問題解決の流れを整理する。第 1.1 章の背景に示した通り、インターネット上の攻撃は年々巧妙化している。特に、不正パケットの量が、脅威レベルに必ずしも対応しなくなるなど、不正アクセス等の脅威の検出は困難になってきている。本研究では、脅威の定義および脅威モデルを明確化することにより、攻撃の進化に影響されない、一般性の高い脅威検知手法を開発することを課題とする。

第 5.2 章に示す通り、脅威モデルを明確化することにより、不正パケットの増加だけでは捕らえられない、本質的な脅威を検出する手法を検討した。具体的には、ワームの感染力に注目し、不正パケットの特徴量から、脅威を評価するための手法として開発、ベイズ推定脅威分析法、グラフ構造脅威分析法を示す。また、ワームの感染力に関係の深い、新種のワームの検出を目的とした異常検知手法として、周期成分異常検知法、自己相関異常検知法、パターンマイニング異常検知法を示す。これらの手法はそれぞれ、検知カバー域、誤検知、必要とされる観測データ量、検知速度の面で問題点を持つ。また、インターネット上では多様なワームによる脅威が同時多発的に存在するため、単一の手法では、脅威検知が困難である。実際のインターネット脅威検知のためには、それぞれの特徴を生かした脅威検知手法を同時適用することで、検知カバー域の広い、高速な検知を行う必要性があることを示す。

以上のような問題解決の流れをまとめたものが図 5.1 である。

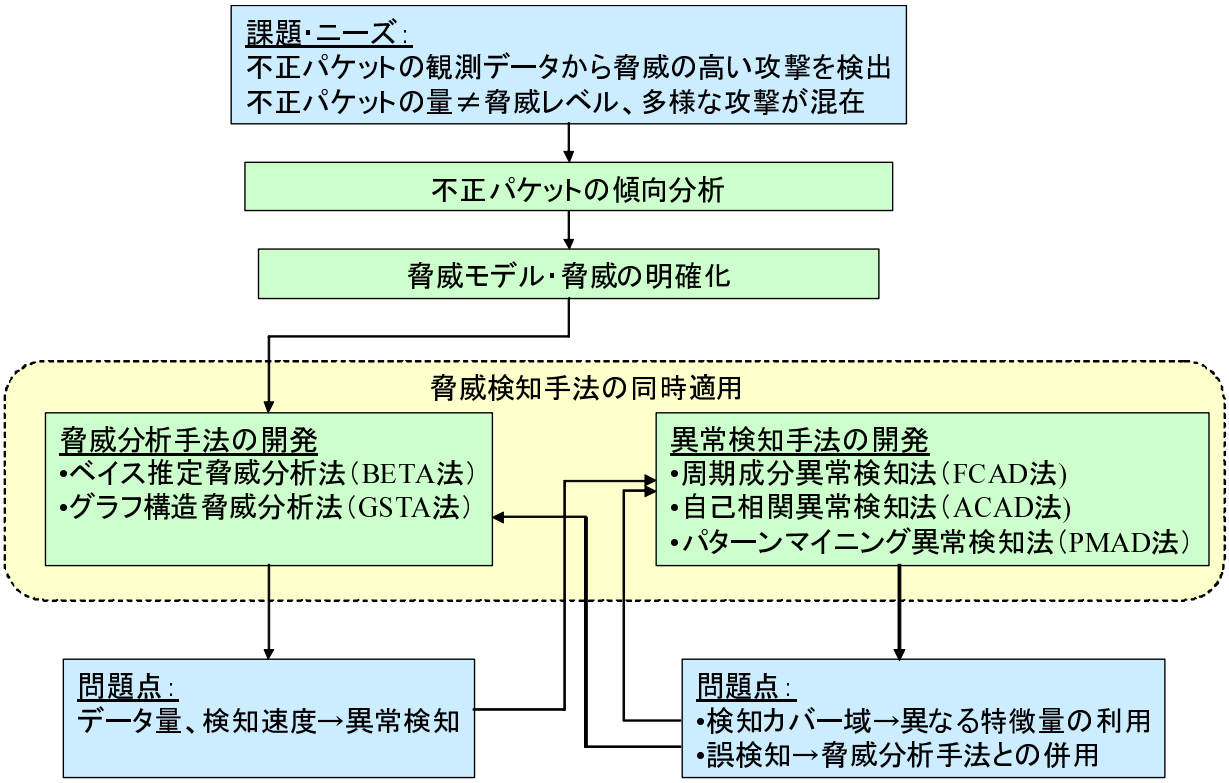

図 5.1: 研究課題の抽出と問題解決の流れ

5.2 脅威モデル

インターネット観測システムで観測される不正パケットには、ワーム、ポートスキャン、ネットワーク機器の設定不備によるパケットなど様々なものがある。警察庁@police が公開するインターネット定点観測のデータ (例：図 5.2) から、これらの不正パケットのうち、年間を通じて、ワームによるパケットが 9 割を占めていることが確認できる。

ワーム以外の不正パケットについては、大多数を占めるワームに混在するため、不正パケットのアクセスパターンが特定される既知の攻撃の検知は可能である。一方、大多数を占めるワームの感染探索戦略は、近傍の IP アドレスを優先し、ランダムに探索する戦略を持つ [19, 1, 5](以下、局所選好ランダム探索戦略と呼ぶ)。そのため、新しいアクセスパターンを持つワーム以外の攻撃を、膨大なワームのランダムな不正パケットの中から分離することは困難である。

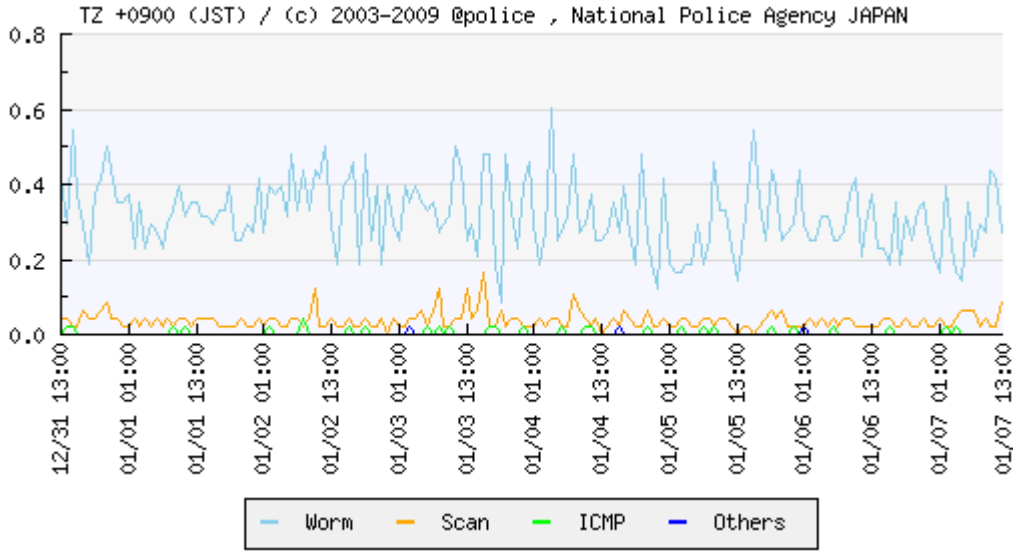

図 5.2: 不正パケットの種類別の時系列変化

そこで、本研究では、インターネット脅威分析システムが検出の対象とする脅威としてワームを中心に考える。

5.2.1 ワーム感染の数理モデル

ワーム感染の生物学的考察は、1911 年の単純感染モデル (Simple Epidemic Model), 1927 の一般化感染モデル (General Epidemic Model) などにより行われている [12, 15]。 インターネット上のワーム感染においても、ワームとなる感染ホストと、被感染対象となる脆弱性を持つホストの関係から、類似の考察をすることができる。ここでは、インターネットワーム感染の数理モデルについて Medleck の文献に基づき [15] まとめる。

単純感染モデル (SI モデル)

まず、時刻 t におけるインターネット上のホストを以下の 2 つのグループに分ける：

- 脆弱なホスト数 $S(t)$
- 感染ホスト数 $I(t)$

この時、以下の前提を考える：

- 全ホスト数は、脆弱なホストと感染ホストの和

$$N = S(t) + I(t) \tag{5.1}$$

- 感染率は、脆弱ホスト数に比例する。つまり、感染率 $\lambda = kI(t)$ ただし、k は、比例定数。

以上のとき、以下の 2 つの微分方程式が感染モデルを定義付ける（図 5.3 参照)。

$$\frac{dS}{dt} = -kI(t)S(t) \tag{5.2}$$
$$\frac{dI}{dt} = kI(t)S(t) \tag{5.3}$$

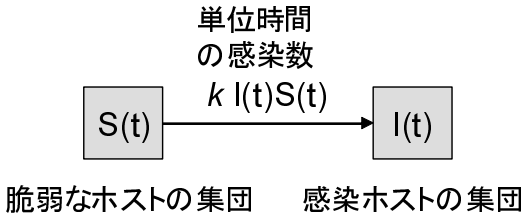

図 5.3: 単純感染モデル

式 (5.1) より、式 (5.2) は、式 (eq:epidemic-model2) の方程式と等価である：

$$S(t) = N - I(t) \tag{5.4}$$
$$\frac{dI}{dt} = kI(t)(N - I(t))$$

この方程式は、ロジスティック成長方程式として知られており、$I(t)$ は、以下のロジスティック式として得られる (図 5.4 参照)。

$$I(t) = \frac{I(0)N}{I(0) + (N - I(0))e^{-kNt}} \tag{5.5}$$

一般化感染モデル (SIR モデル)

一般化感染モデルを用いれば、感染ホストから復旧ホストへの移行を含むより現実的なモデル化が可能である。

まず、時刻 t におけるインターネット上のホストを以下の 3 つのグループに分ける：

- 脆弱なホスト数 $S(t)$

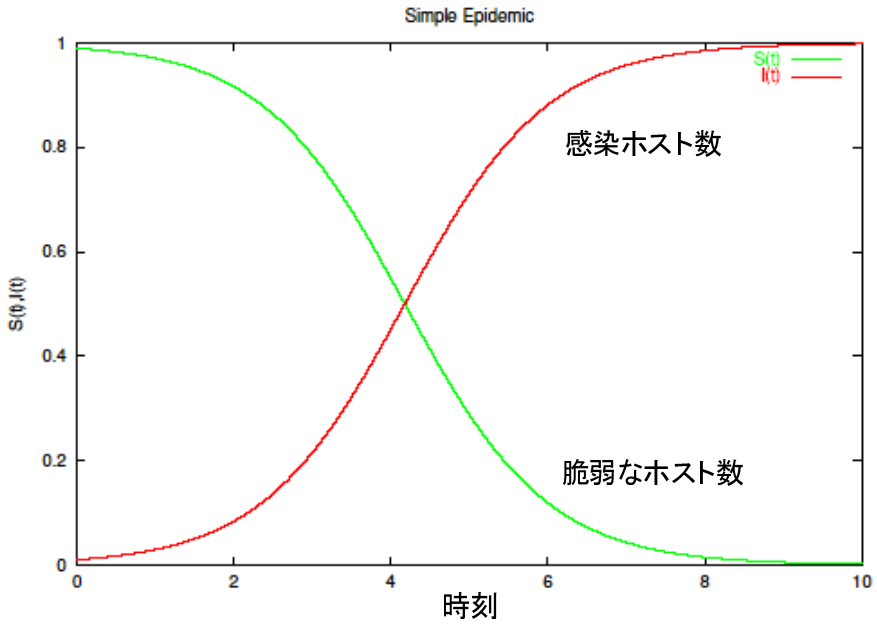

図 5.4: 単純感染モデルの時間推移 (感染ホストの増加)

- 感染ホスト数 $I(t)$
- 復旧ホスト数 $R(t)$

全ホスト数は、$N = S(t) + I(t) + R(t)$ である。SI モデルに対して、感染ホスト $I(t)$ から復旧ホスト $R(t)$ に単位時間当たりの移行数は、I(t) に比例する。

この時、微分方程式 (5.6) でモデルを定義できる (図 5.5)。

$$\frac{dS}{dt} = -k_1 I(t) S(t) \tag{5.6}$$

$$\frac{dI}{dt} = k_1 I(t) S(t) - aI(t) \tag{5.7}$$

$$\frac{dR}{dt} = aI(t) \tag{5.8}$$

式 (5.6) の第 1,3 式から以下が得られる。

$$\frac{dS}{dR} = -\frac{k_1}{a} S(t) \tag{5.9}$$

よって、脆弱なホスト数と復旧ホスト数の関係は以下の通りである。

$$S(t) = S(0) e^{-\frac{k_1}{a}(R(t) - R(0))} \tag{5.10}$$

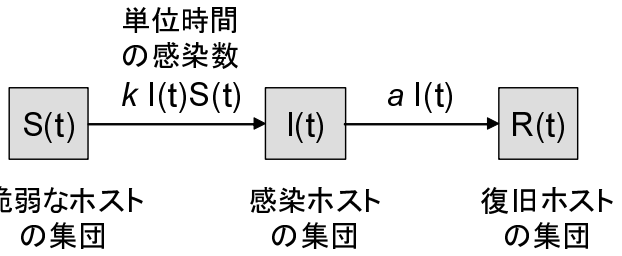

図 5.5: 一般化感染モデル

感染ホストが最初に出現した時の感染ホストの増加率は、式 (5.6) の第 2 式から、以下のようになる。

$$\frac{dI}{dt} \approx (kN - a)I(t) \tag{5.11}$$

$\frac{dI}{dt} > 0$ の時、つまり、

$$R_0 = \frac{kN}{a} > 1 \tag{5.12}$$

の時、感染爆発が発生し、$R_0 < 1$ の時、感染爆発は抑えられる (図 5.6)。

5.2.2 脅威の定義

本研究では、インターネット脅威観測システムが検出の対象とする脅威としてワームを中心に考えることを述べた。

ワームの脅威は、ある時点の感染ホスト数ではなく、感染力で決まると考えられる。なぜなら、新たに感染するホストが多いほど被害は拡大するためである。多くのホストに感染したワームであっても、残るホストの脆弱性に修正が施されていれば、脅威とはみなされない。したがって、そこで本研究では、ワームの脅威は、感染力の高さと定義する。ワームの感染力の高さは、単位時間当たりの感染数を表す式 (5.2) で表現することができる。

この場合、$S(t)$ が、脆弱性を持つホスト数を表し、ワームが感染を行うポートの集合によって決まる。　また、k は、インターネット上の脆弱なホストの分布に対して、どのように効率的に探索を行うかによって決まる。

以上のことより、ワームの脅威を表す感染力は、以下の関係で表現することができる：

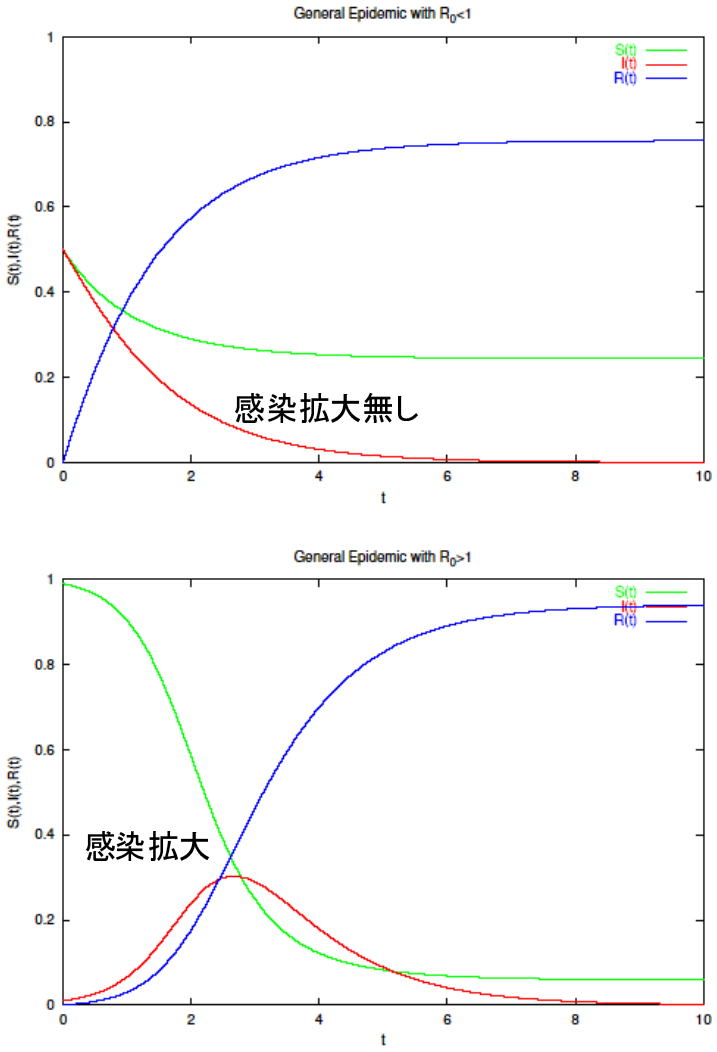

図 5.6: SIR モデルにおける R_0 による感染爆発の有無

$$(\text{感染力}) \propto (\text{脆弱なホスト数}) \times (\text{感染探索の効率性}) \tag{5.13}$$

インターネット上の脆弱なホストは、通常、偏りを持ったランダムな分布をしているため、ワームの感染探索戦略は、局所選好ランダム探索戦略が効率的である。

ワームの脅威を表す感染力を評価するためには、多くの観測データが必要になる。そこで、脅威の評価ではなく、不正パケットのパターンの変化のみに注目した異常検知により、早期に脅威の候補を検出することが考えられる。脅威の高いワームは、新種のワームにより生まれる可能性があるためである。

本章における脅威検知手法は、第 5.3, 5.5 章は、脅威の評価を目的とした手法を、第 5.4,

5.6 章は、異常検知を目的とした手法について示す。

5.3 ベイズ推定脅威検知法

インターネット上の脅威の原因となるワーム等の攻撃元の集団は、時間と共に動的に変化するため、静的なモデルに基づく分析手法では、時間に伴う変化を適切に検知することは難しい。非正規パケットの観測に対して、ベイズ推定に基づき、危険度推定値を繰り返し更新することにより、危険度の動的な変動に対応した推定を行う手法を提案する。

5.3.1 検知手法

特定の IP アドレスにおいて観測されるポートスキャン頻度の時系列データから、インターネット上の広域的な攻撃活動による危険状態の検出を行う手法を示す。

ベイズ推定は、観測をもとに、推定する状態に対する確信度の更新を繰り返す。本手法では、ポートスキャン頻度をそのトレンド (移動平均) からの差として観測することにより、ネットワークの危険状態を推定する。我々が推定したいネットワークの危険状態を、インターネット上の広域的な攻撃活動の活発化により、自サイトへの攻撃による被害が発生する可能性の高い状態とする。

危険度の推定を行う対象時刻を t(図 5.7 中、推定時刻と表示) とする。時刻 t からそれ以前の一定期間 T(確信度更新区間と呼ぶ) のポートスキャン頻度および、その区間の各時刻における移動平均 (トレンドと呼ぶ)、および、各時刻の一定区間のポートスキャン頻度のトレンドとの差の標準偏差 (局所標準偏差と呼ぶ) を観測し、各時刻において、1 つ前の時刻の危険度に対する確信度 (事前確率) を用いて、式 (5.14) に基づきベイズ更新を行い、次の時刻の危険度 (事後確率) を推定する。確信度更新区間 T において式 (5.14) に基づくベイズ更新を繰り返すことで、時刻 t における危険度を推定する。

$$P(s_i|r) = \frac{P(s_i)P(r|s_i)}{\sum_j P(s_j)P(r|s_j)} \tag{5.14}$$

ここで、$s_i(i = 0, 1)$ はネットワーク攻撃に危険状態の有無を表す。

$$\begin{cases} s_0 & : \text{危険状態} \\ s_1 & : \text{安全状態} \end{cases} \tag{5.15}$$

r はポートスキャン頻度のトレンドからの差を表す観測値、$P(s_i|r)$ は、r を観測した時に状態が s_i であると推定される確率を表す事後確率、$P(s_i)$ は、観測前に状態が s_i である確率を表す事前確率、$P(r|s_j)$ は、状態が s_j であるときに r が観測される尤度を表す。尤度関数 $P(r|s_j)$ は、危険状態であるときにネットワーク攻撃につながるポートスキャンが活発化することから設定することができる。本モデルにおいては以下のように定義する。

$$P(r|s_0) = \frac{r}{k\sigma_r + r} \tag{5.16}$$

$$P(r|s_1) = \frac{k\sigma_r}{k\sigma_r + r} \tag{5.17}$$

r は、ポートスキャン頻度のトレンドからの差、σ_r 観測地点での、一定区間のポートスキャン頻度のトレンドからの差から求めた局所的な標準偏差、k は、ベイズ更新における更新のスピードを定めるパラメータ (ベイズ更新偏差係数と呼ぶ) である。式 (5.16) および式 (5.17) は、0 から 1 の区間の実数値をとり、ポートスキャン頻度のトレンドから差が正の方向に大きいほど、危険状態であることを意味する。ベイズ推定に基づく危険度推定の計算手順は、図 3 のようになる。

図 5.8 は、あるポートに対するスキャン頻度時系列データである図 5.7 に対して、各時刻におけるベイズ推定による攻撃状態に対する確信度の推移を示している。

5.3.2 評価実験

本攻撃検知手法を、信号検出理論の分野で有効なことが知られている ROC 分析 (Reciever Operating Characteristics Analysis) [3] により評価する。一般に、正事例、負事例の判別精度の評価には、正答率や相関を用いることができるが、これらの方法は、正事例と負事例の比率強く依存し、一方の比率が極めて高い場合には、無条件に比率の高い方を予測すれば、評価値が高くなる問題点がある。ROC 分析の場合、負事例を誤って正事例と判断する偽陽性率 (False-Positve Fraction) と正事例を正しく正事例と判断する真陽性率 (True-Positive Fraction) の両面を考慮した総合的な尺度による評価が可能になる。真陽性率は、検出の感度に相当するものであり、偽陽性率は、特異な事例の比率を表す特異度に対応するものである。ROC 分析は、検出の閾値や事例の分布に依存しないものである。

本手法は、インターネット上で広域攻撃がある程度活発化した危険状態を検出するものである。このような定義による真の危険状態は、一般に、認識不可能なものであり、性能

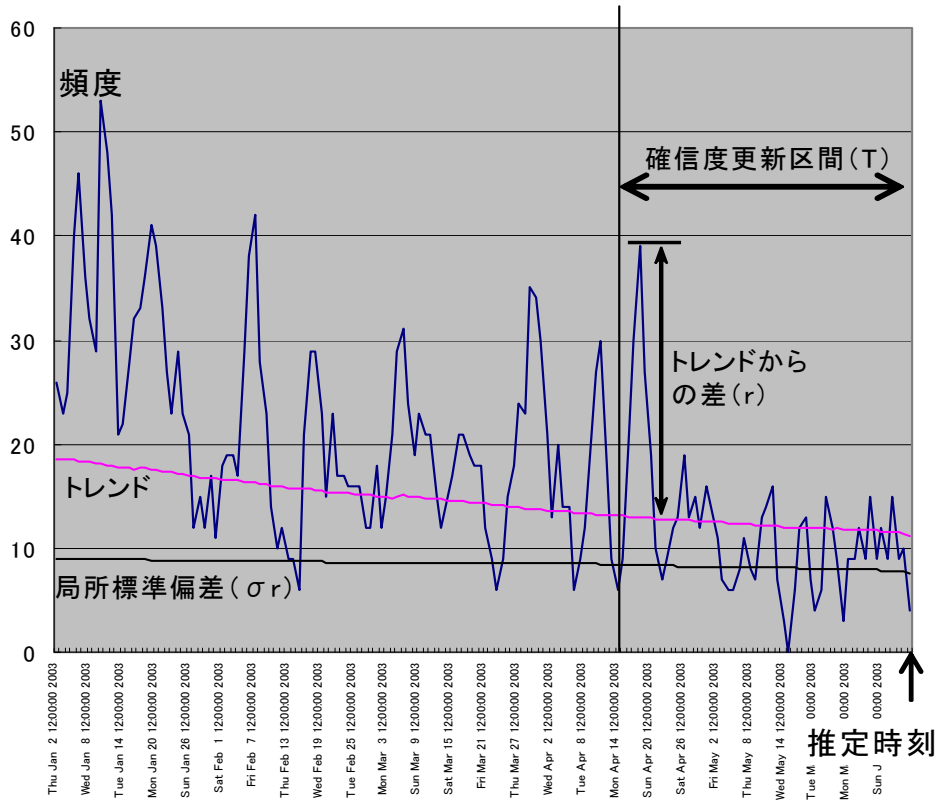

図 5.7: ベイズ推定における観測データとパラメータ

評価を行う上での困難の原因となっている。この評価においては、ポートスキャンデータから、人が見て危険であると見なされる時刻を設定し、それを真の危険状態の近似と見なすことにより、自動検知結果との比較分析を行う。

図 5.10 は、2003 年に JPCERT における緊急報告で注意喚起をされた脆弱性に該当するポートを含む比較的ポートスキャン数の多いポートを選び、ポートスキャン頻度の時系列推移を表示したものである。

ポート 80 は http サーバで使用され、2003 年 3 月 18 日に “Microsoft IIS 5.0 の脆弱性に関する注意喚起”(JPCERT-AT-2003-0003) のあったものである。ポート 135 は、Windows RPC サービスで使用されるもので、2003 年 8 月 15 日に、W32/Blaster ワームによって大規模な被害を生じた “TCP 135 番ポートへのスキャンの増加に関する注意喚起”(JPCERT-AT-2003-0006) のあったものである。ポート 25 は、メールサーバで使用されるもので、2003 年 3 月 31 日に “新たな sendmail の脆弱性に関する注意喚起”(JPCERT-AT-2003-0004) の

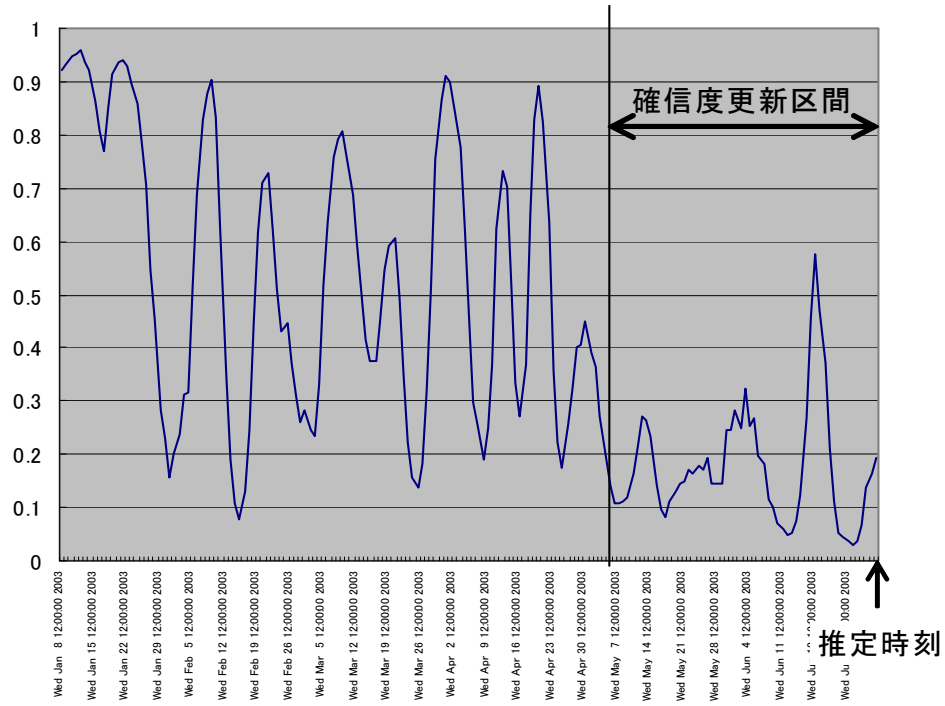

図 5.8: ベイズ更新区間と各時刻のベイズ推定履歴

あったものである。ポート 1434 は、Microsoft SQL Server 2000 で使用されるもので、2003 年 1 月 27 日に “UDP 1434 番ポートへのスキャンの増加に関する注意喚起” (JPCERT-AT-2003-01-27) のあったものである。

図 5.11 は、図 5.10 のポートスキャン頻度に対して、各時刻における危険状態を推定したものである。

これらの結果のうち、データの対象期間である 2003 年において、CERT Advisory 等において、注意勧告のだされた脆弱性のうちの 1 つで、対象期間のポートスキャン頻度の時系列変動の複雑なポート 25(smtp) を対象として ROC 分析を行う。

分析対象データは、1 地点の IP アドレスにおいて 2003 年 1 月 1 日から 2003 年 12 日 1 日までの 11ヶ月に観測された TCP/UDP ポートアクセスである。図 5.12 は、この期間における上記の危険状態 (図中 Positive cases) とそれ以外の安全状態 (Negative cases) に関して、攻撃検知によるベイズ推定値の分布を示したものである。

このグラフより、危険状態におけるベイズ推定値の分布のピークは 0.9 ぐらいの高い値に位置し、安全状態におけるそれは 0.1 ぐらいの低い値に位置しており、両事例の判別指

1. ポートスキャン頻度時系列から、トレンドを求める。

2. 確信度更新区間の初期時刻 T_0 における危険度事前確率の初期値を設定する。

3. 確信度更新区間の時刻 t の事前確率および、トレンドに基づく観測値 (r) をもとに、ベイズ更新式 (5.14) に従い、時刻 $t+1$ の危険度 (事後確率) を求める。

4. 確信度更新区間の最終時刻 T_f まで、ステップ 3 のベイズ更新を繰り返す。

図 5.9: 危険度推定の計算手順

標として使えることを示している。

図 5.13 の横軸 (ベイズ推定値) の閾値に対して、危険状態と安全状態のそれぞれのグラフの閾値より右側の面積を求めることで、True-Psitive Fraction および False-Positive Fraction を求めることができる。ROC 曲線は、この閾値を変化指せたときに得られる True-Psitive Fraction および False-Positive Fraction を 2 次元上に分布させることにより描くことができる。

図 5.13 は、ポート 25 に関して、攻撃検知手法のいくつかのパラメータを変化させた時 (表 5.1)、上記と同様に True-Psitive Fraction および False-Positive Fraction を求めることで得られる ROC 曲線を示している。

ROC 曲線は、$y = x$ 線上に位置し、上方に位置する程判別性能が高いことを示し、その性能評価は、ROC 曲線の下面積 (Az 値) によって行う。 表 5.1 中の Az 値から、トレンド区間 301 日、ベイズ更新偏差係数 0.5、ベイズ更新区間 5 の場合の検知性能が良好であることが確認された。

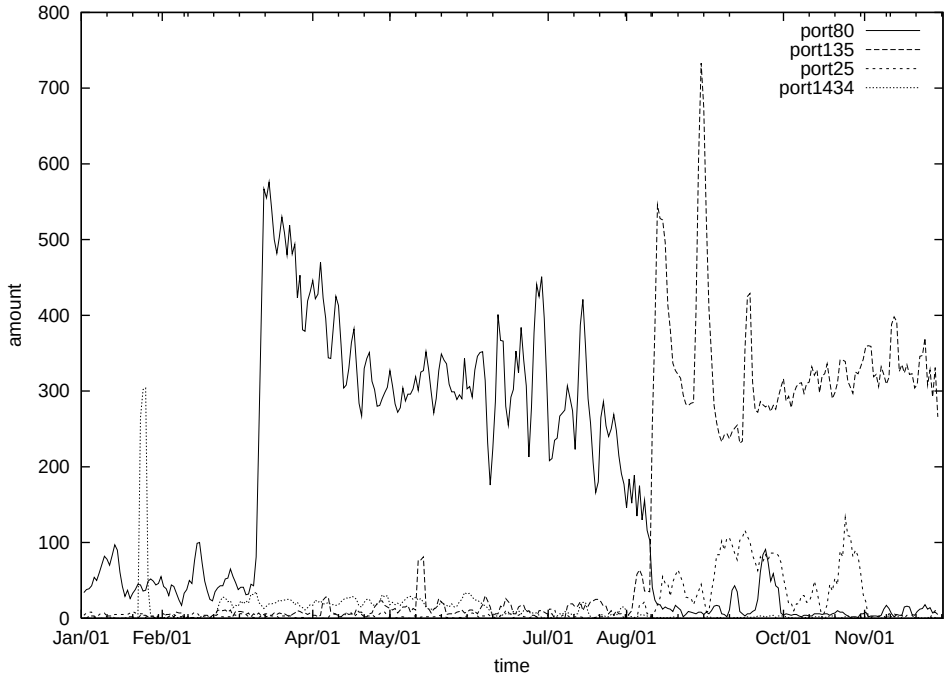

図 5.10: ポートスキャン頻度の時系列推移

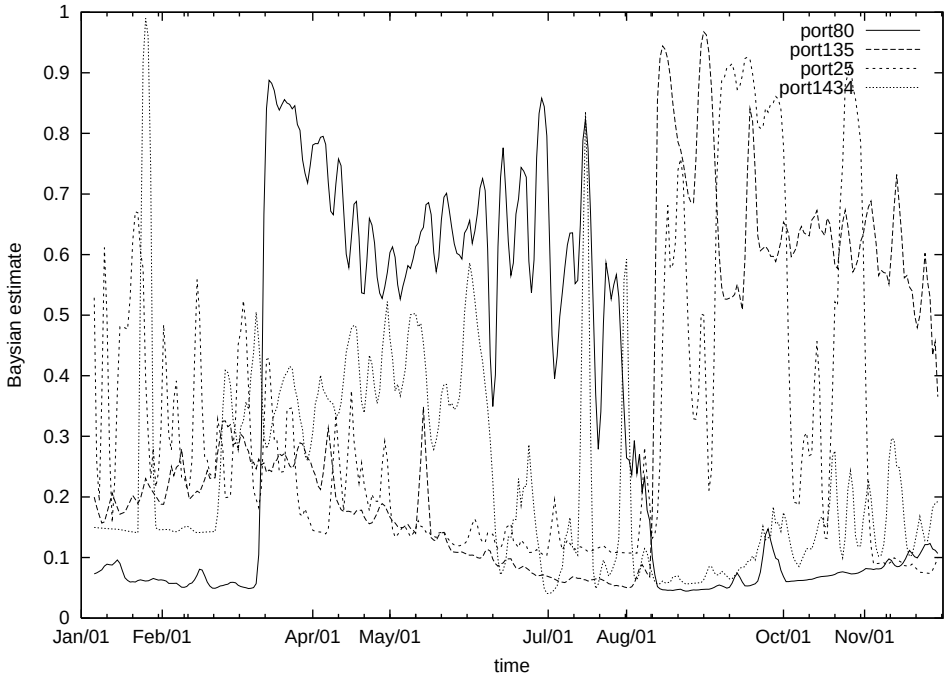

図 5.11: 危険状態推定値の時系列推移

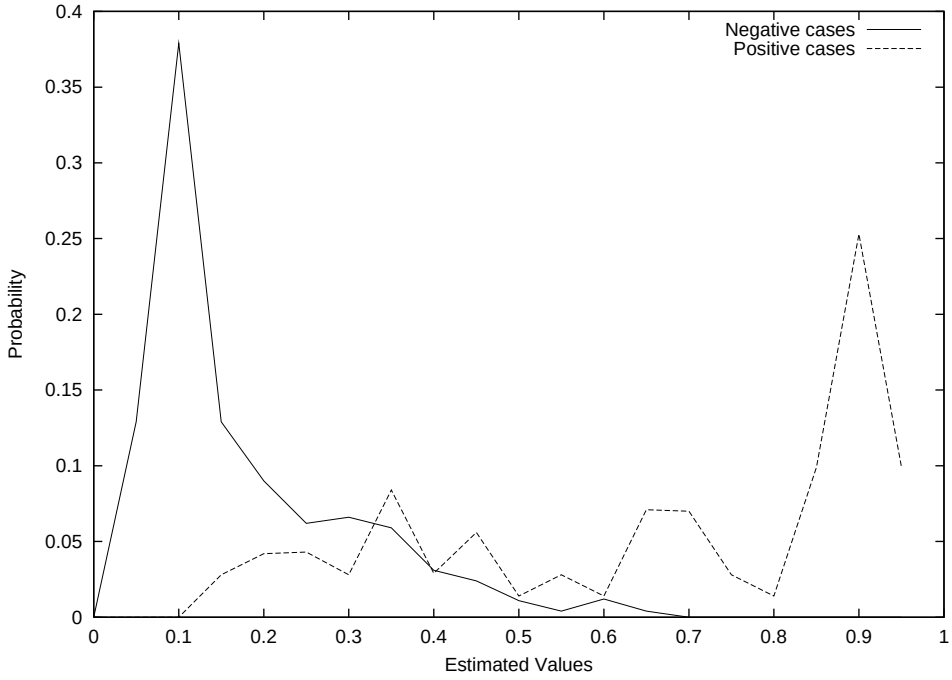

図 5.12: 危険状態の区分によるベイズ推定値の分布

表 5.1: 攻撃検知手法の分析パラメータ値

図 5.13 凡例 ID	ベイズ更新偏差係数	ベイズ更新区間	トレンド区間	Az 値
T301B5k0.5	0.5	5	301	0.95
T100B5k0.5	0.5	5	100	0.79
T100B3k0.5	0.5	3	100	0.80

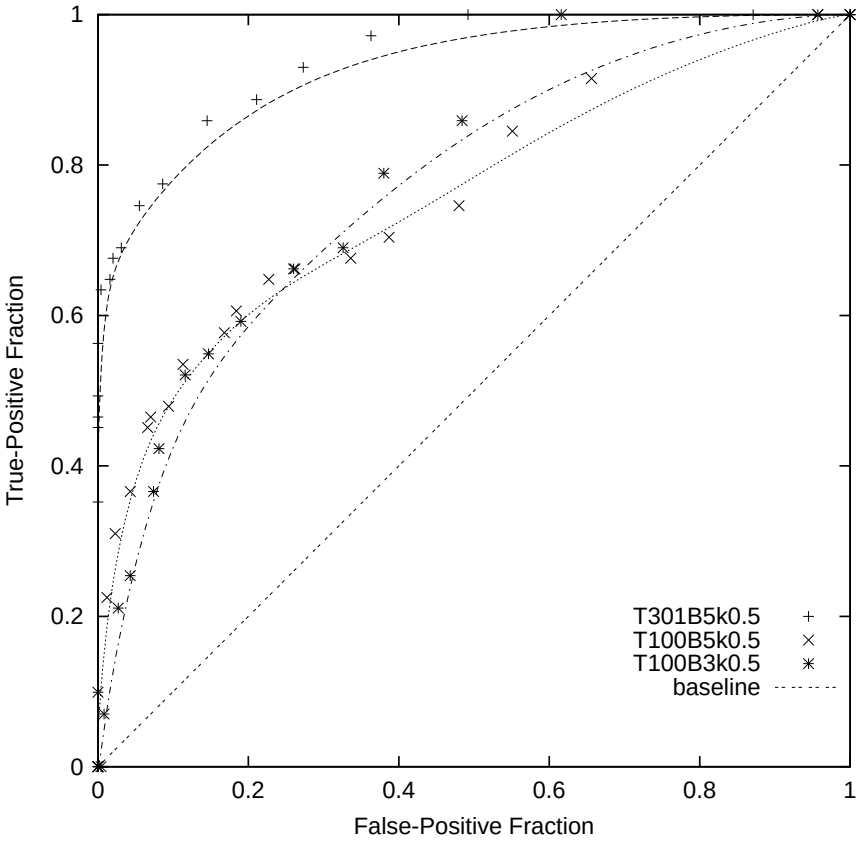

図 5.13: ポート 25 のベイズ推定に対する ROC 曲線

5.4 周期成分変化検出法

ワームの感染活動は時間帯，曜日など周期性をもつものが多い．このため，不正パケット数の時系列データには，さまざまな周期の振動成分 (周波数成分) が観測される．このような周波数成分の変化を捉えることで，インターネット上のワームの構成比率や振舞いの変化など，不正パケット数の増減からでは検出が難しい変化を検知することができる．本手法では、ウェーブレット解析を用いて、時間的に変化する不正パケットの振動成分を検知する。

5.4.1 異常検知手法

ウェーブレット変換を用いて，時系列データに含まれる局所的な振動成分（周波数成分）の変化を時間軸上で検出することができる．離散ウェーブレット変換は，図 5.14 に示す通り，元データに対して，低い周波数を通すフィルタ（スケーリングフィルタ）と高い周波数を通すフィルタ（ウェーブレットフィルタ）を繰返し適用することにより時間周波数成分を求める．図 5.14 左上のグラフは，元データを示す．これに対して，スケーリングフィルタ，ウェーブレットフィルタを適用することにより，レベル 1 の 2 段目に示すスケーリング係数およびウェーブレット係数を求める．得られたスケーリング係数は，元データの 1/2 の時間解像度のトレンドを示し，ウェーブレット係数は，スケーリング係数からの変動を示す．以下同様にして，レベル i のスケーリング係数に対して，スケーリングフィルタおよびウェーブレットフィルタを適用することにより，単位時間に対して 2^{i+1} のスケールのトレンド成分および周波数成分の強度を求めることができる．レベルの値 i が大きい程，低い周波数に対応する．

ウェーブレット係数の絶対値は，局所的な周波数成分の強度を示している．本手法では，標準的なドビッシーウェーブレット (長さ 8) を用いたウェーブレット変換を行う．また，位相のずれに伴うウェーブレット係数の変動はノイズとなるため，ウェーブレット係数を被う包絡線を求め，ウェーブレット係数の変動を抑えた補正済みウェーブレット係数を求める．この係数を用いて評価対象時刻から過去一定期間の局所的なウェーブレット係数の平均値および標準偏差を用いて，評価時刻のウェーブレット係数が近い過去の分布における偏差 (Z スコア) を求めることができる．図 5.15 は，異常検出の対象時刻と，元データに対して得られた特定のレベルのウェーブレット係数のうち，異常検出に用いられる過去のデータの関係を示している．

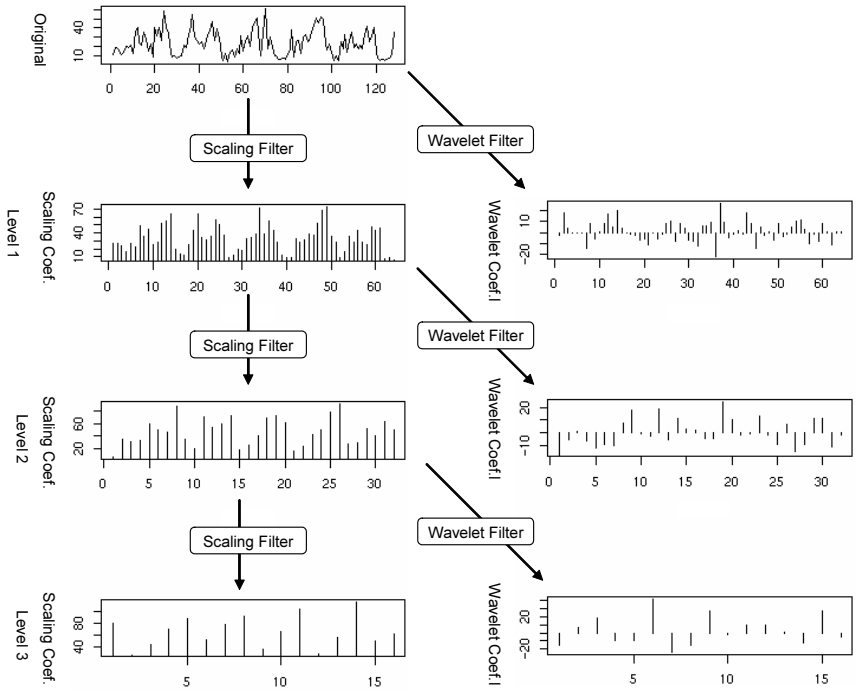

図 5.14: 離散ウェーブレット変換の概略

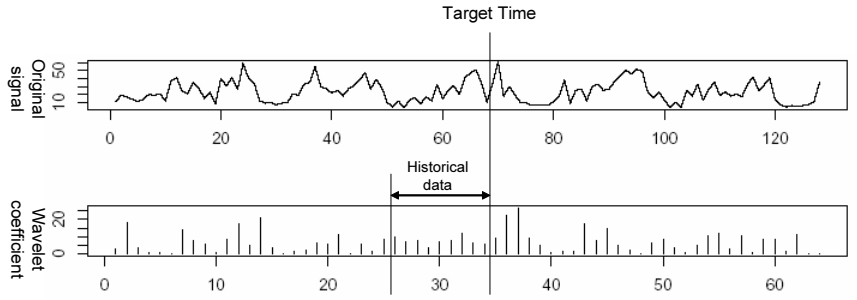

図 5.15: 異常検出におけるウェーブレット係数

レベル i の離散ウェーブレット変換によって得られるウェーブレット係数[1] の系列を $\mathbf{w_i} = \{\mathbf{w_{i,1}}, \cdots, \mathbf{w_{i,t}}\}$ とする．分析対象時刻 t_0 に対して，近い過去の平均値を $M_i(\mathbf{w_i})$ とする．ここで，レベル i の平均は，時刻 $t_0 - 1$ から過去 N_i 個の系列の平均として求めるものとする．

[1] 包絡線により補正したウェーブレット係数

$$M_i(\mathbf{w_i}) = \sum_{\mathbf{t=t_0-1}}^{\mathbf{t_0-N_i}} |\mathbf{w_{i,t}}|/\mathbf{N_i} \tag{5.18}$$

また，レベル i における近い過去の標準偏差 (不遍推定量) を $SD_i(\mathbf{w_i})$ とする．レベル i の標準偏差は，N_i 個の系列から求めるものとする．

$$SD_i(\mathbf{w_i}) = \sqrt{\sum_{\mathbf{t=t_0-1}}^{\mathbf{t_0-N_i}} (|\mathbf{w_{i,t}}| - \mathbf{M_i(w_i)})^{\mathbf{2}}/(\mathbf{N_i - 1})} \tag{5.19}$$

レベル i のウェーブレット係数に対して，対象時刻における周波数成分の強度に関する，近い過去の周波数強度の分布における偏差 (Z スコア) は以下の通り定義することができる．

$$v_{i,t_0} = \frac{|w_{i,t_0}| - M_i(\mathbf{w_i})}{SD_i(\mathbf{w_i})} \tag{5.20}$$

ワーム等の感染活動は多様であるため，特定の周波数に限定せず複数の周波数の Z スコアを用いて異常を検出する．ただし，高周波成分には，ノイズによる変動が多く含まれるため，最も周波数の高いレベル 1(2 時間解像度) の成分は除外し，レベル 2(4 時間)，レベル 3(8 時間)，レベル 4(16 時間)，レベル 5(32 時間) の成分を対象として，いずれかの成分に大きな偏差が見られる時に異常と判断する．1 つの判定法は，時刻 t_0 における各周波数成分から求めた Z スコアの組 $\{v_{2,t_0}, v_{3,t_0}, v_{4,t_0}, v_{5,t_0}\}$ が，各レベル $2 \cdots 5$ に指定した対応する閾値 $\{T_2, T_3, T_4, T_5\}$ のいずれかを越える場合に，異常と判定する．閾値は，発生頻度が稀であることを示す目安として 6.0 などを基準に，必要とする検知の感度から実験的に設定する．また，またもう 1 つの簡易な判定法として，Z スコアの組のうち最大値が指定した閾値 T を越える時に異常と判断する方法を用いる．

5.4.2 事例実験

本章では，具体的な観測データに対して本手法を適用した結果を示し，その特徴を示す．

図 5.16 は，2005 年 4 月 8 日から 4 月 30 日までの 135/TCP ポートの 1 時間単位の不正パケット数の時系列データに対して，本手法を適用した結果である．横軸は，経過日数を示す．左上のグラフは元の不正パケット頻度を表し，その下には，トレンド (s5)，32 時間

(d5)，16 時間 (d4)，8 時間 (d3)，4 時間 (d2)，2 時間 (d1) の周波数成分を抽出した時系列データを示す．右側のグラフは各成分に対応する異常検知結果を示す．

元の不正パケットは開始日から 9 日目ごろから典型的な 24 時間周期の振動がみられ，それより前の時期は不規則な頻度が確認される (グラフ中の三角形の印で示される位置)．

図 5.16 左側 d4 のグラフでは，16 時間 (d4) 周波数成分のグラフで，9 日目以降強い振動成分が確認でき，対応する右側のグラフでは，9 日目ごろに周波数成分の変化を検知したピークが確認され 16 時間周期成分の変化を示している．また，右側の周波数成分 2 時間 (L1) のグラフを除く，各周波数成分のグラフからそれぞれ乖離度が最大のものを抽出すると，L2 グラフから 10 日目，L3 グラフから 8 日目，L5 グラフから 12 日が異常として検知される．これらはそれぞれ異る要因の周波数成分変化と考えられるため，それぞれ異常と判断する．

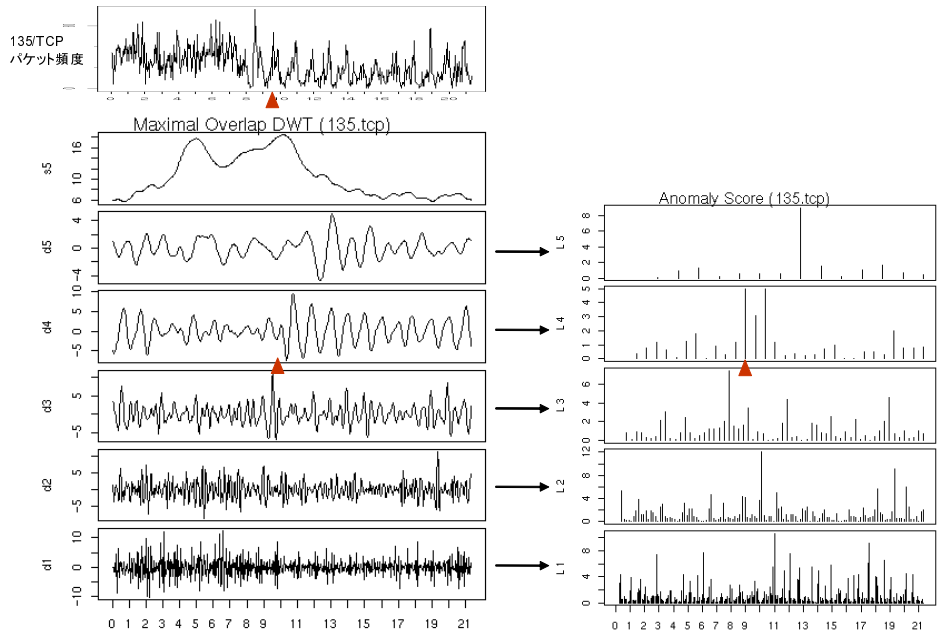

図 5.16: 周波数成分異常検知手法の適用結果 (135/TCP)

図 5.17 は，2005 年 4 月 8 日から 4 月 30 日までの 1433/TCP ポートの 1 時間単位の不正パケット数の時系列データに対して，本手法を適用した結果である．

元の不正パケットの時系列データは全体的に不規則であるが，6 日目前後に変化が見られる．周波数成分を示す右のグラフでは，16 時間 (L4)，8 時間 (L3)，4 時間 (L2) 成分に強いピークが見られる．これらからは異る周期性に関する変化が複数同時に起っている可能

性が推測される．しかし，正確な原因を特定するためには，パケットデータなどの詳細な分析が必要となる．

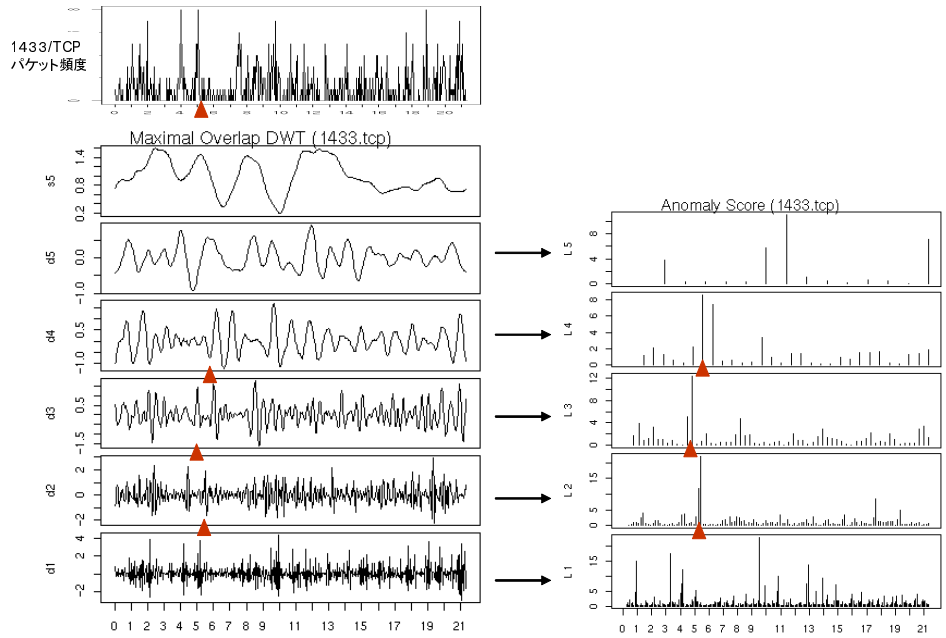

図 5.17: 周波数成分異常検知手法の適用結果 (1433/TCP)

5.4.3 評価実験

本手法の性能を評価するために，実際の観測データを用いて False-Positive Ratio(FPR) および False-Negative Ratio(FNR) を求める．FPR は，誤検知率を表し，FNR は検知漏れ率を表す．FPR と FNR はトレードオフの関係にあり，双方の値が同時に小さい程高い性能を示す．

不正パケットに基づく異常検知では，観測されるものは全て不正なパケットであり，その中から脅威の高いものを検出することを目的とする．そのため，IDS やスパムメール検出のように真の正解を確認することは一般的に困難である．そのため，現実的に行える評価方法として，従来手法の不正パケットの増減により検知される異常を基準として，その前の一定期間あるいはその前後の一定期間で検知できるかによって評価する．

図 5.18 は，検知性能の評価法について示している．(図は，2006 年 8 月の Windows サーバーサービス (ポート 139 番) のインシデント (MS06-040) が発生した時期を例としている．)

横軸は，8 月 10 日からの経過日数を示している．グラフは，下から順に，不正パケット数，不正パケット数の各時刻の Z スコア (従来手法)，本手法の検知スコアを示している．

正解として，インシデントの発生している期間と，インシデントの無い期間を定め，各期間で本手法の検知結果と比較する．まず，不正パケット数の時系列データから Z スコア (「不正パケット数」の Z スコアのことであり,「周波数成分」の Z スコアでは無い) により急激な増加を検知し,「インシデント観測開始時刻」とする．図 5.18 に示すインシデントの例では，横軸の経過日数 2 日目辺りに脆弱性情報が公開され，経過日数 13 日目辺りからパケット数の増加によるインシデントが検知されている．インシデントの正確な発生時期は特定できないが，図の例では，脆弱性情報が公開された経過日 2 日目から，パケット数の増加により検知された経過日 13 日目の間に，異常を検知されることが期待される．

本評価では，(a) 早期検知の性能と (b) 遅い検知を含む緩い性能評価の 2 通りの評価を行う．脆弱性情報の公開から従来手法によるインシデントの検知までの期間のずれ (図の例では約 10 日間) を考慮し，インシデント観測開始時刻前の 7 日間を，早期検知の評価においてインシデントを検知すべき期間 (図中「区間 3」に該当) と定める．また，時間軸を 7 日単位の区間に分け，インシデントを検知すべき期間以外の期間を，危険なインシデントの無い期間と仮定する．このようにして正解を決めた各区間において，本手法の検知結果をもとに評価する．一方，(b) 緩い性能評価では，インシデント観測開始時刻の前後 7 日間を検知すべき期間 (図中「区間 3」と「区間 4」に該当) とし，それ以外を「インシデント無し」とした場合について評価する．

センサーにより不正パケットの増加が確認されない場合や，不正パケットの増加からの検知が難しいインシデントが存在するため，一般には，正解となるインシデントの発生期間を特定することは難しいが，本評価では，統計的な傾向性を確認することを目的として，不正パケットの増加が見られない時期を,「インシデント無し」と見なす．

評価実験は，2006 年 8 月から 12 月までの期間で，5 つのセンサー別に，パケット数が上位 7 つのポートに対して行った．本手法を，42 日単位の時系列データに対して適用し，周波数成分が 4 時間，8 時間，16 時間，32 時間のそれぞれについて，この期間で補正済みウェーブレット係数が最大の値を示すもの異常と判定した．表 5.2 は，センサー全体のインシデント有無と警報有無の事例数を示している．表 5.3 は，センサー全体および各センサーの FPR および FNR を示している．

表 5.3「(a) インシデント観測開始前 7 日間」の結果は，不正パケットの増加によって特定できるインシデント観測開始時刻よりも早期に検知する場合の評価を行っているため，

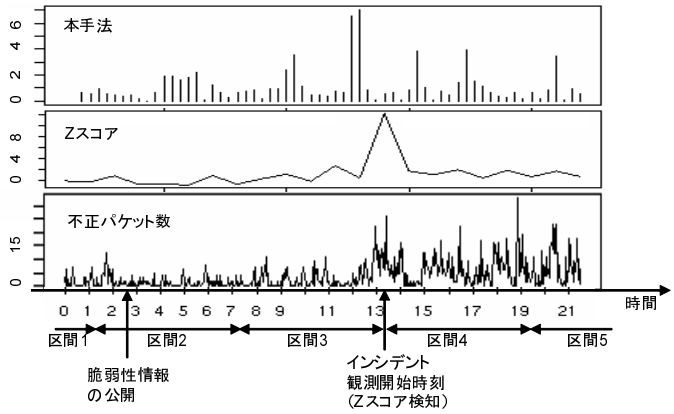

図 5.18: 性能評価方法 (Windows サーバーサービスのインシデントの例)

FPR は 39%，FNR は 14% と検知性能はあまり高くはない．表 5.3「(b) インシデント観測開始前後 7 日間」の結果は，検知評価を前後 7 日間の期間に条件を緩めているため，(a) よりも良い結果を示している．

本手法は，不正パケットの増加からでは検知が困難なものを周波数成分の変化という従来とは異るアプローチで早期に検知することを目的としたもので，不正パケットの増減による検知手法と比べて，検知できるインシデントの性質が異る．したがって，本手法を補完的に利用することで，不正パケットの増加による手法で検知できないインシデントを検知するために利用することが考えられる．

5.4.4 考察

周波数分析に基づく提案手法では，ウェーブレット解析を用いている．これは，不正パケット数の時系列データは，周波数成分が時間とともに変化する非定常的な性質を持つため，分析区間全体で同一の周波数成分を仮定するフーリエ変換よりもウェーブレット解析が適していると考えられることによる．短時間フーリエ変換を用いれば，時間的に局在する周波数成分を分析できるが，窓の区間を手動で与えることが難しく，また，解析区間全体を通じて，窓の区間が固定であるため，時間帯によって分析したい周波数と時間解像度を適応させることが難しい．ウェーブレット変換であれば，時間帯ごとに適切な時間解像度と周波数の分析に適している．

表 5.2: 異常検知事例数 (センサー全体)

(a) インシデント観測開始前 7 日間

		インシデント		
		有り	無し	計
警報	有り	25	43	68
	無し	4	66	70
	計	29	109	138

(b) インシデント観測開始前後 7 日間

		インシデント		
		有り	無し	計
警報	有り	27	27	54
	無し	3	81	84
	計	30	108	138

表 5.3: 異常検知性能

(a) インシデント観測開始前 7 日間

センサ	FPR	FNR
全体	39%	14%
センサ A	41%	18%
センサ B	42%	20%
センサ C	47%	0%
センサ D	33%	17%
センサ E	33%	0%

(b) インシデント観測開始前後 7 日間

センサ	FPR	FNR
全体	25%	10%
センサ A	31%	8%
センサ B	21%	20%
センサ C	27%	0%
センサ D	17%	17%
センサ E	20%	0%

5.5 グラフ構造分析法

本章では，定点観測システムによって観測される不正パケットの送信元および送信先の関係によって構成されるグラフの構造からインターネット上の脅威を分析する方法について示す．ワーム感染ホストの増大は，送信元 IP アドレスの増加によって捉えることができる．さらにそれを拡張し，脆弱性の高いポートへのワーム感染の効率性の観点を考慮した脅威分析法を示す．

5.5.1 分析アプローチ

定点観測システムによって観測される不正パケットの主な原因はワームからの感染パケットと考えられる [21]．ワームによるインターネット上の脅威は，(1) 感染自体の脅威と，(2) 感染後にワームから受ける被害に分けられる．感染後にワームから受ける被害は，ファイルの削除，個人情報の送信など定性的な要素が大きく，数理的に評価することは困難である．また，感染の脅威と感染後の被害の関係は独立性が高く，定点観測システムで観測されるワームの感染活動から感染後の被害を推論することは原理的に困難である．したがって，本研究では，ワームの感染力の強弱によるインターネット上の脅威を評価の対象とする．

通常，感染の脅威の評価においては，観測される不正パケット数自体よりも，不正パケットの送信元の数が重要である．なぜなら，不正パケットの送信元の数が多ければ，実際に多くのホストにワームが感染したことを示しているが，観測されるパケット数自体が多いだけでは，感染力の弱い少数のワームが多くのパケットを送信している場合があるためである．

図 5.19 は，一定時間の観測パケットを，送信元 IP アドレスと送信先ポート番号の間のアクセス関係によりグラフ表示したものである．図の左側 (図中 “Source IP Addresses (Renumbered)”) には送信元 IP に対応するインデックスを並べ，右側 (“Destination Port Numbers”) には，送信先ポート番号を並べたものである．この例では，ポート 1433(MS SQL サービス), 21(ftp), 80(http) は，多くの異なる送信元 IP アドレスからのアクセスを受けていることが分かる．この事例は，ポート 1433 番への感染攻撃を行う spida ワームの感染が活発化し，インターネット上の脅威が高い時期を示している．

このようにあるポートへの不正パケットの送信元の数から脅威を評価することができる．さらにこれを拡張し，仮に送信元ごとにワームの感染力によって決る脅威の高さ (脅威値と呼ぶ) が分かっていると仮定すれば，送信先が受ける脅威は，送信元の脅威値の総和で

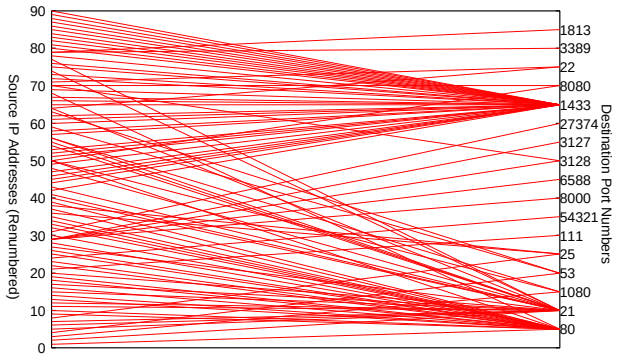

図 5.19: 送信元と送信先ポートのアクセスグラフの構造

評価することができる．図 5.20 は不正パケットの送受信の関係をグラフで示したものとする．送信先ノード d1 が受ける脅威は，送信元 s1, s2, s3 の脅威値の和によって評価できる．

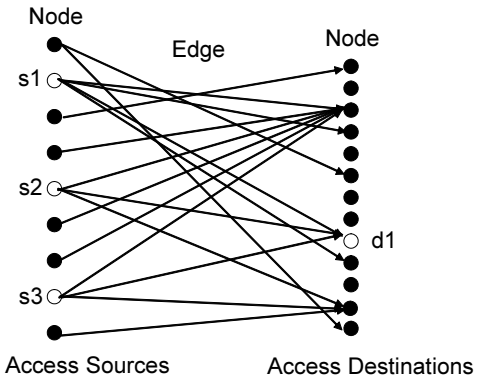

図 5.20: 送信先 d1 と送信元ノードの関係

一方，特定のポートに重大な脆弱性が存在すれば，そのポートを狙うワームの感染数が増大する傾向があることから，脆弱性の高いポートほど，ワームから脅威を受ける傾向が高い．ワームの感染力は，脆弱なポートを攻撃する方が高くなるため，仮に送信先のポートの脆弱性が分かっていれば，送信元のワームの脅威は，送信先が受ける脅威 (脆弱性) の和によって評価することができる． 図 5.21 のグラフで言えば，送信元 s4 の脅威は，送信先 d4, d5, d6 が受ける脅威の和として評価することができる．

以上の 2 つの関係のうち，一方は，送信元の脅威値を仮定して，送信先が受ける脅威値を評価し，もう一方は，送信先が受ける脅威値を仮定して，送信元の脅威値を評価する．こ

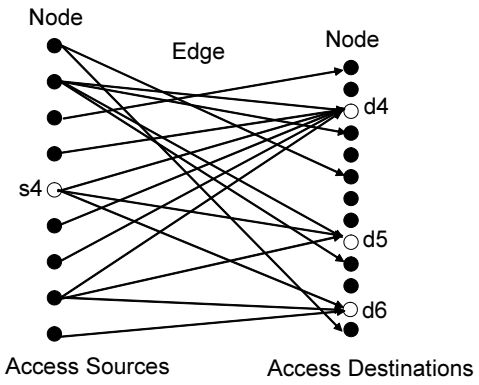

図 5.21: 送信元 s1 と送信先ノードの関係

の関係をまとめると以下のようになる.

送信元の脅威と送信先が受ける脅威の関係：

関係 1 送信先 (のポート) が受ける脅威は，送信元の (ワームの) 脅威の総和で評価できる.

関係 2 送信元の (ワームの) 脅威は，送信先 (のポート) が受ける脅威 (脆弱性) の総和で評価できる.

送信元と送信先に対して任意の脅威値を初期値として設定し，上記の関係 1, 2 を交互に繰返し適用することで脅威値が収束した場合，不正パケットのグラフの構造から評価される脅威値を示している．このようにして求めた脅威は，感染ホスト数の増加とワームによる脆弱なポートへの効率的な攻撃という観点から，ワームの感染力に基づく脅威を評価していると考えることができる．また，以上のようにして評価される脅威値は，インターネット上の一種の脅威の高さを定義していると考えることができる.

次節では，漸化式の関係を固有値問題に変換することにより初期値を与えずに脅威値を計算する方法を示す.

5.5.2 脅威の計算法

不正パケットのネットワーク上でのグラフの構造を図 5.22 のようなグラフによって表現する．送信元はインターネット上の IP アドレスによって決るノードである (図 5.22 中の

“Access Sources”)．送信先は定点観測システムのポート番号によって決るノードである (図 5.22 中の “Access Destination”)．また，送信元から送信先へのアクセスの有無をグラフのエッジ (“Edge”) として表現する．このとき，観測されるアクセスは，センサ外の IP アドレスからセンサのポートへのアクセスのみである．また，送信元と送信先のノードの集合には重複はない．これらのことからアクセスグラフは 2 部グラフとなる．定点観測システムのセンサの IP アドレスが複数ある場合は，図 5.23 で示すようにセンサの IP アドレスごとにポート番号を区別したものを送信先ノードとすることで，自然に拡張できる．

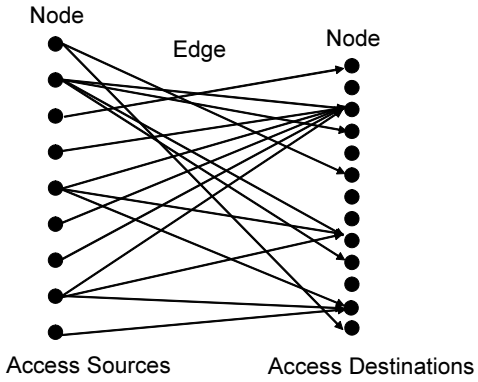

図 5.22: ポートアクセスのネットワーク上のグラフ

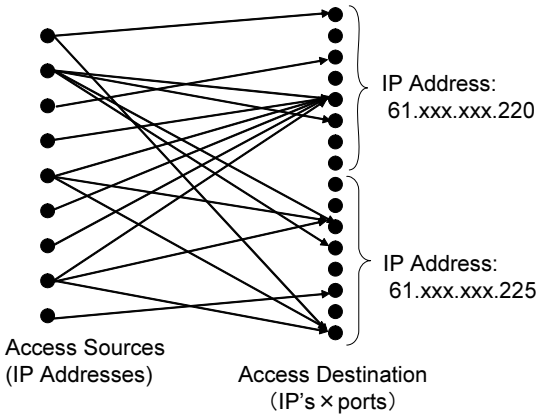

図 5.23: ポートアクセスのグラフ (送信先が複数 IP アドレス)

これらの関係を一般化して定義する．送信元のノード i の脅威レベル t_i を組にしたベクトル，および送信先のノード j がさらされる脅威レベル v_j を組にしたベクトルを考え，そ

れぞれ $\mathbf{t}, \mathbf{v}$ と定義する (以下，送信元脅威ベクトル，送信先脅威ベクトルと呼ぶ).

$$\mathbf{t} = (t_1, t_2, \cdots, t_n) \tag{5.21}$$

$$\mathbf{v} = (v_1, v_2, \cdots, v_m) \tag{5.22}$$

この時，前述の関係 1 から，送信先 j がさらされる脅威 v_j は，送信元 i の脅威 t_i に，送信元 i から送信先 j へのアクセスのエッジの重み $w_{i,j}$ によって重み付けした総和として式 5.23 のように定義できる.

$$\begin{cases} v_1 = c_1(w_{1,1}t_1 + w_{2,1}t_2 +, \cdots, w_{n,1}t_n) \\ \cdots \\ v_m = c_1(w_{1,m}t_1 + w_{2,m}t_2 +, \cdots, w_{n,m}t_n) \end{cases} \tag{5.23}$$

ただし，c_1 は，本章で後述する固有方程式を解くことにより決る係数である.

エッジの重み $w_{i,j}$ は，送信元 i からのアクセスが，送信先 j に与える脅威にどの程度影響するかをもとに定義する．ここでは，ワームからのアクセスの性質を考慮し以下のように定義する．ワームからのアクセスは，同一の送信元から慢性的に送信されるものよりも，感染力の強いワームの感染拡大により新たな送信元から来るアクセスの方が脅威が高い．そのため，連続する 2 つの観測期間を前期と後期とし，前期にはアクセスは無く，後期にアクセスが発生した場合は，新たに感染したワームからの攻撃と見なし 1 とし，それ以外の場合は 0 とする． それ以外の場合とは，前期，後期の両方からのアクセスがあるような慢性的なアクセスがある場合，前期にアクセスがあり，後期にアクセスが止む場合，前期も後期もアクセスが無い場合の 3 通りである.

次に，前述の関係 2 に基づき，送信元のノード j の脅威 t_j について考える．脅威 t_j は，送信先 i がさらされる脅威 v_i に，上記で定義したエッジの重み $w_{j,i}$ をかけたものを総和し，係数 c_2 を用いて，式 5.24 のように定義できる.

$$\begin{cases} t_1 = c_2(w_{1,1}v_1 + w_{1,2}v_2 +, \cdots, w_{1,m}v_m) \\ \cdots \\ t_n = c_2(w_{n,1}v_1 + w_{n,2}v_2 +, \cdots, w_{n,m}v_m) \end{cases} \tag{5.24}$$

係数 c_2 は，c_1 と同様に，後述する固有方程式を解くことにより決る.

関係式 5.23 は，送信元脅威ベクトルから送信先脅威ベクトルを求める漸化式を表し，関係式 5.24 は，送信先脅威ベクトルから，送信元脅威ベクトルを求める漸化式を表している．送信先脅威ベクトルおよび送信元脅威ベクトルの初期値として任意のベクトルを選び，関係式 5.23 および 5.24 を漸化式として交互に用いて計算を繰返す．その結果収束したベクトルが送信先脅威ベクトルおよび送信元脅威ベクトルである．

この繰返し計算を無限回行った結果得られる収束解は，以下のように固有値ベクトルの計算により求められる．上記で定義した送信元から送信先へのアクセスの重みを，行列 5.25(アクセス行列と呼ぶ) として定義する．

$$W = \begin{pmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,m} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,m} \\ \vdots & & & \vdots \\ w_{n,1} & w_{n,2} & \cdots & w_{n,m} \end{pmatrix} \tag{5.25}$$

関係式 5.23, および関係式 5.24 を，行列 W を用いて表現すると以下の通りである．

$$\mathbf{v} = \mathbf{c_1}\,{}^{\mathbf{t}}\underset{\mathbf{m\times n}}{\mathbf{W}}\mathbf{t} \tag{5.26}$$

$$\mathbf{t} = \mathbf{c_2}\underset{\mathbf{n\times m}}{\mathbf{W}}\mathbf{v} \tag{5.27}$$

ただし，行列 ${}^tW$ は，行列 W の転置行列で，各行列の下に記した $m \times n$ 等は，行列の行数と列数を示している．

これらの関係式を変形すると以下の固有方程式が得られる．

$$\mathbf{v} = \mathbf{c_1c_2}\,{}^{\mathbf{t}}\underset{\mathbf{m\times m}}{\mathbf{W}}\mathbf{W}\mathbf{v} \tag{5.28}$$

$$\mathbf{t} = \mathbf{c_1c_2}\underset{\mathbf{n\times n}}{\mathbf{W}^{\mathbf{t}}\mathbf{W}}\mathbf{t} \tag{5.29}$$

この固有方程式より，送信先脅威ベクトル $\mathbf{v}$ は，サイズ m の正方行列 $({}^tW\underset{m\times m}{W})$ に関する固有値 $\frac{1}{c_1c_2}$ の固有ベクトルとなり，送信元脅威ベクトル $\mathbf{t}$ は，サイズ n の正方行列 $(W^t\underset{n\times n}{W})$

に関する固有値 $\frac{1}{c_1 c_2}$ の固有ベクトルとして求めることができる．特に，${}^tW\underset{m\times m}{W}$，$W\underset{n\times n}{{}^tW}$ は，その要素がすべて正であれば，Perron-Frobenius の定理 [38] から最大固有値の固有ベクトルはすべて正となる．よって，送信先の脅威ベクトル $\mathbf{v}$ と送信元の脅威ベクトル $\mathbf{t}$ の解が一意に求まる．したがって，上記の漸化式による脅威値の算出は，固有値問題に変換されるため，漸化式の脅威値に対する初期値の設定は必要なくなる．

インターネット上のワームによる不正パケットは，送信先を確率的に選択するため，すべての送信元から送信先に対してランダムな少数の不正パケットが存在すると仮定できる．これらの不正パケットに対応して，アクセス行列 W のすべての要素に微少量 $\delta(\ll 1)$ を加えれば，すべての要素が正のアクセス行列を定義できる．これにより，式 5.28 を解いた送信先脅威ベクトルの要素はすべて正となる．固有ベクトルとして，単位ベクトルを選べば，送信先脅威ベクトルの要素は 0〜1 の間の値となる．

5.5.3 評価実験

アクセスグラフに基づく脅威の評価法を，実際のインシデントの発生期の観測データに対して適用し，本手法の実用性，有効性を検証する．

本研究では，インターネット上の脅威を，感染力の強いワームによってホストがさらされる脅威として定義した．ワームの感染力は，攻撃対象となる脆弱性を持つホストのインターネット上での分布や，ワームの感染先探索戦略によって決る．インターネット上のこれらの情報全体を知ることは困難である．そこで，本実験では，JPCERT/CC などによって公表された注意勧告から深刻なインシデントが発生した時期をインターネット上の脅威が高いものと仮定し，本手法によって求めた脅威レベルと比較する．

過去のデータからの機械学習を用いた評価手法の場合，学習に用いた訓練データと評価用のテストデータを分けて性能評価を行う必要がある．本手法では，機械学習を用いておらず，評価結果が過去のデータに依存しないため，訓練データとテストデータを分けた評価実験は必要ではない．

(1) MS SQL に関するインシデント

ここで評価対象とするデータは，JPCERT/CC から MS SQL の脆弱性を狙ったポート 1433 番への攻撃に対する注意勧告 (JPCERT-AT-2005-0006) が出された 2005 年 7 月 9 日から 7 月 13 日までの 5 日間の TCP アクセスである．

5 日間の観測データを，図 5.24 に示すように，2 日間のデータを 1 組として，1 日ずつず

らした 4 回分のデータとして本手法を 4 回適用する．各 1 回の適用 (図中 “第 1 回評価用”,“第 2 回評価用”,..,“第 4 回評価用” の各回) においては，前期 1 日 (図中 “前半”) と後期 1 日 (図中 “後半”) を用いて前節で定義した通りアクセス行列を求める．つまり，7 月 10 日の適用結果は，7 月 9 日と 7 月 10 日の 2 日のデータを用いてアクセス行列を求め，7 月 11 日の適用結果は，7 月 10 日と 7 月 11 日の 2 日のデータを用いてアクセス行列を求めるというように順に 4 回適用する．

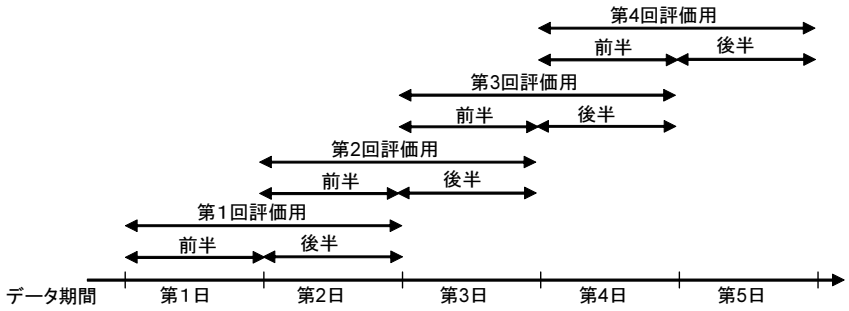

図 5.24: 脅威評価におけるデータ期間の使い方

表 5.4 は，本手法を 4 回適用した結果得られた脅威レベルのうち上位 10 位を示している．表 5.4 中の “port” は，ポート番号， “count” は，ポートへのアクセスパケット数（当日のパケット数から，送信元，送信先が同じ前日のパケット数を差引いたもの），“threat” は本手法で求めたポートがさらされる脅威のレベルを示している．

表 5.4: ポート 1433 インシデント時の脅威計算結果の上位 10 件

July 10			July 11			July 12			July 13		
port	count	threat	port	count	threat	port	count	threat	port	count	threat
135	1031	0.627	135	1038	0.789	135	885	0.792	135	1057	0.636
445	1121	0.472	445	822	0.378	445	820	0.432	1433	346	0.331
12345	10	0.163	139	208	0.160	1433	222	0.233	445	739	0.305
139	232	0.159	1433	159	0.130	139	219	0.195	2745	6	0.148
1433	115	0.132	12345	13	0.109	9898	7	0.089	139	204	0.135
3410	8	0.123	901	14	0.109	1024	2	0.085	2100	3	0.111
901	9	0.123	3410	11	0.087	4899	64	0.078	8080	3	0.111
22	12	0.112	3389	6	0.087	3306	19	0.064	8535	3	0.111
3090	7	0.112	3306	18	0.087	2100	1	0.064	25	6	0.111

表 5.4 では，発生したインシデントに対応するポート 1433 番の脅威のレベルは，7 月 10 日から 7 月 13 日の順に，0.132, 0.130, 0.233, 0.331 と増加している．また，他のポートと

比較した脅威のランクは，7月10日から7月13日にかけて，5位，4位，3位，2位 へと上昇している．

本実験では，ポート1433番に関するインシデントの存在が確認されている時期を対象としているため，定常的に不正パケットの多い445番，135番ポートよりも，1433番ポートの脅威が高くなることが期待される．　　表5.4の7月13日の結果では，不正パケット数については，ポート445番よりポート1433番が少ないにも関わらず，脅威レベルではポート1433番の方がポート445番より高いため，不正パケット数の増加よりも，本手法の脅威レベルの方がポート1433番の脅威を良く捉えている[2]．一方，135番ポートに関しては，1433番ポートよりも高い評価値となっている．これは，本手法では，グラフの構造以外に，不正パケット数自体も脅威値に影響を与えるためであると考えられる．アクセスグラフのウェイトを最適化することで，不正パケット数よりもグラフ構造を重視した評価を行うことで改善できる可能性がある．

表5.4では，7月10日のポート12345番 (Amitis.Bバックドア)，7月12日のポート9898番 (Win32.Dabber.Bワーム)，7月13日のポート2745番 (Bagleワームのバックドアを利用するAgobotボットネットワーム) は，不正パケット数が少ないにもにも拘らず脅威レベルは上位に位置している．これらの事例についても，従来の不正パケット数によるものよりも本手法の脅威値の方がポート12345番，ポート9898番の脅威を高く評価している．

(2) Windowsファイル共有に関するインシデント

ここで評価対象とするデータは，IPAによって公開された注意情報[6]で，Windowsファイル共有で利用されるポートTCP/139番に対してボットネットによる攻撃に関するインシデントで，2005年6月8日から6月12日までの5日間である．本実験でも，前節の実験と同様に，2日間のデータを1組として本手法を4回適用した．表5.5は，本手法を適用した結果で，脅威の高いポートを順に上位10位を示したものである．表中の “port”, “count”, “threat” は，前節に示したものと同様ある．

本結果では，インシデントのWindowsファイル共有に該当するポート139番は，最初の2日の脅威レベルは，順に0.029(ランク20位), 0.055(ランク33位)で，値が小さいため，表5.5の上位10位には現れない．しかし，6月11日，6月12日の適用結果では，脅威レベルは，0.081, 0.106と上昇し，脅威ランクは4位，3位へと上昇している．

表5.5では，7月11日から7月12日にかけて不正パケット数は減少しているにも関わら

[2] 実際には，インターネット上の真の脅威の有無 (正解) は，完全には知り得ないため，あくまでも脅威の高いインシデントが確認された時期の観測データに対する比較結果のみ示すことができる．

表 5.5: ポート 139 インシデント時の脅威推定結果の上位 10 件

June 9			June 10			June 11			June 12		
port	count	threat	port	count	threat	port	count	threat	port	count	threat
135	2551	0.954	135	2174	0.883	135	2834	0.879	135	1906	0.846
445	751	0.209	445	1008	0.227	445	1308	0.244	445	989	0.249
1433	140	0.078	1080	4	0.104	12345	11	0.085	139	242	0.106
4899	43	0.052	44599	8	0.099	139	257	0.081	42857	2	0.102
1521	1	0.052	10589	4	0.099	21	4	0.077	4899	46	0.076
8535	1	0.052	8080	2	0.070	1433	142	0.065	143	1	0.076
8536	1	0.052	4899	47	0.070	44599	3	0.064	3306	9	0.076
2100	3	0.052	22	23	0.070	10589	3	0.064	1256	3	0.076
22	10	0.052	25	10	0.070	11524	2	0.064	2419	1	0.076
143	1	0.052	3306	4	0.070	42857	2	0.064	6346	3	0.076

ず，脅威値は上昇している．脅威の高いインシデントが発生した時期を対象としているため，本実験においても不正パケット数によるものよりも本手法の脅威値の方が良い結果を示している．

5.5.4 考察

グラフ構造分析手法の特徴は，グラフ全体の構造から脅威値が定量的に計算され，一見関係の無いようなグラフ上の離れたノードに関する不正パケットのアクセス関係もグラフ全体のノードの脅威値に波及する点である．不正アクセスの送信元送信先のペア解析 [32] では，グラフ全体の構造から定量的な評価が行われていない点が異る．本手法は，ワームの感染力に関する脅威とポートの脆弱性の関係に基づき，不正パケットが構成するグラフの構造から求められるインターネット上の脅威について一つの定義を提示している．本手法は，送信先のポートに対する脅威をベクトル値として計算することに対して，エントロピー手法 [28] では，不正パケットの送信先に対するランダム性や偏りによる評価値をスカラー値として与えているが，グラフの空間的な構造の違いは反映されていない点が本手法と異る．

実験では，実験期間 (2005 年 6 月から 7 月) において，JPCERT/CC の注意喚起などにより深刻なインシデントと考えられるものを対象とすることで，実験的に本手法の効果を示した．

評価実験では，2 日を 1 組として 1 日単位のデータに対して本手法を適用している．これは，送信元 IP アドレスが，DHCP などにより動的に変化する影響を抑えるためである．1 日単位のパケット観測の場合，DHCP による送信元 IP アドレスの変化による誤差が比較

的小さく抑えられる [17] ことから，本実験では，1 日を単位とした実験を行った．送信元ソース IP アドレスの変化による誤差の影響を低減する方法としては，IP アドレスの論理的な距離が近い/24 のネットワークブロック内のアクセス元のノードを同一のグループと見なして，送信元ノードをグループ化したものに本手法を適用することで，確率的な変動を抑える方法などがある．これは，IP アドレスの近いネットワークには，設定の似たホストが多数存在すると考えられることが理由である．本手法は，ワームの感染力に基づく脅威を [0,1] の区間で標準化して示しているため，過去の履歴に基づく分布からの偏差を用いずに脅威値から直接比較している．

5.6 自己相関変化検出法

非正規パケットは、その原因となるワーム等の活動に対応して、時間周期性が見られる。新種のワームが発生した場合には、時間周期性に変化が生じるため、自己相関分析により時系列相関の変化を検出することで異常検知を行うことが可能である。

5.6.1 検知手法

本手法は、ポート毎に、時刻 t からさか上って一定期間の 1 時間単位の不正パケット数の時系列データに対して、タイムラグ k 時間の自己相関係数を求め、自己相関係数が大きく変化する時刻 t を検知する。第 4.3 章で示したとおり、不正パケットの自己相関関数は、タイムラグが 24 時間でもっとも高い相関を示す。従って、本手法では、$k = 24$ 時間として、不正パケットの自己周期性から外れた時刻に、何らかの異常が発生したと見なして警報を発する。(図 5.25 参照)

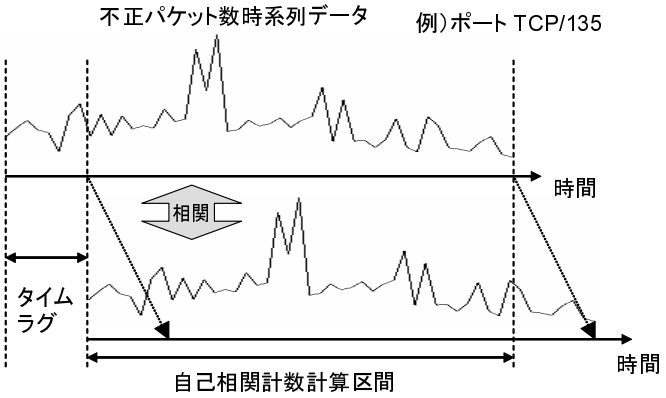

図 5.25: 不正パケット数時系列データの自己相関

5.6.2 実験結果

本手法を実際の観測データに対して適用した結果を示す。図 5.26 は、2007 年 9 月 7 から約 1 週間の不正パケット数の多いポートおよび全ポートの合計に関して自己相関係数の時系列変化を示したものである。横軸は 1 時間単位の経過時間を表している。グラフにおいては、自己相関係数が低下する時点の異常検知を行うことが重要である。なぜなら、通常

周期性を示す不正パケット数の時系列に対して、周期性を乱す変化が生じることを意味しているからである。本実験では、TCP/445 番ポートは、時刻 120 ごろから大きく下落している点がそれに該当する。また、ICMP/0 番ポートに関しても、時刻 50 ごろから下落していることが確認できる。

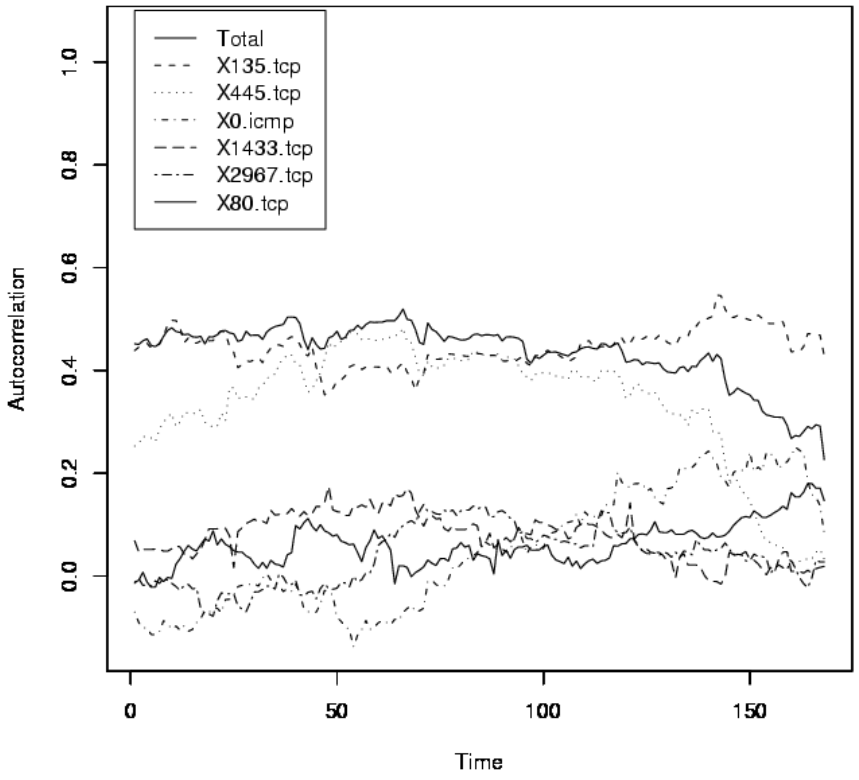

図 5.26: ポート別自己相関係数時系列変化 (2007/9 月)

図 5.27 は、図 5.26 と同時期の不正パケット数が上位のポート間の相関を視覚的に確認するために散布図を表示したものである。このグラフからは、比較的不正パケット数の多いポート間で、相関性の高さを示す対角線上のプロットが多く見られることが確認できる。

図 5.28 は、別の時期 2007 年 8 月 11 日から 1 週間の自己相関係数の時系列変化を示したものである。このグラフでは、TCP/135 ポートが、時刻 50 番から次第に自己相関係数が減少していることが確認できる。また、時刻 160 ごろに急激に自己相関係数が減少していることが確認でき、何らかの要因で周期性が変化したと推定される。

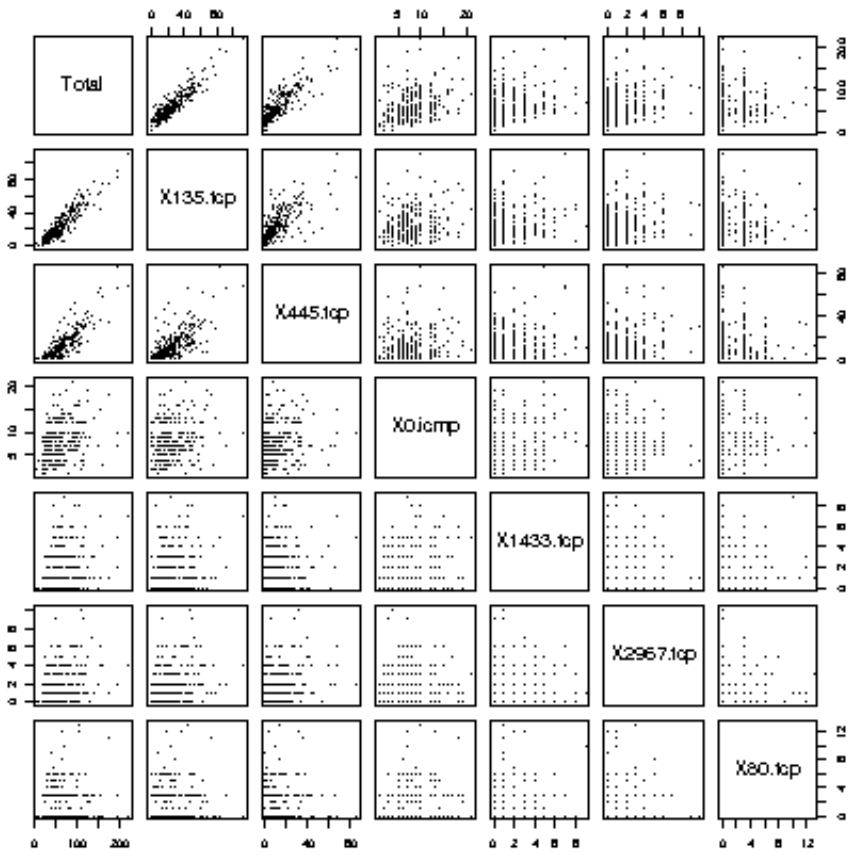

図 5.27: ポート間相関 (散布図)

5.7 脅威の可視化による分析手法

インターネット上の脅威の状況を視覚的に捉えることは、将来の脅威の推定や自動的な検知手法だけでは検知が困難な事象に対して有効な手段である。本章では、インターネット上で観測される不正パケットの国別の送信元・送信先関係を 3 次元地球儀上に表示するシステムを提案する。

5.7.1 全体構成

分析および可視化表示処理の構成は図 5.29 のように示される。

センサーから収集された不正パケット情報は、データサーバから SQL インタフェースを用いて、脅威分析& 3 次元可視化システムに取得する。過去の不正パケットに対して単位時間ごとに不正パケットから求めるセキュリティイベントの時系列データを計算する。次

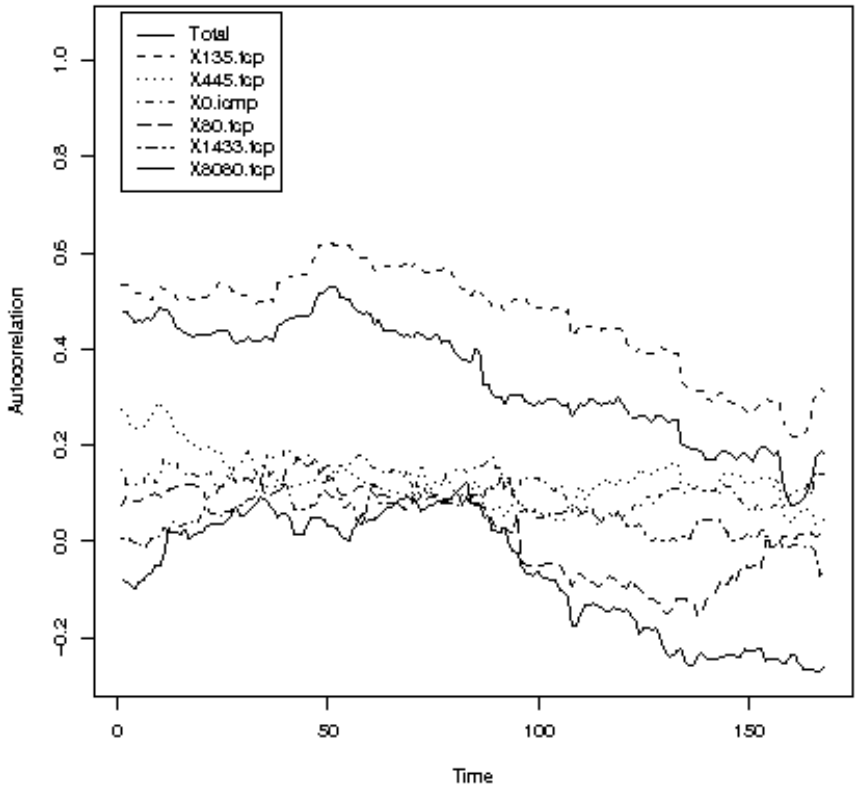

図 5.28: ポート別自己相関係数時系列変化 (2007/8 月)

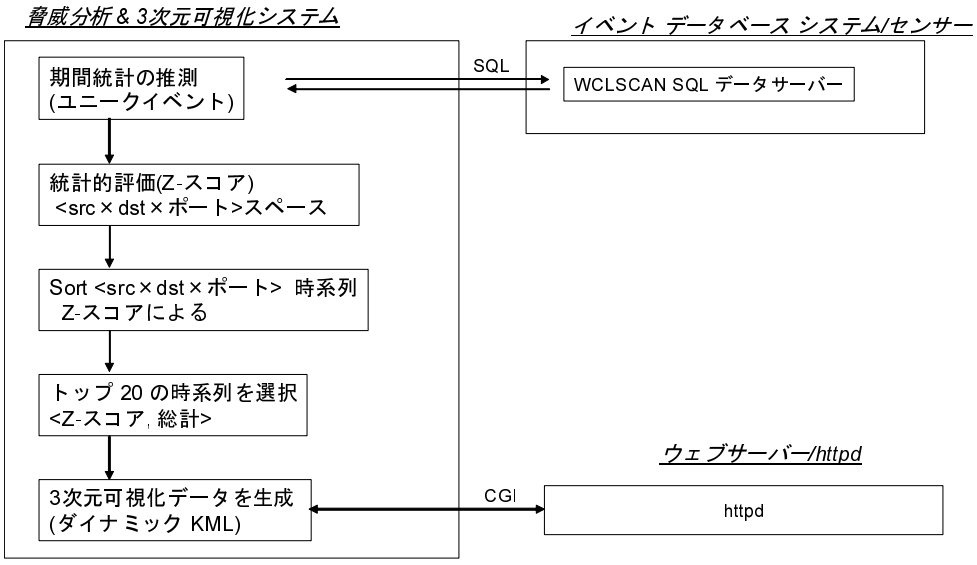

図 5.29: 不正パケットの分析および可視化処理の構成

に、過去のユニークイベント数の分布から、最新の期間のユニークイベント数の乖離 (Z スコア) を求める。クライアントの GoogleEarth からのリクエストに対して、乖離の大きい上位の送信元国・送信先国に関して GoogleEarth 上で 3 次元表示するための可視化データを送信する。

5.7.2 ユニークイベント

インターネット上の脅威は、単位時間当たりの不正パケットのうち送信元、送信先の IP の異なる対のイベント数（「ユニークセキュリティイベント数」または「ユニークイベント数」と呼ぶ）によって評価することができる。ワーム、ボットの脅威は、インターネット上で、効率的に感染拡大によって評価することが出来るためである。図 5.30 はユニークイベントの求め方を事例を用いて示している。

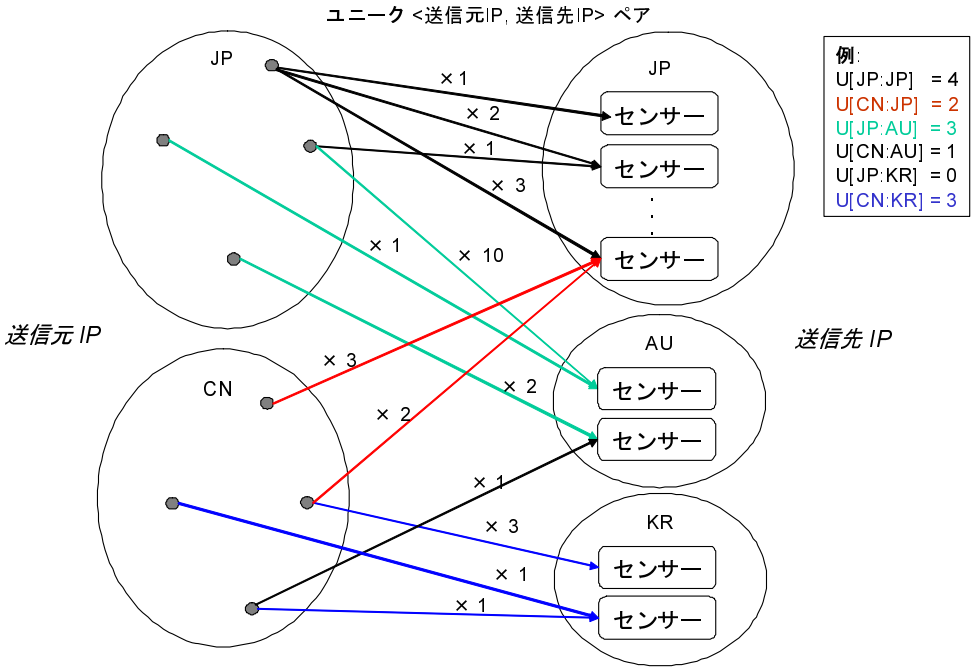

図 5.30: ユニークセキュリティイベント (事例)

不正パケットを、送信元国、送信先国の対でグループ化し、各グループで、送信元 IP アドレス、送信先 IP アドレスのいずれか一方でも異なるイベントをユニークイベントとしてカウントする。図中の送信元から送信先への矢印の途中に書かれている「× (数字)」印は、同じ送信元・送信先対の繰返し回数を示している。これらの繰返しは単一のユニーク

イベントとしてカウントする。

5.7.3 時系列ユニークイベント数

単位時間当たりのユニークイベント数は、送信元国、送信先国、送信先ポート/プロトコル種別の3つ組ごとに計測する（図5.31）。これにより、ユニークイベント数に関する複数の時系列データが生成される。時間間隔は、24時間の整数倍を用いる[3]。

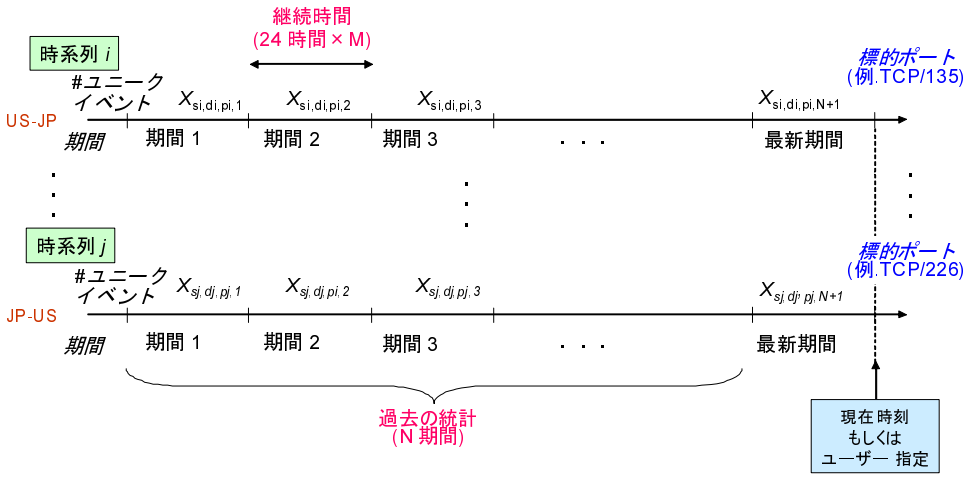

図 5.31: 時系列ユニークイベント数の算出法

5.7.4 統計的偏差分析

ユニークイベント数の時系列データに関して、過去のイベント数の分布に対して、最新の期間におけるイベント数の乖離により、脅威を評価する。図5.32は、過去の時系列データを表している。

図5.32の各単位時間における頻度の分布（図5.33）に対して統計的な乖離を求めるためにZスコアを用いる。

Zスコアは、分布の平均と標準偏差を用いて図中$\sigma(X)$のように定義される。

[3] 不正パケットのユニークイベント数は、1日の内で時間周期性を持つため、24時間の非整数倍の単位時間を用いるためには、時系列データの自己回帰分析による時間帯別のイベント数の推定が必要である。

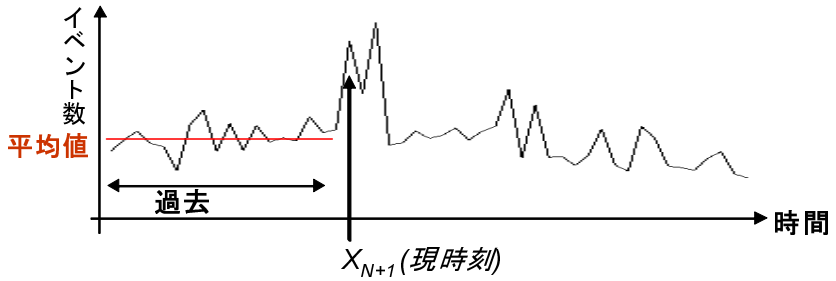

図 5.32: ユニークイベント数の時系列グラフ

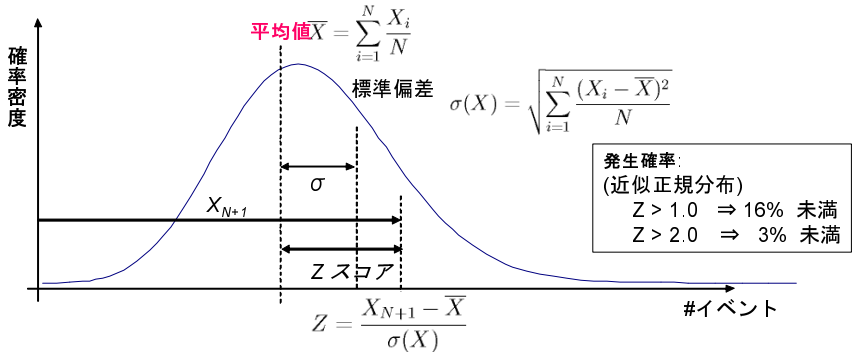

図 5.33: ユニークイベント数の統計的分布における乖離 (Z スコア)

5.7.5 可視化表示例

図 5.34 は、2007 年 12 月 3 日の特定の時間帯の分析結果に関する 3 次元可視化表示を示している。線の色は、Z スコアに基づき、セキュリティユニークイベント数の統計的な変化率を示し、緑から赤に 5 階調で表示する。線の太さは、セキュリティユニークイベント数を表し、log スケールで 10 段階表示している。可視化表示では、マウスを用いて、地球儀を自由に回転したり、視点の位置を自由に設定することができる。

図 5.34 情報に時間スライドバーを表示し、3 次元可視化中の時間帯を示している。時間スライドバーをマウスで移動することにより、自由に可視化中の時間帯を変化させることができる。また、時間スライドバーの右にあるアニメーションボタンを押すことにより、脅威分析結果の時系列データを連続的に再生することができる。

また、送信元・送信先のリンクをクリックすることで、分析結果の数値情報をポップアップ表示させることができる (図 5.35)。

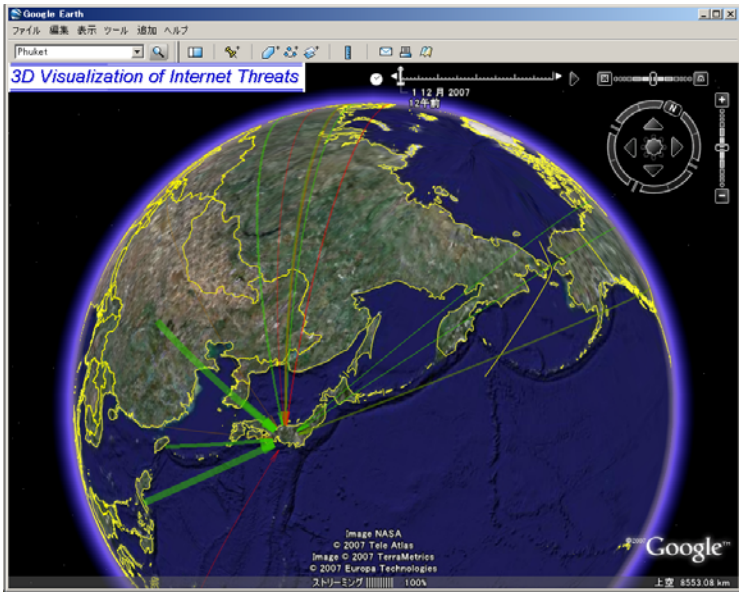

図 5.34: 3 次元可視化・アニメーション表示

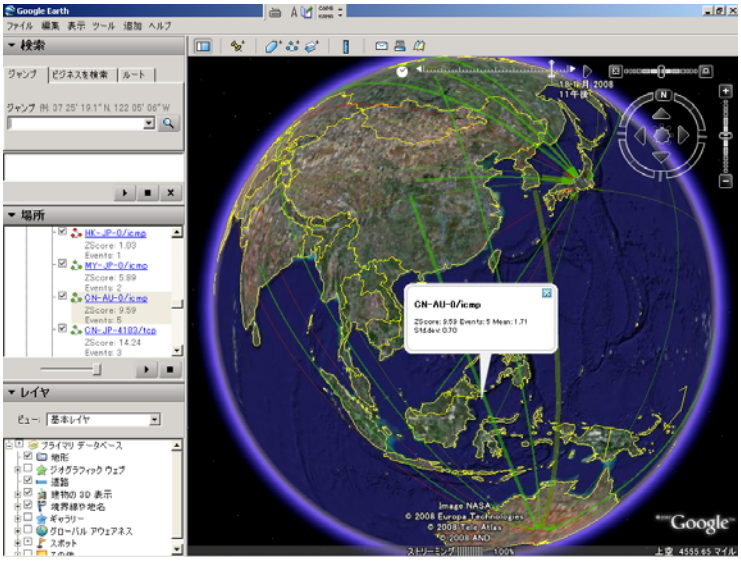

図 5.35: リアルタイム脅威可視化 (数値情報ポップアップ)

第 6章

分析手法の分類と関係

6.1 検知対象と分析手法の関係

本研究では、脅威をワームの感染力と定義した。ワームの感染力は、攻撃するポートに脆弱性を持つホストがインターネット上に存在する数と、脆弱なホストへの感染探索の効率性によって決まる。

$$(\text{感染力}) \propto (\text{脆弱なホスト数}) \times (\text{感染探索の効率性}) \tag{6.1}$$

感染力を決めるこれらの要素は、IP アドレス空間が膨大であり、脆弱なホストや感染ホストが常に変化し続けるため、リアルタイムの計測は現実的ではない。グラフ構造分析法は、不正パケットの送受信関係から、脅威を推定することを目的とした評価手法である。本研究では、このような脅威の推定を目的とした手法を、「脅威分析手法」と呼ぶ。

グラフ構造分析法は、不正パケットの送受信の双方を区別して構造を細かく分析するため、確率的な誤差の影響が大きいなどの問題がある。一方、周期成分変化検出法や自己相関変化検出法は、脅威そのものを評価すること無く、インターネット上の不正パケットのパターンの変化を検出する手法である。脅威の高いワームは、脆弱性を持つホストが多く存在するポートに対して、効率的にアクセスを行う新種の出現によって発生する可能性がある。長期間攻撃が継続されたポートは、パッチ等により脆弱性が修復されたり、感染ホスト数が飽和状態に近づくなど、脅威レベルは下がる。本研究では、このような過去のパターンからの変化を検出する手法を「異常検知手法」と呼ぶ。脅威分析手法と異常検知手法をまとめて、「脅威検知手法」と呼ぶ。

異常検知手法では、変化のみに着目するため、検出された事象が、本当に脅威を示すも

のかどうか、人手による詳しい分析を必要とするが、より少ない観測データから早期に検出できるメリットがある。本研究では、グラフ構造分析法のような脅威評価手法に加え、異常検知手法を組合せることで、早期の脅威検出が可能になる。

開発した脅威検知手法とワームの感染フェーズの関係を示したものが図 6.1 である[1] 。異常検知手法は、早期検知を利用することが適している。また、検出漏れを防ぐために検出特徴量の網羅性が重要である。逆に、誤検知を防ぐために、脅威分析手法と組合わせることで、フィルタリングを行うことが有効である。一方、脅威分析手法は、高精度の脅威検知に適している。ただし、検出にはある程度の観測データを必要とするため、異常検知手法による早期検知と組合わせることが有効である。

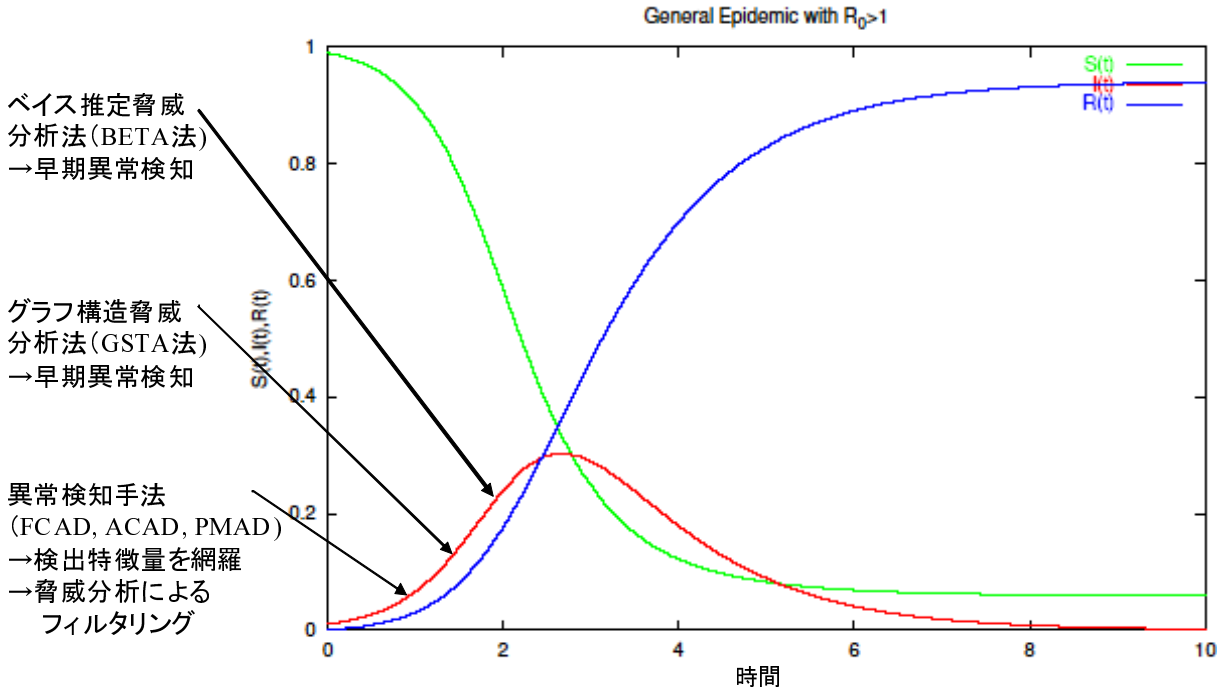

図 6.1: 脅威検知手法とワーム感染フェーズの関係

以上のことから、本研究で開発した脅威検知手法の特徴や関係を整理すると表 6.1 のようにまとめられる。

[1] SIR モデル [15] の感染数時間推移グラフ上に本手法を示したもの。

表 6.1: 提案する脅威検知手法の分類と関係

	手法	分析特徴量	検知対象ワーム	問題点
脅威分析手法（汎用）	ベイス推定脅威分析法（BETA法）	不正パケットの増加（マクロ・空間）	拡散したワーム	拡散後に検出される。
	グラフ構造脅威分析法（GSTA法）	感染の効率性（ミクロ・空間）	拡散開始期のワーム	大量のデータが必要になる。
異常検知手法（早期）	周期成分異常検知法（FCAD法）	周期性の変化（マクロ・時間）	周期性の異なる新種のワーム	脅威の高くない新種のワームの検知（誤検知）
	自己相関異常検知法（ACAD法）	過去のパターンの変化（マクロ・時間）	集団パターンの異なる新種のワーム	脅威の高くない新種のワームの検知（誤検知）
	パターンマイニング異常検知法（PMAD法）	攻撃元別パターンの変化（ミクロ・時間）	攻撃パターンの異なる新種のワーム	脅威の高くない新種のワームの検知（誤検知）

6.2 特徴量による手法の分類

前節の脅威分析手法と異常検知手法は、共に検出する特徴量に基づき分類することができる。第 3 章では、インターネット脅威分析手法に関する関連研究を、マクロ分析・ミクロ分析および時間特徴量分析・空間特徴量分析を基準として分類した。インターネット上の脅威は、異なる特徴・振舞いを持つ攻撃が同時多発的に発生しているため、様々な攻撃に対応して幅広く脅威を検知するためには、このような分類軸で表される異常検知手法を幅広く実装し、同時に適用することが必要である。

ここでは、本研究で提案した手法 (表 6.1) を含めて分類整理する。

マクロ分析は、観測パケットの送信元の母集団を推定し、その変化を検知する手法である。ミクロ分析は、観測パケットを送信元別に区別し、それぞれ送信元の不正パケットの振舞いパターンの変化を検出する方法である。

空間特徴量分析と時間特徴量分析は、明瞭には分かれないが、個々の特徴量自体に時系列特徴量をもつものを時間特徴量分析とし、主にＩＰアドレスの空間的な特徴を対象としているものを空間特徴量分析と分類する。

本研究で提案したベイズ推定脅威検知法、周期成分変化検知法、グラフ構造分析法、自

表 6.2: 脅威分析手法の分類

	空間特徴量分析	**時間特徴量分析**
マクロ分析 (**母集団推定**)	ベイズ推定脅威検知法 統計偏差 (Z スコア) 法 ポート間相関分析 送信元エントロピー カルマンフィルター感染率推定 主成分分析による異常検知	周期成分変化検知法 自己相関変化検知法 送信先エントロピー
ミクロ分析 (**振舞い分析**)	グラフ構造分析法 送信先 IP/ポート集合分析	アクセスパターン分析

己相関変化検知法は、表 6.2 のように位置付けられる。

グラフ構造分析法は、送信元 IP アドレスごとの送信先ポートの多様性と送信元 IP の集団としての多様性の両方を分析しているため、ミクロ分析とマクロ分析の両方を融合した手法とも考えることができるが、便宜上、前者に力点を置くためミクロ分析に分類した。

6.3 分析手法の利用者と利用法

本研究で示した脅威検知手法は、通信会社等の SOC(Security Operation Center), 組織内の CSIRT などのセキュリティ専門家から、組織のネットワーク管理者、セキュリティ担当者や、一般のネットワークユーザまで、インターネット上のセキュリティインシデントに影響を受ける様々な人に利用されることが考えられる。一般のユーザに対しては、CSIRT が、脅威検知手法によるネットワーク状況に関する情報を公開することで、利用することが可能になる。

インターネット上には，性質や振舞いの異るさまざまなワームによる脅威が存在する．これら特徴の異る多様な脅威に対応するためには，本研究で提案するような複数の手法を組合わせて，多角的に分析することが必要である．

提案した脅威検知手法のうち、脅威分析手法は、ワームの感染力に基づく脅威を評価するため、検出されたものは高い確率で危険であることを示す。一方、異常検知手法は、脅威とは関係なく、過去の不正パケットからの変化を早期に検出するためのものであるため、

検出されたインシデントに対して、個々の不正パケットを人手で詳細に分析することが必要になる。いずれの手法も、それぞれ異なる特徴量を対象に、脅威評価あるいは異常検知を行っている。したがって、それぞれの手法の独立性が高く、各手法を組み合わせて利用することにより、6.1で示した各手法で検出できない対象をお互いに補完する効果がある。

また、第5.7節で示した3次元可視化手法と併用することにより、地球上の脅威の変化の全体像を捉えて、個々の脅威検知手法の結果の理解に役立てることができる。このような複数の手法を組合わせることで，インターネット上のインシデントの状態を深く把握することが可能となり，脅威の性質や原因の分析に役立てることができる．

インシデント対応における開発手法の利用の流れをまとめたものが図6.2である。まず、異常検知手法により早期警戒を行い、それにより異常が検知された場合には、マニュアルによる観測データの原因分析を行うと共に、CSIRT, SOCなどの専門機関に早期警報を伝達する。また、監視体制を強化し、脅威分析手法による高精度検知を利用する。　脅威検知手法においても脅威レベルが高いことが確認された場合には、一般向けに警報を発すると共に、ベンダーによる脆弱性パッチ対策の推進を図る。

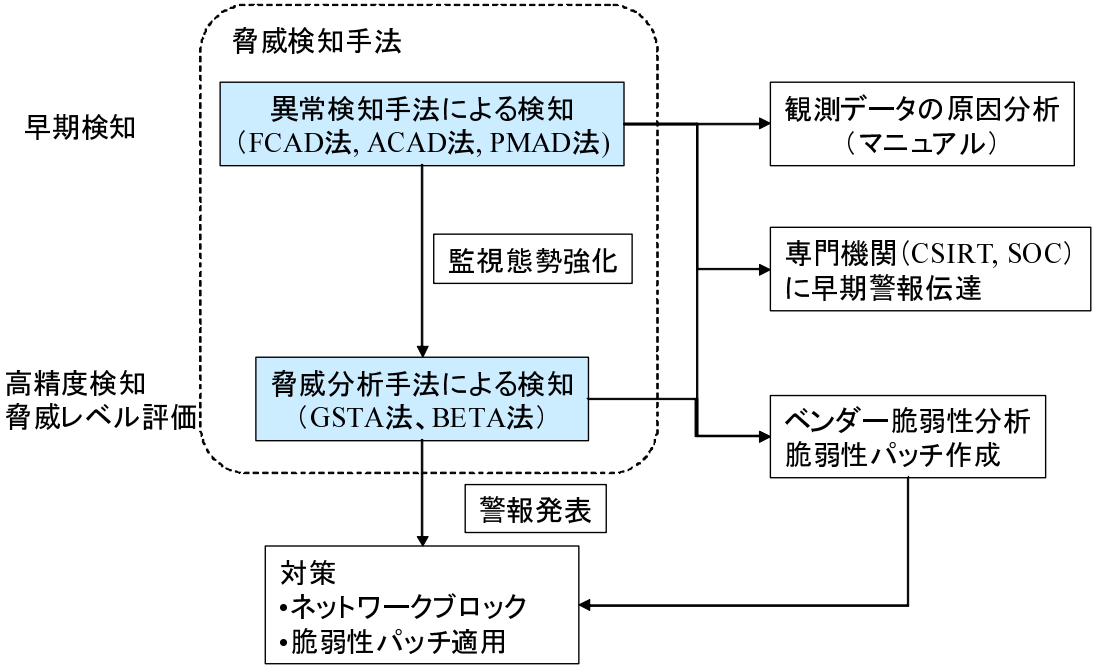

図 6.2: インシデント対応における脅威検知手法の利用の流れ

提案した各脅威検知手法は、特定のワームの特徴を前提とはしていないため、将来に渡っ

て比較的汎用に用いられる。しかし、近年のボットでは、攻撃を目立たないように行うスピア型が増えている。このようなボットにおいても、感染のための不正パケットは、局所選好ランダム探索戦略が有効である限り、本提案手法の検知できると考えられる。一方で、表 6.2 に示した脅威検知手法は、各分類領域をカバーしているように見えるが、それぞれの領域の特徴量は、ここで挙げたもの以外に無数に存在する。異常検知手法については、異なる特徴量についての検知はできないため、将来新たな特徴を持つ脅威が発生した場合には対応できない可能性がある。一方、脅威分析手法の場合には、脅威の定義であるワームの感染力を評価しているため、異なる特徴を持つワームが出現しても、感染力の高さとして検出ができるという汎用性を持つ。ただし、脅威の評価に、多くの観測データが必要になるといった課題の解決が必要になる。

第 7章

結論

本研究では、インターネット上のワームなどから送信される非正規パケットの分布や原因について考察し、脅威の種類に応じた複数の脅威検知手法を示し、観測データによる評価実験を行った。インターネット上の脅威は、異なる特徴・振舞いを持つ攻撃が同時多発的に発生しているため、単一の分析手法で、あらゆる脅威を検知する万能な手法は存在しない。様々な攻撃に対応して幅広く脅威を検知するためには、様々な特徴量をベースとする脅威検知手法を同時に適用することが必要である。本研究では、従来の非正規パケット数の変化だけからでは検知が困難な脅威を検知する手法を開発することで、インターネット上の脅威検知に貢献する基礎的な環境を構築した。

謝辞

本研究を行なうに当たり, 終始御指導を賜わった篠田陽一教授に深謝致します.

また、本論文をまとめるに当たって御協力いただいた村瀬一郎氏、鈴木裕信氏、早稲田大学の後藤滋樹教授、群馬大学の浅香緑准教授に厚く御礼申し上げます.

参考文献

[1] CERT/CC. Cert advisory ca-2001-26 nimda worm. http://www.cert.org/advisories-/CA-2001-26.html.

[2] DShield.org. Distributed intrusion detection system. http://www.dshield.org/-index.html.

[3] Richard O. Duda et al. *Pattern Classification*. John Wiley & Sons, 2001.

[4] SANS Institute. Worm propagation and countermeasures, 2004.

[5] IPA. 新種ワーム「w32/sasser」に関する情報. http://www.ipa.go.jp/security/topics-/newvirus/sasser.html.

[6] IPA. インターネット定点観測での観測状況について. http://www.ipa.go.jp/security/txt/2005/documents/TALOT2-0507.pdf, 2005.

[7] Masaki Ishiguro, Shigeki Goto, Hironobu Suzuki, and Ichiro Murase. Analyses on distribution of malicious packets and threats over the internet. In *Proceedings of APAN Network Research Workshop*, pages pp.9–16, Aug. 2007.

[8] Masaki Ishiguro, Hironobu Suzuki, Ichiro Murase, and Hiroyuki Ohno. Internet threat detection system using bayesian estimation. In *16th Annual FIRST Conference on Computer Security Incident Handling*, 2004.

[9] Masaki Ishiguro, Hironobu Suzuki, Yoichi Shinora, Ichiro Murase, and Shigeki Goto. An internet threat evaluation method based on access graph of malicious packets. In *19th Annual FIRST Conference on Computer Security*, 2007.

[10] Masaki Ishiguro, Hideyuki Tanaka, and Kanta Matsuura. The effect of information security incidents on corporate values in the japanese stock market. In *The Workshop on the Economics of Securing the Information Infrastructure (WESII)*, 2006.

[11] JPCERT/CC. インターネット定点観測システム internet scan data acquisition system (isdas). http://www.jpcert.or.jp/isdas/.

[12] W. Kermack and A. McKendrick. A contribution to the mathematical theory of epidemics. In *Proc. R. Soc.*, volume A 115, pages 700–721, 1927.

[13] Ramana Rao Kompella, Sumeet Singh, and George Varghese. On scalable attack detection in the network. In *4th ACM SIGCOMM conference on Internet measurement*, pages 187 – 200, 2004.

[14] A. Lakhina, M. Crovella, and C. Diot. Characterization of network-wide anomalies in traffic flows. In *Proc. Internet Measurement Conference*, October 2004.

[15] Jan Medlock. Mathmatical modeling of epidemics. manuscript, may 2002.

[16] David Moore, Colleen Shannon, Geoffrey M. Voelker, and Stefan Savage. Network telescopes: Technical report. Technical report, CAIDA, 2004.

[17] David Moore, Collen Shannon, and Jeffry Brown. Code-red: A case study on the spread and victims of an internet worm. In *Internet Measurement Workshop*, pages 273–284, 2001.

[18] R. Pang, V. Yegneswaran, P. Barford, V. Paxson, and L. Peterson. Characteristics of internet background radiation. In *Proceedings of ACM Internet Measurement Conference*, 2004.

[19] @police. W32/codered.f. http://www.cyberpolice.go.jp/server/virus/pdf/W32CodeRed_F_jp_20030302.pdf.

[20] @police. インターネット定点観測. http://www.cyberpolice.go.jp/detect/observation.html.

[21] @police. インターネット定点観測. http://www.cyberpolice.go.jp/detect/observation.html.

[22] The Honeynet Project. Tools for honeynets. http://www.lucidic.net/.

[23] Moheeb Abu Rajab, Fabian Monrose, and Andreas Terzis. On the effectiveness of distributed worm monitoring. In *14th USENIX Security Symposium*, pages 225–237, 2005.

[24] SANS Institute. Internet storm center. http://isc.sans.org/.

[25] Stuart Schechter, Jaeyeon Jung, and Arthur W. Berger. Fast detection of scanning worm infections. In *7th International Symposium on Recent Advances in Intrusion*, 2004.

[26] Kenneth Theriault, Daniel Vukelich, Wilson Farrell, Derrick Kong, and John Lowry. Network traffic analysis using behavior-based clustering. BBN Technologies Technical Paper.

[27] University of Michigan. Internet motion sensor (ims). http://ims.eecs.umich.edu/index.html.

[28] Arno Wagner and Bernhard Plattner. Entropy based worm and anomaly detection in fast ip networks. In *14th IEEE International Workshop on Enabling Technologies: Infrastructures for Collaborative Enterprises*, 2005.

[29] K Yamanishi and J Takeuchi. A unifying approach to detecting outliers and change-points from non stationary data. In *8th ACM SIGKDD International Conference on Data Mining and Knowledge Discovery*, pages 676–681, 2002.

[30] Cliff C. Zou, Don Towsley, and Weibo Gong. On the performance of internet worm scanning strategies. *Perform. Eval.*, 63(7):700–723, 2006.

[31] Cliff Changchun Zou, Lixin Gao, Weibo Gong, and Don Towsley. Monitoring and early warning for internet worms. In *the 10th ACM conference on Computer and communications security*, pages 190 – 199, 2003.

[32] 久保田和己, 鳥居悟, and 小谷野修. 不正アクセスシナリオの導出に向けた検知ログ解析. In **情報処理学会第** *64* **回全国大会**, pages 379–380, 2002.

[33] 竹内 純一, 佐藤 靖士, 力武 健次, and 中尾 康二. 変化点検出エンジンを利用したインシデント検知システムの構築. In *SCIS2006*, 2006.

[34] 寺田 真敏, 高田 真吾, and 土居 範久. ネットワークワーム動作検証システムの提案. **情報処理学会論文誌**, 46(8):2014–2024, 2005.

[35] 石黒 正揮, 鈴木 裕信, 村瀬 一郎, and 篠田 陽一. インターネット上の脅威分析を支援する空間および時間的な特徴量に基づく分析手法. **情報処理学会論文誌** *Vol.48, Number 9, pp.3148-3162, Sep. 2007*, 48(9):pp.3148–3162, Sep. 2007.

[36] 石黒正揮. It 事故および情報セキュリティ対策の企業価値に及ぼす影響. 3 2007. リスク定量化ワークショップ.

[37] 中尾康二, 松本文子, 井上大介, 馬場俊輔, 鈴木和也, 衛藤将史, 吉岡克成, 力武健次, and 堀良彰. インシデント分析センタ nicter の可視化技術. 情報処理学会研究報告 Vol.2006, No.81, 2006-CSEC-034, pp.313-319.

[38] 日本数学会. **数学辞典**. 岩波書店, 1985.

[39] 鈴木裕信, 石黒正揮, 村瀬一郎, and 大野浩之. インターネット早期広域攻撃警戒システム wclscan. In **ソフトウェアシンポジウム** *2004* **予稿集**. http://www.clscan.org, 2004.