

我が国におけるインターネット治安情勢について

1 はじめに

今般、インターネットは社会生活において不可欠な存在となりつつある。インターネットにはウイルスやワーム、不正アクセスといった様々なリスクが存在しており、これらリスクを小さくするために情報セキュリティ対策が重要な社会的課題となってきた。ゆえに、インターネット上で起こっている事象についての的確に把握するとともに、適切な対処を行っていく必要がある。

警察庁技術対策課では、現在、全国の警察機関のインターネット接続点（57 か所）において、侵入検知装置（Intrusion Detection System：IDS）や定点観測ポイントを設置し、インターネット上で発生する様々な事象について調査・分析を行っている。

平成 15 年におけるインターネット上で発生した事象に関して侵入検知装置のアラート情報や定点観測ポイントのログを基に分析、検討を行った。

2 アラート検知に見るサイバー攻撃の特徴

平成 15 年における外部ネットに起因するアラートの総検知件数は約 398,000 件、検知ホスト数は約 94,000 件であり、発信元の国や地域は 176 か国（発信元不明を含む）に及んでいる。

図 2.1 に平成 15 年中に検知した国/地域別の累積検知件数を示す。図 2.1 に示すように全世界から何らかの攻撃が行われていることがわかる。特に日本を含む東アジア地域、ヨーロッパ地域、北アメリカ地域からの攻撃が顕著となっている。

2.1 発信元 IP アドレスによる分析

(1) 国/地域別の検知状況

平成 15 年中に検知されたアラートの発信元 IP アドレスに基づいて国別に分類した上位 10 か国のアラート検知比率を図 2.2 に示す。

検知したアラート件数が最も多い国はアメリカ合衆国であり、全体の約 34%を占めている。次いで中華人民共和国の約 15%、大韓民国の約 7%の順となっており、日本国内からのアラートの検知件数は約 4%程度に留まっている。

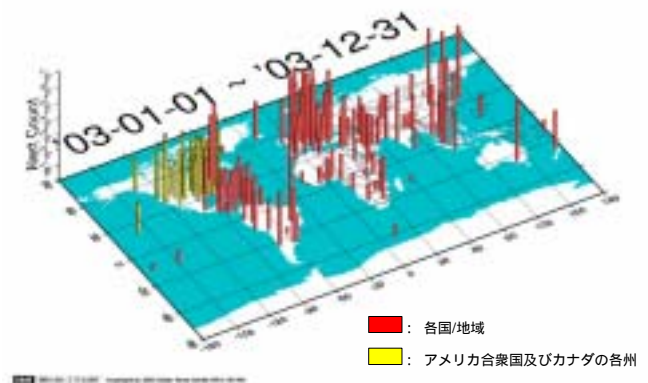


図 2.1 平成 15 年中における国/地域別の累積検知件数

チェコ共和国では検知したアラートの内、ウィンドウサイズ 55808 の TCP パケットが約 95%を占めており、分散型ポートスキャナー「Stumbler」やワーム「Randex.C」などによる影響が考えられ、その送信元 IP アドレスは詐称されている可能性が高いと推察される。

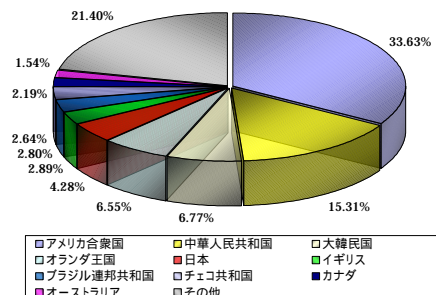


図 2.2 上位 10 か国のアラート検知比率

(2) ホスト数の検知状況

平成 15 年中に検知されたアラートの発信元 IP アドレスに基づいて国別に分類した上位 10 か国のホスト比率を図 2.3 に示す。

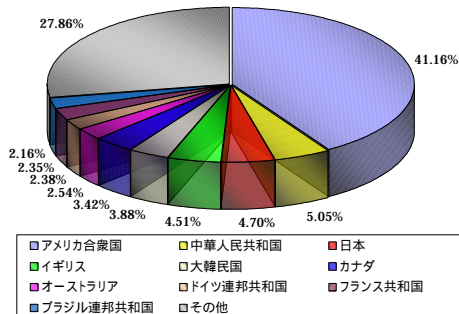


図 2.3 上位10か国のホスト比率

アラート検知比率と同様にアメリカ合衆国が約 41% と最も多く、次いで中華人民共和国の約 5% の順となっている。

アラートの検知比率では5位であった日本はホスト比率では 3 位となっている。また、アラートの検知比率で上位であったオランダ王国(NL)やチェコ共和国は 18 位及び 20 位となっており、小数のホストから多くの攻撃を行っていたことがうかがえる。

2.2 アラート種別による分析

(1) アラート検知の推移状況

平成 15 年中に検知したアラート種別の推移状況を図 2.4 に示す。

平成 15 年にインターネットにおいて社会的影響の大きかった事案の 1 つに 1 月に発生した SQL Slammer ワームが挙げられる。ワーム発生後には沈静化の傾向にあったが、2 月下旬に再び活動が活発となり 12 月においては 1 日当たり約 700 件前後で推移しており一向に沈静化の気配はない。図 2.4 中のワーム(Worm)系のアラートの推移状況はこの SQL Slammer ワームのみの検知件数である。

3 月には、アメリカ合衆国やイギリスの各サイトに対してイラク戦争に反対する旨の Web ページ改ざん事件が発生している。日本国内においても同様の Web ページ改ざん事件が数件発生している。

4 月からポートスキャン(Scan)系のアラートが

増加してきている。これはオランダ王国の特定ドメインからの TCP1080 ポートへのスキャンが主な原因である。

この特定ドメインからのポートスキャンはプロキシサーバで使用されるポート (80、3128、6588、8080 等) に対してインターネット中をスキャンしており、多くの HTTP プロキシソフトウェアの脆弱性 (CERT/CC Vulnerability Note VU#150227) を悪用して、スパムメールの送信を行っていた。この特定ドメインとスパム業者との関わりがイギリスの記者によって明らかにされたことを受け、7 月 1 日付で上位プロバイダからサービスを打ち切られたことを最後に検知されなくなった。

7 月には Web 改ざんコンテスト「The Defacers Challenge(TDC)」が行われた。この Web 改ざんコンテストに伴い国内においても不正アクセスが増加する恐れがあり、警戒を強化していたが国内の被害は数件程度に留まった。このコンテスト開始直前には中華人民共和国の特定サイトから定点観測ポイントの 1 ホストに対して全ポートスキャンが行われている。

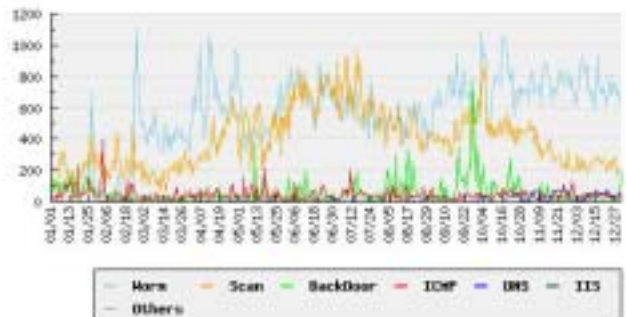


図 2.4 平成 15 年におけるアラート種別の推移状況

8 月には「RPC インターフェイスのバッファ オーバーランによりコードが実行される (823980) (MS03-026)」の脆弱性を悪用した Blaster ワーム及び Blaster ワームの亜種である Welchia ワームが発生した。

これらワームは侵入検知装置では検知できなかったものの、定点観測によるトラフィックの急増によりその把握を行ってきた。Welchia ワームに関しては ICMP パケットの急増をいち早く検知し注意喚起を行っている。

これらワームの発生後、Blaster ワームの TCP135 ポートへの攻撃や Welchia による ICMP エコーが依然高い件数で推移していた。

(2) アラート種別の検知状況

平成 15 年中に検知されたアラート種別による検知比率を図 2.5 に示す。

図 2.5 に示すように検知されたアラートの種別には「Worm」系が約 53%、「Scan」系に関して約 36%と上位 2 つのアラートで全体の約 90%を占めている。

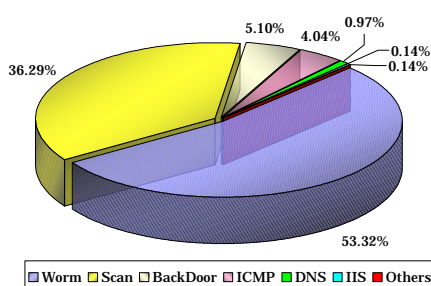


図 2.5 アラート種別比率

この要因は、サイトに侵入するための準備段階の攻撃に関連するアラートが中心となっており、サーバ等への脆弱性を悪用した直接的な攻撃等は約 10%程度に留まっている。これは、各接続点でのセキュリティを強固に設定しているために、一方的にパケットを送りつけてくる SQL Slammer ワームやポートスキャン等の準備行為のみであきらめてしまう攻撃者が多いためと推察される。

(3) 宛先ポートと発信元 IP アドレス

図 2.6 に平成 15 年に検知したアラートの宛先ポートと発信元 IP アドレス別の累積件数を示す。X 軸は発信元 IP アドレス、Y 軸は宛先ポート番号 (TCP/UDP)、Z 軸は累積検知件数である。

発信元 IP アドレスでは、クラス A 第 1 オクテットが 60 及び 80 付近、クラス C の第 1 オクテットが 210 付近からの攻撃が、宛先ポートでは、21 番、1080 番、1434 番、27374 番ポートへの攻撃が顕著となっている。

また、52076 番ポートへのアクセスに関しては他のポートと比較し多く存在するが、発信元 IP アドレスがチェコ共和国のウィンドウサイズ

55808 の TCP パケットである。

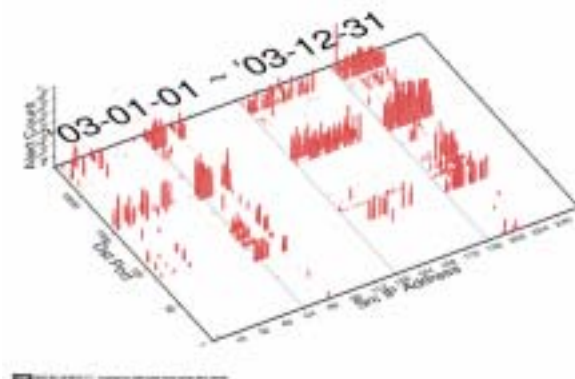


図 2.6 宛先ポートと発信元 IP アドレスの累積検知件数

2.3 アラート件数による主成分分析

前節で分析を行った、国/地域別及びアラート種別のアラート件数を基に主成分分析により各国の特徴に関して調査を行った。

(1) 主成分分析とは

主成分分析とは、相関関係にあるいくつかの要因から複数の成分を抽出し、その背後に隠れている総合力や特性などを求める方法である。主成分分析では、選択された 数個の変数からなる行列の相関行列または分散共分散行列から、固有値、固有ベクトルを求める。(1)~(2)

(2) 主成分の算出

表 2.1 に主成分分析を行った結果 (固有値、固有ベクトル及び寄与率(%)) を示す。

主成分を抽出後、いくつの主成分を採用するか判断する必要がある。主成分の採用の目安は固有値が 1 以上で累積寄与率が 70%以上の水準を満たしていなければならない。このため、本結果においては第二成分までを採用することとする。

次に、各主成分の固有ベクトルに着目し採用した 2 つの主成分の持つ意味について考えてみる。

第 1 主成分は、全ての固有ベクトルの値がプラスであり、「IIS」及び「Others」以外の固有ベクトルがほぼ同じ値であることから、攻撃の活動の度合いを表す指標と解釈できる。

表 2.1 主成分の抽出結果

		主 成 分						
		第1成分	第2成分	第3成分	第4成分	第5成分	第6成分	第7成分
固有値		4.350	1.209	0.913	0.309	0.142	0.072	0.005
固有ベクトル	Worm	0.454	0.051	0.201	0.031	-0.639	-0.258	0.525
	Scan	0.428	-0.214	0.060	-0.560	0.468	-0.484	-0.007
	BackDoor	0.426	-0.318	-0.184	-0.199	0.056	0.776	0.195
	ICMP	0.409	-0.207	-0.056	0.802	0.351	-0.142	-0.002
	DNS	0.451	0.206	-0.220	-0.039	-0.356	-0.009	-0.760
	IIS	0.200	0.468	0.775	0.017	0.237	0.276	-0.093
	Others	0.138	0.739	-0.520	-0.026	0.252	-0.001	0.316
寄与率		62.15%	17.28%	13.04%	4.41%	2.02%	1.03%	0.06%
累積寄与率		62.15%	79.43%	92.46%	96.88%	98.90%	99.94%	100.00%

また、第2主成分に関しては固有ベクトルの値にプラスとマイナスの符号が入り混じっており、その符号と絶対値に着目してみる。まず、マイナスの項目の固有ベクトルには、「BackDoor」や「Scan」、「ICMP」のアラートである。「BackDoor」は、バックドアを仕掛ける攻撃ではなく、バックドアを仕掛けられたホストを探索する行為であり、「Scan」、「ICMP」などもポートスキャンやホスト探索などの行為である。逆にプラスの項目の固有ベクトルには、「Others」や「DNS」、「IIS」のアラートがある。「Others」の固有ベクトルが0.739と高いのは、「DoS」関連のアラートが「Others」に含まれているためと推察される。これらのアラートは脆弱性を悪用した直接的な侵入行為であると考えられる。

したがって、第2主成分はプラスであれば脆弱性を悪用した侵入行為を行っていることを、また、マイナスであればポートスキャンなどの侵入に対する準備行為を行っていることを示す指標と解釈できる。

(3) 各国の主成分得点

表2.1の結果から第1主成分及び第2主成分における各国/地域の基準化した主成分得点を算出し、第1主成分の上位10か国及び第2主成分の上位/下位10か国を表2.2に示す。

第1主成分は攻撃の活動を示す指標であり、最も活発な国はアメリカ合衆国、次いで中華人民共和国の順となり、2.1節の上位10か国のアラート検知比率と同様の結果が得られている。

第2主成分は侵入に対する準備行為か実際の侵入行為かの指標であり、ポートスキャン等の準備行為が顕著な国にはアメリカ合衆国、大韓民国、オランダ王国が上位となっており、TCP1080ポートに対するスキャンが多かったことが要因として挙げられる。

また、実際の侵入行為が顕著な国は日本、オーストラリア、ポーランド共和国の順となっている。これは、日本やオーストラリアなどではスキャン系のアラートは少なくワーム関連のアラートが殆どを占めているためである。

表 2.2 各国の主成分得点

第1主成分			第2主成分					
上位10か国			上位10か国			下位10か国		
国/地域名	ID	得点	国/地域名	ID	得点	国/地域名	ID	得点
アメリカ合衆国	US	11.434	日本	JP	0.983	アメリカ合衆国	US	-10.231
中華人民共和国	CN	5.147	オーストラリア	AU	0.294	大韓民国	KR	-5.992
大韓民国	KR	2.071	ポーランド共和国	PL	0.280	オランダ王国	NL	-4.816
オランダ王国	NL	2.016	メキシコ合衆国	MX	0.256	チェコ共和国	CZ	-1.865
日本	JP	1.283	オーストリア共和国	AT	0.237	ラトビア共和国	LV	-0.751
イギリス	GB	0.806	マレーシア	MY	0.198	国名不明	?	-0.710
ブラジル連邦共和国	BR	0.790	ノルウェー王国	NO	0.196	カナダ	CA	-0.657
チェコ共和国	CZ	0.696	ハンガリー共和国	HU	0.195	中華人民共和国	CN	-0.557
カナダ	CA	0.554	ベトナム社会主義共和国	VN	0.191	イタリア共和国	IT	-0.400
オーストラリア	AU	0.348	エルサルバドル共和国	SV	0.191	イギリス	GB	-0.341

第一主成分と第二主成分の各得点の散布図を図 2.7 に示す。

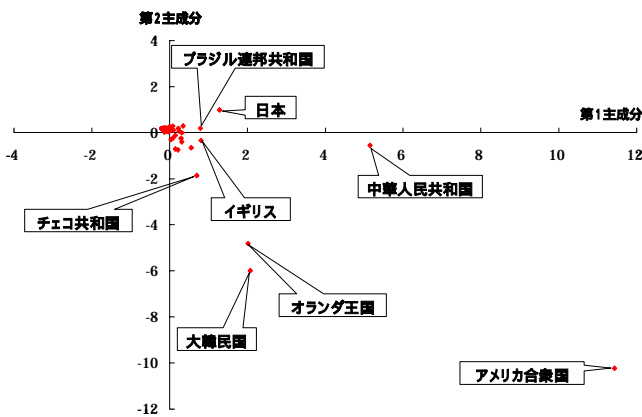


図 2.7 第 1 主成分と第 2 主成分得点の散布図

図 2.7 に示すように、第 2 主成分がマイナス領域に属する国や地域が顕著であり、前節で述べたようにホスト探索やポートスキャンなどの準備行為であきらめてしまう攻撃者が多いためと推察される。

3 Blaster ワームと SQL Slammer ワーム

平成 15 年のインターネットにおける最大の事案は、1 月に発生した SQL Slammer ワーム、8 月に発生した Blaster ワームの 2 つのワームが挙げられる。これら 2 つのワームに関して時系列解析を行った。

3.1 Blaster ワーム

(1) Blaster ワームの概要

Windows の Remote Procedure Call(RPC) を使用する Distributed Component Object Model (DCOM) インターフェイスの脆弱性(マイクロソフト社、「RPC インターフェイスのバッファオーバーランによりコードが実行される (MS03-026)」) を悪用した Blaster ワームが 8 月 12 日に発生し、また 18 日にも同脆弱性を悪用した Welchia ワームが発生し社会的にも大きな影響を与えた。

脆弱性が発見される以前は、TCP135 ポートへのアクセス件数は 1 日当たり 30 件程度で推移し

ていたが、Blaster ワームの発生により 12 日は 9,220 件に急増した。その後、17 日までは緩やかに減少傾向にあったが、18 日の Welchia ワームの発生により再び急増している。これは、Welchia ワーム発生当時、いくつかの定点観測ポイントにおいて ICMP に応答する設定であったため、Welchia ワームに感染したホストからの TCP135 番ポートへのアクセスが増加したためである。

Blaster ワーム発生時の定点観測ポイントにおけるポート別のアクセス状況を図 3.1 に、国/地域別のアクセス状況を図 3.2 に示す。

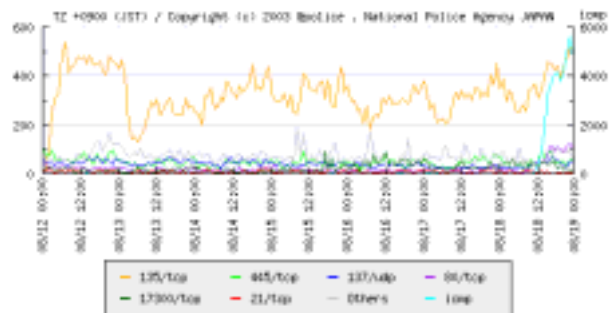


図 3.1 Blaster ワーム発生時のポート別アクセス状況



図 3.2 Blaster ワーム発生時の国/地域別アクセス状況

(2) テンドの抽出

常に変化を繰り返す時系列データにおいて増加傾向にあるか逆に減少傾向にあるのかトレンドを的確に把握しておく必要がある。このトレンドは把握する一つの手法に移動平均がある。移動平均線により時系列データの不規則変動が除去されトレンドや周期変動を抽出することができる。

時系列データのトレンドの抽出には単純移動平均線や加重移動平均線などがあるが、タイムラグが小さく、小さなトレンドに迷われないという点で単純移動平均線より有効な指数平滑移動平均線を用いる。指数平滑移動平均線の計算式は次式で表される。(3)~(4)

$$y(t) = y(t-1) + (P - y(t-1))$$

：平滑化定数 = $2 / (n + 1)$; 0 1

y(t) : 時間 t における平均

y(t-1) : 時間(t-1)における平均

P : 時間 t における観測値

n : 平均する期間

使用したデータは Blaster ワームに感染した端末数の状況を把握するため、アクセス件数でなく 1 時間当たりの TCP135 番ポートへアクセスのあったホスト数の時系列データによりトレンドの抽出を行った。

り、その後勾配が緩やかになってきておりワームの感染が抑えてきていることが分かる。8 月 18 日 午前 11 時頃発生後には再びトレンドの傾きが急になっている。これは、定点観測の幾つかの拠点において ICMP に応答する設定になっていたためであり、Blaster ワームの亜種である Welchia に感染したホストによるものと考えられる。

また、8 月 25 日以降、ホスト数が激減しているのは、これら拠点において行ったポリシー変更を行った結果、Welchia ワームの影響を受けなくなったためと考えられる。

(3) 時系列データの変動性

トレンドの把握とともにその事象がどの程度変動しているか、変動性を客観的に把握する必要もある。ここでは、株式などで一般的に用いられている変動性を測る指標の一つである RSI (Relative Strength Index) を適用して時系列データの変動性の検討を行った。

RSI の指標値は、次式で表される。(5)

$$RSI(\%) = U / (U + D) \times 100$$

U : 指定期間内における増加データの平均

D : 指定期間内における減少データの平均

RSI で得られた指標は、80%以上または 20%以下を異常な増減である目安とされている。

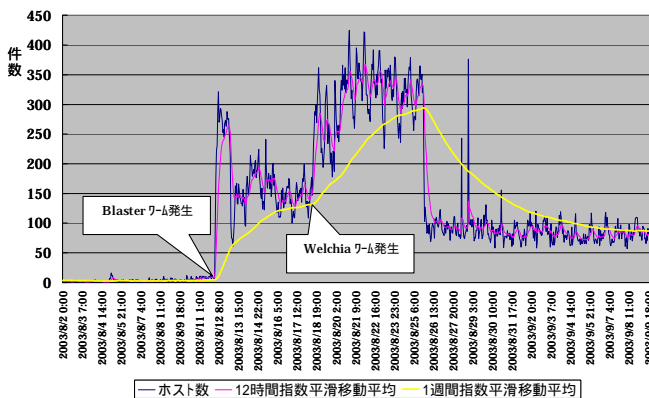


図 3.3 Blaster ワーム発生前後のトレンド

ここでは短期的なトレンドには 12 区間(12 時間)を、長期的なトレンドには 168 区間(1 週間)を期間として指数平滑移動平均線を求めた。図 3.3 にその結果を示す。

図 3.3 に示すように、短期的トレンドの場合、時系列データに即応しすぎてしまい多少判断しづらい結果であった。しかし、長期的トレンドの場合では、移動平均線が直線的になりその傾向を把握しやすい結果となった。

Blaster ワーム発生直後の長期的トレンドの勾配は急になっており多くのホストが感染しつつあ

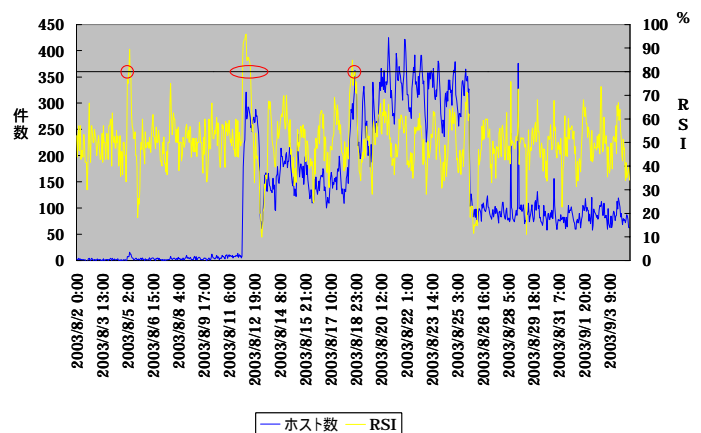


図 3.4 Blaster ワーム発生前後における RSI 値

このRSIを用いてBlasterワーム発生前後の感染したホスト変動性について分析を行ってみた。なお、指定期間を12区間(12時間)として計算を行い、その結果を図3.4に示す

図3.4に示すようにRSI値が80%を超えている箇所は7月28日にBugtraqに攻撃コードが投稿され警戒していた8月5日の4時から5時、Blasterワームが発生した8月3日の2時から13時とWelchiaワームが発生した8月18日の17時から18時と22時の4か所あり、それぞれ通常ではない異常なトラフィックの増加であると判断できる。

警察庁では、上記トラフィックが急増した8月5日に「TCP135番ポートに対するアクセスの急増について」、8月12日に「Windowsの脆弱性を利用したワームの蔓延について」及び8月18日に「ICMPトラフィックを急増させたワームについて」を警察庁セキュリティポータルサイト@police (<http://www.cyberpolice.go.jp>)に掲載し広く注意喚起を行っている。

3.2 SQL Slammer ワーム

(1) SQL Slammer ワームの概要

平成15年1月に発生したSQL Slammerワームは、Microsoft SQLサーバの持つ脆弱性(SQL Server 2000 解決サービスのバッファのオーバーランにより、コードが実行される(323875)(MS02-039))を悪用して感染し、感染後はランダムにホストを選択し同様の脆弱性を持つホストへと拡散していく。このワーム発生時には、多くのホストが一度に感染したため、トラフィックの急増により通信に支障をきたしたことは記憶に新しい。

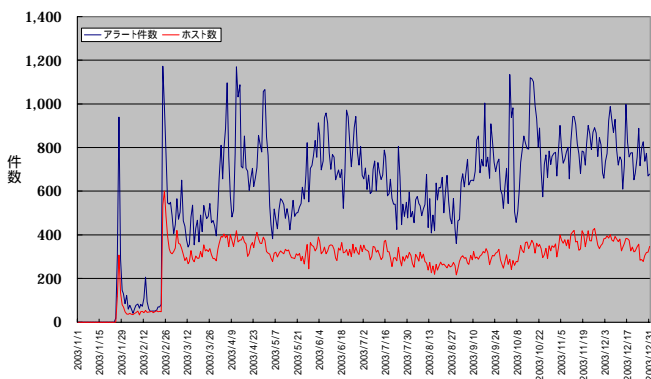


図 3.5 SQL-Slammer ワームの推移状況

SQL Slammer ワーム発生後の推移状況を図3.5に示す。

ワーム発生後には一時的に沈静化の傾向にあったが、2月下旬に再び活動が活発となり12月においては1日当たり約700件前後で推移している。日本国内においても発生後、1日あたりのアラート件数で約40件、検知ホスト数で約14ホスト前後とほぼ一定数で推移しており沈静化の気配はない。

また、平成15年に検知した発信元IPアドレスを基に国別の検知比率及び国別のホスト比率を図3.6、図3.7に示す。

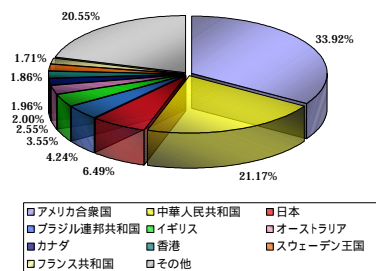


図 3.6 国別アラート件数比率

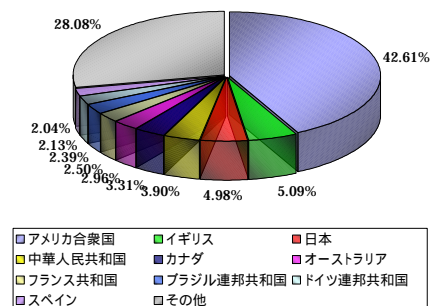


図 3.7 国別ホスト比率

国別の検知比率ではアメリカ合衆国が約34%と最も多く、次いで中華人民共和国の約21%、日本の約6%の順となっている。国別のホスト比率においてもアメリカ合衆国が約43%と最も多く、ホスト数が少なく検知数が多い中華人民共和国では、感染したホストが長期間に渡り適切な処置がなされていないためと推察される。

(2) 時系列モデル

SQL Slammer ワームの時系列解析に際し、10月1日から12月21日までの侵入検知装置で検出した1日当たりのホスト数(発信元IPアドレス)を標本データとして時系列モデルを作成し、12月22日以降の検知ホスト数の予測を行った。

時系列モデルにはデータが定常時系列の場合に用いられる、自己回帰モデル(AR(p)モデル)、移動平均モデル(MA(q)モデル)や自己回帰移動平均モデル(ARMA(p,q)モデル)があり、次式で表される。⁽⁶⁾

AR(p)モデル

$$y(t) = \mu + \alpha_1 y(t-1) + \alpha_2 y(t-2) + \dots + \alpha_p y(t-p) + u_t$$

$y(t-1), y(t-2), \dots, y(t-p)$: 時間(t-1), (t-2), ..., (t-p)における観測値

$\mu, \alpha_1, \alpha_2, \dots, \alpha_p$: パラメータ定数

u_t : 時間 t における攪乱項(ホワイトノイズ)

MA(q)モデル

$$y(t) = \mu + u_t + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \dots + \alpha_q u_{t-q}$$

$u_t, u_{t-1}, \dots, u_{t-q}$: 時間 t, (t-1), ..., (t-p)における攪乱項(ホワイトノイズ)

$\mu, \alpha_1, \alpha_2, \dots, \alpha_q$: パラメータ定数

ARMA(p,q)モデル

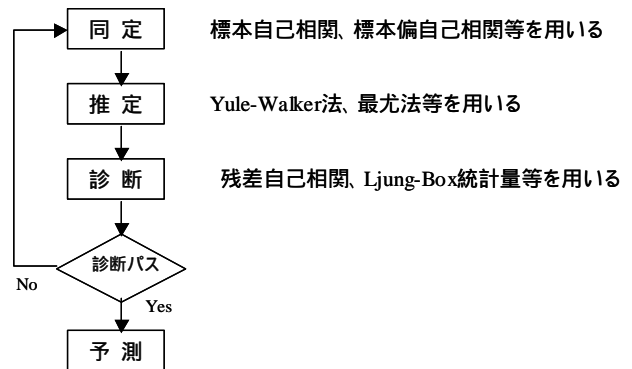
$$y(t) = \mu + \alpha_1 y(t-1) + \alpha_2 y(t-2) + \dots + \alpha_p y(t-p) + u_t + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \dots + \alpha_q u_{t-q}$$

AR モデルは現在の自分を過去の自分の加重和として表すもので、その影響大きさは時間に依存せず異時点間の自分に一定の規則性に従った係わり合いを持ったモデルであり、また、MA モデルは現在の自分を過去の攪乱(ホワイトノイズ)の加重和として表すもので、同様にその影響力も時間に依存しておらず異時点間の自分に対して一定の規則性が生じるモデルである。ARMA モデルは、AR モデル及び MA モデルを同時に含むモデルである。

一方、データが非定常時系列の場合には自己回帰和分移動平均モデル(ARIMA(p,d,q)モデル)や季節変動を考慮した季節性自己回帰和分移動平均モデル(SARIMA(p,d,q)モデル)などが用いられている。

非定常時系列モデルにおいては、差分や季節差分を取ることによって非定常時系列から定常時系列に変換を行い、前者の3モデルに当てはめて分析を行うことになる。

時系列モデルの作成に当たっては、図 3.8 に示す順序に沿って作業を進めていくことになる。



(出典元:山本 拓「経済の時系列分析」)

図 3.8 Box-Jenkins による時系列モデルの作成手順

図 3.5 に示すようにアラート件数と同様にホスト数もトレンドを持ち非定常時系列である。また、大きな季節変動が見受けられないことから自己回帰和分移動平均モデル(ARIMA(p,d,q)モデル)を用いて分析を行った。

非定常時系列では、まず定常時系列に変換する必要がある。このため与えられた時系列データの1次階差(d = 1)を取ることによりトレンドを除去し定常時系列への変換を行った。

その結果を図 3.9 に示す。

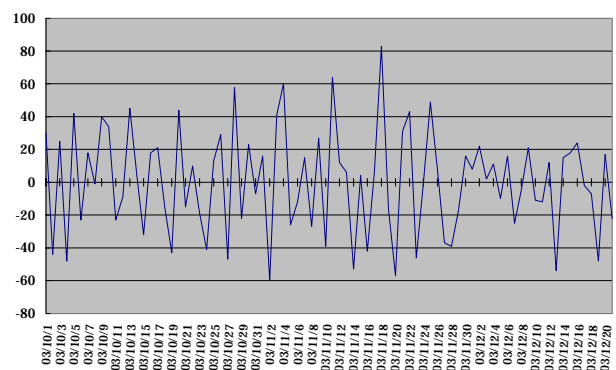
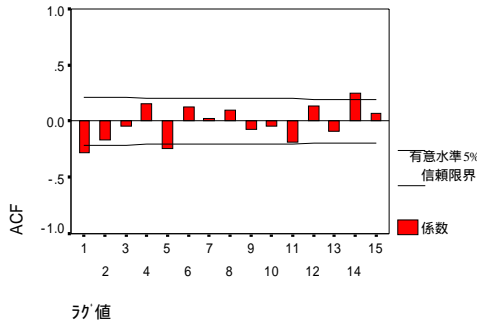


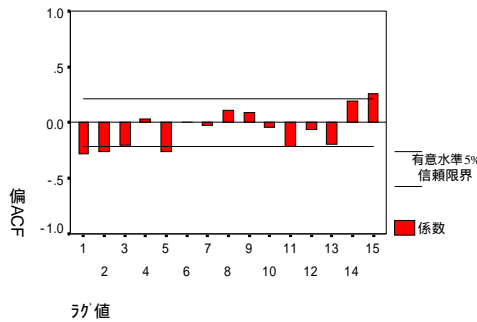
図 3.9 ホスト数の1次階差

次に、与えられた 1 次階差データから自己相関及び偏自己相関を使用してモデルの p,q を同定する必要がある。この 1 次階差データに対する自己相関及び偏自己相関を図 3.10 に示す。

図 3.10 に示すように自己相関は 1 次(-0.277)、5 次(-0.238)が有意水準 5%で有意であり、偏自己相関は 1 次(-0.277)、2 次(-0.260)、5 次(-0.258)が有意水準 5%で有意となっている。



(1) 自己相関



(2) 偏自己相関

図 3.10 階差データの自己相関と偏自己相関

また、各モデルの自己相関及び偏自己相関には表 3.1 に示すような特徴がある。

表 3.1 各モデルの自己相関と偏自己相関の特徴

	自己相関	偏自己相関
AR(p)モデル	幾何級数的減衰	ある次数以降切断
MA(q)モデル	ある次数以降切断	幾何級数的減衰
ARMA(p,q)モデル	幾何級数的減衰	幾何級数的減衰

以上のことからこのモデルには AR 部分及び MA 部分が含まれていることが予想され、幾つかの低次の ARMA モデルを作成し検討を行った。

各モデルの推定では、最小二乗法を用いて各係数の算出を行い、モデルの診断には推定結果より算出した残差の自己相関について Ljung-Box 検定を行った。

Ljung-Box 検定は、次の仮説が成立するかどうかを調べる検定方法である。

$$\text{仮説 } H_0 : \quad (1) = (2) = (3) = \dots = (m) = 0$$

(1), (2), (3)・・・, (m)は自己相関係数

また、Ljung-Box 検定量 Q は次式で表される⁽⁷⁾。

$$Q = t(t+2) \left\{ \frac{\rho(1)^2}{t-1} + \frac{\rho(2)^2}{t-2} + \frac{\rho(3)^2}{t-3} + \dots + \frac{\rho(m)^2}{t-m} \right\}$$

この検定量は自由度(m - p - q)の χ^2 分布に従うことから、

$$Q < \chi^2_{m-p-q, 0.05}$$

となる時に、仮説 H_0 が棄却されなければ良いことになる。

その結果を表 3.2 に示す。

表 3.2 Ljung-Box 検定量と AIC 値

ARIMA モデル			Ljung-Box検定量	情報量基準
p	d	q	Q ₁₅	AIC
1	1	1	15.585	6.732
1	1	2	18.836	6.759
1	1	3	17.704	6.779
2	1	1	19.736	6.759
2	1	2	16.014	6.779
2	1	3	13.122	6.721

表 3.2 に示す Ljung-Box 検定量(Q)は、Ljung-Box 検定において、各モデルについても残差に系列相関が無いという帰無仮説は有意水準 5%で棄却されなかったため、いずれのモデルにおいても妥当性を有している。

また、情報量基準 AIC(Akaike's Information Criterion)を用いて算出した値が最小なモデルは ARIMA(2,1,3) モデルとなり、下記の

ARIMA(2,1,3)モデルを最適モデルとして選択した。

$$Y(t) = 2.9624 - 1.0074Y(t-1) - 0.6747Y(t-2) - Ut + 0.6298U(t-1) - 0.0327U(t-2) - 0.6911U(t-3)$$

選択した ARIMA(2,1,3)モデルを用いて予測を行った結果を図 3.11 に示す。

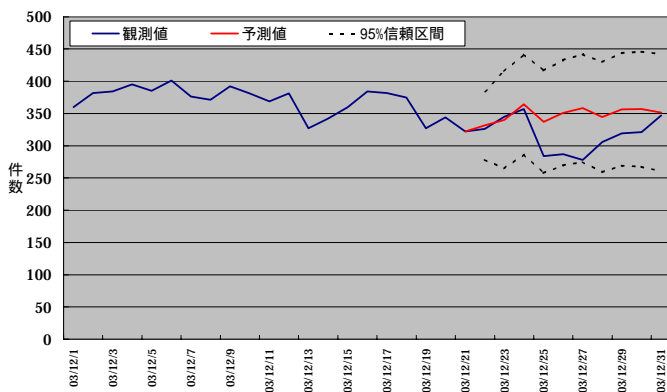


図 3.11 ARIMA(2,1,3)モデルによる予測

図 3.11 に示すように、1 期先から 3 期先においては観測値に近い値が得られているが、4 期先以降ではかなりのずれが生じている。しかし 95%信頼区間を考慮するならば、予測は概ねあてはまっていると考えられる。

また、予測期間が長くなるにつれ推定期間の標

本平均に収束していくことが分かっており⁽⁶⁾、本モデルによる長期予測では 350 件付近に収束する。

4 おわりに

今回、検討を行ったような統計情報や定点観測の 1 時間単位の状況、各種セキュリティ情報などは警察庁セキュリティポータルサイト@police (<http://www.cyberpolice.go.jp>) に適宜掲載しているので、今後のセキュリティ対策を実施するうえの資料の一つとして活用していただきたい。

参考文献

- (1) 山口 和範 「よくわかる統計解析の基本と仕組み」秀和システム(2003.4)
- (2) 「やさしい金融エンジニアリング講座 イールドカーブ」<http://www.telerate.co.jp/market/lecture/yield/yield01.html>
- (3) 「移動平均線」<http://www.soumasa.com/kikaku/semima/ma1.htm>
- (4) 「MACD」<http://www.toyokeizai.co.jp/data/chartcd/chart26.html>
- (5) 「RSI」http://www.toyokeizai.co.jp/data/chartcd/weekly_chart/10rsi.html
- (6) 山本 拓 「経済の時系列分析」創文社(1988.2)
- (7) 石村 貞夫,ステファニー・リチャルト 「Excel でやさしく学ぶ時系列」東京図書(2002.2)