2004－ＤＰＳ－117　(13)
2004－CSEC－　24　(13)
2004／3／4

ウィルスの感染アルゴリズムの違いによる伝搬状況のシミュレーション

小畑 直裕*　宮地 玲奈*　川口 信隆*　重野 寛*　岡田 謙一*

昨年，Blaster という多くのクライアント・サーバ端末を対象としたワームが流行した．そのため Blaster のような特徴を持つ新種のワームを考慮したシミュレーションが求められている．Blaster の性質上，多くのクライアント端末が感染対象となった．そしてクライアント端末には稼働時間が存在する．しかし従来のワームの感染シミュレーションではホストの稼働時間を考慮したものはなかった．本稿ではホストの稼働時間を考慮したワーム感染シミュレーションを提案する．さらに，ワームの感染アルゴリズムを変化させ，その結果も考察する．

Worm Propagation Simulation Considering Online Hosts

Naohiro OBATA *　Reina MIYAJI*
Nobutaka KAWAGUCHI*　Hiroshi SHIGENO*　Kenichi OKADA*

In this paper, we propose Internet worm propagation simulation considering “Online Hosts” and worm scanning algorithm.There is no simulation model for the propagation of a worm that considers “Online Hosts” as a parameter and we beleive that this is the first attempt.

1　はじめに

昨年 2003 年は国内のワーム被害届け出件数が 752 件と過去最多となり，Blaster や Slammer 等新種のワームが流行した. こういった状況を受けてワームの被害予測に対する要求が高まっている. 被害予測を行う方法としてワームの感染シミュレーションがあるが，従来のシミュレーションでは端末の稼働時間をパラメータとして考慮したものはなかった. 今まではサーバのサービスを対象としたワームが多く，クライアント端末で動作してるサービスを対象としたワームがほとんどなかったため，24 時間稼動のサーバのみ考慮すればよかった.

しかし，昨年の Blaster のようにクライアント端末で動作しているサービスのセキュリティホールを突いて感染するワームが出現し，多くのクライアント端末が攻撃対象となった．ここでいうクライアント端末とは，Windows XP や Windows 2000 などのデスクトップ PC 用の OS がインストールされている端末を指す．クライアント端末はサーバとは異なり，使用しているユーザが電源の on/off を行うため，端末が稼働している時間帯がある．

そこで本稿では稼働時間を考慮したワーム感染シミュレーションを提案する．本研究ではプロバイダなどの一つのネットワーク管理組織 (AS:Autonomous System) を対象とし，一般家庭の稼動時間帯，企業の稼動時間帯と従来の 24 時間稼動の時間帯の 3 つの時間帯を設定しシミュレーションを行った．また，ワームの感染アルゴリズムを変化させ，シミュレーションを行った.

本稿では，一つのネットワーク管理組織において稼働時間帯の考慮と感染アルゴリズムの違いによる感染状況の変化を考察する.

以下， 2.2 章において Blaster の感染方法について述べ，3 章でワーム感染シミュレーションの関連研究に関して述べる．4 章では本提案について述べ，5 章でシミュレーション結果を考察する．そして 6 章で結論を述べる.

* 慶應義塾大学 理工学部 情報工学科
Department of Instrumentation(Information), Faculty of Science and Technology, Keio University

2 ワーム

2.1 ワームとは

ワームとは，ネットワーク経由で端末に感染する独立したプログラムである．ワームが感染を行うときに，他の端末に侵入し，自分自身をその端末にコピーする．このような動作は通常ではできないが，その端末上で動作しているソフトウェアにセキュリティホールがある場合は可能となる場合がある．その場合は当該ソフトウェアが動作しているだけでその端末に侵入できることになる．

ワームは感染活動を行うときに感染先の IP アドレスを決める．IP アドレスを決定するときのアルゴリズムは現在感染しているホストの IP アドレスを用いて連番で IP アドレスを決める方法と，ランダムな IP アドレスを生成してそれを感染先とする方法の 2 種類がある．前者をシーケンシャルスキャン，後者をランダムスキャンと呼ぶこととする．

2.2 Blaster

Blaster は Windows OS に見つかったセキュリティホール [1] を利用して端末に侵入する．このセキュリティホールは Windows NT4.0/2000/XP/Server 2003 でデフォルトで動作しているサービスに発見されたため，多くの個人用 PC も感染対象となった．

Blaster のように感染対象となる端末数が多く，かつ侵入しただけで自動的に発病してしまうワームは今まではなかった．そのため感染が拡大する前から，各報道陣で取り上げられ大騒ぎとなった程である．

Blaster のスキャンアルゴリズムはシーケンシャルスキャン率 40%，ランダムスキャン率 60%である．

3 関連研究

ワームの感染シミュレーションの研究において数式モデルで行うものと仮想的なネットワークをモデルとして構築してシミュレーションを行うものと主に 2 種類存在する．数式モデルで行う研究として [2] がある．
この研究は

- 離散時間モデルを用いている．
- ワームが行うスキャンとしてランダムスキャン，シーケンシャルスキャンの 2 つを用いている．

という, 2 つの特色を持つ AAWP モデルという数式モデルを提案している．また, 感染したホストに「パッチを当てる」という行為も考慮している．

また仮想的なネットワーク上でワームの感染状況をシミュレートする研究としては [3] がある．ネットワークの帯域をパラメータとして考慮しているのが特徴である．複数のネットワークで構成されるインターネットモデルを用い, 各ネットワークごとに帯域幅を設定し, またワームのサイズも考慮した. その結果，帯域差のある様々なネットワークの組み合わせにおける感染状況の変化，及び感染拡大開始時に帯域の広いネットワークで開始した場合のワームの感染速度がが 5〜6 倍速くなることもシミュレーションで示した．

どちらの研究も本稿で述べる端末の稼働時間を考慮したものはなく，本研究は既存研究とは異なり Blaster などの新しいワームを考慮したシミュレーションであると言える．

4 提案

本提案では 2.2 章で述べた Blaster と同様な特徴を持つワームを想定したシミュレーションを行う．つまり以下の性質を持つワームを前提としたシミュレーションとする．

- ワームの感染及び発病が人の操作を介さずに自動的に行われる
- サーバ･クライアント端末ともに感染対象である
- ネットワークに接続しているだけで感染対象端末に感染可能である

そのため，メールを読む，特定のサービスをインストールするといった人間の行動は一切考慮しないものとする．モデルを単純化し，パラメータとして設定が難しい人間の行動はシミュレーションを複雑にするため本稿では考慮しないこととする．唯一人間の行動として考えられるもので端末の稼働時間はパラメータ化しているが，企業などある程度一定であるし，また一般家庭の稼働時間は [4] より統計データがあり，それを元に設定を行った．

本研究では主に以下の項目をパラメータとして導入する．

- 端末の OS
- 端末の稼働時間帯

- ワームの感染アルゴリズム
- 端末の合計台数
- 一つのスタブネットワークに含まれる端末の台数

本提案の稼働時間とワーム感染アルゴリズムを考慮したシミュレーションを行うことによって，1つのネットワーク管理組織においてワームの感染状況予測の一助となることと思う．

5　シミュレーション環境

本章では，シミュレーションの実装環境及び，ネットワーク構成，ワームの感染アルゴリズム，端末の稼働時間を具体的にどう実装したかを述べる．

5.1　実装環境

システムは以下の環境に実装した．開発環境も同様である．

開発言語 JDK 1.4.2

OS Windows XP

CPU Pentium4 3GHz

RAM 2G

5.2　ネットワーク構成

本シミュレーションで用いるネットワーク構成を図 1 に示す．
同心円上に配置されたルータがあり，末端のルータからエンドホストがスター型で接続されている．

図 1 で丸で囲っている末端 LAN をスタブネットワークと呼ぶ．スタブネットワーク内にあるエンドホストの数は全て同一である．また，全てのスタブネットワークの合計台数，すなわちエンドホストの合計台数を 16000 台とした．

5.3　ワームの感染アルゴリズム

本モデルではワームの感染アルゴリズムとして考慮するのは，Blaster のようにシーケンシャルスキャン率とランダムスキャン率の 2 つである．そして 2 つのスキャン率を変化させることによってワームの感染アルゴリズムを変化させる．スキャンは 2 種類のためどちらか一方のスキャン率が決まればもう一方も決まる．例えばシーケンシャルスキャン率 40%としたらランダムスキャン率は 60%となる．本研究では，シーケンシャルスキャン率を 10%から最大 80%まで変化させてシミュレーションを行った．

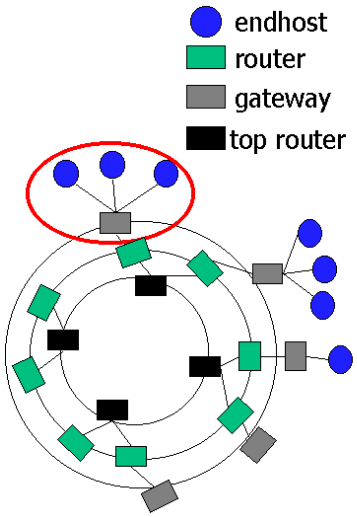

図 1: ネットワーク構成

5.4　端末の稼働時間

本研究では, 端末が全て一般家庭であることを想定した「一般家庭ネットワーク」と企業のネットワークであることを想定した「企業ネットワーク」の 2 つのネットワークに加えて従来の 24 時間稼働端末だけで構成されるネットワークの 3 つのネットワークを設定しシミュレーションを行った．

一般家庭ネットワークの稼働時間帯は端末の稼働率を以下のとおりに設定した．

0-2 時	30%	12-14 時	28%
2-4 時	9%	14-16 時	32%
4-6 時	8%	16-18 時	32%
6-8 時	17%	18-20 時	38%
8-10 時	28%	20-22 時	58%
10-12 時	36%	22-24 時	62%

同様に企業ネットワークの稼働時間帯を以下のように設定した. このネットワークでは 8 時〜18 時において一般家庭ネットワークより高い稼働率を設けた．

0-2 時	30%	12-14 時	90%
2-4 時	9%	14-16 時	90%
4-6 時	8%	16-18 時	90%
6-8 時	17%	18-20 時	38%
8-10 時	90%	20-22 時	58%
10-12 時	90%	22-24 時	62%

以上のようなシミュレーション環境でシミュレーションを行った.

6 結果と考察

本章では5章で述べたシミュレーション環境で行ったシミュレーションの結果を述べ，考察する.

6.1 スタブネットワーク台数の違いによる感染速度の比較

まず, 24時間稼働ネットワークでスタブネットワーク台数を変化させて行ったシミュレーションの比較結果を図 2 に示す. スタブネットワークの台数を 100 台，200 台, 250 台と変化させた. グラフは各ネットワークにおいて，シーケンシャルスキャン率を 10%〜80%まで変化させたときに，一定数 (8000 台) 感染するまでにかかる時間を比較したものである. 稼働時間を考慮した，企業ネットワークと一般家庭ネットワークのスタブネットワーク台数は 100 とした.

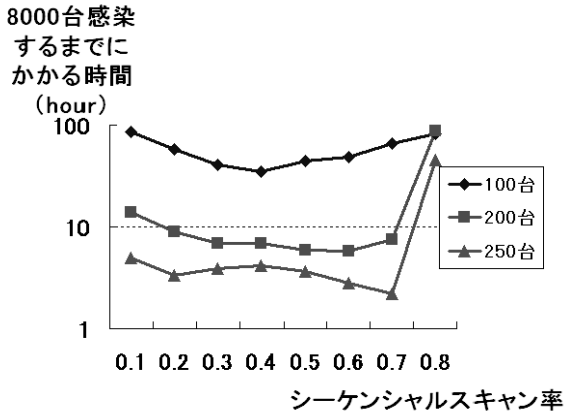

図 2: スタブネットワーク台数の違いによる感染速度の比較

図 2 より以下の事が言える.

- スタブネットワークの台数が増えるほど, 感染速度が速くなる
- スタブネットワークの台数が増えるほど, 各曲線の最小値がグラフの右側に推移する傾向がある

スタブネットワークの台数が多くなるほど, 感染速度が速くなるのは図より明らかである. また, スタブネットワークの台数が 100 台の時は 200 台の時と比べて 10 倍程度の時間がかかるということが分かる.

また, スタブネットワークの台数が多くなるほど最小値が高い値となっているのはシーケンシャルスキャンの性質を考えると説明することができる. シーケンシャルスキャンは一つのスタブネット内の IP アドレスを順次スキャンしていく. そのためスタブネットワークの台数が多いほど, 一度のシーケンシャルスキャンによって感染する台数が増えることになる. ゆえにスタブネットワークの台数が増えるほど, より高いシーケンシャルスキャン率の時が最も感染速度が速くなるということが言える.

6.2 稼働率の違いによる感染速度の比較

次に，図 3 に 24 時間稼働ネットワークにおいて, シーケンシャルスキャン率 10%〜80%まで変化させたときの感染速度の比較を示した. ここで各シーケンシャルスキャン率の違いによる，グラフにばらつき（分散）があることが分かる. すなわちシーケンシャルスキャン率の違いによって，感染数の増加の仕方に差が大きく見られるということである. また, 同様に図 4, 図 5 に企業ネットワーク，一般家庭ネットワークのグラフを示す.

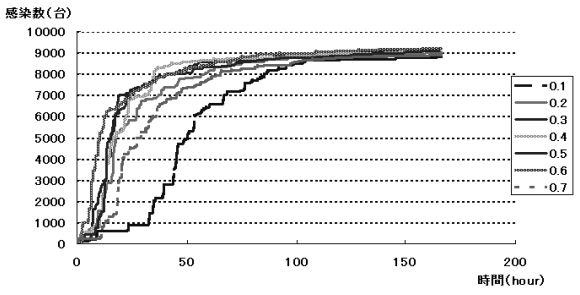

図 3: 24 時間稼働ネットワークにおける感染速度の比較

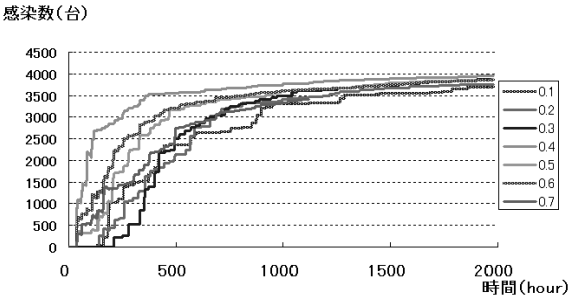

図 4: 企業ネットワークにおける感染速度の比較

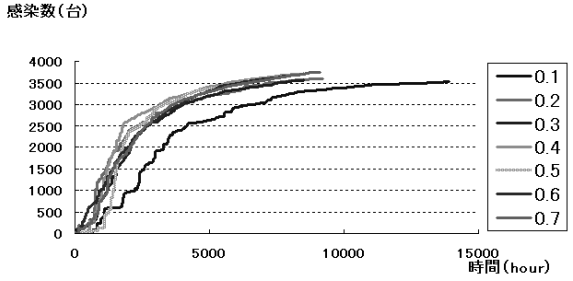

図 5: 一般家庭ネットワークにおける感染速度の比較

図 3, 図 4, 図 5 の分散に着目し，さらに企業ネットワーク，一般家庭ネットワークも分散を求めて，3 つのネットワークで分散を比較した．分散は以下のように求めた．各シーケンシャルスキャン率の感染数の平均値を M，感染数を n_i(i はシーケンシャルスキャン率) とすると，

$$\sigma^2 = \frac{\sum_{i=0.1}^{0.7} (n_i - M)^2}{7}$$

として計算した．上記の計算を各時間ごとに行って，3 つのネットワークで分散を比較したものを図 6 に示す．

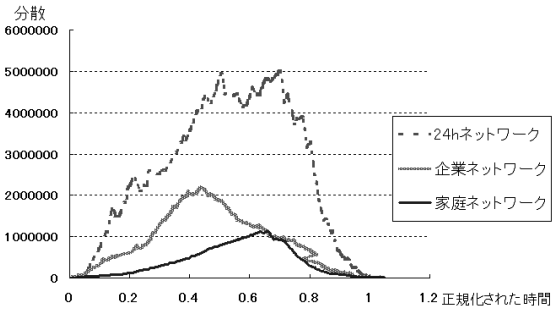

図 6: 3 つのネットワークにおける分散の比較

図 6 を見ると,24 時間稼動ネットワーク, 企業ネットワーク, 一般家庭ネットワークと稼働率の高い順に分散が大きいという傾向が見られた. 分散が大きいということはワームの種類* によって感染速度が大きく異なるということになる. 逆に分散が小さいということはワームの種類が異なってもそれほど感染速度に大きな差はないということになる. すなわち既存の 24 時間稼働シミュレーションと比較して, 実際はもっと感染速度の予想がしやすいということが言える．これらの要素はワーム感染の被害予測を行う上で有用な情報になると思われる.

7 おわりに

本提案ではワームの感染シミュレーションに新しく稼働時間をパラメータとして導入した．その結果稼働時間を考慮した場合としていない場合でシミュレーション結果に差が出ることが分かった．まず，稼働時間を考慮したことにより，感染拡大にかかる時間が 5〜10 倍になるということ．また，稼働率の高いネットワーク，低いネットワークと比較した所，稼働率が低いほど，ワームの感染アルゴリズムの変化に対して感染状況に差が出ないということが分かった．ワームの感染アルゴリズムが変化しても差が出ないということは，未知のワームが出現した時に被害予想がしやすいネットワークであるということが言える．今までは端末が 24 時間稼働を前提としたシミュレーションしか行ってこなかったが，稼働率を考慮することによって従来よりも予想がしやすくなるということが本研究で分かった.

* スキャンアルゴリズムが異なるワーム

今後の課題としては，現在ではまだデータが少ないため，様々な稼働時間帯でシミュレーションを行い稼働率と分散の関係を数学的に導き出すことである．

参考文献

[1] 「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」．
http://www.microsoft.com/japan/technet/security/bulletin/ms03-026.asp, December 2003.

[2] Zesheng Chen, Lixin Gao, Kevin Kwiat: Modeling the Spread of Active Worms, *IEEE INFOCOM* (2003).

[3] Arno Wagner, Thomas Dübendorfer, Bernhard Plattner, Roman Hiestand: Experiences with Worm Propagation Simulations, *Proceedings of the 2003 ACM workshop on Rapid Malcode* pp34-41 (October 2003).

[4] インターネット利用時間調査
http://www.netratings.co.jp/nnr/PDF/0820_2003Ranking_J_data_final.pdf.