コンピュータウィルス感染モデルのシミュレーション

2006MI075 鬼頭 英一　　2008MI127 松原 宗太　　2008MI276 山内 勇樹
指導教員　石崎 文雄

1　はじめに

近年，情報通信技術の発展やインターネットの普及にともない，多くの人々がコンピュータ等を通して気軽にネットワークにアクセスできるようになった．しかしそれと同時に，コンピュータウィルスの種類の多様化，被害の悪質化や拡大が社会問題となっている. 図１は情報処理推進機構が調査したコンピュータウィルス被害の届け出件数の年別推移グラフである．これを見ると被害件数の増加が，コンピュータやインターネットの普及にともなって上昇していることが見て取れる [1].

コンピュータウィルスとは，プログラムに寄生する極めて小さなプログラムであり，自分自身を勝手に他のプログラムファイルにコピーする事により増殖し，コンピュータウィルス自身にあらかじめ用意されていた内容により予期されない動作を起こす事を目的とした特異なプログラムのことをいう．情報処理推進機構の定める「コンピュータウィルス対策基準」においては，コンピュータウィルスの定義を『第三者のプログラムやデータベースに対して意図的に何らかの被害を及ぼすように作られたプログラムであり，次の機能を一つ以上有するもの』と規定している．この次の機能とは，(1) 自己伝染機能，(2) 潜伏機能，(3) 発病機能の三つのことである [1]．近年におけるコンピュータの利用者の増加や，常時接続回線の普及によるウィルスの拡散速度の上昇などにともない，コンピュータウィルスが深刻な社会問題として認識されるようになった [2].

本研究では，ネットワークによる繋がりをもったコンピュータ間において，ウィルスの感染がどのように広がるのかをシミュレーションにより研究する．シミュレーションにおいては，コンピュータウィルスの直接的な感染は，ネットワークで接続された隣接するコンピュータ間でのみ発生すると仮定し，時間経過とともにネットワーク全体にどのようにコンピュータウィルスが広がっていくか観察する．本研究では，コンピュータネットワークのトポロジーのモデルとして，ランダムネットワークモデルと Barabasi-Albert（バラバシ＝アルバート）モデル [3] を考える．ランダムネットワークモデルとは，ノードとノードの間のリンクが指向性もなく，規則性もなく，ランダムに張られるネットワーク生成モデルのことである．Barabasi-Albert モデルとは，1999 年に考案されたスケールフリー性を持つネットワークを構築できるネットワーク生成モデルである．

スケールフリー性とは，ノードから出ているリンク数の分布がべき乗則に従っているという性質である．スケールフリー性を持つネットワークでは，ごく一部の少数のノードが他のたくさんのノードとリンクで繋がっている一方で，大多数のノードはごくわずかなノードとしか繋がっていないようなネットワークトポロジーになる．現実世界のインターネットのようなネットワークのトポロジーもスケールフリー性を持っていることが知られている．スケールフリーなネットワークの弱点として、特定の重要なノードをピンポイントで狙った攻撃に対しては脆弱であるということが挙げられる [4]．本研究では，コンピュータウイルス感染の観点から，スケールフリー性を持つネットワークのこのような脆弱性が観察できるかどうかを，ランダムネットワークと比較することで論じる．

本研究のシミュレーションには，artisoc と呼ばれるシミュレータを使用する．artisoc は，人間同士の相互作用をコンピュータ上で誰もが簡単に再現することができ，ダイナミックに変化する社会現象を生きたまま分析することのできるマルチエージェント・シミュレータである [5]．マルチエージェントとは，各々が自律的に動作可能なエージェントと呼ばれる主体が集まって，全体として高度なシステムを実現する方法，もしくはそのようなシステムをモデル化・理解する方法のことである．そのモデル化・理解の方法のことを MAS（マルチエージェントシミュレーション）という．

2　ウィルス感染シミュレーション

本節では，artisoc によるコンピュータウィルス感染モデルのシミュレーションの概要について述べる．このシミュレ―ションでは artisoc を利用してランダムネットワークを用いた仮想的なネットワークモデルと，Barabasi-Albert モデルを参考にした実際に近いネットワークモデルの二つのモデルを作成し，その二つのモデルに関してウィルス感染の様子をシミュレーションする．

二つのネットワークモデルの設定は次のようになっている．

2.1　ランダム感染モデル

本節ではランダムネットワークを用いて作成した仮想的なネットワークモデル（以後，ランダム感染モデル）と，その根幹であるランダムネットワークについて述べる．

そもそもランダムネットワークとは，ノードとノードの間のリンクが指向性もなく，規則性もなく，ランダムに張られているネットワークのことである. ノード間のリンクがランダムで決まるので，いずれのノードともリンクしていない孤立したノードが発生する場合がある．

本研究ではランダム感染モデルと，後述する BA 感染モデルとの比較によるシミュレーションを行う．

このとき、BA 感染モデルでは孤立したノードが発生

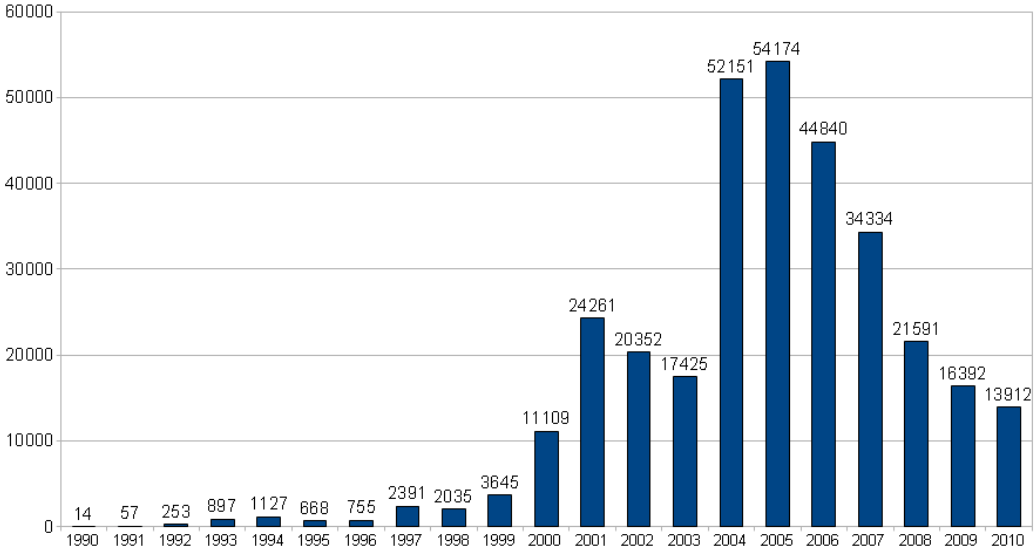

図 1　ウィルス被害届件数の年別推移 [1]

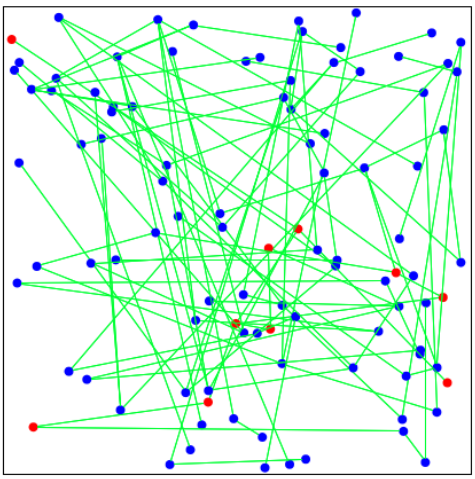

図 2　ランダム感染モデル図

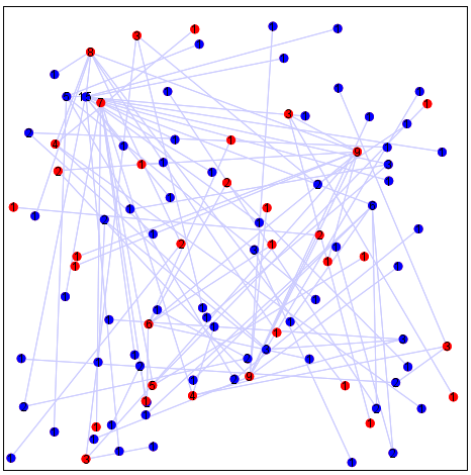

図 3　BA 感染モデル図

しないのに対し，ランダム感染モデルにおいてランダムネットワークモデルをそのまま用いると孤立したノードが発生してしまう．よって BA 感染モデルとの比較をより対等な条件で行うために “ 全てのノードから最低でも一本以上のリンクを出し，リンクの張り方のみランダムとする ” といった条件を設定した．

実装したランダム感染モデルの動きは以下のようになる．

(a)　リンクを張るノードを 1 つ選択し, ランダムにリンク先を繋げていく．
(b)　1 度選択されたノードは，2 度目以降には選択されないようにする．
(c)　以上のことを, 指定したネットワークサイズになるまで続ける．

2.2　BA 感染モデル

本節では Barabasi-Albert モデルを参考にして作成した実際に近いネットワークモデル（以後，BA 感染モデル）と，その根幹である Barabasi-Albert モデルについて述べる．そもそも Barabasi-Albert モデルとは，A.L. バラバシと R. アルバートが考案したスケールフリー性を持つネットワークモデルのことである．バラバシらはこのモデルを構築するために次の二つのことを考慮した．

1. 成長 (growth)
   ネットワークは固定のものではなく, 常に成長 (又は変化) する．
2. 優先的選択 (preferential attachment)
   ノード同士はランダムにリンクを張り合うのではなく, 影響力の強いノードほどリンクを張られや

すい傾向にある.

また, 上記の二つを考慮して次のアルゴリズムを持つモデルを考えた.

1. m_0個のノードからなるネットワークを用意する (初期条件, m_0は十分小さい自然数).
2. ノードを一つずつ追加していく (成長). そしてノードを追加していくごとに, その追加するノードからネットワークにすでに存在するノードに対して m 本のリンクを張ってゆく. ここで, すでにネットワーク上に存在するノード i にリンクが張られる確率は, そのノードの次数 k に比例する (優先的選択).
3. 以上のことを, 欲しいネットワークサイズになるまで続ける [6].

実装した BA 感染モデルの動きは以下のようになる.

(a) 初期値 m_0は 1 , 追加するノードからのリンク数は 1 .
(b) ノード数, ウィルスの感染確率, 実行ステップ数をこちらで指定可能.
(c) 指定したノード数のバラバシ・アルバートモデルを作成する.

2.3 ウィルスについて

本節ではウイルスの感染についての設定について説明する. 本研究ではコンピュータウイルスの感染拡大の様子をシミュレーションし, その様子を観察することを目的としている. そのため本研究でのコンピュータウイルスは, 特定のウイルス (ワームやトロイの木馬など) のような種類を設定するようなことはせず, 単純にネットワークで繋がったコンピュータ間を感染していくだけのものとする. また, ウイルスはエージェントとしてではなく, ノードエージェントの属性として指定した.

また, ウイルスの動作は次のようになっている.

(a) ネットワークモデルが完成したら, ランダムに一つのエージェントを選んで感染状態にする (色が変わる).
(b) 感染したエージェントのリンクを確認し, 次のステップで感染したエージェントとリンクしている未感染のエージェントに一定の確率で感染する (今回のシミュレーションでは 20% の確率とした).
(c) (b) の動作を繰り返す (今回は 20 ステップ繰り返す).

3 シミュレーションの結果と考察

本節では, 本研究において行ったシミュレーションの説明と, その結果についての考察を行う.

今回のシミュレーションは, ランダム感染モデルと BA 感染モデルといった二つのモデルを作成し, その二つのモデルに関してウイルス感染の様子をシミュレーションした. この二つのモデルにおいて

1. ノード, リンクの数をそれぞれ 100 とする.
2. 一回のシミュレーションは, ウイルスが感染してから 20 ステップで終了とする.
3. ウイルスの感染確率は 20% とする.
4. シミュレーションは一万回行う.
5. シミュレーション一回毎に, 全体の感染割合を記録していく.

といった共通の設定をした.

また, 上記の設定でのシミュレーション後 “ リンク数の多いノードにウィルス感染対策を施したシミュレーション ” も行った. そのシミュレーションにおいては

1. リンク本数が多い上位三つのノードを選択.
2. 選択されたノードのウィルス感染確率を全体の 5 分の 1(4%) にする.

の条件を二つのモデルに共通して追加した.

図 4 はウィルス感染対策を施していないモデルのシミュレーション結果で, 図 5 はウィルス感染対策を施したモデルのシミュレーション結果である. それぞれ赤色のバーがランダム感染モデル, 青色のバーが BA 感染モデルを表している. このグラフはシミュレーション一回毎の全体の感染割合を集計したもので, X 軸が感染割合を 10% 刻みで表し, Y 軸がその感染割合になったシミュレーションの回数を表している.

また表 1・表 2 はそれぞれ, ウィルス感染対策を施していないモデルとウィルス感染対策を施したモデルの結果の数値である.

まず図 4 のグラフを見ると, ランダム感染モデルでは 10% から 19% 台の感染割合になったシミュレーション結果が 4433 回と圧倒的に多い. それに対して BA 感染モデルでは 1% から 9% 台の感染割合になった結果の回数が一番多く 1471 回となっているものの, それ以外の感染割合になった回数と大きな差は見られない.

ランダム感染モデルでは確定していないが, BA 感染モデルにはリンクが集中しているノードが必ず存在する. そのためリンクの集中しているノードがウィルスに感染すると, 火急的にウィルスが広がってしまうためこのような結果が得られたと考えられる.

また図 5 のグラフをみると, ランダム感染モデルの結果は図 4 のウィルス感染対策を施していないものの結果とほとんど変化が見られない. しかし BA 感染モデルの結果をみると, 著しく変化していることが見てとれる. とくに 1% から 9% 台では 4000 近い回数の増加があることが分かる.

これはランダム感染モデルではウィルス感染対策を施したノード以外にもリンクが数本出ているノードが多く存在するため, 別の感染ルートから容易く感染してしまうからだと考えられる。しかし, BA 感染モデルではリンクの集中しているノードに感染対策を施して感染確率

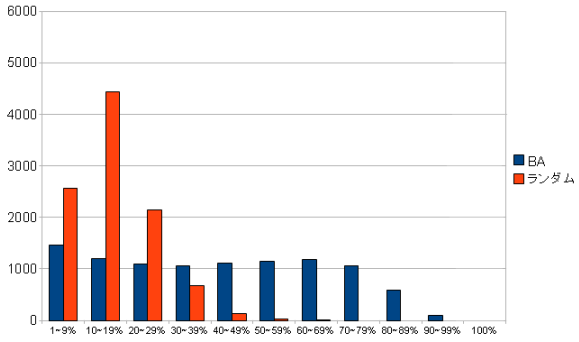

図 4　ウィルス感染対策を施さないシミュレーション結果

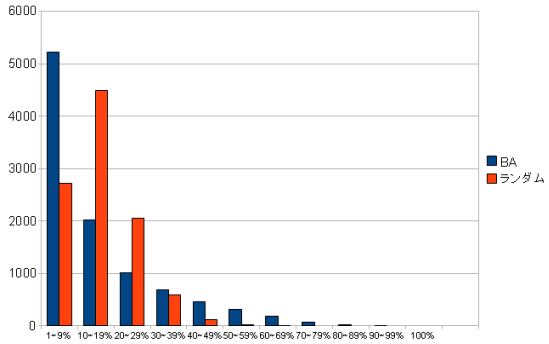

図 5　ウィルス感染対策を施したシミュレーション結果

表 1　ウィルス感染対策を施さないシミュレーション結果

	BA 感染	ランダム感染
1〜 9%	1471	2570
10〜19%	1201	4433
20〜29%	1095	2149
30〜39%	1053	672
40〜49%	1113	138
50〜59%	1143	34
60〜69%	1176	4
70〜79%	1068	0
80〜89%	589	0
90〜99%	91	0
100%	0	0

表 2　ウィルス感染対策を施したシミュレーション結果

	BA 感染	ランダム感染
1〜 9%	5217	2717
10〜19%	2020	4482
20〜29%	1020	2060
30〜39%	691	599
40〜49%	456	123
50〜59%	315	16
60〜69%	181	3
70〜79%	71	0
80〜89%	26	0
90〜99%	3	0
100%	0	0

を下げた場合，感染しやすいルートがかなり限られてくる．このようなことから，BA 感染モデルのほうが感染対策の効果が高いと考えられる．

4　まとめ

シミュレーション結果から，コンピュータウィルスがネットワーク上で拡大する様子が観察できた．ウィルス感染対策をまったく施さずにランダム感染モデルと BA 感染モデルとの二つのネットワークモデルでコンピュータウィルスの感染シミュレーションを行うと，BA 感染モデルの方がよりウィルスが拡大しやすいという結果が得られた．また図 4 と図 5 の結果の比較から，Barabasi-Albert モデルのようなネットワーク体系の方が効率よくウィルス感染の対策ができるということも分かった．

本研究ではランダム感染モデルと BA 感染モデルでコンピュータウィルスの感染拡大についての実験を行ったが，今後の課題としてはコンピュータウィルスの種類を確定してシミュレーションを行うことが挙げられる．今回の研究を基盤としてワームやトロイの木馬，ボットなど特定のウィルスに対応した感染対策のシミュレーションを組み上げれば，より効率的なコンピュータウィルスの対策案が検討できるのではないかと考えられる．

参考文献

[1] 独立行政法人情報処理推進機構：独立行政法人情報処理推進機構 HP, http://www.ipa.go.jp/(accessed 2011.7)

[2] 総務省：総務省の情報通信政策に関するポータルサイト，http://www.soumu.go.jp/main_sosiki/joho_tsusin/joho_tsusin.html(accessed 2011.8)

[3] A.-L. Barabasi and R. Albert：“Emergence of scaling in random networks”, Science 286, pp. 509-512, 1999.

[4] R. V. Sole and J. M. Montoya：“Complexity and fragility in ecological networks”, Proceedings of the Royal Society B: Biological Sciences 268, No. 1480, pp. 2039-2045, 2001.

[5] 構造計画研究所：構造計画研究所 MAS コミュニティ, http://mas.kke.co.jp/(accessed 2011.6)

[6] 複雑ネットワークについて，http://f39.aaa.livedoor.jp/~hukuryu/index.html (accessed 2011.7)