数理解析研究所講究録
第1706巻 2010年 26-34

ネットワークにおけるコンピュータウイルスの危険度

菅原啓・**芹田政晃・吉田信介・****松本悠希**

*大阪大学大学院生命機能研究科生命機能専攻，**静岡大学工学部システム工学科，
大阪大学歯学部，*大阪大学大学院情報科学研究科バイオ情報工学専攻

***Hiroshi Sugahara, **Masaaki Serita,**
*****Shinsuke Yoshida, ****Yuki Matsumoto**

**Graduate School of Frontier Biosciences, Osaka University*
***Department of Systems Engineering, Shizuoka University*
****School of Dentistry, Osaka University*
*****Graduate School of Information Science and Technology, Osaka University*

ネットワークの研究が盛んになった事もあり、コンピュータウイルスの感染をモデル化した研究が盛んになった。しかし多くの研究が1つのコンピュータの状態について感染しているか、いないかの2状態しか考えない物である。そこで本研究では1つのコンピュータ内のウイルスの数を明示的に取り入れたモデルについて解析した。最初に単一のコンピュータ内でのウイルス数のダイナミクスについて、その後BAモデルで生成されたスケールフリー・ネットワーク上のコンピュータ内のウイルス数について議論した。

1.導入

近年コンピュータやインターネットにより人々の生活が豊かになる一方で，コンピュータウイルスによる被害も拡大している。コンピュータウイルスによる被害を拡大させないためにインターネット上での感染をシミュレーションによる研究も進められている。先行研究ではコンピュータウイルスの伝搬も生物のウイルスと同様に確率的なSISモデルなどを用いた研究が盛んであり、インターネットの幾何学的な側面に注目された研究も報告されている。例えば格子モデルでは感染率がある一定の閾値でコンピュータウイルスが絶滅する相としない相で相転移が起こるが、実際のインターネットと同様のスケールフリー性を持たせると閾値が無くなり、有限の感染率ではインターネットウイルスが絶滅しない事が報告されている[1]。しかし実際にコンピュータが感染する場合は潜伏期間が存在したり1つのコンピュータに1種類のコンピュータウイルスが複数個のファイルに感染したりする事が多い[2]。本研究ではこの事実に着目し，モデル化する事で従来のSISモデルなどのモデルとの違いやワクチンソフト

を導入した時の振る舞いについて調べる。

2.モデル

まずコンピュータ内でどのようにウイルスが増殖するかを考える。単位時間当たりにウイルスがある一定数λ自分のコピーを生み出すとすると、コンピュータ内のウイルスの数Nは

$$\frac{dN}{dt} = \lambda N$$

という微分方程式で記述する事が可能となり、その解は初期値をN_0とすると$N(t) = N_0 \exp(\lambda t)$と書く事が出来る。

コンピュータウイルスの感染経路はメールやファイル共有ソフトを通じた感染が一般的であると言われるが、中にはインターネットに接続しただけで侵入するようなウイルスも存在する。そこで単位時間当たりC個のウイルスが新有する事を考えると微分方程式は

$$\frac{dN}{dt} = \lambda N + C$$

と書ける。その解は

$$N(t) = N_0 \exp(\lambda t) + \frac{C}{\lambda}(\exp(\lambda t) - 1)$$

となる。

このままだとコンピュータ内のウイルスの数は発散してしまうので常駐のワクチンソフトにより単位時間当たりμのウイルスが駆除されると考えるとウイルスの数は

$$\frac{dN}{dt} = \lambda N + C - \mu$$

という微分方程式で表す事が出来て、その解は

$$N(t) = N_0 \exp(\lambda t) + \frac{C-\mu}{\lambda}(\exp(\lambda t) - 1)$$

と解く事が出来る。これよりウイルスが爆発的に増殖するか絶滅するかは$N_0 + \frac{C-\mu}{\lambda}$の符号によって決まる。

ここで常駐のワクチンソフトのウイルス除去は単位時間当たりに一定数駆除するのではなく、コンピュータ内のファイルの中からランダムにファイルを選んで、そのファイルがウイルスに感染していたら駆除するという場合を考える。単位時間当たり

にυ個の割合で除去が可能すると

$$\frac{dN}{dt} = \lambda N - \upsilon N + C$$

という微分方程式で書き表せる。これは見かけのウイルスの増加率が$\lambda - \upsilon$と表す事が出来る。

これまでのワクチンソフトに関しては常駐、即ち常にワクチンソフトが働き続けてコンピュータ内をスキャンしている状態を考えたが、実際には月に1回や週に1回くらいのペースでスキャンするのが一般的である。そこで遅れ効果を取り入れたインパルシブな形によりウイルス数を計算する事を考える。遅れ効果を取り入れた場合の微分方程式は

$$\frac{dN}{dt} = \lambda\left(1 - \frac{N(t-T)}{K}\right)N(t)$$

と表せる。これは一般的に解析的に解く事は出来ない。そこで1次のオイラー法により解いた。パラメータは$\lambda = 1.1, K = 30$とした。微小な時間間隔$h = 0.01$とした。このときTの値を2.0と2.5にした場合を図1に示す。

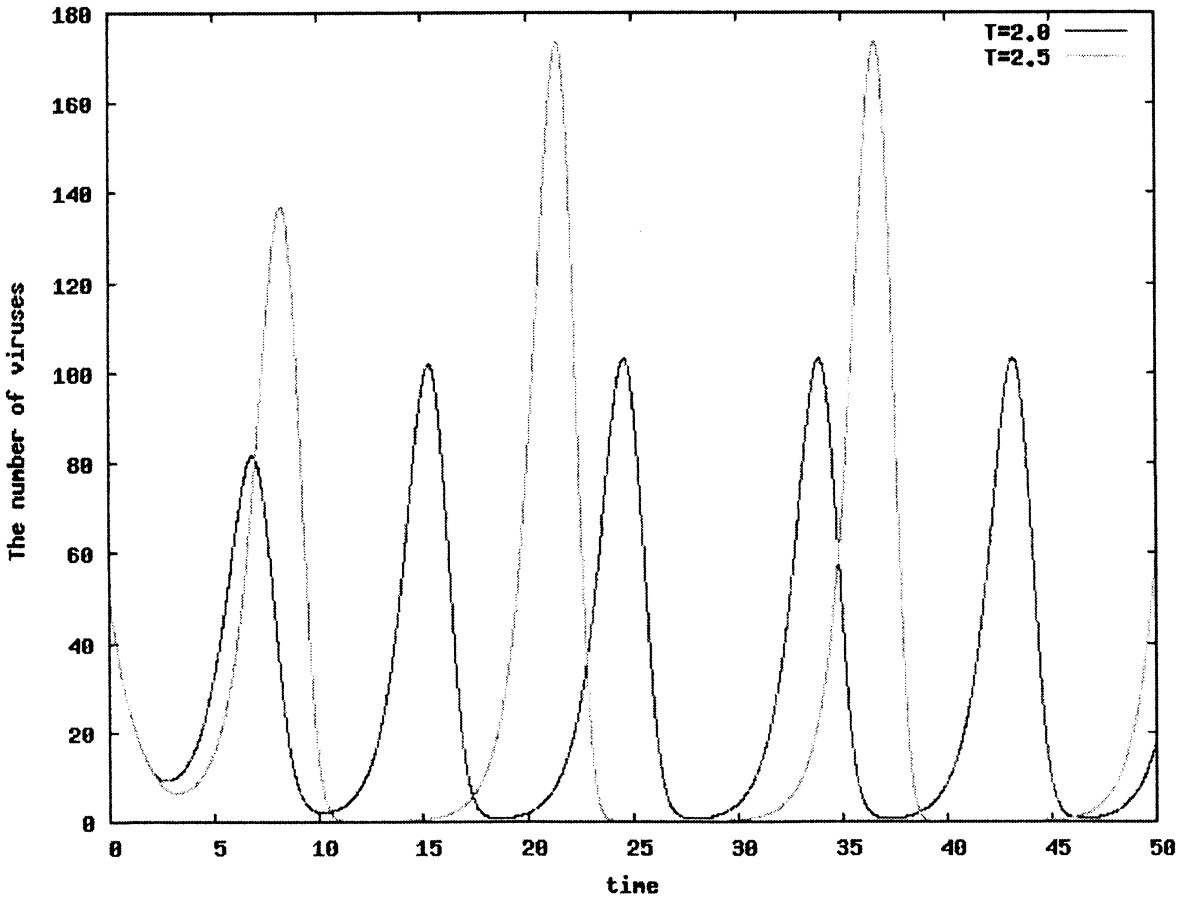

図1：遅れ効果を取り入れたウイルス数のプロット

この図を見ると確かに定期的にワクチンソフトでスキャンした様子を定性的に表しているように見える。しかしスキャンする周期がより長い$T=2.5$の方がウイルス数の最小値が小さいというのは我々の直感に反する。そこでこの遅れ効果のある関数を検証する為に$T=0$の場合について考えてみよう。微分方程式は

$$\frac{dN}{dt}=\lambda\left(1-\frac{N(t)}{K}\right)N(t)$$

と書ける。これはロジスティック成長と呼ばれる物で解析的に解く事が可能となり、その解は

$$N(t)=\frac{KN(0)\exp(\lambda t)}{K+N(0)\left(\exp(\lambda t)-1\right)}$$

となる。ここで$N(0)=80,10$の関数を遅れ効果があるものと共にプロットした物を図2に示す。

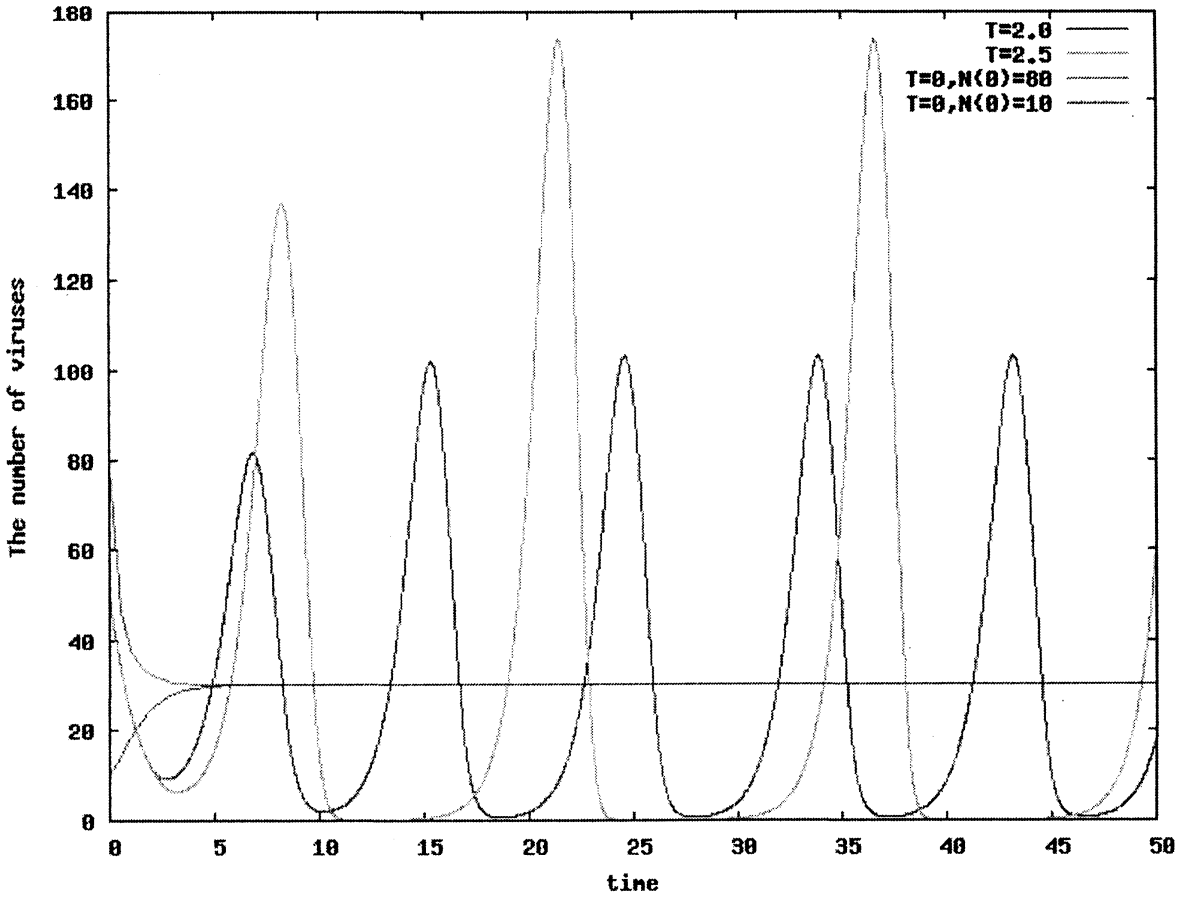

図2：ロジスティック成長と遅れ効果の比較

$T=0$というのは常にセキュリティソフトがウイルススキャンをしているという事である。しかしこの場合コンピュータ内のウイルス数は$K=30$で安定している。これも遅れ効果と比較すると$T=0$の方がウイルス数が多いときがあるので、遅れ効果を取り入れた関数はセキュリティソフトを導入したコンピュータ内のウイルス数を表す関数としては不適当だと考えられる。

ここで連続な式を用いた定期的なウイルススキャンがある場合のウイルス数の表現ではなく、不連続な式を用いた表現を考える。実際にウイルススキャンをかけるときはファイル数にもよるがコンピュータ全体だと数時間程度かかる。これを簡単のためウイルススキャンに必要な時間はウイルススキャンを書ける周期Tに対して無視出来るほどに短いと仮定しよう。最も単純に考えると初期ウイルス数が$N(0)$の場合ではウイルス数が指数関数的に増殖し時間Tが経過するとウイルス数は

$$N(T)=N(0)\exp(\lambda T)$$

となる。このときにウイルススキャンにより割合pでウイルスを駆除出来ると仮定すると、$1\geq p\exp(\lambda T)$が成り立つならばウイルス数は発散しない。さらに単位時間当たりにCのウイルスが侵入する場合は時間T後のウイルス数が

$$N(T)=N_0\exp(\lambda T)+\frac{C}{\lambda}(\exp(\lambda T)-1)$$

と表せるので、ウイルス数が発散しない為には

$$\exp(\lambda T)\geq\frac{1+\dfrac{C}{\lambda N}}{p+\dfrac{C}{\lambda N}}$$

という関係が成り立っていれば良い。値によってウイルス数が発散する場合と発散しない場合を図3に示す。

これまでは単一のコンピュータ内のウイルス数について考えてきた。しかし我々が求めたいのはネットワークにおけるコンピュータウイルスの危険度なので、これをネットワークまで広げる必要がある。そこで本研究ではバラバシ＝アルバートモデル（BAモデル）[3]を用いた。現実世界のネットワークにはスケールフリー性やスモールワールド性、クラスター性の性質を持つ事が知られているが、BAモデルはスケールフリー性を持っている事が知られている。これは以下のアルゴリズムで生成する[4]。

1. m個のノードからなる完全グラフK_mをスタートとする。
2. 新しいノードを1個追加する。そのノードから既に存在しているm個のノードに対してエッジを張る。このとき、エッジが張られる確率は、それぞれのノードの

その時点での次数kに比例するものとする。

3. 手順2をノードが所定の数になるまで繰り返す。

この手順を踏む事でグラフはスケールフリー性、即ち次数がkとなる次数分布$p(k) \propto k^{-3}$となり$\gamma = 3$となる。BAモデルでネットワークを構築し、微分方程式で自動的に侵入してくるウイルスの数をCとしていたものをネットワークを介して繋がっ

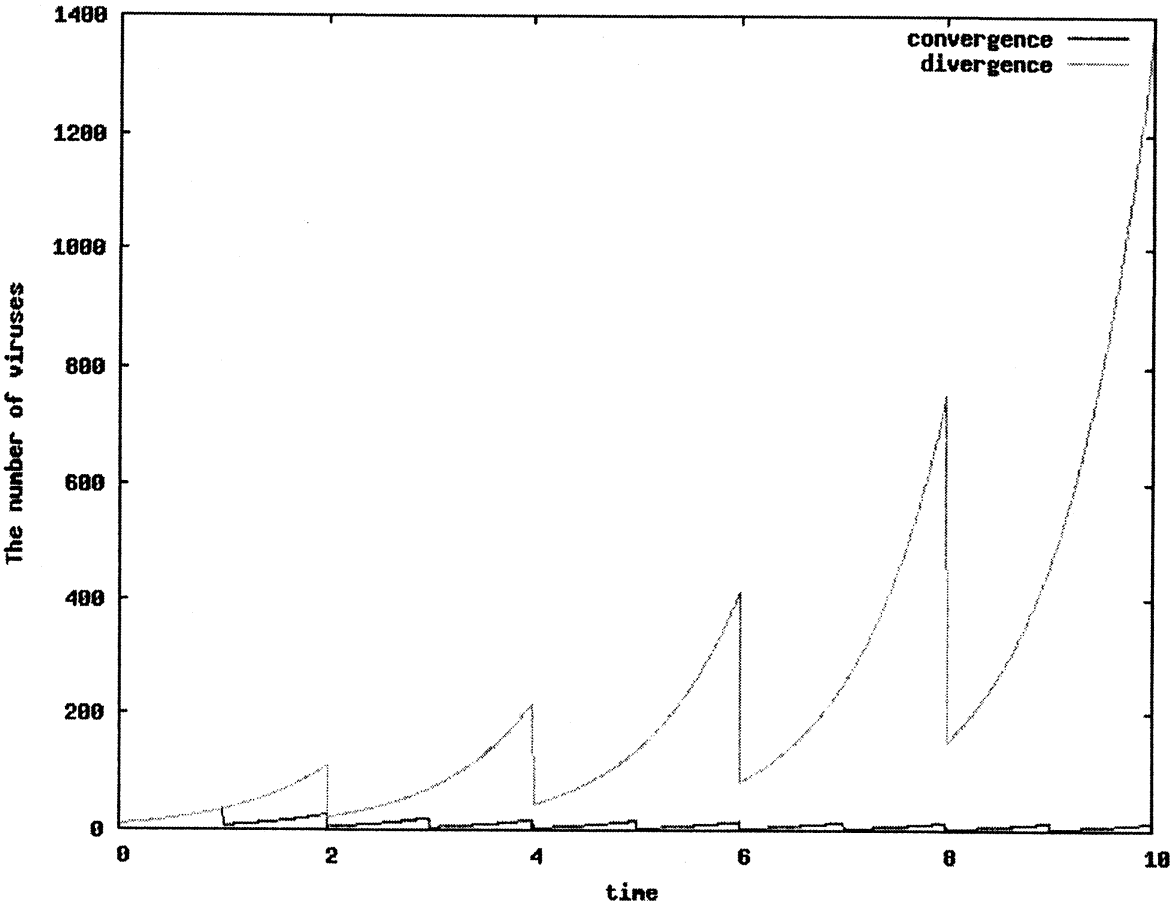

図3：ウイルス数が収束する場合と発散する場合

ているコンピュータのウイルス数のうち割合pのものに置き換える。このときi番目のコンピュータ内のウイルス数N_iのダイナミクスを表す微分方程式は

$$\frac{dN_i}{dt} = \lambda N + p\sum N_j$$

となる。ここで$\sum$はi番目のコンピュータと繋がっているコンピュータjについて和をとることを表している。

3.結果

BAモデルでのダイナミクスを示した物を図4に示す。赤色の線がBAモデルでの

ウイルスの数をプロットしたもの、緑色の線がネットワークを考えずに指数関数的に増加するものを示したグラフである。ここで縦軸のウイルス数を対数で表示した場合にグラフが直線になっているので，ネットワークを考慮した場合でもウイルス数が指数関数的に増加する事が分かるが、ネットワークを考慮しない場合と比べると傾きが圧倒的に高い。

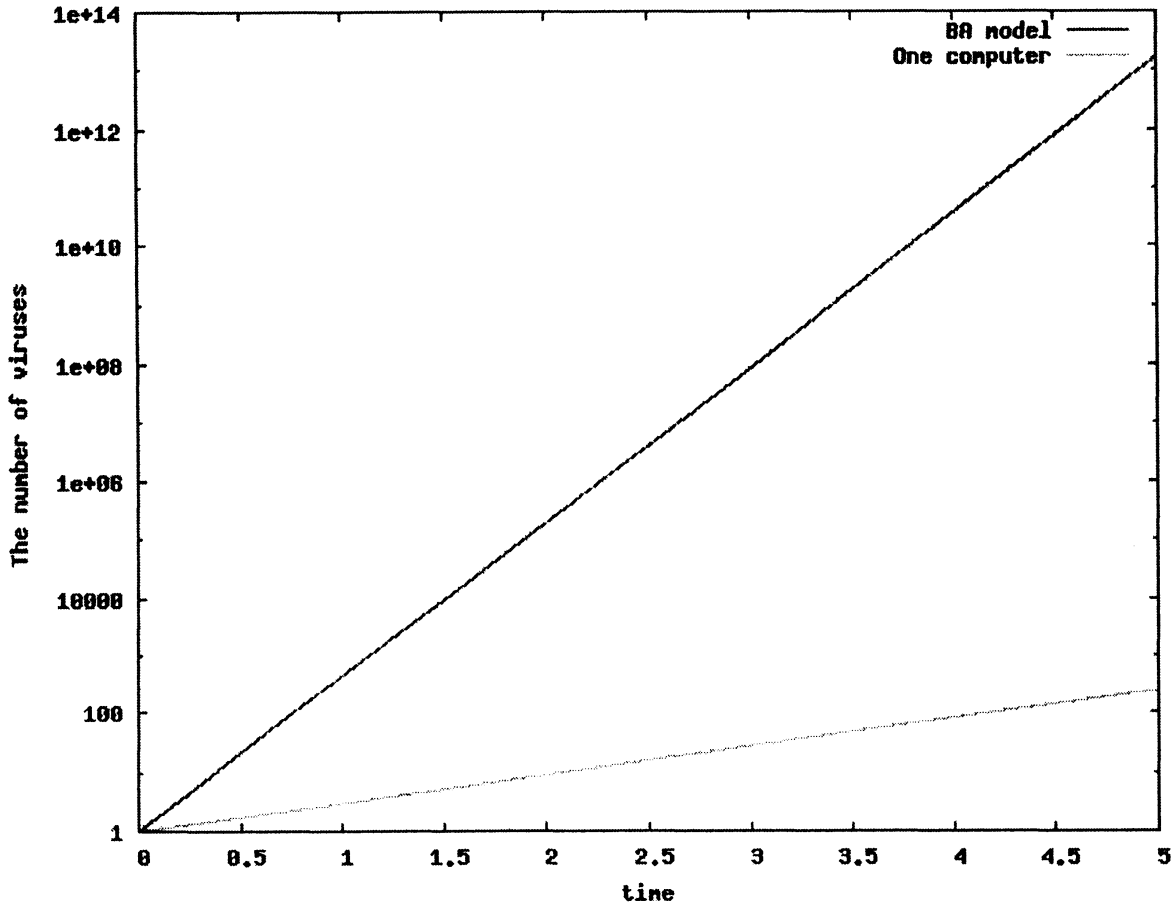

図 4：BA モデルでネットワークを考慮した場合のプロット

次にこれにウイルススキャンをかける場合を図 5 に示す。スキャンをかけるタイミングは一斉にかけるようになっている。この図での$T = 0.1$は図 3 の収束している場合と同じ値を、$T = 0.2$は発散している場合と同じ値を用いている。ウイルススキャンを考えない場合は単独の場合よりもネットワークを考慮した場合の方がより増殖速度が大きかったのに対して、ウイルススキャンがある場合は単独の場合よりもネットワークを考慮した場合の方がより増殖速度が小さくなるのは興味深い。これは単独の場合は常に一定量のウイルスCが侵入してきたのに対して、ネットワークでは他のコンピュータも一斉にスキャンをすることによって侵入するウイルス数が急激に減少する為だと考えられる。しかしこの場合でもスキャンの周期Tが大きすぎるとウイルス

の増殖を押さえきれなくなり、ウイルス数が発散する。

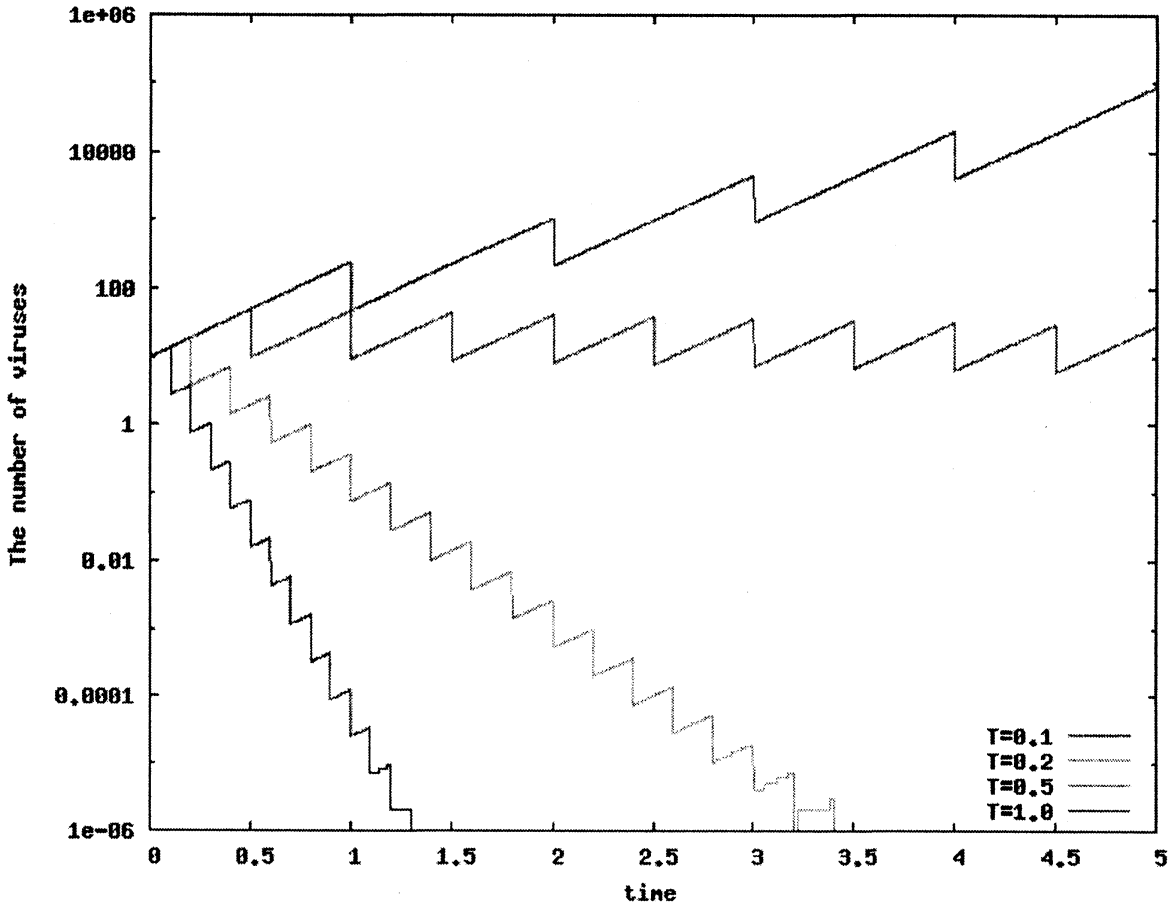

図5：BAモデルのネットワーク上でウイルススキャンをかけた場合

4.議論

本研究ではウイルスの数を定量的に表す事でより現実的なモデルの構築を目的としたものであり、ウイルスの増殖速度は一貫して指数関数的な増殖を考えた。一方ウイルス対策としてのセキュリティソフトについては常駐型、あるいは周期的なスキャンを考えた。

これまで議論してきたウイルス対策は、基本的にはファイル感染型、ブートセクター型マクロ型、添付ファイル型などのクライアントを狙ったウイルスやワームの脅威から身を守るものであるが、サーバーを狙うリモートからの不正アクセスの検知には適さない(例えば[5])。何故なら、ウイルス対策ソフトによる常駐検査により、ワームが本体となるファイルをターゲットのサーバーに送信した時点で検出する事は可能だが、Webサーバーのように外部から頻繁にアクセスのあるホストで行う事は、レス

ポンスの遅延に繋がるため現実的ではない。従ってサーバーを狙うワームから身を守る為には、サーバーのセキュリティーを向上し、さらにそのセキュリティーレベルを維持することが重要である。その維持のためにはパッチの適用やバージョンアップといった情報収集をいかに迅速に行うかが重要である。このモデル化としては、ウイルススキャンと同様の考え方が適用できるので、ウイルスの種類（並びにそれに対する対策方法）の変数を増やす事により、より現実に即したモデルへ改良する事が考えられる。

また、サーバーのセキュリティー向上と言った場合、外部に公開されているサーバーのみでなく、内部ネットワークに存在するサーバーの存在も重要である。これはクライアントを狙ったウイルスの場合にも同様にいえる事であり、ネットワークモデルの導入により検討すべき部分である。

コンピュータウイルスの危険度からその適切な対策を求めるというのが本研究の動機の一つである。実際には様々な角度からのウイルス対策が必要である。クライアントを狙ったものに対しては、本稿で重点的にモデル化したウイルス対策ソフトの導入に加え、ユーザー教育を含むクライアント対策が重要である。サーバーを狙ったものに対しては、侵入検知システムの導入等も含むセキュリティーレベルの維持が必要である。どの部分の対策にコストをさくべきかは各種ウイルスの量比などによって異なると思われる。本モデルはその検討に入る前の準備段階として、クライアントをねらったウイルスに対する対策の定量的評価を示したものであるといえるだろう。

5.参考文献

[1]R. Pastor-Satorras, A. Vespignani, *Phys Rev Lett* **86**, 14, 2001
[2]渡部章, 『コンピュータウイルス辞典』, オーム社, 1993
[3]Barabási, A.L., and Albert, R., *Science* **286**, 1999
[4] http://ja.wikipedia.org/wiki/複雑ネットワーク
[5]三輪信雄,『ウイルスの原理と対策』, ソフトバンクパブリッシング株式会社, 2002