平成 24 年度卒業論文

論文題目

Android 端末を対象とした
ウイルス拡散モデルの評価

神奈川大学　工学部 電子情報フロンティア学科
学籍番号　200902880
古屋　智規

指導担当者　木下宏揚　教授

目 次

図 目 次

第1章　序論

近年，スマートフォンが急速に普及してきている．スマートフォン利用者は2011年と2012年で比較すると利用率が1年で倍増して，日本でも4人に1人がスマートフォンを所有しているという状況である[1]．またスマートフォン利用者の4割が2012年からスマートフォンの利用を開始している．さらに，非利用者のうち6割強が利用を検討していて，スマートフォンの需要は衰ろえが見えていない．OSのシェアを比較するとiOSの割合は3割程度でAndroidの割合が6割強である．前年比を見るとAndroidユーザーは増加傾向にあるという結果が出ている[2]．今後もAndroidは世界中でシェアを伸ばしていくと考えられる．Androidが圧倒的なシェアを持つ理由として挙げられるのが，自由さである．iOSではApple Storeの厳重な審査が必要で，Apple Store経由でなければアプリの導入が不可能だ．一方，Androidはその点が大きく異なる．Android Marketにはアプリに対して審査を行っておらず，アプリを開発する側も自由な発想で開発を行うことができる．自由な場所で競争して本当にいいものが評価され，利用される環境にあるのだ．また，Androidがスマートフォンの市場を拡大させたのも理由の1つだ．iOSがApple社からしか発売されていないのに対し，Androidはオープンであるので様々な開発会社から多種多様な端末が登場する[3]．

Androidユーザーが増えていく中で，Android端末を狙ったウイルスも急増している．その数は，Androidが発売されてからの数年で10倍以上とされている．Android端末を狙ったウイルスが急増している原因は，Androidが標的にしやすい環境におかれているからである．AndroidはOSが統一されていることで全世界で利用者が急増している．それゆえ，1つのウイルスを作成して撒くことで多くの端

末に感染させることができるのだ．また歴史が浅く，知識に乏しい人が多いことも狙われやすい 1 つの要因だ．Android が販売開始されてからまだ数年しか経過していない．それゆえ，未完成の部分が多くセキュリティーも十分でないうえに，ウイルスや対策法などの情報も乏しいのが現状だ．また携帯電話にウイルスが存在しなかったため，携帯電話から Android に乗り換えた利用者のウイルスに対する意識も低いのである [4].

Android のウイルスの感染経路とされているのはアプリのインストールである．Android Market における自由さは，Android の最大の魅力である．しかし，同時に自由さが脆弱性の原因となってしまっている．また，Android Market 以外からもインストールできることが感染の一因となってしまっている [5]．アプリにウイルスが混入されていても，ゲームや実用系のアプリとして公開されているためウイルスに気付く事が困難である．しかし，偽の口コミといった情報操作により，感染させることは容易である．

これまでに PC におけるウイルス拡散モデルが研究されてきた．同じ端末間で電子メールによるウイルス感染を様々な形でモデル化し，実データと比較する研究が一般的であった．同じ端末間での感染であれば，モデルに当て嵌めることは可能である．Android 端末のみの感染については従来手法のウイルスの拡散モデルから，ウイルスの特性を定量的に評価することは可能である．

Android が普及した近年，Android から PC への異なる端末間でのウイルス感染も確認されている．データのやり取りでの感染はもちろん報告されている．そして単に充電のために会社の PC に Android を接続した際に感染した例も報告されている [6]．この事例は企業からの報告であるが，学校，一般家庭など様々な日常の場面で起こりうることである．スマートフォンは電池が切れやすく，PC で充電する機会はとても多い．それゆえ，日常の中でウイルスを無意識に拡散してしまう危険性は高いのだ．このように，Android から PC に異なる端末へと感染する “2 次感染” が起こるようになった．感染経路が全く異なるためウイルスに感染，拡散する率は Android と PC で

は異なると考えられる．従来研究されてきた拡散モデルでは複数の異なる感染率を考えることは，想定されていない [8] [9]．そのため，従来の拡散モデルで同時に 2 つの端末の感染を考えることは難しい．Android から PC へのウイルス感染を考えるのであれば，なおさらである．

本論文では，Android から PC への “2 次感染” に注目し，異なる 2 つの端末間でのウイルス感染を想定する．従来のウイルス拡散モデルを基にし，PC 間感染，Android 間感染，PC-Android 間感染を含めたウイルス拡散過程のモデル化を行う．具体的には想定するモデルの状態遷移図をモデリングし，定常状態を仮定して各状態における状態確率を求める．求めた状態確率から状態確率の期待値を導く．最終的には，提案したモデルにおけるウイルスの特性を過渡状態において定量的に評価することを目標とする．求めた値から PC 間感染，Android 間感染，PC-Android 間感染，それぞれ拡散の様子にどの程度の差が出るか，比較検討を行う．

第2章　基礎知識

2.1　確率分布

2.1.1　確率分布概略

数学的な変数 X の各値に，その値の確率が組み合わさっている場合，X を確率変数といい，確率の集まりを確率分布という．

期待値と分散

確率変数の出る値はその名の通りバラバラでランダムである．そのため，1つの代表的な値に集約させる必要が生じる場合がある．1つの値へ集約は，確率の平均をとる計算を行えばよい．そのようにして算出した値を期待値と呼ぶ．

離散型確率分布の場合，確率変数の値 x その確率 $f(x)$ の和を $E(X)$ とし，X の確率論的期待値といい，以下の式で表す．

$$E(X) = \sum_{x} xf(x)$$

また確率論の期待値は，ランダムに出る各値を1つの平均的値にまとめた指標尺度である．しかし，期待値だけではランダムさの程度そのものの指標尺度も考えておかなければ，不十分である．そのために用いられるのが，確率変数の分散である．分散の値が大きいことは，ばらつきが大きいことを意味する．分散 $V(X)$ は期待値 $E(X)$ からのばらつきであり，以下の式で表す．

$$V(X) = E(X^2) - E(X)^2$$

2.2 離散時間型確率分布

確率分布の中でも，時間域が $T = 1, 2, 3, \cdots$ と飛び飛びであるときに，離散型確率分布という．一般に確率変数 X が値 x をとる確率 $P(X = x)$ は x の関数であり，$f(x)$ で表記される．

$f(x)$ が離散型確率分布であるとき，以下の 2 条件を満たす．

$$(1)\quad f(x) \geqq 0$$
$$(2)\quad \sum_{x} f(x) = 1 \qquad (\text{全確率} = 1)$$

2.2.1 幾何分布

一般に，確率 p の試行である事象が起こるまで繰り返し試行し，$n-1$ 回待って n 回目に初めてその事象が起こる確率は試行は独立であるとすれば

$$(1-p)^{n-1}p$$

で表される．始めて起こる回数 n を確率変数 X とした離散確率分布のことを，幾何分布という．

幾何分布の期待値は，

$$E(X) = \frac{1}{p}$$

と，確率 p の逆数となる．

2.2.2 二項分布

1 回の試行である事象が確率 p で起こり，その試行を n 回繰り返すことで現れる二項分布の確率は，$(a+b)^n$ の二項展開

$$(a+b)^n = \sum_{k=0}^{n} {}_n\mathrm{C}_k a^{n-k} b^k$$

で，$a = p, b = 1 - p(= q)$ とおいたときの各項によって得られる．このとき，n 回の試行で起こった確率 p で起こる事象の回数を X とすると，X は確率変数で

$$P(X = x) = {}_n\mathrm{C}_x p^x (1-p)^{n-x} \qquad (x = 0, 1, \cdots, n)$$

で表される．二項分布はパラメータ (n, p) によって表されるので $Bi(n, p)$ で表す．

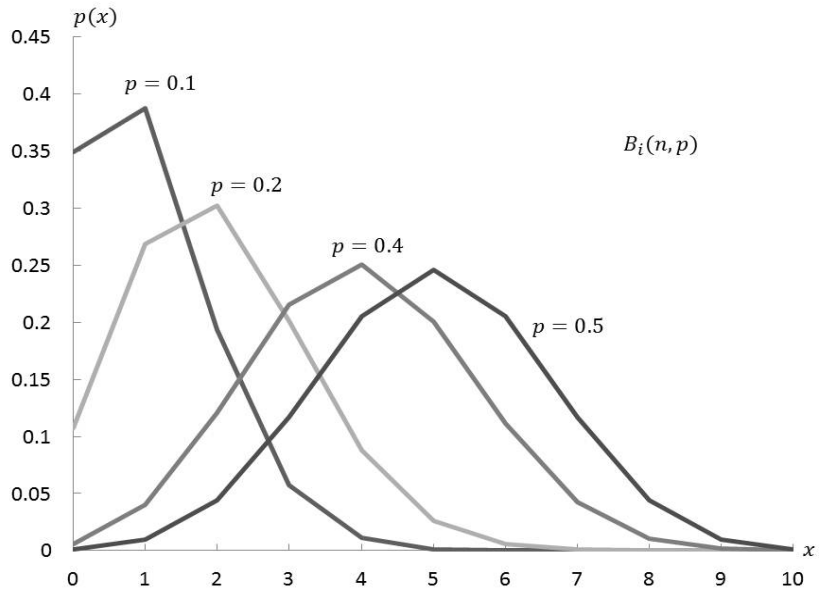

図 2.1: 二項分布

期待値と分散

二項分布における期待値は，1 回あたりの成功率 p で n 回試行するので，

$$E(X) = np$$

であることがわかる．また，分散については

$$V(X) = E(X^2) - E(X)^2$$

を用いると，X が 0 か 1 しかとらないので，$E({X_k}^2) = E(X_k)^2 = p$ であり，

$$V(X_k) = E({X_k}^2) - E(X_k)^2 = p - p^2 = p(1-p)$$

と，任意の1つの分散が求められる．これと $X = X_1 + X_2 + \cdots + X_n$ であることから，分散 $V(X)$ は，

$$V(X) = np(1-p)$$

となる．

2.2.3 ポアソン分布

二項分布 $Bi(n,p)$ において n が極めて大きく，p が極めて小さい値の時，確率は非常に計算しにくくなる．その場合，二項分布の確率をより計算しやすく近似したものをポアソン分布という．

具体的には，$n \to \infty$，$p \to 0$ で $np =$ 一定であれば，$np = \lambda$ とし，二項分布 $Bi(n,p)$ を

$$P(X = x) = {}_n\mathrm{C}_x p^x (1-p)^{n-x} \to e^{-\lambda} \cdot \frac{\lambda^x}{x!} \qquad (x = 0, 1, 2, \cdots)$$

と近似することができる．これをパラメータ λ のポアソン分布といい，$Po(\lambda)$ で表す．

時間 t を導入する場合，$[0,t]$ で起こる回数を $X(t)$ とし，$X(t)$ の確率分布は $Po(\lambda t)$ となり

$$P[X(t) = x] = e^{-\lambda t} \cdot \frac{(\lambda t)^x}{x!} \qquad (x = 0, 1, 2, \cdots)$$

となる．この $X(t)$ をポアソン過程という．

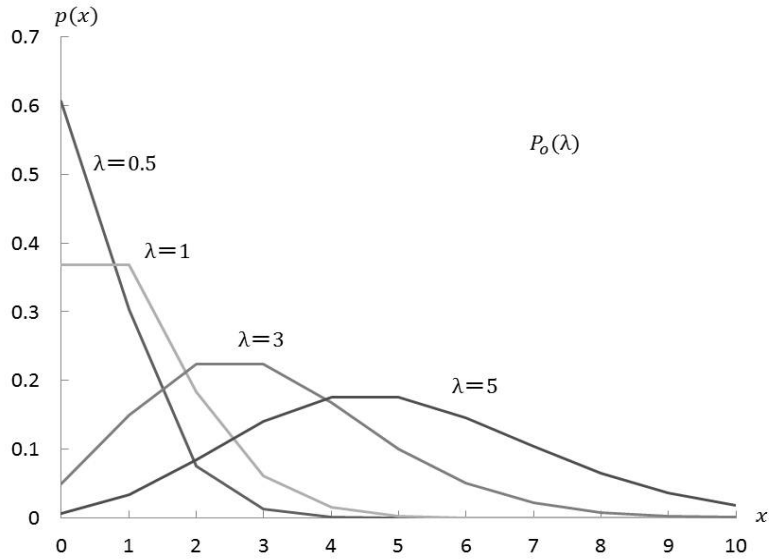

図 2.2: ポアソン分布

期待値と分散

ポアソン分布の期待値と分散は，二項定理の期待値と分散から求められる．二項分布の期待値と分散は $E(X) = np$，$V(X) = np(1-p)$ であった．これらに $p \to 0, np \to \lambda$ であることを適応すると

$$E(X) = \lambda, \qquad V(X) = \lambda$$

が得られる．すなわち，ポアソン分布は特徴として期待値，分散はポアソン分布のパラメータ λ と等しい．

2.3 連続型確率分布

確率変数 X が連続的に値をとるときは，関数 $f(x)$ を範囲で積分して，その範囲の確率を求める．このとき，$f(x)$ を確率密度関数という．密度関数で表される確率分布を連続型確率分布という．

一般に連続型分布では，範囲 A の確率は

$$P(A) = \int_A f(x)dx$$

で表される.

連続型確率分布の密度関数の条件は，以下の 2 つである.

$$(1)\quad f(x) \geqq 0$$
$$(2)\quad \int_{-\infty}^{\infty} f(x)dx = 1 \qquad (\text{全範囲の確率} = 1)$$

2.3.1 指数分布

指数分布は指数関数 e^x を用いて定義される連続型分布であり，パラメータ λ を含む密度関数

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geqq 0) \\ 0 & (x < 0) \end{cases}$$

により定められる．指数分布は，ある条件のもとで決められた事象がおこるまでの待ち時間の分布として知られる.

確率過程の理論では，ポアソン過程で $X(t)$ でカウントされる事象の起きる時間が指数分布に従う.

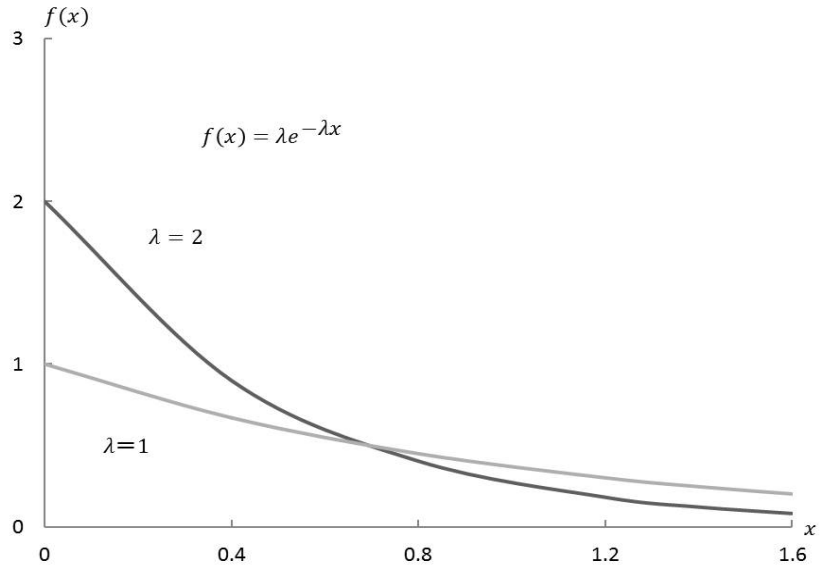

図 2.3: 指数分布

期待値と分散

期待値は,

$$E(X) = \int_0^\infty x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

となる．これより λ が小さいほど期待値が大きくなる事がわかる．

さらに，分散は

$$E(X^2) = \int_0^\infty x^2 \cdot \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

であるから

$$V(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

となる [7].

2.4 マルコフ過程

2.4.1 マルコフ性

マルコフ過程は，確率過程の1つである．マルコフ過程は定義として，“過去の状態に依るが，過去の過去には依らない” という性質がある．

これがマルコフ過程におけるマルコフ性である．

2.4.2 マルコフ連鎖

状態空間 S 上の離散時間確率過程 $\{X_n; n = 0, 1, 2, \cdots\}$ が，任意の時刻列 $0 \leq n_1 < \cdots < n_k < n < m$ と状態 $i_1, \cdots, i_k, i, j \in S$ に対して，

$$P(X_m = j \mid X_{n1} = i, \cdots, X_{nk} = i_k, X_n = i) = P(X_m = j \mid X_n = i)$$

を満たすとき，$\{X_n\}$ は S 上のマルコフ連鎖と呼ばれる．

式では過去の経緯をもとに時刻 m で状態 j をとる確率（左辺）は，直近の過去の状態にだけ依存して決定するという意味を持つ．すなわち，これはマルコフ性を表している．

2.4.3 推移確率

マルコフ連鎖の解析のためには，条件付き確率を詳しく調べる必要がある．その基本となるものは，

$$P(X_{n+1} = j \mid X_n = i)$$

で表現される．これを時刻 n における i から j への推移確率という．一般に推移確率は n に依存して決まるが，それが n に依らず一定である時，そのマルコフ連鎖は定常的であるという．

状態空間 S 上の定常的なマルコフ連鎖 $\{X_n\}$ に対して，推移確率

$$p_{ij} = p(i, j) = P(X_{n+1} = j \mid X_n = i), i, j \in S$$

を成分とする行列

$$P = [p_{ij}]$$

をマルコフ連鎖の推移確率行列という.

2.4.4 状態遷移図

2 状態 $\{0, 1\}$ 上のマルコフ連鎖は，推移確率

$$p(0,1) = p, \qquad p(0,0) = 1-p, \qquad p(1,0) = q, \qquad p(1,1) = 1-q$$

で決まる．ここで，p, q は $0 \leq p \leq 1, 0 \leq q \leq 1$ を満たす定数である．推移確率行列は,

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$$

で与えられる．これをマルコフ情報源と呼ぶこともある.

マルコフ連鎖はグラフ化すると便利である．2 状態 i, j が $p(i,j) > 0$ を満たすときに i から j に有向辺を引いてできるグラフを状態遷移図という．図 2.4 のような状態遷移図にはマルコフ連鎖の本質的な情報はすべて反映されている [10].

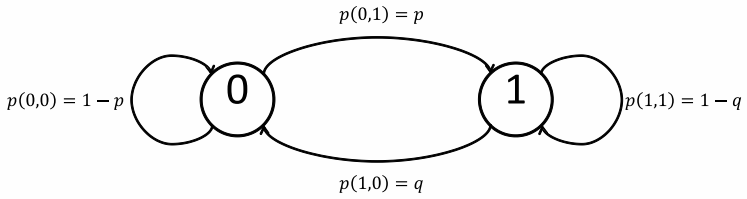

図 2.4: 2 状態の状態遷移図

2.4.5 定常分布

状態の確率分布が時間変化しても変わらない時，定常分布と呼ぶ.

具体的には，状態空間 S と推移確率行列 $\{p_{i,j}\}$ を持つマルコフ連鎖に対して，S 上の確率分布 $\{\pi_i\}$ が，

$$\pi_j = \sum_{i \in S} \pi_i p_{i,j} \qquad (\forall j \in S)$$

を満たすときに，$\{\pi_i\}$ を定常分布，上式を定常方程式と呼ぶ [11].

2.5 待ち行列

指数分布を用いると，普段の生活に身近な順番を待つ状況を確率的に表すことができる．

2.5.1 $M/M/1(\infty)$

日常生活の例として，ATM の平均的な待ち時間について考える．ATM の待ち行列には，お客の到着（ATM への参加）とサービスの終了の 2 つの指数分布が関係している．ここでは，お客が到着する割合を λ，サービスが終了する割合を μ と表す．λ と μ は互いに指数分布である．待ち行列にいる人数は，お客が到着すれば $n \to n+1$ のようになり，サービスが終了すれば $n \to n-1$ のように変化する．

ここで，$\rho = \lambda/\mu$ とする．これは，変化として $n \to n+1$ となる確率が λ/μ であることを表す．

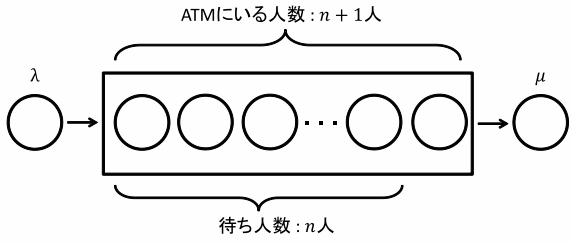

図 2.5: ATM 待ち行列

平衡方程式

状態の集合 C, D に対して，単位時間あたりに，C から D に訪れる回数の期待値を $\alpha(C, D)$ と表す．離散時間型の確率過程では単位時間ごとに時間が進むので，定常分布 $\{\pi_j\}$ を持つマルコフ連鎖では，

$$\alpha(C, D) = \sum_{i \in S} \pi_i \sum_{j \in D} p_{i,j}$$

である．

$$\pi_j(1 - p_{j,j}) = \sum_{i \in S - \{j\}} \pi_i p_{i,j}$$

であるから，

$$\alpha(\{j\}, S - \{j\}) = \alpha(S - \{j\}, \{j\})$$

が得られる．この式は，j から出る率と j へ入る率が等しいことを表している．この式を平衡方程式と呼ぶ．

ATM にお客が 0 人，1 人，2 人，$\cdots$, n 人, $\cdots$ いる確率を求める．

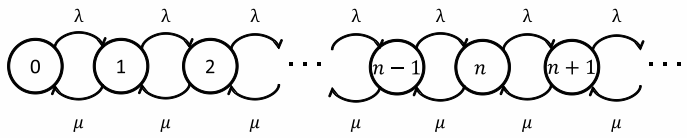

図 2.6: $M/M/1(\infty)$ 待ち行列

まず待ち人数 0 における平衡方程式は，

$$\mu p_1 = \lambda p_0$$

で表すことができる．また，待ち人数 1 における平衡方程式は，

$$\begin{aligned} \lambda p_0 + \mu p_2 &= (\lambda + \mu) p_1 \\ \mu p_2 &= (\lambda + \mu) p_1 - \lambda p_0 \\ p_2 &= (\rho + 1)\rho p_0 - \rho p_0 \\ &= \rho^2 p_0 \end{aligned}$$

で表すことができ，同様にして考えると待ち人数 n における平衡方程式は，

$$p_n = \rho^{n+1} p_0$$

となる.

ここで,

$$\sum_{n=0}^{\infty} p_n = 1$$

であるので,

$$\begin{aligned}
\sum_{n=0}^{\infty} p_n &= p_0 + p_1 + p_2 + \cdots + p_n + \cdots \\
&= p_0 + \rho p_0 + \rho^2 p_0 + \cdots + \rho^n p_0 + \cdots \\
&= p_0 \frac{1}{1-\rho} \qquad (\rho < 1)
\end{aligned}$$

したがって,

$$p_0 = 1 - \rho$$

となり,

$$p_n = \rho^n p_0 = \rho^n (1 - \rho)$$

と平衡方程式を解いて，ATM にお客が 0 人，1 人，2 人，$\cdots, n$ 人，$\cdots$ いる確率を求めることができる [12].

2.5.2 $M/M/c/c$

図 2.7 のような状態遷移で表される待ち行列を考える.

平衡方程式

状態 k に対する平衡方程式は次式で与えられる.

$$\begin{aligned}
\lambda p_o &= \mu p_1 \\
(\lambda + k\mu) p_k &= \lambda p_{k-1} + (k+1)\mu p_{k+1}, \quad k = 1, 2, \cdots, c-1 \\
\lambda p_{c-1} &= c\mu p_c
\end{aligned}$$

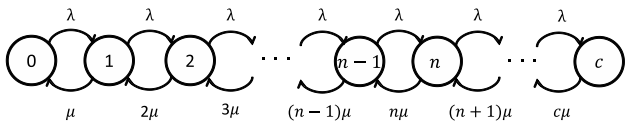

図 2.7: $M/M/c/c$ 待ち行列

$k = 0, 1, \cdots, n-1 (n \leq c)$ について辺々，足し合わせて $n = k$ とすれば

$$\lambda p_{k-1} = k\mu p_k, \quad k = 1, 2, \cdots, c$$

となる．すなわち

$$p_k = \frac{\lambda}{k\mu} p_{k-1} = \frac{\rho}{k} p_{k-1}, \quad k = 1, 2, \cdots, c$$

がすべての $k (= 1, 2, \cdots, c)$ について成立する．よって

$$p_k = \frac{\rho}{k} p_{k-1} = \frac{\rho^2}{k(k-1)} p_{k-2} = \cdots = \frac{\rho^k}{k!} p_0, \quad k = 1, 2, \bullet\bullet\bullet, c$$

を得る．上式は $k = 0$ の場合も成立することに注意する．ここで，確率の総和は 1 であることから

$$\sum_{k=0}^{c} p_k = \sum_{k=0}^{c} \frac{\rho^k}{k!} p_0 = 1$$

が任意の ρ に対して成立する．よってこの待ち行列は定常状態が存在し，

$$p_0 = \left[\sum_{k=0}^{c} \frac{\rho^k}{k!} \right]^{-1}$$

となるので，定常状態確率は

$$p_k = \frac{\rho^k/k!}{\sum_{i=0}^{c} \rho^i/i!}, \quad k = 0, \cdots, c$$

となる 2.7.

2.6 関連研究

2.6.1 Goel-Okumoto モデル

GO モデルは，ソフトウェアの信頼性を定量的に把握する数理モデルである．単位時間あたりに発見される期待エラー数はソフトウェア内に残存する期待エラー数に比例すると仮定し，指数系ソフトウェア信頼度成長モデル

$$\frac{dR(x)}{dx} = b\{a - R(x)\}, \quad a > 0, \quad b > 0, \quad X(t) = x$$

を提案した．ここで，a は極限で観測されるソフトウェアの期待エラー数，b は残存エラー 1 単位あたりのエラー発生率としている．上式を初期条件 $R(0) = 0$ の下で解くと，

$$R(x) = a(1 - e^{-bx})$$

と平均値関数が与えられる．ここで，$R(\infty) = a$ である．さらに，

$$r(x) = abe^{-bx}$$

と強度関数が与えられる．[14][15]

拡散モデルの研究 [9] においては，ソフトウェア故障率を $\lambda(t)$，時刻 t までのソフトウェアの累積故障数を $N(t)$，時刻 t までに検出されるソフトウェアの期待故障数を $m(t) = E[N(t)]$ とし，

$$\begin{aligned} m(t) &= a(1 - e^{-bx}) \\ \lambda(t) &= abe^{-bx} \end{aligned}$$

としている．さらに $N(t)$ が $m(t)$ のポアソン過程であることから，$N(t) = n$ となる確率を

$$P\{N(t) = n\} = \frac{[m(t)]^n \cdot e^{-mt}}{n!}, \quad n = 0, 1, \cdots$$

としている．

2.6.2 NHCTMC を用いた拡散モデル

GO モデルでは感染台数 n は無限数で表されていた．しかし，実際には端末台数は有限であり，感染台数に無限数を用いると実用的でないとされた．そこで感染台数を有限数で表し NHCTMC に拡張された [16].

NHCTMC を用いたウイルス拡散モデル [9] では，マルコフ過程によりウイルスの拡散を図 2.8 の状態遷移図で表している．ここでは，情報システムを構成する総 PC 台数を N 台とし，時刻 t においてウイルスに感染していない PC 台数，感染している PC 台数をそれぞれ $\{X(t), t \geq 0\}, \{Y(t), t \geq 0\}$ の連続時間確率過程で定義している．また，ウイルスの感染率を $\lambda(t)$，駆除率を $\mu(t)$ とし，

$$\begin{aligned} \lambda(t) &= abe^{-bt} \\ \mu(t) &= \mu \end{aligned}$$

と仮定している．ここで，極限で観測される期待感染台数を a，感染発生率を b と定義し，駆除率は時間に依存せず一定の μ としている．

時刻 t においてウイルスに感染していない PC 台数が i 台，ウイルスに感染している PC 台数が j 台である確率を $\pi_{i,j}(t), 0 \leq i+j \leq N$ としている．さらに，その確率ベクトルを

$\pi(t) = [\pi_{0,0}(t), \pi_{1,0}(t), \cdots, \pi_{i,j}(t), \cdots, \pi_{1,N-1}(t), \pi_{0,N}(t)]$ とすると，

$$\pi(t) = \pi(0)e^{\int_0^t Q(x)dx}$$

と表現している [17]．ここで $Q(x)$ は時刻 x での NHCTMC の無限小生成作用素（推移率行列）である．

また，総 PC 台数の K 台以上がウイルスに感染している状態をハザードと定義し，時刻 t において情報システムの信頼度 $R(t)$ を以下のように表現している，

$$R(t) = 1 - \sum_{j=K}^{N} \pi_{i,j}(t)$$

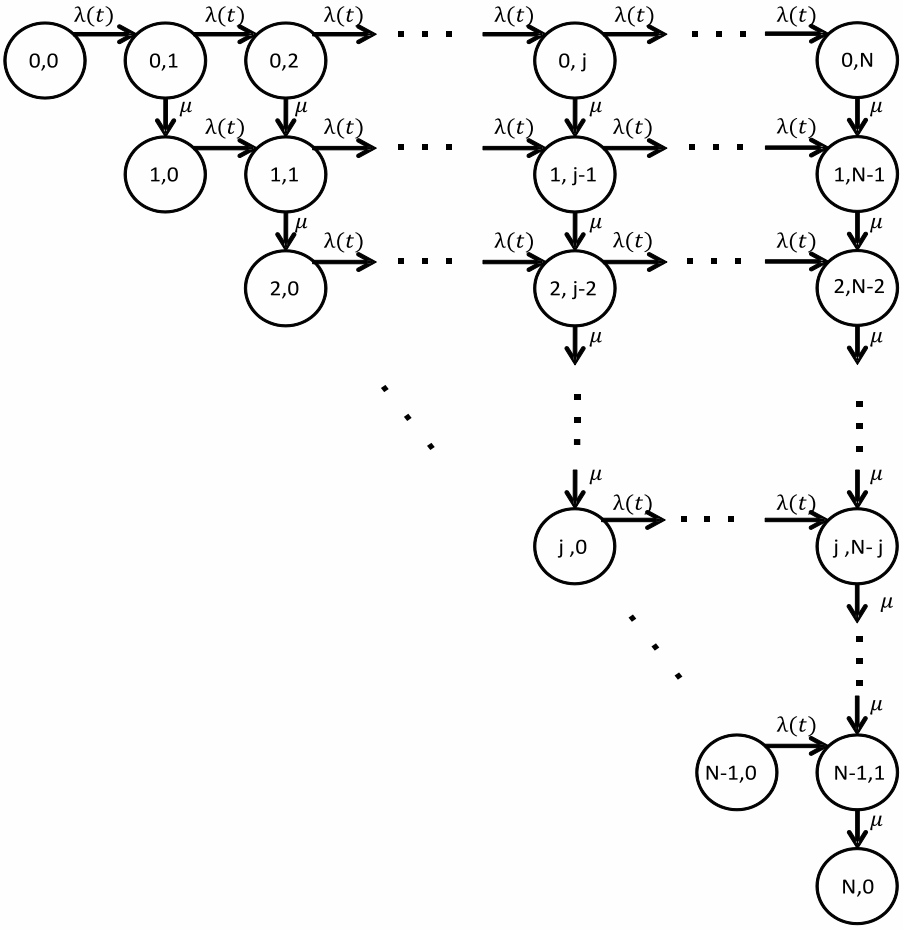

図 2.8: NHCTMC を用いたウイルス拡散モデル

第3章　2次感染拡散モデル

3.1　感染経路の違い

PC と Android には，ウイルスの拡散方法について感染経路に違いがある．それゆえ，ウイルスの発生率や感染率，駆除率は各端末ごとに異なると考えられる．

同じ端末間でのウイルス拡散については従来手法の値を変えることで，各端末のウイルス拡散について評価することができる．

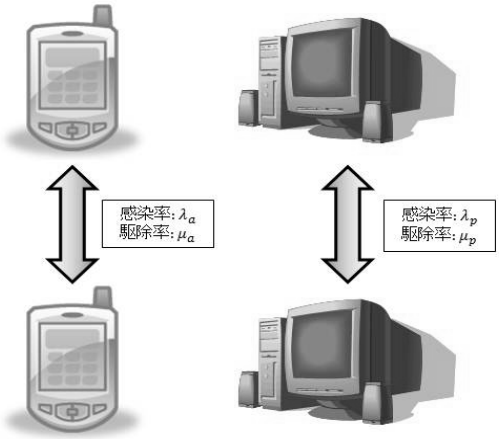

図 3.1: Android，PC のウイルス感染

図 3.1 では Android 端末における感染率，駆除率をそれぞれ λ_a, μ_a，PC 端末における感染率，駆除率をそれぞれ λ_p, μ_p としている．

3.2 従来手法の問題点

先に述べた通り，同じ端末間での感染については従来手法の値を変えればそれぞれの端末について評価することは可能である．しかし，Android から PC への 2 次感染については，2 次感染の感染率を導入しなくてはいけない．従来手法では，1 つのモデルにつき 1 つ端末の感染率，駆除率を設定している．それゆえ，従来手法を用いたモデル化は難しくなる．1 つのモデルで，異なる 2 つの端末の感染，拡散を考えなくてはいけないためである．

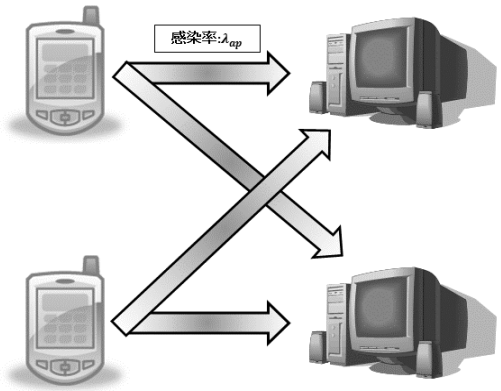

図 3.2: Android から PC への 2 次感染

図 3.2 では，Android から PC への感染率を λ_{ap} としている．

3.3 提案モデル

3.3.1 状態遷移図

従来手法の問題点を考慮して従来モデルに Android を導入し，Android からの 2 次感染を導入したモデルを提案する．

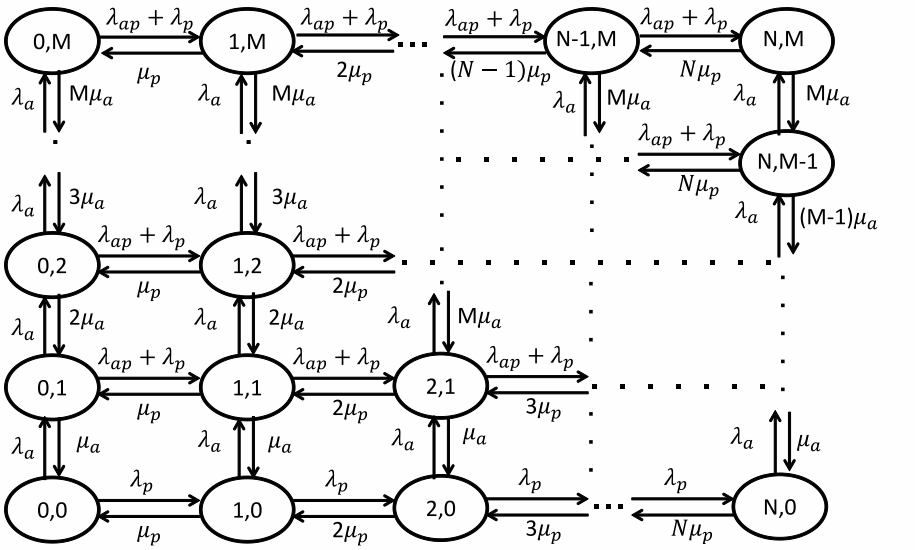

図 3.3: 2 次感染拡散モデル

提案モデルの状態遷移図を図 3.3 のようにウイルスの拡散，駆除をモデリングする．図 3.3 では，総 PC 台数を N 台，総 Android 台数を M 台としている．ここで PC のウイルス感染率を λ_p，Android の感染率を λ_a とし，Android から PC への感染率を λ_{ap} のポアソン分布としている．

Android 感染台数 j が $j \geq 1$ であるときの PC への感染確率は，PC 間感染，2 次感染が想定できる．したがって，$\lambda_{ap} + \lambda_p$ と表す．さらに PC のウイルス駆除率を μ_p，Android の駆除率を μ_a の指数分布としている．また $\{p_{ij}\}, 0 \leq i \leq N, 0 \leq j \leq M$ を状態 i, j となる確率と定義する．

時間が十分経過したとき，PC と Android の感染台数がどのような値をとるのかを求めていきたい．したがって，ここでは時間が十分

経過した定常状態であることを仮定する.

3.3.2 平衡方程式

図 3.3 より 2 次感染モデルにおいて，$j = 0$ の場合と $j \geq 1$ の場合で PC への感染率が異なる．そこで平衡方程式を求めるにあたり，$j = 0$ と $j \geq 1$ で場合分けをする必要がある．平衡方程式は以下のようになる.

・$j = 0$ の場合

$$\begin{cases} (\lambda_p + \lambda_a)p_{00} = f(1)\mu_p p_{10} + f(1)\mu_a p_{01} \\ \lambda_p p_{(i-1)0} + f(i+1)\mu_p p_{(i+1)0} + f(1)\mu_a p_{i1} = (\lambda_a + \lambda_p + f(i)\mu_p)p_{i0} \\ (\lambda_a + f(N)\mu_p)p_{N0} = \lambda_p p_{(N-1)0} + f(1)\mu_a p_{N1} \end{cases} \tag{3.1}$$

・$j \geq 1$ の場合

$$\begin{cases} \{\lambda_a + (\lambda_{ap} + \lambda_p) + f(j)\mu_a\}p_{0j} = f(1)\mu_p p_{1j} + f(j+1)\mu_a p_{1(j+1)} + \lambda_a p_{1(j-1)} \\ \{\lambda_a + (\lambda_{ap} + \lambda_p) + i\mu_p + j\mu_a\}p_{ij} \\ = \lambda_a p_{i(j-1)} + (\lambda_{ap} + \lambda_p)p_{(i-1)j} + j\mu_a p_{(i+1)j} + i\mu_p p_{i(j+1)} \\ (\lambda_a + f(N)\mu_p + f(j)\mu_a)p_{Nj} = (\lambda_a p + \lambda_p)p_{(N-1)j} + f(j+1)\mu_a p_{N(j+1)} + \lambda_a p_{N(j-1)} \end{cases} \tag{3.2}$$

また PC 間感染，Android 間感染についての平衡方程式は

$$\begin{cases} \lambda_p p_0 = f(1)\mu_p p_1 \\ \{\lambda_p + f(n+1)\mu_p\}p_{n+1} = \lambda_p p_n + f(n+2)\mu_p p_{n+2} \\ \lambda_p p_{N-1} = N\mu_p p_N \end{cases} \tag{3.3}$$

$$\begin{cases} \lambda_a p_0 = f(1)\mu_a p_1 \\ \{\lambda_a + f(m+1)\mu_a\}p_{m+1} = \lambda_a p_m + f(m+2)\mu_a p_{m+2} \\ \lambda_a p_{M-1} = M\mu_a p_M \end{cases} \tag{3.4}$$

でそれぞれ表される.

3.3.3 2 次感染モデルの期待値

まず$j=0$の場合について考える．式3.1において，$i=0,1,\cdots,s-1(s\leq N)$について辺々足し合わせれば

$$\lambda_p p_{(s-1)0}+\lambda_a\sum_{c=0}^{f(s)-1}p_{c0}=f(s)\mu_p p_{s0}+\mu_a\sum_{c=0}^{f(s)-1}p_{c1}$$

が得られる．これはすべての$s=1,2,\cdots,N$について成立する．よって

$$p_{s0}=\frac{\rho_p}{f(s)}p_{(s-1)0}+\frac{1}{\mu_p f(s)}\left(\lambda_a\sum_{c=0}^{f(s)-1}p_{c0}-\mu_a\sum_{c=0}^{f(s)-1}p_{c1}\right)$$

ここで

$$f(f(s))=\lambda_a\sum_{c=0}^{f(s)-1}p_{c0}-\mu_a\sum_{c=0}^{f(s)-1}p_{c1}$$

とすれば

$$\begin{aligned}
p_{s0} &= \frac{\rho_p}{f(s)}\left\{\frac{\rho_p}{f(s-1)}p_{(s-2)0}+\frac{f(f(s-1))}{\mu_p f(s-1)}\right\}+\frac{f(f(s))}{\mu_p f(s)}\\
&= \frac{\rho_p^2}{f(s)f(s-1)}p_{(s-2)0}+\frac{\rho_p f(f(s-1))}{\mu_p f(s)f(s-1)}+\frac{f(f(s))}{\mu_p f(s)}\\
&= \cdots\\
&= \frac{\rho_p^s}{\prod_{l=0}^{s}f(l)}p_{00}+\frac{1}{\mu_p}\sum_{d=1}^{s}\frac{\rho^{d-1}f(f(s-d+1))}{\prod_{l=s-d+1}^{s}f(l)} \qquad (3.5)
\end{aligned}$$

となる．次に$j\geq 1$についても同様に考えれば，

$$\begin{aligned}
p_{s,j} =& \frac{1}{\prod_{l=0}^{s}f(l)}\left\{\frac{(\lambda_{ap}+\lambda_p)+f(j)\mu_a}{\mu_p}\right\}^s p_{00}\\
&+\sum_{d=1}^{s}\frac{\{(\lambda_{ap}+\lambda_p)+f(j)\mu_a)^{d-1}f(f(s-d+1))}{\mu_p^d\prod_{l=s-d+1}^{s}f(l)} \qquad (3.6)
\end{aligned}$$

となり，PC 感染の方向における状態方程式を解くことができた．

式3.6 を Android 感染の方向に平衡方程式を解き p_{st} を導きたい．p_{st} が求ることができれば確率の総和は 1 であるので，

$$\sum_{s=0}^{N}\sum_{t=0}^{M} p_{st} = 1 \tag{3.7}$$

を解くことで，2 次感染モデルの平衡方程式を解くことができると考えられる．

平衡方程式の解を用いれば

$$E_{ap} = \sum_{s=0}^{N}\sum_{t=0}^{M}(s+t)p_{st}$$

として 2 次感染モデルにおける期待感染台数を求めることができると考えている．

3.3.4 PC 間，Android 間感染の期待値

PC 間感染，Android 間感染の感染台数の期待値 E_p, E_a を求める．PC 間感染について式3.3 において，$n = 0, 1, \cdots, k-1 (k \leq N)$ について辺々足し合わせると

$$\lambda_p p_{k-1} = f(k)\mu_p p_k$$

が得られる．すなわち

$$\begin{aligned} p_k &= \frac{\rho_p}{f(k)}p_{k-1} = \frac{\rho_p^2}{f(k)f(k-1)}p_{k-2} \\ &= \cdots = \frac{\rho_p^k}{\prod_{l=1}^{k} f(l)}p_0 \qquad (k = 1, 2, \cdots, N) \end{aligned}$$

である．ここで$k = 0$のときもp_kが成り立つことに注意し，確率の総和が 1 であることを用いれば

$$\begin{aligned}\sum_{k=0}^{N} p_k &= p_0 + \sum_{k=1}^{N} p_k \\ &= p_0 + \sum_{k=1}^{N} \frac{\rho_p^k}{k!} p_0 \\ &= \sum_{k=0}^{N} \frac{\rho_p^k}{k!} p_0 = 1\end{aligned}$$

よって

$$p_k = \frac{\rho_p^k}{k!}\left\{\sum_{k=0}^{N} \frac{\rho_p^k}{k!}\right\}^{-1}, \quad k = 0, 1, \cdots, N \tag{3.8}$$

Android 間感染についても式 3.4 を用いて同様にして考えれば,

$$p_q = \frac{\rho_p^q}{q!}\left\{\sum_{q=0}^{N} \frac{\rho_p^q}{q!}\right\}^{-1}, \quad q = 0, 1, \cdots, M \tag{3.9}$$

が得られる.

以上より感染台数の期待値を求める．式 3.3 より，PC 間感染における感染台数の期待値は以下のようになる 2.7.

$$\begin{aligned}E_p &= \sum_{k=0}^{N} k p_k \\ &= 0 \cdot p_0 + \sum_{k=1}^{N} k \cdot \frac{\rho_p^k}{k!}\left\{\sum_{k=0}^{N} \frac{\rho_p^k}{k!}\right\}^{-1} \\ &= \sum_{k=1}^{N} \frac{{\rho_p}^k}{(k-1)!}\left\{\sum_{k=0}^{N} \frac{{\rho_p}^k}{k!}\right\}^{-1}, \qquad k = 0, 1, \cdots, N\end{aligned}$$

$$\begin{aligned}E_a &= \sum_{q=0}^{M} q p_q \\ &= \sum_{q=1}^{M} \frac{{\rho_a}^q}{(q-1)!}\left\{\sum_{q=0}^{M} \frac{{\rho_a}^q}{q!}\right\}^{-1}, \qquad q = 0, 1, \cdots, M\end{aligned}$$

第4章　結論

本論文では従来の拡散モデルにおいて2次感染を考慮し，マルコフ過程を用いて2次感染拡散モデルを提案した．

3章の提案モデルでは，時間が十分経過している状態を仮定して時間に依存しない定常状態における状態遷移を提案した．その際にウイルス感染率をポアソン分布，ウイルス除去率を指数分布で表記した．本論文では定常状態において各状態における平衡方程式を解くことで，任意の状態における状態確率を求めた．さらに任意の状態確率を用いて状態確率の期待値の式を導いた．

今後の課題としては，定常状態における状態確率を計算する．さらに，システムが過渡状態である場合についてを考える．それによって，ウイルスの拡散の様子が時間経過につれてどのように変化していくかをグラフ化できる．グラフ化することで2次感染が起きた場合，同じ端末のみの感染の場合とウイルス拡散の特徴を直感的に視覚するためである．

謝辞

本研究を進めるにあたり多くの助言とご指導をして下さった木下宏揚教授，宮田純子氏に感謝致します．また公私にわたり良き研究生活を送らせていただいた木下研究室の方々に感謝致します．

2013年2月
古屋　智規

参考文献

[1] リッチマンコンテンツ・マーケティング情報局
http://www.richcontent.jp/researchdata/rd45-01.html

[2] impress R&D
http://www.impressrd.jp/news/121120/kwp2013

[3] Android Smart
http://android-smart.com/2011/08/android-vs-iphone.html

[4] Android のウイルスに最も効果的なウイルス対策ソフト比較
http://www.hs-works.com/

[5] ジェネラルインフォ
http://www.general-info.org/mobile/599

[6] TREND MICRO, “TrendLabs SECURITY BLOG, ”
http://blog.trendmicro.co.jp/archives/4485

[7] 松原　望, 入門確率過程, 東京図書株式会社,2003.

[8] 小林　尚志, 岡村　寛之, 土肥　正,“確率モデルに基づいたコンピュータウイルスの挙動解析,” 日本オペレーションズ・リサーチ学会秋季研究発表会アブストラクト集 2003, pp.122-123, 2003-09-10.

[9] 植手　大輔，奥田　隆史，井手口　哲夫，“非同時連続時間マルコフ連鎖を用いたコンピュータウイルス拡散モデルの評価，” 電子情報通信学会総合大会講演論文集，2004 年 情報・システム (1), pp．132，2004 - 03 - 08.

[10] 尾畑　伸明, 確率モデル要論 -確率論の基礎からマルコフ連鎖へ-, 牧野書店,2012.

[11] 宮沢　政清, 確率と確率過程, 近代科学社,1993.

[12] 松原　望, 松原望の確率過程　超!入門, 東京図書株式会社,2011.

[13] 滝根　哲哉，”初等通信トラヒック理論，” http://www2b.comm.eng.osaka-u.ac.jp/ takine/tmp/v3.4.pdf

[14] A.L.Goel and K.Okumoto,“Time-dependent error-detection rate model for software reliability and other performance measures,” IEEE Trans. Realiability, vol.R-28,pp.206-211,1979.

[15] 土肥　正, 富岡　恒雄, 海生　直人,“エラー保全時間を考慮した最適ソフトウェア・リリース政策に関する一考察,” 経済科学研究,9(1),pp.45-53,2005-09-30.

[16] S.S.Gokhale, P.N.Marinos, M.R.Lyu, K.S.Tribedi, “Effect of Repair Polices on Software Reliability”, The 2000 Pacific Rim International Symposium on Dependable Computing, Los Angels, CA, 2000.

[17] K.S.Tribedi. Probability and Statistic with Reliability, Queueing, and Computer Science Apprications, John Wiley & Sons Ltd, New York. 2002.

質疑応答

Q. 感染率について，PC 間の感染率と Android 間の感染率どちらが強いのか.

A. 今回は PC 間感染，Android 間感染モデルと，2 次感染モデルについて比較を行った．PC 間感染モデルと Android 間感染モデルについて比較は行っていない．それゆえ，PC 間の感染率と Android 間の感染率の大小についての比較は行う必要がないと考えている.