

MASによるブログシミュレーション

2006MI011 藤川 千絵美

2006MI199 山田 諭史

指導教員 石崎 文雄

1 はじめに

近年、インターネットの普及に伴い、多くの人々が様々な方法で情報を発信することができるようになった。その情報発信の手段の一つとして、ブログが人気を集めている。図1示されているように、日本におけるブログの利用者数は年々増え続け、2007年3月末には782万人に達している。

ブログとは、ウェブログ (Weblog) の略称で、自分の考えや社会的な出来事に対する意見、物事に対する論評、他のウェブサイトに対する情報などを公開するためのウェブサイトのことである。ブログが人気を得ている理由は、ウェブサイトを作成するために最低限必要となるHTMLなどの専門知識を必要とせずに電子掲示板に書き込む要領で利用できることや、携帯電話などのモバイル通信端末のインターネット機能を用いることで、いつでもどこにいても更新できるなど誰もが簡単に利用できるようなになっていること等が理由として考えられる。さらに最近では、タレントのブログ利用も増加しており、自身でブログを作成していない人でも閲覧者としては利用している人も多くなっている。そのせいか個人の利用だけでなく、ブログを利用した販売促進支援を行うサービスも続々と登場しており、企業がビジネスとして利用することも多くなった。これらの理由により、ブログの利用者数は年々増え続けている。

ブログの先行研究として、大衆にとって多くの選択肢の中から自由に選べるシステムでは、全体のごく一部分が不釣り合いなほどのトラフィック (注目や収入など) を得ることがある。これはモラルの弱さや裏切りやその他の心理的な理由とは全く関係がなく、選択するという行為自体が広範にかつ自由に行われる結果、ベキ法則分布ができあがるのである。ベキ法則分布とは、よく言われるように、80/20の法則とか勝者の独り占めのような状態を生じる形であり、昨今は十分有効なものであると理解されている。ブログでは、リンクによって繋がったブログ同士のネットワークが構築されている。そのネットワークも、そのようなベキ法則に従っていると思われる [2]。ベキ乗則に従うものとして、例えば、ウェブサイトのリンクの数、ウェブサイトのヒットの件数、株価の変動とその頻度の関係などが知られている。

最近ではある他人のブログの記事に自身のブログへのリンクを作成するトラックバック機能など、リンク機能も充実してきた。また、あるブログのページの閲覧者が、他のどのサイトに表示されているリンクから来たかをカウントし、それをランキングとして表示するためのツールが多数開発され、ブログ利用者間で人気となっ

ている [3]。

このようなランキングを自分のページに掲載すると、掲載されたランキングからも他のブログへのリンクがはられることになる。このようなツールの利用によりブログ間では、動的、自動的にリンクがはられることが多いため、通常のWebページのリンクによって繋がったネットワークとは異なる性質を持つ可能性が考えられる。

本研究では、ブログ同士のリンクの繋がりで形成されたネットワークがどのような性質を持っているのかをシミュレーションにより研究する。特に、ブログ内にランキングが掲載され、そのランキングから他のブログへのリンクがはられていることを仮定し、ランキングのリンクの存在が、構築されるネットワークの性質に及ぼす影響を調べる。図2はブログを利用している人にアンケートをとったものである。図2を見てみると実際のブログの閲覧数は数字を伸ばしているのは極一部であり大多数は閲覧数は100以下という事実がある。特に、本研究のシミュレーションでは、ブログ内にランキングを掲載することで、ブログの閲覧数に影響を及ぼすかどうか、影響するのであればどのように影響するのかを考察する。

本研究のシミュレーションには、*artisoc* というシミュレータを利用し、仮想的に構築したブログのネットワークのシミュレーションを行う。

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できるマルチエージェント・シミュレータである。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、または、そのような複雑なシステムをモデル化・理解する方法のことである [4]。そのモデル化・理解の方法をMAS (マルチエージェントシミュレーション) という。

2 ブログネットワークシミュレーション

2.1 ブログネットワークシミュレーションの概要

本節では *artisoc* によるブログネットワークのシミュレーションの概要について述べる。MASによるシミュレーションということで大きく分けてブログページと閲覧者という2つのエージェントを作成する。ブログページは初期設定で作成し、閲覧者は毎ステップ作成し動きをみせる。まずそれぞれのエージェントの動きを次のように構築する。

2.1.1 ブログページと閲覧者の設定

1. ブログページの設定

- (a) 任意の数のブログページを作る (今回のシ

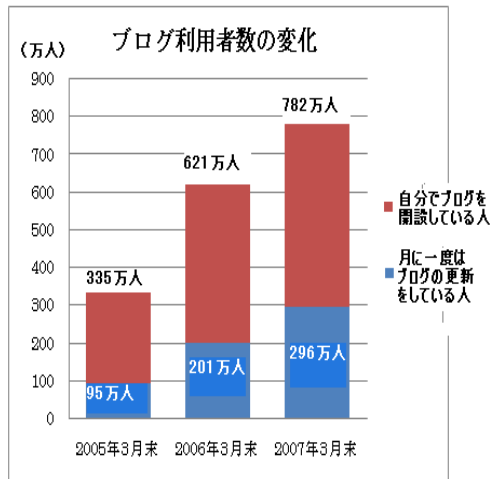


図1 ブログ利用者数の変化 [5]

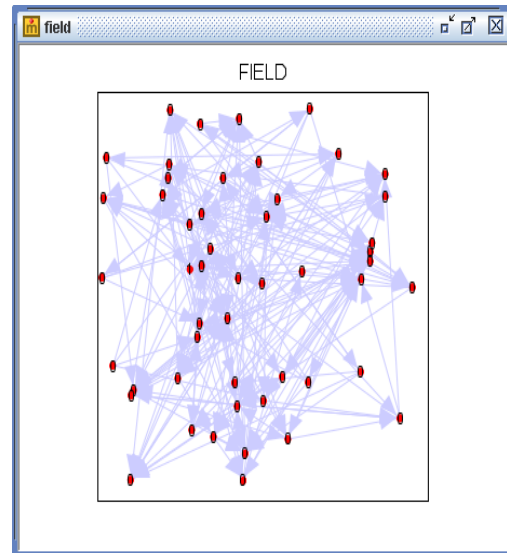


図3 artisocによるブログページのネットワーク図

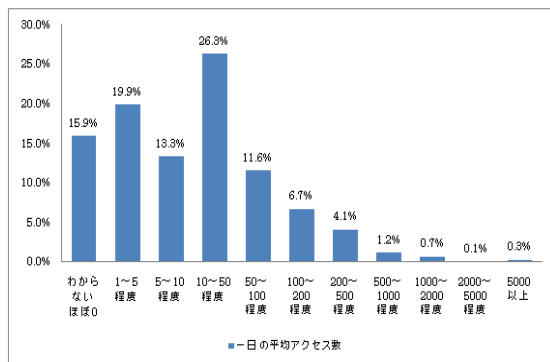


図2 ブログページの1日の平均アクセス数調査 [6]

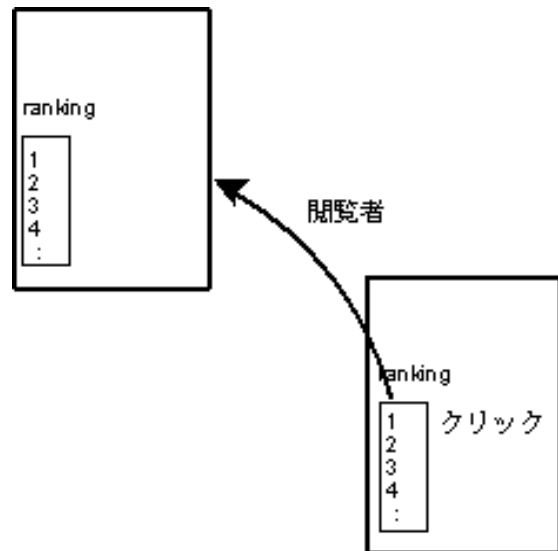


図4 閲覧者の動きのイメージ

- ミュレーションでは任意の数は100とした)
- (b) それぞれのページには自分のブログページ以外のすべてのページへのリンクを作る
 - (c) それぞれのブログページには閲覧者が次に閲覧するページを決定するランキングというものを付ける
2. 閲覧者の設定
- (a) 本研究では100人の閲覧者を作成しブログページを移動させる。
 - (b) 閲覧者は次のステップで次に見るべきページを今閲覧しているページのリンクの中からランキングを参考にして選ぶ
 - (c) 選ばれたページへのリンクを辿り、閲覧者はそのページへ移り閲覧する
 - (d) 以下 b と c の作業を繰り返す

2.1.2 ランキングの設定

本節では閲覧者の動きを決定するうえで重要になってくるランキングの設定について述べる。

ランキングは例えば図4のようにブログページのパーツのひとつであり、それぞれのブログページからリンク先のブログページへ移動した閲覧者の数をランクづけしたものである。ランキングはそれぞれのブログページに設けてあり、ブログページからリンク先のブログページへの移動した閲覧者の数をカウントし、降順で並び替えたものである。今回のシミュレーションではこのランキングのみで閲覧者の動きを決定することとする。

本研究では、ランキングに表示されたリンクを閲覧者

がどう選ぶかについて以下のことを仮定する．

- 閲覧者がランキング内のどのリンクを選ぶかは確率的に決まる．
- 閲覧者はランキング内の順位が高いリンクほど高い確率で選ぶ．
- ランキングの順位の動向はリンク先の閲覧数の多さによって決められる．
- ランキングには同率順位を設けている．閲覧数の同じリンク先が複数個あった場合そのリンク先へは同じ遷移確率を与える．(初期設定ではすべて同じ閲覧数のため遷移確率は同じ)

これらの仮定は実際の状況を適切に反映していると考えられる．仮定にしたがった閲覧者のリンク選択を行わせるために以下の乱数 A を考える．

$$A = \text{INT}[\text{Rnd}()^x \times \text{ブログページのリンクしている数}] \quad (1)$$

ここに、 $\text{Rnd}()$ は $[0,1)$ 上の一様分布に従う疑似乱数系列を発生させる関数で、 INT は整数部分を取り出す関数で、 x ($x \geq 1$) はランキング重度と呼ばれるパラメータである．閲覧者は順位 A に対応するリンク先を移動先として選ぶこととする．式 (1) に含まれるランキング重度 x が 1 のときは、リンクの選ばれやすさランキングの順位と無関係になり、 x が 1 より大きくなればなるほど、ランキングで高順位のリンクほど選ばれやすくなることに注意．式 (1) で決められる順位をもとに 100 位までの確率を求めて閲覧数の多いリンクほど高い確率を割り振る．そして今回のシミュレーションでは同率順位を作る．同率順位であるリンク先にはその順位に準ずる同じ遷移確率を与えることにする．例えば 2 位が 3 個ある場合、あらかじめ式 (1) に決められた 2 位から 4 位までの遷移確率を足して 3 で割りそのリンク先への遷移確率を割り振る．

本研究のシミュレーションでは、各ブログを状態と考える有限状態マルコフ連鎖を考える．このとき、ブログ 0 のランキングにブログ 1 が入っているとき (言い換えると、ブログ 0 からブログ 1 にリンクがはられているとき)、マルコフ連鎖の状態 0 から状態 1 への遷移確率は、ランキングにおいてブログ 1 が選ばれる確率となる．リンクがはられていないときは遷移確率は 0 となる．各閲覧者のブログ間の移動は、このような有限状態マルコフ連鎖のサンプルパスを表わしていると考えられる．ただし、マルコフ連鎖の遷移確率は時間とともに変化する．特に、閲覧者がランキングでどのブログを選んだかにより、変化する．遷移確率行列 (各状態間の遷移確率) が、時間で変化するのでこのマルコフ連鎖の各状態の定常確率も変化していく．ある時刻での定常確率は、今後遷移確率が変化しなかったとしたときに、閲覧者が各状態 (各ブログ) を訪問している確率 (長期間観察した時に、その状態にいる時間割合) に相当する．よって、各時点での各状態 (各ブログ) の定常状態確率を求めれば、その時点でそのブログが全体の中でどのぐらいアクセスを

集められるブログになっているかが分かる

2.1.3 有限状態マルコフ連鎖

ある時点から将来の振舞いがたがう確率法則が、現在の状態のみに依存して過去の状態推移には無関係に決まる性質を無記憶性といい、この性質をもった確率過程をマルコフ過程という．その中で状態が離散的なものを一般的にマルコフ連鎖という．ここで、状態空間が有限ならば遷移確率は行列で表現され遷移行列となる．時間的に一様なマルコフ連鎖で p_{ij} で記述でき、ベクトル π の要素 π_j の和が 1 で、次を満たすとす．

$$\pi_j = \sum_{i \in S} \pi_i p_{ij} \quad (2)$$

式 (2) の場合にはベクトル π を定常分布という．非周期的であり、そのすべての状態が再帰的ならば、定常分布を持つ．定常分布とはそのノードに留まっている確率を出したものである．

本研究の定常分布を遷移確率行列を 300 回乗算することにより定常分布を求めている．

3 シミュレーションの結果と説明

本節ではシミュレーションの結果を示し、結果の考察を行う．図 5-8 それぞれランキング重度 $x = 1, 1.5, 2, 3$ でシミュレーションを行ったものである．今回のシミュレーションでは 100 あるブログページの定常分布を全て出力してしまうとグラフが見づらくなるためその定常分布から 10 個 (ノード ID0~90 の 10 個おき) を選抜してグラフに出力した．シミュレーションは 200 ステップさせたところで終了させている． x 軸は時間を表し y 軸は定常状態確率を示しているまた、 y 軸は artisoc の仕様で 0.5 以下の数値を適切に表現できないため数値を 100 倍してある (実際の y 軸の表記は 0.01, 0.02, 0.03) ．

図 5-8 において、定常状態確率の数値の高いグラフとそのほかの差を見ると、ランキング重度の変化によって差が見られた．もしランキングがブログページの閲覧数に影響を与える大きな要因であるのなら x の値が増えることによってそれぞれブログページの閲覧数の差がつきやすくなるはずである．また時間の経過に伴う定常状態確率の順位変動をに着目し、人気のブログページを決定するのははじめの段階で決まっていることがわかる．

4 まとめと考察

シミュレーション結果から、ランキングがブログネットワークに及ぼす影響が観察できた．シミュレーションではランキングの順位による閲覧者の移動を決定する確率を変化させて、ランキングの順位の反映度によるそれぞれのブログページの定常状態確率の変化をみたが、順位がよいものがより選ばれやすくとそれぞれのブログページの閲覧数の差が大きくなる．よってランキングによる閲覧数の変化はあるのではないかと考えた．閲覧数を多いブログはブログの文章の優劣だけではなく他のブログのランキングなどに干渉を受けて閲覧数が多く

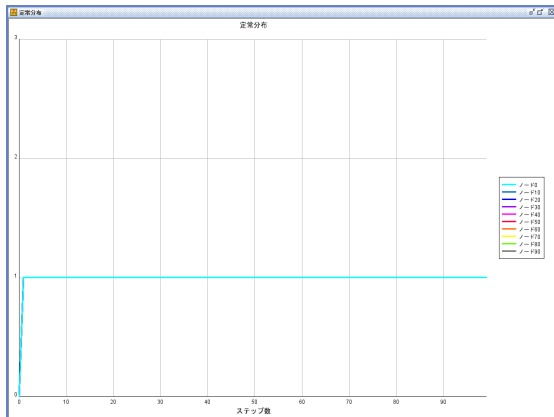


図5 $x = 1$ (ランキング反映なし)

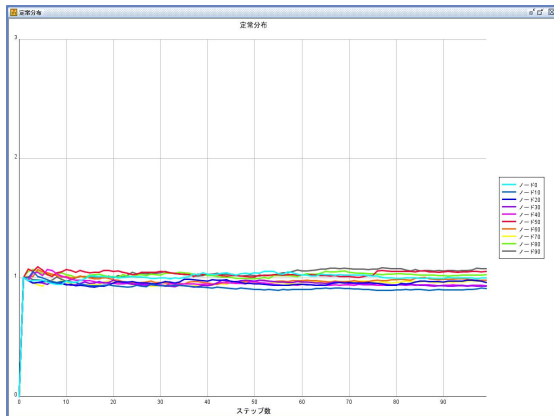


図6 $x = 1.5$

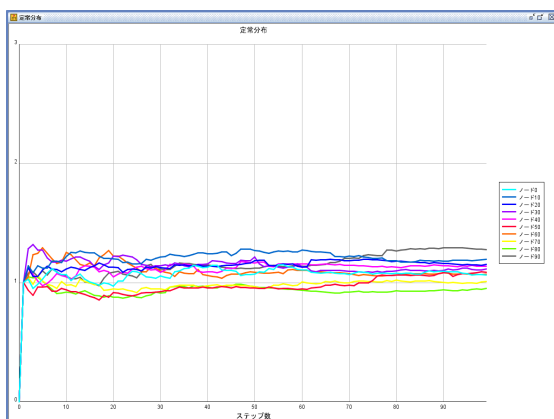


図7 $x = 2.0$

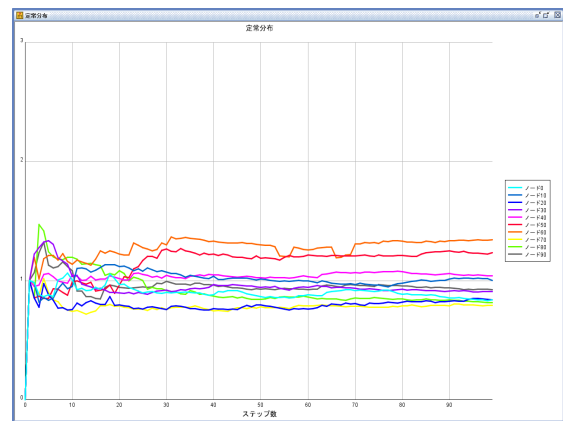


図8 $x = 3.0$

なっているのではないかと考えられる。そして、初めの段階で人気のブログページは決定してしまい、初めの段階で定常確率が高いブログページはその後高い定常確率を維持したままであることがわかる。また閲覧数を初めの段階で多く獲得できなかったブログページの定常確率は低いままでその後も閲覧数を獲得できないままになる。つまりランキングがシミュレーションの結果のように初めの段階で決まった人気によってブログページの閲覧数の増加に影響を与え続けるのではないかと考えられる。

参考文献

- [1] 山影 進：人工社会構築指南，書籍工房早山（2007）。
- [2] Clay Shirky：Power Laws, Weblogs, and Inequality, (2003）。

http://www1.ocn.ne.jp/~kobakan/contents/powerlaws_inequality.html

- [3] ACR WEB-CGI レンタル ブログパーツ:

<http://www.ziyu.net/> .

- [4] MAS コミュニティ:

<http://mas.kke.co.jp/> .

- [5] 調べ学習で使えるワンポイントグラフ集:

<http://www.yomiuri.co.jp/nie/note/kids/200703/02/suji01.htm> .

- [6] iMi リサーチバンク:

<http://www.imi.ne.jp/blogs/research/2008/05/080521blog.html> .