

# 世論形成におけるハブ・ターゲティング戦略の特性

森田 大<sup>†</sup> ディミトリ イラチェット<sup>‡</sup> 生天目 章<sup>†</sup>

<sup>†</sup> 防衛大学校情報数理専攻

<sup>‡</sup> サンシール陸軍士官学校

## 1. 序 論

現在では、誰もが容易に大量の情報にアクセスできる社会になっているが、その分世論に影響を与えるメッセージを伝播することは難しくなっているといえる。ここでは、多くの企業や組織においてどのように広告情報を伝播するか、という事は、関心を持つべき問題である。

この問題に関して、人工社会の構築や社会ネットワークにおける情報伝播などの多くのモデルが考案されており、そこで求められるものは、広告効果を最適にするようなマーケティング戦略である。

本研究では、コンピュータシミュレーションを用いて、マーケティング戦略の一つであるハブ・ターゲティング戦略について、その効果を明らかにするものである。

要領としては、第一に、「二段階モデル」とハブ・ターゲティング戦略の要旨を述べる。

次に、シミュレーションのためのプログラム実装、及びシミュレーションの流れの細部について述べる。

最後に、「二段階モデル」をシミュレーションするため、及びハブ・ターゲティング戦略の効果を理解するためのプログラムをどのように組み立てるかを説明し、結果と結論について述べる。

## 2. 理論モデル

### 2.1 二段階モデルとハブ・ターゲティング戦略

ハブ・ターゲティング戦略は、katz and Lazarsfeld(1955)によって初めて提唱された世論形成モデルを基にしている。彼らはこれを「二段階モデル」とした。

このモデルは、人間の中には2つの異なる種類の個人があるとして、多くの人々を「平凡な人」と分類し、一部の人々を「影響力のある人」と分類した。

二段階モデルでの最も大きな仮説は、世論形成に必要なステップは2ステップであるということである。人々が情報を受け取るとき、その情報はメディアから、まず影響力のある人に伝播し、影響力のある人から平凡な人に伝播していくという構造である。

このモデルが、世論形成において正しいとすれば、マーケティング戦略にはこの影響力のある人に焦点を当てるのが適当である。これはすなわち、影響力のある人をハブとした、ハブ・ターゲティング戦略ということである。つまりこのマーケティング戦略は、最良の情報伝播効果を得るために、オピニオン・リーダー（ハブ・エージェント）をターゲットにする戦略である。

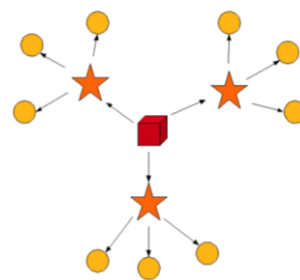


図1：二段階モデル

## 2.2 形式モデル

しかし、二段階モデルだけでは、ハブ・ターゲティング戦略の効果を確認するには不十分であるので、Watts and Dadds(1999)のモデルを参考にする。

### 2.2.1 ランダムネットワーク

本研究では、エージェント数  $N$  でシミュレーションを行い、この  $N$  を人口サイズと呼ぶ。エージェント  $i$  は、リンク数  $n_i$  を持ち、そのリンク先エージェントを一様に選択し、時間の経過により変化することはない。このようにして作られるランダムネットワークは、リンク数  $n_i$  からリンクを求めるといふ点で、各エージェントに対して確率  $p$  でリンクを張る Erdős-Rényi モデル(1959) とは異なる。

ここで、各エージェントのリンク数  $n_i$  から、エージェント間の平均リンク数  $n_{avg}$  を導く。この平均リンク数  $n_{avg}$  は、得られるランダムネットワークの構造を表し、併せてハブ・ターゲティング戦略の効果を表すための相互関係の強さを表す。

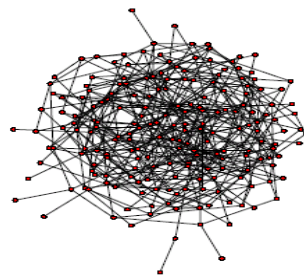


図2：ランダムネットワーク

### 2.2.2 ハブ・エージェント

二段階モデルの肝である影響力のある人を定義する。

影響力のある人という概念を採用するため、最大リンク数  $n_{max}$  を定義する。  $n_{max}$  は本研究のシミュレーション間における最大リンク数となる。

次に、各エージェントにおいて、どのエージェントが影響力のある人であるかを定めるため、定数  $q(0 \leq q \leq 1)$  を定める。この定数  $q$  に基づき、エージェント  $i$  のもつリンク数が  $n_i > qn_{max}$  であるならば、そのエージェント  $i$  を影響力のある人と定める。影響力のある人を「ハブ・エージェント」と定める。

ここで、本ネットワークはランダムネットワークなので、その次数分布はポアソン分布に従う。よって、人口サイズに占めるハブ・エージェントの割合は以下の式で示される。

$$Q(x, \lambda) = \sum_x^{\infty} f(t, \lambda) \quad (2.1)$$
$$x = q * n_{max}, \lambda = n_{avg}$$

### 2.2.3 情報伝播ルール

エージェント個々を見れば、情報を採用するか採用しないかの二種類の状態を遷移するものとなりそうだが、本研究では、各ステップにおける状態変化も重要になるため、三種類の状態を用意する。「情報採用者」「情報未採用者」「情報非伝播者」である。情報採用者は情報を採用し、他に対して影響を与えるものであり、情報未採用者は情報を採用しておらず、情報採用者からの影響をうけるものである。情報非伝播者は、情報を採用してはいるが、他に対して影響を与えないものとする。

各エージェントの初期状態は、全て情報未採用者である。ここから、情報を伝播していく初期選択エージェントを選択する。初期選択エージェントを情報採用者に変更し、シミュレーションを開始する。

初期選択エージェントから情報伝播していくルールとして、「閾値ルール」と「SIR モデル」を用いる。それぞれについて、以下で説明する。

### 2.2.3.1 閾値ルール

閾値ルールは、エージェント  $i$  がリンクを張っているエージェント全てにおいて、情報採用者の割合で、意思決定をするルールである。

エージェント  $i$  が張っているリンク先のエージェントの情報採用者数を  $b_i$  とする。  $b_i$  はシミュレーション間において、一定または増加することになり、減少または  $n_i$  を超えることはない。

次に、エージェント  $i$  が持つ閾値を  $\phi_i (0 \leq \phi_i \leq 1)$  とする。  $\phi_i$  は配分によって決定される。また人口サイズ全体の平均閾値  $\phi_{avg}$  を定義する。

以上の定義から、閾値ルールにおいて、以下の条件を満たした場合、状態遷移が起きる。

$$b_i > \phi_i n_i \quad (2.2)$$

閾値ルールの状態遷移は確定的ルールに基づいており、各エージェントのリンク先エージェントの情報採用者率が閾値を越えたならば確実に状態遷移する。よって、三種類目の状態について言及しなくてもよい。

また、本シミュレーションにおけるハブ・エージェントのリンク数  $n_i$  は他と比べて大きくなるので、ハブ・エージェントが情報採用者になるには、より多くの情報採用者がリンク内にいなくてはならない。

### 2.2.3.2 SIR モデル

SIR モデルは伝染病の感染モデルでよく知られたもので、「感染可能者」、「感染者」、「免疫獲得者」からなる。本研究では、このルールを情報未採用者、情報採用者、情報非伝播者に当てはめて使用する。

ここで、情報伝播して状態遷移する確率を  $\phi_i$  とし、情報採用者から情報非伝播者に状態遷移する確率を  $r_i (0 \leq r_i \leq 1)$  とする。

情報未採用者エージェントのリンク先に情報採用者がいたならば、  $\phi_i$  の確率で情報採用者に状態遷移する。

情報採用者エージェントは、  $r_i$  の確率で情報非伝播者に状態遷移する。この際、情報が伝播したエージェント数は情報採用者と情報非伝播者を加えたものとする。

閾値ルールの  $\phi_i$  と SIR モデルの  $\phi_i$  は値は同じであるが、その意味合いは全く違う。閾値ルールが、エージェントのリンク先の情報採用者の割合に基づく確定的モデルであるのに比して、SIR モデルはエージェントのリンク先の情報採用者数に基づき、リンク内の全ての情報採用者から確率的に影響を受けるということである。つまり、閾値ルールではハブ・エージェントが情報採用者になるには、リンク先に多くの情報採用者を要したのに比して、SIR モデルの場合は、ハブ・エージェントであれば、より容易に情報採用者に状態遷移するものであり、閾値ルールのそれと相反する。

## 3. モデル作成

### 3.1. artisoc

本シミュレーションには、マルチエージェントシミュレータ artisoc を用いた。図3は、本シミュレーションのクラス図である。

### 3.2. 状態遷移図

図4に、本シミュレーションの状態遷移図を示した。その細部を説明する。

#### 3.2.1 エージェント作成

人口サイズ  $N$  に基づき、エージェントを作成する。その際、初期状態を情報未採用者と設定する。

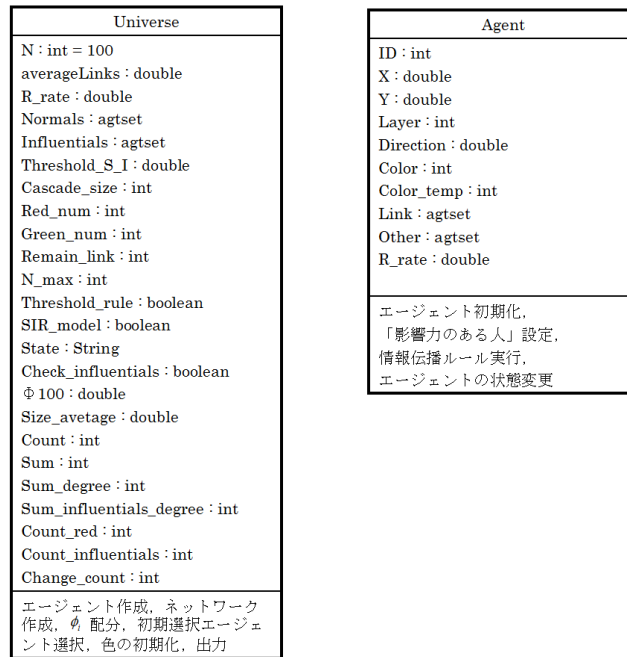


図 3 : クラス図

また, グローバル変数として  $n_{avg}$ ,  $\phi_{avg}$ ,  $r_{avg}$  も併せて設定しておく.

ここで, シミュレーション動作を速めるために, 各エージェントが持つ  $r\_rate$  を全てのエージェントで同値( $r_{avg}$ ) をとることにする.  $\phi_{avg}$  と同様に配分してもいいのだが, シミュレーションが遅くなるとともに, 結果に違いが見られないからである.

### 3.2.2 $\phi$ 配分, ネットワーク作成

シミュレーションに先立ち, 各エージェントに  $\phi_{avg}$  を配分し, ランダムネットワークを作成する.

当初  $\phi_i$  は全て 0 とする.  $\phi_{avg}$  を用いて, 全エージェントの  $\phi_i$  の総量は  $\sum \phi_i = N\phi_{avg}$  であることがわかる.  $\phi_i$  を一様分布にするため, 人口サイズ  $N$  から一体のエージェント  $i$  をランダムに抽出し, 抽出したエージェント  $i$  の  $\phi_i$  に 0.01 を加算していき, それを  $100\sum \phi_i$  回繰り返す.

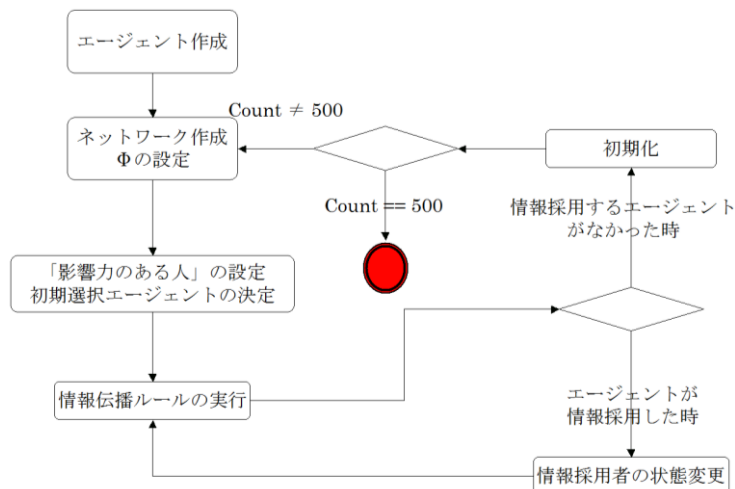


図 4 : 状態遷移図

ネットワーク作成では、平均リンク数  $n_{avg}$  からネットワークの総リンク数の2倍値  $\sum n_i = Nn_{avg}$  が導けるので、これを利用して、各エージェントにリンクを張らせる。すなわち、あるエージェント  $i$  を選択し、エージェント  $i$  とリンクを張っていないエージェント  $j$  をランダムに取り出し、相互にリンクを張る。この操作を  $\sum n_i / 2$  回繰り返すとランダムネットワークが作成できる。

### 3.2.3 ハブ・エージェントの設定、初期選択エージェントの決定

次に初期選択エージェントを決めるため、ハブ・エージェントを設定する。ハブ・エージェントは、 $n_i > qn_{max}$  を満たすエージェント  $i$  である。その割合は(2.1)の通りである。

ここから初期選択エージェントを決定する。初期選択エージェントは、ハブ・エージェントから選ぶのと、比較用にランダム選択を用いる。

### 3.2.4 情報伝播ルール、情報採用者の状態変更

「情報未採用者」、「情報採用者」、「情報非伝播者」をそれぞれ Cyan, Red, Green で表現する。

当初の各エージェントの色は Cyan になり、初期選択エージェントの色は Red になる。

また `artiscoc` はそのまま実行しても同期がとれないので、一時的な変数として各エージェントに `color_temp` という変数を持たせる。

各情報伝播ルールによって、各エージェントは自分の色を変えるかどうかを意思決定する。「情報採用者」は一時的に情報を採用し、そのステップが終わったならば、一括して色を変更する。これは各エージェント間で同期を取るためである。

### 3.2.5 初期化

各変数を初期化し、また  $\phi$  配分及びネットワーク作成から実施し、これを 500 回繰り返す。

### 3.2.6 出力

初期選択エージェントから、各情報伝播ルールに従って、情報の伝播する様子を、その情報の広がり具合をもって判断する。すなわち、情報伝播ルールのループにおいて、新規情報採用者がいなくなったならば、その際の情報伝播者数を出力する。

また、「影響力のある人」たちの平均リンク数を求め、これと併せてネットワークの最大リンク数も出力する。これは、リンク比によるカスケードサイズ比は効率的かどうかを調べるためである。

次に各ステップにおける「情報採用者」の平均リンク数を出力して、 $n_{avg}$  及び  $qn_{max}$  と比較する。

## 4. シミュレーション

情報伝播ルールは閾値ルールと SIR モデルの二通り、また初期選択エージェントに「影響力のある人」から選択するハブ・ターゲティング戦略、ランダムに選択するランダム戦略の二通り、併せて四通り実施する。 $n_{avg}$  を変化させて、異なるネットワークにおいてどのように伝播していくのかを観察する。

### 4.1. 初期条件

グローバル変数を  $N = 100$ ,  $\phi_{avg} = 0.18$ ,  $r_{avg} = 0.4$ ,  $q = 0.9$  と設定する。

### 4.2. 結果

シミュレーションを開始すると、初期選択エージェントという一つのエージェントから次々と情報が伝播していく様子が観察できる。このような現象を「カスケード現象」という。

また、シミュレーションの結果、各エージェントは Cyan, Red, Green のどれかの色を表示するが、情報採用したものは Red と Green であり、この総数を「カスケードサイズ」と呼称する。このカスケードサイズをエージェント数で割ったものが情報伝播率になる。

「カスケード現象」が一部にしか起きず、全体に情報伝播しない状態を「ローカルカスケード」、全体の情報伝播していき、情報伝播率が 1 に近づく状態を「グローバルカスケード」と表現する。

#### 4.2.1 閾値ルール

初期選択エージェントにハブ・エージェントを設定するハブ・ターゲティング戦略とランダム選択のそれぞれの情報伝播率について、平均リンク数を変化させてプロットした。

図5に現れているように、グローバルカスケードが起こるのは、平均リンク数 $n_{avg}$ の特定の範囲内に限られている。この範囲を「カスケードウィンドウ」と表現する。

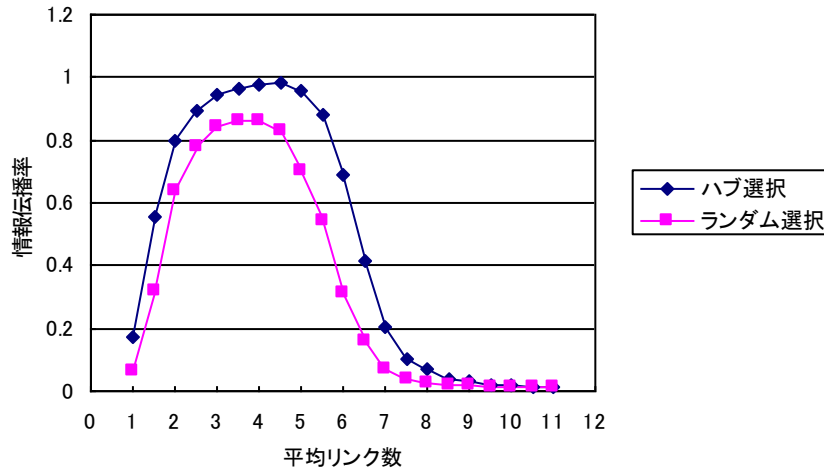


図5：閾値ルールの情報伝播率

$n_{avg}$ が大きくなるにつれ、情報伝播率が0に近づくのがわかる。これは、閾値ルールの特徴的なものであり、エージェントが持つリンク数が大きくなると、エージェントが持つ閾値を超えることが難しくなる。つまり、影響力のある人であればあるほど、情報採用者にはなりにくく、ローカルカスケードしか起こらない状態になる。

また、平均リンク数 $n_{avg}$ が1以下であると、各エージェントが情報伝播するために十分に繋がらず、グローバルカスケードは理論的に起こらない。

ここで、ハブ・ターゲティング戦略とランダム選択の双方を比べると、二つの情報を得る。

一つは、両戦略ともカスケードウィンドウが存在すること。もう一つは、そのカスケードウィンドウの平均リンク数 $n_{avg}$ の範囲は両戦略とも大きく変わらないことがあげられる。

図5から、グローバルカスケードが起こるかどうかは、初期選択エージェントがハブかランダムかという点よりも、平均リンク数 $n_{avg}$ が大きく影響していることがわかる。 $n_{avg}$ は各エージェントのリンク数 $n_i$ の平均値であり、ネットワークの構造を示す。つまり、ネットワーク構造は、それ自体がカスケードサイズに大きな影響を与えているといえる。

とするも、ハブ・ターゲティング戦略が各平均リンク数 $n_{avg}$ において、ランダム選択よりも大きいカスケードサイズを持つのは図5から明らかである。そこで、初期選択エージェントをハブ・エージェントにする効率性について言及するため、それぞれのカスケードサイズ比及びハブ・エージェントの平均リンク数 $n_{influential}$ と全体の平均リンク数 $n_{avg}$ の比を比べてみる。

図6に示された線はすべて比であるので、比較のために値1に水色の線を引いた。

まずハブ・ターゲティングとランダム選択のカスケードサイズ比を青線で示した。図6から見て取れるように、常に1以上の値をとっている。これは、情報伝播では、ハブ・ターゲティングはランダム選択より効果的であることを示す。ハブ・エージェントは他のエージェントと比べてリンク数が多く、カスケードサイズが大きくなることは想像できる。しかし、ハブ・ターゲティング戦略の基となる二段階モデルでは、ハブ・エージェントは単純に、そのリンク数の違いにより、非常に大きな役割を持つ。そこで、エージェントが持つリンク数の違いが、情報伝播にどれほどの影響を与えているのかを確認するため、ハブ・エージェン

トの平均リンク数  $n_{influential}$  と全体の平均リンク数  $n_{avg}$  の比を求める。赤線で示したものがそれぞれの平均リンク比にあたる。

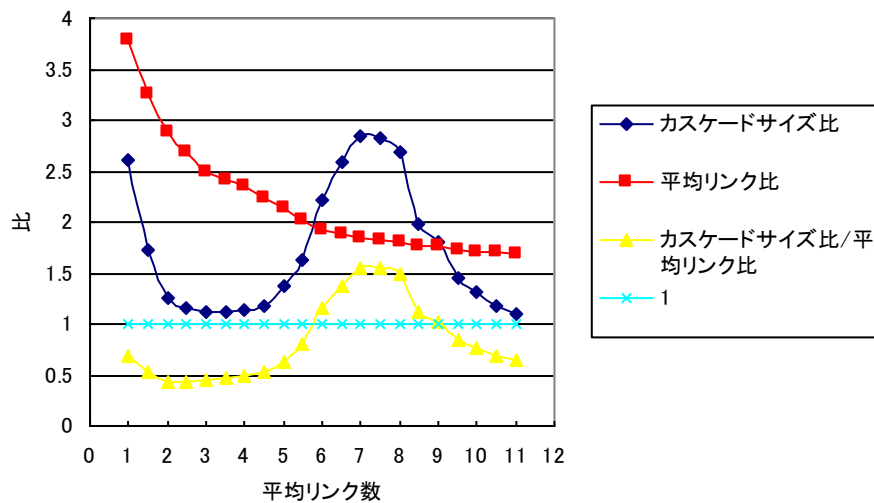


図6： $n_{avg}$ によるハブ・ターゲティング戦略の効率性

得られたカスケードサイズ比（青線）と平均リンク比（赤線）の比を取ったのが、黄線であり、これはエージェントの持つリンク数がカスケードサイズにどれほど影響を与えているかを示す。図6で確認できるように、黄線はほぼ1未満を推移している。これは、ハブ・ターゲティングがランダム選択よりも、カスケードサイズの点においてよい効果を表しているとしても、そのためにかかるコストは得られる効果よりも大きいということである。つまり、初期選択エージェントとして、ハブ・エージェントは必ずしも適切ではないと言える。

ただし、唯一  $6 \leq n_{avg} \leq 8.5$  の時に、ハブ・ターゲティング戦略は効率的であるといえる。この範囲は、先に示したカスケードウィンドウの上限値付近であることも注目すべきことである。以上のことから、ハブ・ターゲティング戦略は、閾値ルール下で、平均リンク数がカスケードウィンドウの上限値付近において、効率的であるといえることができる。

ところで、ハブ・ターゲティングが初期選択エージェントとして適切でないとしても、情報伝播の途中においては、以前その役割は大きい。二段階モデルでは、初めに情報を伝播するのはエージェントではなく、メディアである。しかし、ハブ・エージェントがいるおかげで、全体に情報を伝播することができるのである。

ここで、「早期採用者」について定義する。早期採用者とは、シミュレーション初期に情報を採用したエージェントを指す。もし二段階モデルが正しければ、シミュレーション初期の情報採用者は、ハブ・エージェントである可能性が高い。そこで、本シミュレーションにおいて、情報採用者のリンク数  $n_{informed}$  の平均を  $n_{avg}$  及び  $qn_{max}$  と比較する。ハブ・エージェントが早期採用者であるならば、シミュレーション初期の  $n_{informed}$  の値は  $qn_{max}$  に近づくはずである。

図7に  $n_{avg} = 1.5$  時、及び  $n_{avg} = 5.75$  時のそれぞれの平均リンク数の時間的変化を示す。この際、初期選択エージェントは全てのエージェントからランダムに選択するものとする。

双方とも、 $n_{informed}$  の平均値はシミュレーション初期から  $n_{avg}$  付近を示している。早期選択者であれば、シミュレーション初期において  $qn_{max}$  に近づくはずであるが、常にこの値を大きく下回っている。つまり、ハブ・エージェントは、早期採用者としての特徴を有していない

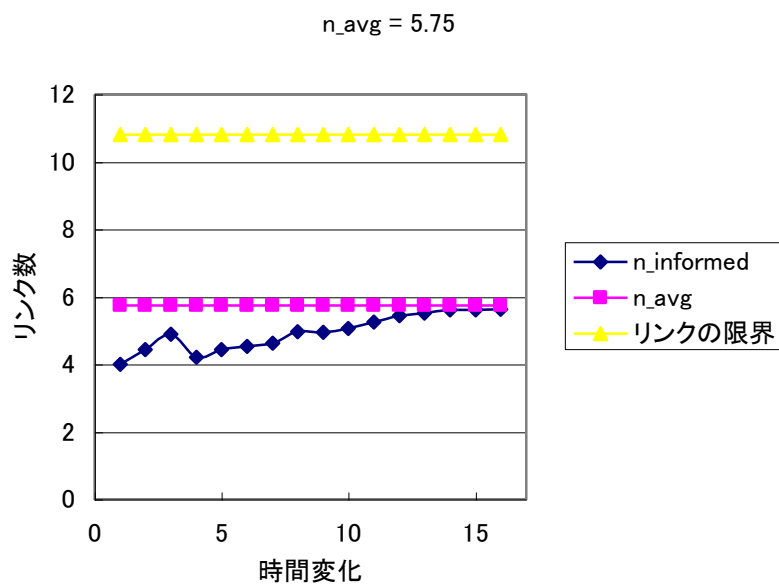
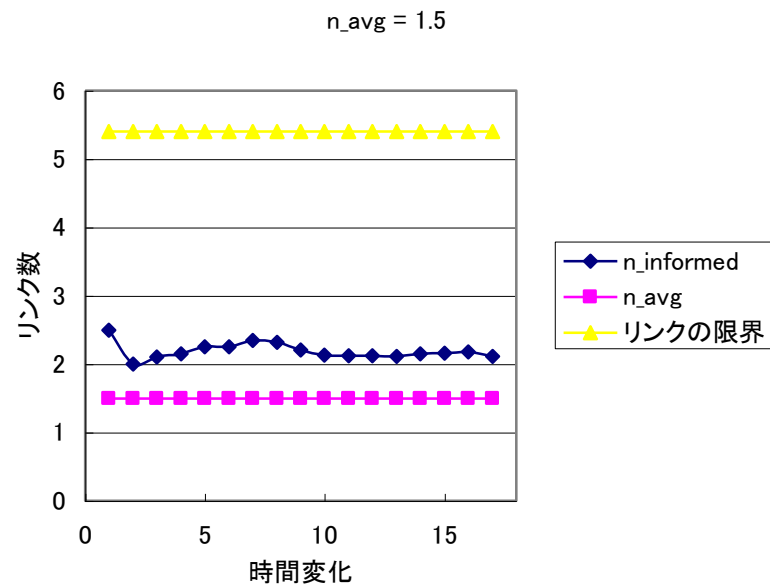


図7：情報採用者の平均リンク数

ということである。以上のことから、ハブ・エージェントは、社会ネットワークにおける情報伝播において、特定の役割を持ってはいないということがいえる。

#### 4.2.2 SIR モデル

閾値ルールについては以上である。次に SIR モデルについて同様に分析する。



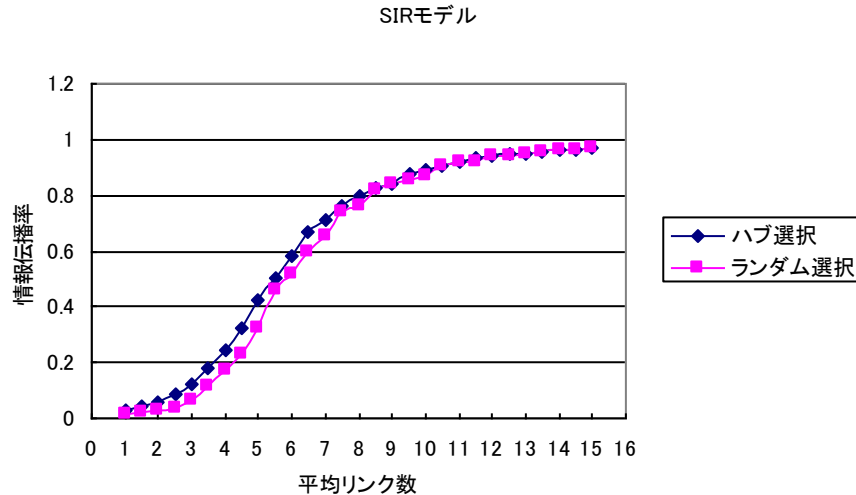


図 8 : SIR モデルの情報伝播率

図 8 で見られるカスケードウィンドウは、閾値ルールで見られたものとはかなり異なる。まず、カスケードウィンドウに上限がなく、また、ネットワークが十分に繋がっていない場合、グローバルカスケードは生起しない。図 8 から、 $n_{avg}$  が大きくなればカスケードサイズも大きくなるし、ハブ・エージェントであれば、容易に情報を採用することがわかる。

しかし、カスケードウィンドウが異なったとしても、グローバルカスケードが起こるか否かは、ほぼネットワークの構造によるものである。

図 8 から、ハブ・ターゲティング戦略は、カスケードサイズにおいて、やや効果的であると取れるが、実際に初期選択エージェントを選択することの効率性を、図 6 と同様の手法で分析する。

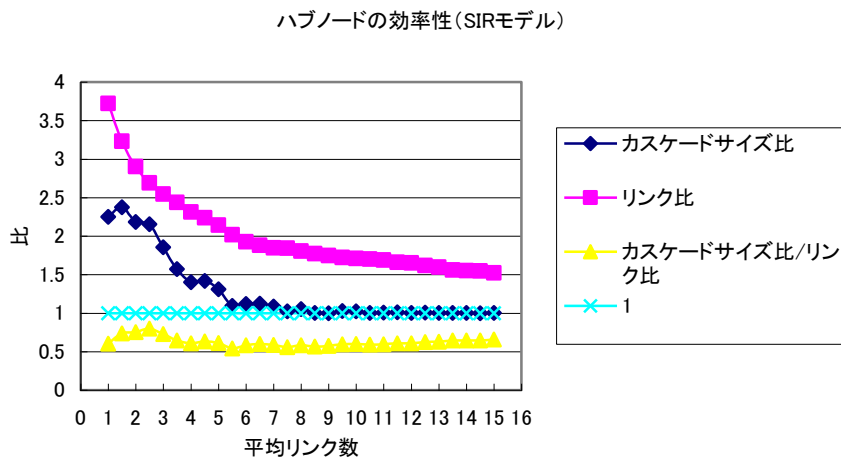


図 9 :  $n_{avg}$  によるハブ・ターゲティング戦略の効率性

SIR モデルにおいても、初期選択エージェントにハブ・エージェントを選択することでの効率性は見られなかった。図 9 の結果より、カスケードサイズの効果よりも、それに関わるリンクコストの方が上回り、黄線は常に 1 を下回った。SIR モデルにおけるハブ・ターゲティング戦略は、閾値ルールにおけるハブ・ターゲティング戦略よりも効率性が劣るものであるといえる。これはカスケードウィンドウの上限がないことに由来するものと考ええる。

次に図 7 と同様に、SIR モデルにおけるハブ・エージェントが、早期採用者であるかどうか調べる。

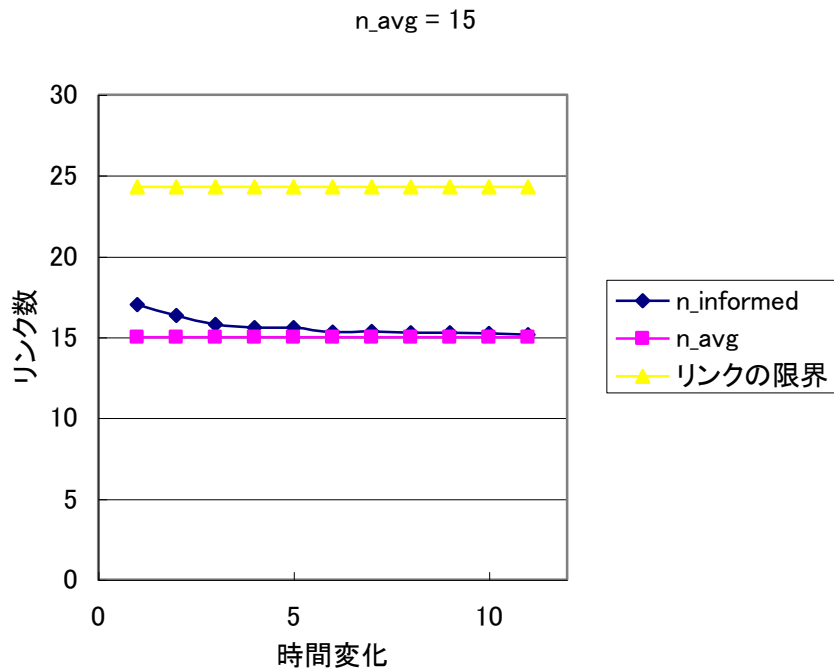
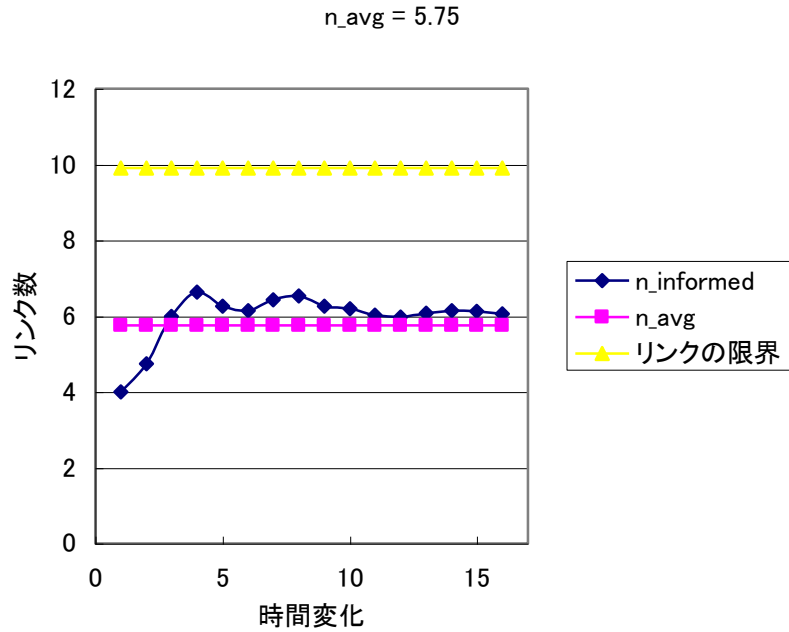


図 1 0 : 情報採用者の平均リンク数

情報採用者の平均リンク数の時間的変化を図 1 0 に示した。平均リンク数  $n_{avg}$  はそれぞれ、5.75 と 15 を用いた。閾値ルールと同じ値を用いない理由は、カスケードウィンドウが違うことと、 $n_{avg} = 1.5$  では、SIR モデルにおいて、グローバルカスケードが生起しないからである。

図 1 0 の二つのグラフから、閾値ルールと同様に、情報採用者の中で、ハブ・エージェン트가多くを占めていないことを示している。

SIR モデルにおいてもハブ・エージェン트는早期採用者としての特徴を有していないことがわかる。

## 5 結論

ランダムネットワーク上で、情報伝播の過程をシミュレーションすることで、マーケティング戦略における、ハブ・ターゲティングの効率性について分析し、三つの知見を得た。

一つは、マーケティング戦略の結果は、概ねそのネットワークの構造自体、特に、エージェント相互の関係の強さに影響を受けるということである。

二つ目は、ハブ・ターゲティング戦略は、ハブ・エージェントとそれ以外のエージェントのリンク比と比べて、効率的ではないということがわかった。唯一効率的であるのは、閾値ルールにおいて、平均リンク数  $n_{avg}$  がカスケードウィンドウ上限値付近にあるときのみである。

そして三つ目は、情報伝播における二段階モデルは、グローバルカスケードを導く資となることである。

本研究は、ランダムネットワークという限定されたネットワークでのみのシミュレーションであったが、現実社会にはスケールフリーネットワークやスモールワールドネットワーク等があり、それらについてシミュレーションすると、また新しい知見が得られるだろう。

また、本シミュレーションの情報伝播ルールは閾値ルールと SIR モデルのみを使用したがる、これだけが情報伝播ルールであることはないので、その他のルールについても言及する価値がある。

## 参考文献

- [1] Duncan J. Watts, Peter Sheridan Dodds.(2007), “Influentials, Networks, and Public Opinion Formation”, *Journal of consumer research Inc.* Vol. 34
- [2] 岡本隆(2002)「個人間の地理的要因による製品普及特性」
- [3] 金春姫(2005)「消費者相互作用からブランド態度の社会的共有へのダイナミックなプロセス」
- [4] 電通消費者研究センター(2007)「互いに共振する消費者『鏡衆』が台頭」
- [5] 山影進(2007)「人口社会構築指南 artisoc によるマルチエージェント・シミュレーション入門」書籍工房早山