

Peer-to-Peer コミュニケーションにおける 情報流通ネットワークの分析

松山 科子, 國上 真章, 寺野 隆雄

筑波大学院 ビジネス科学研究科

本論文では、エージェントベースシミュレーションで行った、Peer-to-Peer でのコミュニケーションにおける情報流通のネットワークの分析について述べる。シミュレーションにおいては、パラメータとなるべき事項をルール記述ファイルとしてシミュレータに与える事でパラメータとその影響を明確に知ることができる仕組みを構築した。

エージェントシミュレーションでは、エージェントと流通するコンテンツの性質を定義した後、コミュニケーションを行う場としての社会を定義しその中で各エージェントが情報流通を行う。情報流通のネットワーク分析は主にスモールワールド性、スケールフリー性について行い、スモールワールドの性質を満たすこと、カットオフが存在するスケールフリーの性質を持つことが示された。

Network analysis for Peer-to-Peer communications by agent-based simulation

Shinako Matsuyama, Masaaki Kunigami and Takao Terano
Graduate School of Systems Management, The University of Tsukuba

In this paper, we describe the results for analyzing the network on Peer-to-Peer communications by agent-based simulation. We propose a system, which can easily know the effects of parameter for the simulation. In this simulation, we define the characteristic of the agents, contents and then define the society, which is the place for communication. We mainly analyze the characteristic of the small world structure and scale-free one, which have recently been discussed actively. Through the simulation, we have obtained the result that this network is Small world and Scale-free with some conditions.

1 はじめに

昨今、Blog(Weblog), FOAF, Orkut といったソーシャルネットワークサービスの出現により Peer-to-Peer コミュニケーションが注目されている。また、携帯電話やウェアラブルコンピュータの発達により個人が蓄積したデータの個人間での流通も今後増えると予測される。情報提供者にとっては、このようなコミュニケーションにおける情報流通の仕組みを知ることは重要である。

Peer-to-Peer コミュニケーションにおける情報流通の研究としては、(竹内),(吉田 01)などがある。(竹内)では、情報に対する評価をユーザ間の関係を元に算出しそれに基づき情報流通を行う情報伝播モデルを提案している。(吉田 01)では口コミネットワークを電子コミュニティ上に実現しそこの情報伝播の性質を分析している。

一方で、昨今スモールワールドやスケールフリーといったネットワークの性質の研究も盛んであり、ヒューマンネットワークに関しても共演俳優のリンク、交友関係などいくつかの検証がなされ

ている((Amaral 00) ,(Barabasi 99))。

本研究では、エージェントシミュレーションにより Peer-to-Peer コミュニケーションでの情報流通の性質を分析し、スモールワールド、スケールフリーの性質を調べた。

一般にエージェントシミュレーションの問題点として、モデル内で設定しているルールの結果における効果を分析しにくい点が挙げられる。

(倉橋01)では、人工社会を形成しそこのコミュニケーションを分析している。先天的属性により社会を形成しその後の後天的属性でコミュニケーションを行っていくものである。先天的属性からの社会形成段階においては、逆シミュレーションという手法を提案しており、得たい結果を得るためのパラメータがどのような性質があるかを分析するアプローチを取っている。それにより上記で挙げた問題点の一点目を解決しているが、後天的属性におけるプログラム内部で設定しているルールに関しては依然、結果におけるその効果の分析がしにくいという問題が残されている。

本研究では、パラメータとして設定したルール

の効用をシミュレータに対して外部から調整可能な構造のシミュレーションソフトを開発した。

本論文の構成は、2章でエージェントシミュレーションソフトウェアの構成について述べる。3章で行ったシミュレーションの概要、4章で実験環境と実験結果を述べる。5章で考察を行い、6章でまとめる。

2 シミュレーションソフトの構成

2.1 ソフトウェアの構成

本ソフトウェアは、(寺野 03)「シミュレーション実験に対する要請」にて述べられている、プログラムの正当性検証、研究者の責任範囲の明確化(責任範囲外のもので利用している事項の明確化)を満たすことを目的とする。具体的には、エージェントが用いるすべての意思決定ルールをルール記述ファイルに記述する手法を採った。一般にルールはさまざまな社会現象を参照して決定されるが、その参照した参考文献をhtmlやW3C RDF(Resource Description Framework)などの形式的な形で表現し組み込むことも可能である。

図1に、ルール記述ファイルの記述例を示す。各意思決定項目毎にユニークな識別子を付け、その中のルールは、識別子:項目の識別子+ユニークな番号、ルールの説明、ウェイトを記述する。項目毎のウェイトの合計は常に1になる必要がある。

また、ランダムな事項を記述することが可能であり、本項目を利用することで検証者はランダムな事象とある一定のルール下での意思決定での事象の現象を比較することができる。

```
#####
# send/receive
#####
# Send の From
E: SendFromIndicator
E1 [客]短期寿命のコンテンツが多い      0.5
E2 [客]送受信回数が多い Agent          0.5
E3 Random                                  0.0

# Send コンテンツ
F: SendConentIndicator
F1 短期寿命のコンテンツ                  0.5
F2 自分の評価値(my_eval 値)が高い        0.5
F3 Random                                  0.0
```

図1 ルール記述ファイルの例

本仕組みを用いることで、シミュレーションソフトの開発者以外のモデルに対する検証者は、社会現象を分析した文献、ルール記述ファイル、実験結果を以って検証が可能になる。

2.2 ソフトウェアの機能

本ソフトウェアのパースは以下を行う。

- 1) ルール記述ファイルの各意思決定項目毎のルールのウェイトの合計値のチェックを行う
- 2) 意思決定毎のルール関数とウェイト取得関数をソースとして生成する
以下に例を示す。

```
// 意思決定毎の Rule 関数
double Rule_A1() {
    return(Rule_A1()*get_weight("Rule_A1")
           +Rule_A2()*get_weight("Rule_A2")+...+Rule_An()*get_weight("Rule_An"));
}
// weight 取得関数
double get_weight(String rule_s) {
    // return to the weight value;
}
// 各ルール関数フレーム
double Rule_A1() {
    // should be filled by programmer;
}
```

2.3 ソフトウェアの利用

図2に本ソフトウェアを用いたシミュレーションの流れを示す。

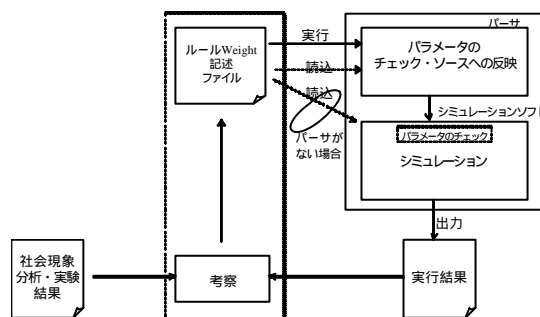


図2 シミュレーションソフト利用の流れ

以下に一連のシミュレーションにおける役割毎の業務をまとめる。役割であるのでそれぞれは同一人物が行っても良い。また本ソフトウェアでのルール記述ファイル利用の規則に沿っていれば本パーサを利用せずプログラマがパーサが生成するソースを記述しても本ソフトウェアを使ったシミュレーションは可能である。

- 1) 設計者
意思決定項目とそのルールを決定
ルール記述ファイルを生成

2) 開発者

ソフトウェアのパーサが生成した各ルール関数の実体を記述する

```
// 各ルール関数フレーム
```

```
double Rule_A1() {  
    /* 記述 */  
}
```

意思決定項目の利用側のソースを記述する

```
Func() {  
    double value_Rule_A = Rule_A(); ...  
}
```

本処理はプログラマに任されるが、重要な点は Rule_A 関数の結果を用いてエージェントあるいはコンテンツを選択するようなコードにすることである。ソースをコンパイル・リンクしシミュレーションソフトウェアを生成

3) 評価者

シミュレーションを実行し、結果に対する分析を行う。

必要に応じて 1) に意思決定項目の変更のフィードバックを行う。

3 シミュレーション概要

3.1 全体構成

Peer-to-Peer の対象をエージェントとしその性質をいくつか定義した。また、情報流通されるコンテンツにもいくつかのカテゴリを導入し性質を定義した。そしてエージェントが存在する社会（エージェント間の関係）を定義し、一定期間情報通信をさせた。

3.1.1 エージェントの定義

各エージェントは情報流通の活発度として3段階のレベルを持つ。各エージェント間の関係は以下で述べる社会の定義により決まる。初期時の隣接の概念を実現するためエージェントはユニークな順序付けられた ID を持つ。

3.1.2 コンテンツの定義

各コンテンツは1つのカテゴリに属する。カテゴリには長期のものと短期のものが存在する。エージェントはランダムに選択したいいくつかのカテゴリを「興味コンテンツリスト」として持つ。(石井 02) によると、コンテンツは各人にとっての長期、短期にも分類できるため、ここでも長期コンテンツに対してはエージェント毎に長期、短期

をランダムに設定している。短期コンテンツはシミュレーション期間でランダムに消失する。1つのカテゴリが消失した場合は新しいカテゴリが生成されエージェント間で利用される。

また、各コンテンツグループ毎にそのグループのコンテンツを平均数以上送受信しているエージェントを「専門家リスト」に登録する。

3.1.3 社会の定義

以下3つの社会を定義した。この社会はエージェントが情報流通を行う相手を選択する範囲の一つの指標である（基本的にはその範囲で条件に合うエージェントを検索する）。

- ALL：皆が知り合いの状態（誰に対してもメッセージの送受信ができる環境）
- SUBGROUP：いくつかかにグループ化されており初期状態ではエージェントはグループ内のエージェントに通信を行う
- NONE：初期状態では知り合いが居ない状態（一人ずつ知り合いになっていく環境）この場合初回は隣接したエージェントが通信相手として選択される

(倉橋 01) によると、リーダーが存在する社会、階層型、二人だけが会話するなど存在するが、実社会の Peer-to-Peer コミュニケーションに注目して考えた場合、以上の3つがメインとして挙げられるため、今回は3つを採用した。

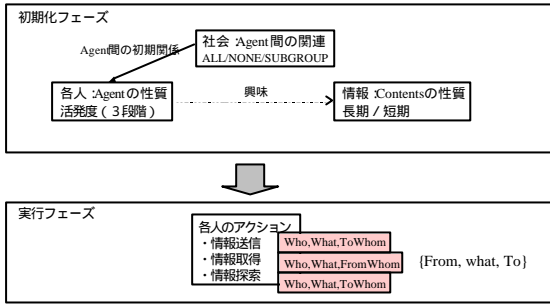
エージェントは、社会に応じて他のエージェントに対する知り合いリストを持つ。

3.1.4 トランザクション

各人（エージェント）のアクション（トランザクション）は以下の3つから成る。

- 送信：自ら自分の持っている情報を特定の相手に送信する
- 取得要求：自分が知っている人に情報取得要求を出し情報をもらう
- 問い合わせ：自分が知っている人に知人を紹介してもらい、その人に情報取得要求を出し情報をもらう (NS98)

図 3 にシミュレーションの流れを示す。



$$\text{アクション } A_i = \{I_j\}$$

$$\text{指標 } I_j = \text{Rule}_{jk} \times W_{jk} \quad (W_{jk}=1)$$

図 3 シミュレーションフロー

各エージェントは、情報を受け取ると情報に対する自分の評価値を計算し保持する。その情報を別エージェントに送付する際には評価値と共に送付する。図 4 にエージェントを中心として見たエージェントのアクションを示す。

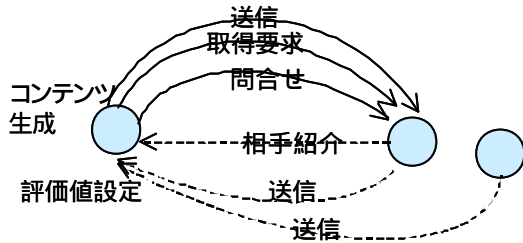


図 4 エージェントのアクション

3.1.5 意思決定に対するルール

3.1.4 で述べた各トランザクションは送信者 (From)、送信コンテンツ (What)、受信者 (To) の組み合わせで表現される。送信者、送信コンテンツ、受信者各々を決定するための指標を以下に示す。基本的には、各エージェントは知り合いの中からコンテンツの評価の類似度をベースに通信相手を決定する。

(1) 送信

From	客観評価) 短期寿命のコンテンツが多い 客観評価) 送受信回数が多い
What	短期寿命のコンテンツ 自分の評価値が高い
To	知り合いエージェントが存在しない場合 隣接エージェントの中で以下の指標値が高いもの 隣接エージェントの中で活発なもの 知り合いエージェントが存在する場合 主観評価) エージェントの興味リストに対象コンテンツのカテゴリがある

	客観評価) グループ内コンテンツの送受信回数が多い
--	---------------------------

(2) 取得

From	客観評価) 活発なエージェント 客観評価) 生成・送受信コンテンツが多いエージェント
What	ランダムに抽出
To	知り合いエージェントが存在しない場合 隣接エージェントの中で以下の指標値が高いもの 隣接エージェントの中で活発なもの 知り合いエージェントが存在する場合 主観評価) グループ内コンテンツの評価の類似度が高いエージェント 客観評価) グループ内コンテンツの送受信回数が多いエージェント 客観評価) そのコンテンツグループの専門家であるエージェント

(3) 検索 (問い合せ)

From	活発なエージェント 生成・送受信コンテンツが多いエージェント
What	ランダムに抽出
To	知り合いエージェントが存在しない場合 隣接エージェントの中で活発なエージェント ランダムに活発なエージェントを選択 知り合いエージェントが存在する場合 知り合いエージェントに対して(2)の取得の To の指標で取得の対象を決定

(4) その他

コンテンツの評価

基準	コンテンツ自身の評価 流通回数 (cnt 値) が高い
	送信者の評価 送信者がそのコンテンツのグループの専門家である
	送信者が同一グループのコンテンツを発信している率
	送信者が同一グループのコンテンツを受信している率
	送信者が生成したコンテンツの流通回数の合計

4 実験環境と結果

4.1 実験環境

以下の環境で実験を行った。

エージェント数：150/300/600/1200

1日当たりのトランザクション数：エージェント数×7¹

コンテンツ種類数：6

総コンテンツ数：約3000

社会タイプ：NONE

実行回数：10日分

意思決定ルール：3.1.5節で述べた各ルールを均等ウェイトで指定した

N: エージェント数、b: 1エージェント当たり、一日のトランザクション生成率、 λ : 1エッジ当たり、一日のエッジ消滅確率、 $E(t)$: 時刻tにおける総エッジ数とすると成長曲線は以下式1で定義できる。今回のシミュレーションでは、エージェント数300に対し、 $E(t)_t$ は約600であるため、b: 1エージェント当たり、一日のトランザクション生成率を1とした場合、 λ は約0.5となる。よって、一日当たりエッジを0.5の確率でランダムに消滅させた。

$$\frac{dE(t)}{dt} \cong b \cdot N \cdot \left(1 - \frac{E(t)}{N \cdot C_2}\right) - \lambda \cdot E(t)$$

式 1 エッジの成長曲線

4.2 実験結果

4.1で述べたトランザクション数毎に総エッジ数 $E(t)$ 、スモールワールドネットワークでの特徴量として使われるクラスタリング係数C、パス長 L (Watts 98)、エッジ数の分布を求めた。表1に $E(t)$ 、C、Lを、図5にエッジ数の分布をまとめる。図中Agの数字はエージェント数、linear-logはエッジの累積分布の片側対数値、log-logは同両側対数値を示す。

表 1 実験結果(C,L 値)

エージェント数	総エッジ数(E)	クラスタリング係数(C)	パス長(L)
150	303	0.24	3.2
300	637	0.26	3.4
600	1385	0.21	3.8
1200	2830	0.21	3.8

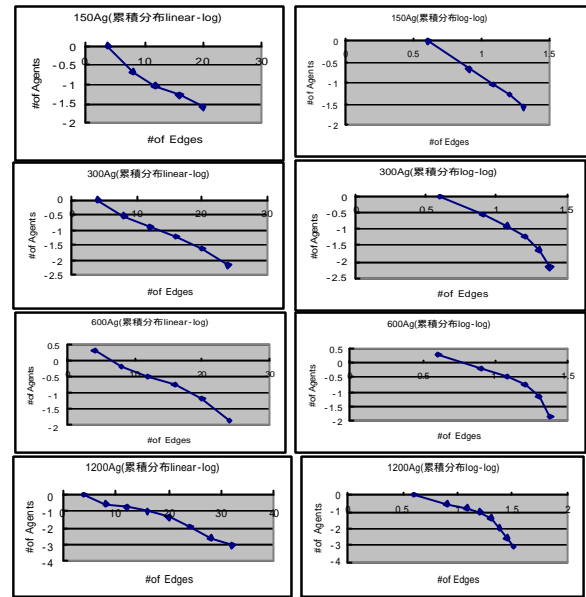


図 5 エッジ数の分布

5 考察

5.1 スモールワールド性

4.2で得た結果に対し、表2に同一程度のエッジを持つランダムネットワークでの値、スモールワールドであるとされているグラフでの値(Caenorhabditis elegansの線虫のニューロンのネットワーク(Watts 98))との比較を示す。

表 2 リンク特性

L	Random	3.9
	C.elegans	2.65
	本実験値	3.4
C	Random	0.06
	C.elegans	0.28
	本実験値	0.26

この結果より、本ネットワークはノード数 $n=300$ に対しエッジの最大数 $k_{max}=24$ であり、 $n \gg k_{max} \gg 1$ であり、ランダムネットワーク間で $L_{rand} \sim L$ かつ $C \gg C_{rand}$ となりスモールワールドの性質

¹ エージェント数に対するトランザクション数は、(イブシ 03)での携帯メールの一日の平均利用回数の値を参考に決定した(メール7回/1日)

を満たすものと言える。これは本シミュレーションでのトランザクションはコンテンツの評価値の類似度をベースに知り合いでの通信を行って行くものであるため、知り合いでのリンクは高くなっていると言える。

5.2 スケールフリー性

スケールフリーの性質については、図 5 よりエッジ数 16 までの範囲では、 $P(k) \sim k^{-1.5}$ < 2.1 となっており、カットオフが存在するスケールフリーネットワークであることがわかる。

このようなカットオフが存在する理由として文献(Amaral 00)では、頂点の年齢とリンクへの追加のコストを挙げているが、本シミュレーションの場合、コンテンツの評価に応じた情報通信をしており、ある一定期間後にはエージェントあるいは興味を持つコンテンツの種類が変化しない限りエッジ数の増加に上限ができると考える。よって今回の場合は(Amaral 00)における Slow aging の性質を持つ Broad Scale ネットワークであると言える。

5.2.1 カットオフ位置に関する仮説

本実験においては通信対象エージェントの選択基準として初期時はエージェントの活発度、保持するコンテンツの種別そしてコンテンツに対する評価の類似度としている。従って、初期時に選択される活発なエージェントに対して選択し得る対象エージェント(エッジとなり得るエージェント)の最大値は以下で予測できる。

$$\text{総エッジ数予測値} = \frac{\text{総トランザクション数}}{\text{活発なエージェント率}} \times \text{コンテンツ種別一致率} \times \text{評価値一致率}$$

式 2 エッジ数予測式

今回の実験値では、総トランザクション数/一日=7, 活発なエージェント率=1/4, コンテンツ種別一致確率=約 0.58, 評価値一致確率=約 0.5 であったため、総エッジ数予測値は約 8 となる。また 4.1 節で述べたとおり今回は 0.5 の確率でエッジを消滅させている。よって $8/0.5 = 16$ がカットオフ位置になるという仮説を立てられる。

5.2.2 カットオフ位置に関する検証

5.2.1 節の仮説を検証するため、トランザクション数を倍に変化させたケース、そしてそのトランザクション数においてコンテンツ種別の一致度を 1/2 にしたケースでの実験を行った。図 6 にそ

れぞれの結果を示す。

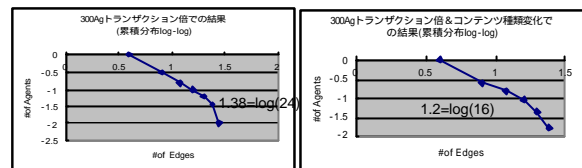


図 6 トランザクション数・コンテンツ種別変化結果

総エッジ数を一定に保ちつつトランザクション数を倍にすると、対応するエッジの消滅率は 0.5 から 0.34 となる。それを考慮すると、カットオフエッジの位置は、式 2 により「24」となる。またコンテンツ種別一致率を 1/2 (約 0.29) にすると、再度「16」になる。実験結果ではこれが示されており、5.2.1 節の仮説が裏付けられる。したがって、本ネットワークはカットオフを持つスケールフリーネットワークであると言える。

6 まとめと今後の課題

本論文では、エージェントベースシミュレーションにより Peer-to-Peer でのコミュニケーションの情報流通の分析を行った。その際、シミュレーションにおいてパラメータとなるべき事項をルール記述ファイルとしてシミュレータに与える事でパラメータとその影響を明確に知る仕組みを構築した。

情報流通のネットワーク分析においては、コンテンツの性質を組み込みその影響をも反映した上で情報の流通性を分析した。

初期状態では誰も知り合いが居ない状態の社会(NONE)での情報通信により形成されるネットワークに関して、スモールワールドの性質を満たすこと、カットオフを持つスケールフリーの性質を持つ事を示した。エージェントシミュレーションでこのような分析を行うことで、エージェントに対してどういう場(社会)を提供し、どういうエージェントに対してどういう性質のコンテンツを流通させるのがネットワークの発達に導くのかといった検証が可能になる。

今後の課題としては、他の社会の性質(全員が初期状態から知り合いである ALL、サブグループ化された SUBGROUP)によるネットワークの性質の差を分析する点、エージェント数の増減、興味を持つコンテンツの種類の変化による影響を見る点がある。また、実世界ネットワークやネットワーク上のバーチャルな世界でのデータによる実験を行いその結果との比較も行う予定である。

参考文献

- (吉田 01) 吉田匡志, “口コミによる分散型情報収集システム”, MACC2001
- (竹内) 竹内亨他, “ユーザの関連性を用いた情報伝播モデルの評価実験”
- (青木 02) アルバート・ラズロ・バラバシ著、青木薫訳「新ネットワーク思考」2002年
- (寺野 03) 寺野隆雄, “エージェントベースモデリング: KISS 原理を超えて”, 人工知能学会誌 18 巻 6 号 2003 年 11 月
- (倉橋 01) 倉橋節也, “人工社会 ネットワーク社会におけるリーダーシップと共同分配規範”, 2001
- (石井 02) 石井淳蔵・厚美尚武編「インタネット社会のマーケティング」2002年
- (NS98) N.S. Contractor, “IKNOW: A Tool to Assist and Study the Creation, Maintenance, and Dissolution of Knowledge Networks”, LNCS 1519, pp.201-217, 1998.
- (イプシ 03) 株式会社イプシ・マーケティング研究所, 「第3回コンシューマレポート 携帯電話の利用に関する調査結果」2003年,
http://www.ipse-m.com/report_csmr/report_c3/report_c3_su_m.htm
- (Watts 98) D.J.Watts and S.H.Strongatz, “Collective dynamics of ‘small world’ networks”, Nature, Vol.393, pp.440-442 1998.
- (Amaral 00) Amaral, L. A. N., Scala, A., Barthélemy, M., and Stanley, H. E.: Classes of small-world networks, Proceedings of the National Academy of Sciences, Vol. 97, No. 21, 2000
- (Barabasi 99) A.L.Barabasi, R.Albert, H.Jeong, “Mean-field theory for scale-free random networks”, arXiv:cond-mat/9907068 v1 5 Jul 1999