

多重グループ所属をノード状態とする投票者モデルのグループサイズ分布

Group Size Distributions of the Voter Model with Multiple Group Affiliations as the Node States

石川 孝
Takashi Ishikawa

日本工業大学
Nippon Institute of Technology

Social media with the group functionality exhibits the power law distribution of the group sizes. Pollner et al reproduced the power law distribution using a group size preferential attachment model for growing networks, however for the case of constant node size the mechanism of keeping the group size distribution is not known. The paper shows the emergence of the power law distribution of the group sizes as the equilibrium state of constant size networks using a voter model with multiple group affiliations as the node states.

1. はじめに

さまざまなソーシャルネットワークにおいて内部が外部とよりも密に結合するコミュニティ構造がどのようなメカニズムによって生じるかを明らかにすることは、複雑ネットワークの重要な研究課題の1つである。これまでの研究において提案されたいくつかのモデルによって、コミュニティ構造が一定の条件の下で創発することが明らかにされてきた[石川 2014]。しかし、実世界のソーシャルネットワークのコミュニティ構造がもつ統計的な性質に対する理論的な解明はまだ研究途上にある。例えば、グループ機能をもつソーシャルメディアではグループサイズがべき乗分布することが知られているが、この統計的な性質を明解に説明するネットワークダイナミクスモデルはまだ提案されていない。そこで本研究は、ソーシャルネットワークにおいて所属関係が明示的なグループのサイズがべき乗分布するコミュニティ構造を創発するネットワークダイナミクスモデルを構成することを目的とする。

本論文はつぎのように構成されている。第2章は関連研究の要点を整理し、その理解を基に第3章で本論文が提案する多重グループ所属モデル(MGAM)について述べる。そして第4章で提案モデルによるシミュレーションの結果について述べ、第5章で本論文の結論を述べる。

2. 関連研究

2.1 ソーシャルメディアのグループサイズ分布

グループ機能(ユーザが特定のグループに所属してメンバー同士でコミュニケーションする)をもつ YouTube や Facebook のようなソーシャルメディアでは、所属関係が明示的なコミュニティに当たるグループのサイズ(所属するユーザ数)がべき乗分布することが知られている。

[Zheleva 2009]は、ユーザが関心のグループにつながり、グループがそのメンバーにつながる2モードの所属ネットワークが共存している Flickr, LiveJournal, および YouTube からの大きなデータセットを使って、それらのネットワークにおけるグループサイズがべき乗分布することを明らかにした。また、これらのネットワークでは、多数のシングルトン(グループ内に友人をもたな

いメンバー)が存在し、特定の次数のユーザに対するグループ所属数が指数分布することも明らかにした。この論文は、ソーシャルネットワークと所属ネットワークの相互発展を捉えるモデルを提案して、実世界のネットワークがもつこれらの統計的な性質を説明した。このモデルは、ユーザがソーシャルネットワークに参加する確率と関心グループに参加する確率を2種類のパラメータとするミクروسコーピックなネットワーク発展モデルである。ユーザは一定の関数に従う確率によってソーシャルネットワークに参加し、各ユーザはその寿命の間に他のユーザとの友人リンクを複数回形成する。さらに、ユーザはあるグループに所属する友人の数に比例した確率でそのグループに参加する。このネットワークダイナミクスモデルは、ネットワーク成長とグループ参加が混在しており、実世界のネットワークがもつ統計的な性質がどのようなメカニズムによって生じるかを明解には説明していない。

[Ferrara 2012]は、Facebook のユーザが作る大規模なオンラインソーシャルネットワークからサンプルを収集し、そのコミュニティ構造を既存のアルゴリズムによって抽出して、コミュニティ構造の統計的な性質とそのソーシャルダイナミクスを解析した。その結果、Facebook のコミュニティ構造におけるコミュニティサイズがべき乗分布すること、つまり多数の小さいグループと少数の大きいグループを形成するユーザの傾向を明らかにした。この結果は2つの異なるコミュニティ抽出アルゴリズムに共通しており、所属関係が明示的でないコミュニティ構造においてもグループサイズが一般的にべき乗分布することを示唆する。

2.2 グループサイズ優先選択モデル

[Pollner 2006]は、コミュニティをノードとし、コミュニティ間の共有メンバーによる結合をリンクとするコミュニティグラフに対するネットワーク成長モデルを使って、そのコミュニティサイズ(メンバー数)がべき乗分布することを示した。このモデルは、基礎となるノードのネットワークに対する次数優先選択による成長モデルとのアナロジーで、新しくコミュニティが作られるときにそのリンクがコミュニティサイズに比例した確率で形成されることを仮定する。この論文は、このグループサイズ優先選択モデルによって、Los Alamos cond-mat e-print archive の共著ネットワークから k -クリークパーコレーション法によって抽出された重なりのあるコミュニティについて、そのサイズがべき乗分布することを再現した。しかし、このモデルは、基礎となるノードのネットワークがどのように形成されるかについて何も説明していない。

連絡先: 石川 孝, 日本工業大学 情報工学科, 〒345-8501
埼玉県南埼玉郡宮代町学園台 4-1, tisikawa@nit.ac.jp

2.3 適応的投票者モデル

エージェント集団の意見形成に対する投票者モデル (voter model) において、エージェントの相互作用プロセスと結合したリンクの適応メカニズムによって、エージェントのネットワーク構造が時間的に変化するモデルは、適応的投票者モデル (adaptive voter model) または相互発展投票者モデル (co-evolving voter model) と呼ばれる [Vazquez 2008]。このモデルは N エージェントのシステムを考え、各エージェントは変化する 1 つの意見 a (A または B) をもつ。エージェント間の相互作用は全体で L 個のリンクによって特徴付けられ、各エージェントに対しては平均次数 $\langle k \rangle = 2L/N$ に対応する。最初、各エージェントは意見 a と $\langle k \rangle$ 個の隣人 (相互作用の相手) がランダムに与えられ、初期ネットワークは一様次数 $\langle k \rangle$ のランダムネットワークである。モデルの時間発展は、つぎのダイナミクスに従う。

- 各時間ステップ t において、1 つのエージェント i とその隣人である 1 つのエージェント j がランダムに選ばれる。
- もしエージェント i と j が同じ意見をもてば、エージェント i はその意見を維持して何もしない。
- もしエージェント i と j が違う意見をもてば、エージェント i は確率 p で j とのリンクを切り、システム全体の非隣人の中からランダムに選んだ同じ意見のエージェント k にリンクをつなぎ替える。あるいは、確率 $1 - p$ で隣人 j の意見を取り入れて意見を変更する。

このモデルでは、隣人を変更するリンクつなぎ替えの動作と意見を変更する動作の両方が隣人エージェントの対が同じ意見をもつように作用する。リンクつなぎ替えは逆の意見のエージェントをつながりのないグループに分離しようとし、意見の変更はつながったグループ内での意見一致の調整を表している。これらのメカニズムによって適応的投票者モデルは、その時間発展の結果として、確率 p の値によって分離状態 (p が大きいとき) または意見一致状態 (p が小さいとき) への吸収転移を示す。

[Ji 2013] は、適応的投票者モデルの従来の解析的扱いを改良して、ノードの次数相関を実ネットワークのデータから経験的に近似することによって、吸収転移についてシミュレーション結果との良い一致を得た。この結果は、従来の平均場近似がなぜ失敗するかを明らかにし、相互発展プロセスのダイナミクスに対する発展方程式において相関をどのように解析するかについての示唆を与える。

[Silk 2014] は、適応的投票者モデルの解析的扱いにおいて、常微分方程式系 (ODE) による近似における不均質モーメント展開、ODE を 2 次元偏微分方程式 (PDE) に写像する生成関数、および PDE 理論のツールによる解法を提案し、シミュレーション結果との良い一致を得た。この論文のアプローチは、ネットワーク科学と PDE 理論とのつながりを確立し、離散的なノード状態をもつネットワークのダイナミクスに広く応用できる。

3. 多重グループ所属モデル

本論文で提案する多重グループ所属モデル (Multiple Group Affiliations Model; MGAM) は、既存の投票者モデルのノード状態が 2 値であるのに対して、ノード状態がノードの所属する複数のグループの集合であること (多重グループ所属という) が特徴である。

3.1 ネットワーク

MGAM のネットワークは N 個のノード (エージェント) からなり、各ノードは固有の状態をもつ。MGAM のノード状態は有限個 a_i (i はノード番号) のグループからなる所属グループ集合 g_i と

する。所属グループとは、あるノードが他の特定のノードと相互作用する場としての有限個 G のグループ集合の要素である。また、各ノード i の所属グループ数 a_i は一定またはある確率分布に従うものとする。NGAM のネットワークにおけるノード間のリンクはノード同士の相互作用を表し、2 つのノードが少なくとも 1 つの同じグループに所属するとき、それらのノード間にリンクが存在すると考える。従来の適応的投票者モデルではリンクがノード状態とは独立に明示的に設定されているのに対して、NGAM ではリンクがノード状態によって定まり、ノード状態だけによってネットワークダイナミクスを探索できる。

3.2 ダイナミクス

NGAM のネットワークは、ノード間の相互作用によって変化する。相互作用は所属グループの 1 つが同じ場合に起こり、相互作用によってノード状態、すなわち所属グループ集合が変化する。各ノード i の所属グループ数 a_i が変化しない場合は、所属グループ集合 g_i の変化は既存の 1 つを削除して代わりに新しい 1 つを追加するグループの置換だけが起こる。ただし、所属グループの自発的な変化はないと仮定する。ノード間の相互作用がノード状態だけによって決まるとすると、相互作用によって既存のグループを置換する新しいグループは相互作用する相手のノードに存在している必要がある。したがって、NGAM の相互作用とは所属グループのコピー (上書き) によるノード状態変化である。コピーする新しいグループと削除する (上書きされる) グループの選択には任意性があり、この選択がダイナミクスを生じる駆動要因の 1 つとなる。

シミュレーションの初期ネットワークは、各ノード i について 1 以上の所属グループ数 a_i を指定の確率分布で設定し、グループ集合全体から重複がない a_i 個のグループをランダムに選択して所属グループ集合 g_i の要素として割り当てる。簡単化のために各ノードの所属グループ数 a_i は時間的に一定とする。

ネットワークの時間変化はつぎのアルゴリズムによって起こる。

1. 各時間ステップ t で 1 つのノード i をランダムに選択する。
2. 所属グループの少なくとも 1 つがノード i と同じである 1 つのノード j (i の隣人, $g_i \cap g_j \neq \emptyset$) をランダムに選択する。
3. ノード i と j のそれぞれの所属グループ集合 g_i と g_j が等しい ($g_i = g_j$) または片方が他方を包摂する ($g_i \subset g_j$ または $g_j \subset g_i$) ならば、何もしない。
4. さもなくば、 g_i と g_j から一致しないグループをランダムに 1 つずつ選択してどちらか一方で他方を置換 (コピー) する。
5. ネットワークが一定時間 (下記の実験では計測間隔 m) の間に変化しなくなるまで以上を繰り返す。

3.3 ダイナミクスの駆動要因

人間社会において周りの人に影響されて意見や態度を変容させる現象を一般に同調 (conformity) という。社会心理学では、同調とは現実あるいは想像上の集団圧力の結果として行動や信念が集団に向かう変化である [Kiesler 1969]。NGAM では、この集団圧力を集団全体ではなく個人間の相互作用の累積によるものとして扱う。その理由は、本研究がシステムのミクロレベルの相互作用からマクロレベルの性質を導くという統計物理学からのアプローチ [Sen 2013] を基礎とするためである。このアプローチはシステムの構成要素が均質である場合に成功しており、NGAM においても各エージェントの性質 (機能と構造) が均質であれば有効であると考えられる。

NGAM における同調は 2 つのパラメータによって制御される。1 つは同調の強さを表す同調度 (conformity) で、シミュレーションにおいてはステップ 4 で状態のコピーが起こる確率 p として

表される。もう一つはコピーするグループ g の選好を表す選好度 (preference) で、シミュレーションにおいてはグループサイズ s_g (グループ g に所属するノード数) に比例して選択する確率 q として表される。このグループサイズによる優先選択は集団全体の影響を間接的に反映すると考えられる。

4. 実験

以下のシミュレーションでの共通の実験条件は、PC で計算可能なノード数 $N=10^4$ 、グループ数 $G=10^4$ 、最大時間ステップ数 $T=10^8$ 、計測間隔 $m=10^4$ 、実験繰り返し回数 $n=10$ である。

4.1 大群化現象

まず、グループ機能をもつソーシャルメディアにおいてグループサイズがべき乗分布する性質を MGAM のシミュレーションによって再現する。このグループサイズ分布は、少数の大きいグループ (つまり大群) と多数の小さいグループの混在であるため、グループサイズ s の変動係数 V_s (標準偏差を平均値で割ったもの) と最大グループサイズ比率 R_s (最大グループサイズをノード数で割ったもの) によって特徴付けられる。これらの特徴量は、以下の実験ではアルゴリズムが停止したときの値とする。

最も単純な MGAM は、同調度 $p=1$ 、選好度 $q=0$ で、所属グループ数 a_i が一定の場合であるため、この条件で $\langle a \rangle = 3$ についての V_s と R_s の時間変化 (最大値で規格化) の 1 例を図 1 に示す。この実験の初期ネットワークの平均次数は 10.0 で、グループサイズの平均値は 3.0 である。また、このシミュレーションの最大時間ステップ $T=10^8$ でのグループサイズ分布を図 2 に示す。この最終ネットワークの平均次数は 10.0 で変わらず、グループサイズの平均値は 10.0 に増加している。これらの結果から、上記の実験条件において NGAM のダイナミクスによって、グループサイズのべき乗分布が平衡状態として創発する。時間変化によって起こるこの動的相転移をここでは大群化現象 (herding) [Easley 2010] と呼ぶ。NGAM において V_s が大きくなる大群化は、図 3 に示すようにグループ数 G が大きい (>1000) ときに、平均所属グループ数 $\langle a \rangle$ の極めて狭い範囲 (3~7) で起こる。

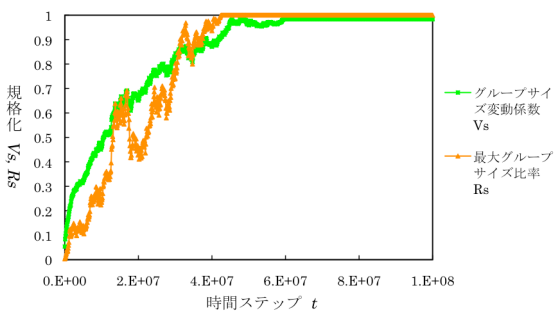


図 1 グループサイズ特徴 V_s と R_s の時間変化

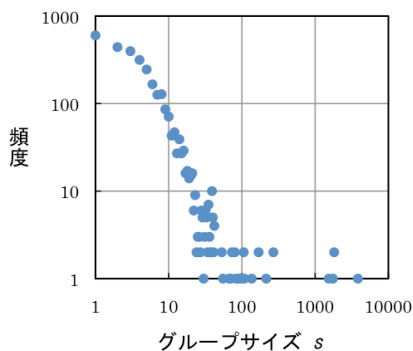


図 2 最大時間ステップ T でのグループサイズ分布

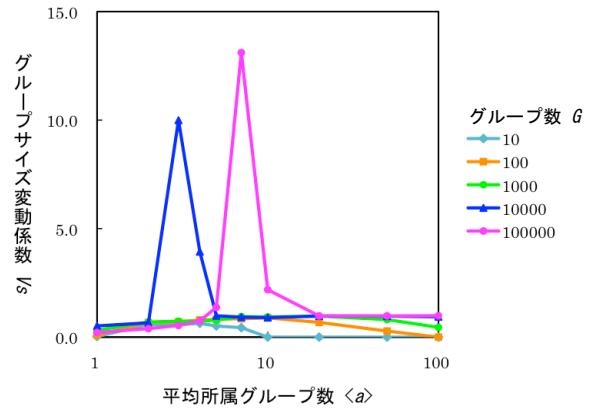


図 3 大群化が起こる条件

4.2 大群化転移

つぎに、3.2 節のアルゴリズムのステップ 3 におけるグループのコピーが起こる確率 (同調度) p の変化によって、大群化が 1 次相転移 [Kadanoff 2000] として起こることを示す。この性質をここでは大群化転移と呼ぶ。4.1 節と同じ条件 (ただし $n=20$) で p を 0~1 の間で変化させたときの最大グループサイズ比率 R_s の変化を図 4 に示す。この図から明らかのように、 R_s は臨界値 $p_c=0.2$ 付近で不連続的に変化しており 1 次相転移が起こっている。 $p < p_c$ である領域はネットワークが小さいグループに分離する分離相で、 $p > p_c$ である領域は大きいグループが存在する大群化相である。 $p > p_c$ で最大グループサイズは p に比例して増加する。同じ条件でノード数 N に対する依存性を調べた結果、大群化は N が大きいとき ($\geq 10^4$) に起こる集団現象である。ネットワークが変化しなくなるまでの収束時間は N に対して指数関数的に増加し、 $N = 1.2 \times 10^4$ 付近で最大時間ステップ T を超えるため、グループサイズ変動係数 V_s はこの付近にピークをもつ。

同調度 p によって大群化転移が生じる理由はつぎのように定性的に説明できる。シミュレーションの最初にほぼ均等であったグループサイズがある特定のグループにおいて増加するためには、そのグループに所属するノードたちからその隣人たちへの同グループのコピーがまとまって起こらなければならない。そうでないと、そのグループが一旦コピーされても他のグループによって書ききされてしまい、グループサイズの増加にはつながらず。つまり同調度の臨界値 p_c 以下ではランダムなグループのコピーが起こっているのに対して、 p_c を超えると特定のグループのコピーが連鎖的に起こるためにそのグループサイズの増加が起こると考えられる。これは一種の雪崩現象であると考えられる。

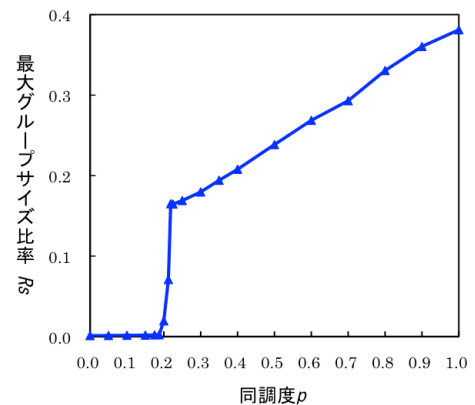


図 4 同調度 p に対する最大グループサイズ比率 R_s の変化

4.3 大群化の促進

さらに、グループのコピーにおけるグループサイズによる選好度 q を大きくすることによって大群化が促進されることを示す。

4.1 節と同じ条件(ただし $n=20$)で q を $0 \sim 1$ の間で変化させたときのグループサイズ変動係数 V_s の変化を図 5 に示す。この図から明らかなように、 V_s が q に対してほぼ線形に増加しており、グループサイズ s_g に比例してコピーするグループ g を選択することが大群化を促進する。

この大群化の促進が起こる理由は、つぎのように定性的に説明できる。大群化が起きるときはある特定のグループのグループサイズが増加するが、グループのコピーにおいてグループサイズによってグループを優先的に選択すると、さらにそのグループサイズが相乗的に増加すると考えられる。この効果は、グループサイズのべき乗分布を創発させる直接の要因ではないが、同調との相乗効果によってべき乗分布の創発を促進すると考えられる。

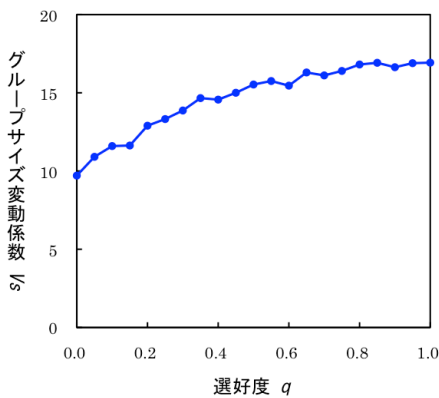


図 5 選好度 q に対するグループサイズ変動係数 V_s の変化

4.4 所属グループ数分布の効果

以上の実験ではノードの所属グループ数 a_i を一定としたが、 a_i の確率分布の違いによる効果を図 6 に示す。実験は、所属グループ数の平均値 $\langle a \rangle$ が同じ条件で a_i がポアソン分布とべき乗分布(指数は-1)に従う場合と a_i が一定の場合についてグループサイズ変動係数 V_s を比較した。この結果、所属グループ数の最大値がより大きい確率分布ほど V_s 大きくなり、グループサイズのべき乗分布がより顕著になる傾向が見られた。この傾向は、所属グループ数の大きいノードがハブになって特定のグループをその多くの隣人へとコピーするためと考えられる。

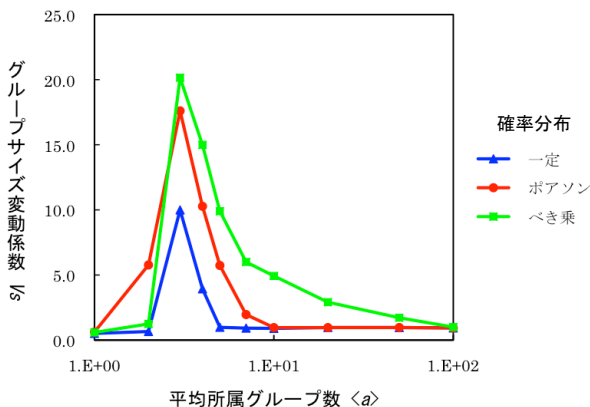


図 6 所属グループ数の確率分布による V_s の変化

5. おわりに

本論文は、既存の投票者モデルにおけるノード状態を 2 値から多値の集合に拡張した、多重グループ所属をノード状態とする投票者モデル MGAM を提案し、そのグループサイズ分布の性質について述べた。MGAM は、ノード状態を等しくしようとする同調度が臨界値より大きいとき、グループサイズがべき乗分布する大群化現象を生じる。この大群化は同調度に対しての 1 次相転移であり、さらにグループサイズによる選好度の増加によって大群化が促進される。これらの MGAM の性質は、社会における同調というメカニズムによって、グループ機能をもつソーシャルメディアにおいてグループサイズがべき乗分布する事実をうまく説明する。

本論文で提案した MGAM は、従来の適応的投票者モデルのようにネットワークのダイナミクスとネットワーク上のダイナミクスを分けて扱う必要がないので、解析的な扱いがより容易になると考えられる。また、MGAM がネットワークにおける雪崩現象や非線形現象を呈することから、本論文で調べたグループサイズ分布以外にも有用な性質をもつ可能性がある。これらの課題を追求することはネットワークが創発する知能の解明につながると期待される。

参考文献

- [Easley 2010] Easley, D., Kleinberg, J.: *Networks, Crowds, and Markets - Reasoning about a Highly Connected World*. Chapter 16 Information Cascades. Cambridge University Press, 2010.
- [Ferrara 2012] Ferrara, E.: A Large-scale Community Structure Analysis in Facebook. EPJ Data Science, 1(9), 2012.
- [Ji 2013] Ji M, Xu C, Choi, Chi Wun, Hui Pak Ming.: Correlations and Analytical Approaches to Co-evolving Voter Models. New Journal of Physics, vol.15, p.113024 (17pp), 2013.
- [Kadanoff 2000] Kadanoff, L. P.: *Statistical Physics. Statics, Dynamics and Renormalization*. Part IV. Phase Transitions. World Scientific, 2000.
- [Kiesler 1969] Kiesler, C. A., Kiesler, S. B.: *Conformity*. Reading, MA. AddisonWesley, 1969. 早川昌範訳:『同調行動の心理学』, 誠信書房, 1978.
- [Pollner 2006] Pollner, P., Palla, G., Vicsek, T.: Preferential attachment of communities: the same principle, but a higher level. Europhys. Lett. 73, pp.478-484, 2006.
- [Sen 2013] Sen, P., Chakrabarti, B. K.: *Sociophysics: An Introduction*. Oxford University Press, 2013.
- [Silk 2014] Silk, H., Demirel, G., Homer, M., Gross, T.: Exploring the adaptive voter model dynamics with a mathematical triple jump. New Journal of Physics, Vol.16, No.9, pp.93051-93063(13), 2014.
- [Vazquez 2008] Vazquez, F., Eguiluz, V. M., San Miguel, M.: Generic absorbing transition in coevolution dynamics. New J. Phys. 10, 063011, 2008.
- [Zheleva 2009] Zheleva, E., Hossam Sharara, H., Getoor, L.: Co-evolution of social and affiliation networks. In Proc. of the 15th ACM SIGKDD, pp.1007-1016, 2009.
- [石川 2014] 石川 孝: 複雑ネットワークの自己組織化メカニズム—同質原理の正体. JSAI2014, 2J4-OS-16a-2, 2014.