Complex Networks, Proc. 3rd Workshop on Complex Networks, CompleNet, Melbourne, FL, March 7-9, 2012, Studies in Computational Intelligence, vol. 424, Springer, Berlin, Germany, 2013, pp. 161-168.

Tipping Points of Diehards in Social Consensus on Large Random Networks

W. Zhang, C. Lim, B. Szymanski

Abstract We introduce the homogeneous pair approximation to the Naming Game (NG) model, establish a six dimensional ODE for the two-word NG. Our ODE reveals how the dynamical behavior of the NG changes with respect to the average degree $< k >$ of an uncorrelated network and shows a good agreement with the numerical results. We also extend the model to the committed agent case and show the shift of the tipping point on sparse networks.

1 Introduction

The Naming Game(NG) has become a very popular model in analyzing the behaviors of social communication and consensus [1]. In this model, each node is assigned a list of names as its opinions chosen from an alphabet S. In each time step, two neighboring nodes, one listener and one speaker are randomly picked. The speaker randomly picks one name from its name list and sends it to the listener. If the name is not in the list of the listener, the listener will add this name to its list, otherwise the two communicators will achieve an agreement, i.e. both collapse their name list to this single name. The variations of this game can be classified as the “Original” (NG), “Listener Only” (LO-NG) and “Speaker Only” (SO-NG) types [2] regarding the update when the communicators make an agreement, and as the “Direct”, “Re-

W. Zhang
Department of Mathematics, Rensselaer Polytechnic Institute, 110 8th Street, Troy, New York 12180-3590, USA, e-mail: zhangw10@rpi.edu

C. Lim
Department of Mathematics, Rensselaer Polytechnic Institute, 110 8th Street, Troy, New York 12180-3590, USA, e-mail: limc@rpi.edu

B. Szymanski
Department of Computer Science, Rensselaer Polytechnic Institute, 110 8th Street, Troy, New York 12180-3590, USA, e-mail: szymansk@cs.rpi.edu

verse" and "Link-updated" types regarding the way that the two communicators are randomly picked. These variations have different behaviors but can be analyzed in the way. In this paper we mainly focus on the "Original" "Direct" version.

Mean field approach has been applied to the NG and a lot of interesting results have been obtained. They reveal the essential difference of the NG compared with other communication models, such as the voter model, in reproducing important phenomena in real world social communications. One of the most significant results is a phase transition at a critical fraction of the committed agents in the network, the tipping point [6]. Above the tipping point, the minority committed agents will persuade the majority to achieve global consensus in a time growing with the logarithm of network size, while below the tipping point, the committed agents would require the time exponential in the network size [7], so practically never, for networks of non-trivial size. However, as most applications of the mean field approximation, these theoretical predictions deviate from the simulations on complex networks especially when the network is relatively "sparse". In many studies, the dynamical behavior of the network given its average degree or the degree distribution is very important.

Recently, a so-called homogeneous pair approximation has been introduced to voter model [5], a model simpler than the NG, which improves the mean field approximation by taking account of the correlation between the nearest neighbors. Their analysis is based on the master equation of the active links, the links between nodes with different opinions. Although it shows a spurious transition point of the average degree, it captures most features of the dynamics and works very accurately on most uncorrelated networks such as ER and scale free networks.

In this paper, we apply this idea to the NG, especially the two-word NG case. Different from the voter model case, there are more than one type of active links, so we have to analyze all types of links including active and inert ones. As a consequence, instead of a one dimensional ODE in voter model case, we have a six dimensional one. We derive the equations by analyzing all possible updates in the process and write it in a matrix form with the average degree $<k>$ as a explicit parameter. The ODE clearly shows how the NG dynamics changes when $<k>$ decrease to 1, the critical value for ER network to have giant component, and converges to the mean field equations when $<k>$ grows to infinity. Then we show the good agreement between our theoretical prediction and the simulation on ER networks. Finally, we show the decrease of tipping point value in low average degree networks, i.e. we need fewer committed agents to force a global consensus in a loosely connected social network. The results of a detailed analysis of this model will be reported in another paper.

2 The Model

Consider the NG dynamics on an uncorrelated network (the presences of links are independent) together with the following assumptions which are the foundation of

the homogeneous pair approximation:
1. The opinions of neighbors are correlated, while there is no extra correlation besides that through the nearest neighbor. To make this assumption clear, suppose three nodes in the network are linked as 1-2-3 (there is no link between 1 and 3), their opinions are denoted by random variables X_1,X_2,X_3 correspondingly. Therefore this assumption says: $P(X_1|X_2) \neq P(X_1)$, but $P(X_1|X_2,X_3) = P(X_1|X_2)$. This assumption is valid for all uncorrelated networks (Chung-Lu type network [4], especially the ER network).
2. The opinion of a node and its degree are mutually independent. Suppose the node index i is a random variable which picks a random node. The opinion and degree of node i, are X_i and k_i. Mathematically, this assumption means $E[k_i|X_i] =< k >$, $P(X_i|k_i) = P(X_i)$ and $P(X_i|X_j,k_i,k_j) = P(X_i|X_j)$ where j is a neighbor of i. This assumption is perfect for the networks in which every node has the same degree (regular geometry) and is also valid for the network whose degree distribution is concentrated around its average (for example, Gaussian distribution with relatively small variance or Poisson distribution with not too small $< k >$). We will show later this assumption is good enough for ER network.

In other words, the probability distribution of the neighboring opinions of a specific node is an effective field. This field is not uniform over the network but depends only on the opinion of the given node. For an uncorrelated random network with N nodes and average degree $< k >$, the number of links in this network is $M = N < k > /2$. We denote the numbers of nodes taking opinions A,B and AB as n_A, n_B, n_{AB}, their fractions as p_A,p_B, p_{AB}. We also denote the numbers of different types of links as $\mathbf{L} = [L_{A-A}, L_{A-B}, L_{A-AB}, L_{B-B}, L_{B-AB}, L_{AB-AB}]^T$, and their fractions are given by $\mathbf{l} = \mathbf{L}/M$. We take $\mathbf{L}$ or $\mathbf{l}$ as the coarse grained macrostate vector. The global mean field is given by:

$$\mathbf{p}(\mathbf{L}) = \begin{pmatrix} p_A \\ p_B \\ p_{AB} \end{pmatrix} = \frac{1}{2M} \begin{pmatrix} < k > n_A \\ < k > n_B \\ < k > n_{AB} \end{pmatrix} = \frac{1}{2M} \begin{pmatrix} 2L_{A-A} + L_{A-B} + L_{A-AB} \\ L_{A-B} + 2L_{B-B} + L_{B-AB} \\ L_{A-AB} + L_{B-AB} + 2L_{AB-AB} \end{pmatrix}.$$

Suppose X_i, X_j are the opinions of two neighboring nodes. We simply write $P(X_i = A|X_j = B)$, for example, as $P(A|B)$. We also represent the effective fields for all these types of node in terms of $\mathbf{L}$:

$$\overrightarrow{P(\cdot|A)}(\mathbf{L}) = \begin{pmatrix} P(A|A) \\ P(B|A) \\ P(AB|A) \end{pmatrix} = \frac{1}{2L_{A-A} + L_{A-B} + L_{A-AB}} \begin{pmatrix} 2L_{A-A} \\ L_{A-B} \\ L_{A-AB} \end{pmatrix}$$

$$\overrightarrow{P(\cdot|B)}(\mathbf{L}) = \begin{pmatrix} P(A|B) \\ P(B|B) \\ P(AB|B) \end{pmatrix} = \frac{1}{L_{A-B} + 2L_{B-B} + L_{B-AB}} \begin{pmatrix} L_{A-B} \\ 2L_{B-B} \\ L_{B-AB} \end{pmatrix}$$

$$\overrightarrow{P(\cdot|AB)}(\mathbf{L}) = \begin{pmatrix} P(A|AB) \\ P(B|AB) \\ P(AB|AB) \end{pmatrix} = \frac{1}{L_{A-AB} + L_{B-AB} + 2L_{AB-AB}} \begin{pmatrix} L_{A-AB} \\ L_{B-AB} \\ 2L_{AB-AB} \end{pmatrix}.$$

To establish the ODE for NG dynamics, we calculate the variable $E[\Delta\mathbf{L}|\mathbf{L}]$. It takes a long discussion to consider of all possible communications in this network. We just take one case as example: listener taking opinion A and speaker taking opinion B. The probability for this type of communication is $p_A P(B|A)$. The direct consequence of this communication is that the link between the listener and speaker changes from A-B into AB-B, so L_{A-B} decreases by 1 and L_{B-AB} increases by 1. Besides, since the listener changes from A to AB, all his other related links change. The number of these links is on average $<k>-1$ (here we use the assumption 2, $E[k_i|X_i] = <k>$). The probabilities for each link to be A-A, A-B, A-AB before the communication is given by $\overrightarrow{P(\cdot|A)}$ (here we use assumption 1). After the communication, these links will change into AB-A, AB-B, AB-AB correspondingly and change the value of $E[\mathbf{L}]$ by

$$(<k>-1)\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\overrightarrow{P(\cdot|A)}.$$

Similarly, we analyze all types of communications according to different listener's and speaker's opinions, and sum these changes into $\Delta\mathbf{L}$ weighted by the probability that the corresponding communication happens and obtain:

$$E[\Delta\mathbf{L}|\mathbf{L}] = \frac{1}{M}\left[D + (<k>-1)R\right]\mathbf{L}$$

where D is a constant matrix, matrix R is a function of $\mathbf{L}$, given by:

$$D = \begin{pmatrix} 0 & 0 & \frac{3}{4} & 0 & 0 & \frac{1}{2} \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{3}{4} & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & -1 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & \frac{1}{4} & -1 \end{pmatrix}, Q_A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad Q_B = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\begin{aligned} R = \big(\mathbf{0}, \frac{1}{2}[Q_A\overrightarrow{P(\cdot|A)} + Q_B\overrightarrow{P(\cdot|B)}], Q_A[\frac{1}{4}\overrightarrow{P(\cdot|A)} - \frac{3}{4}\overrightarrow{P(\cdot|AB)}], \\ \mathbf{0}, Q_B[\frac{1}{4}\overrightarrow{P(\cdot|B)} - \frac{3}{4}\overrightarrow{P(\cdot|AB)}], -(Q_A + Q_B)\overrightarrow{P(\cdot|AB)}\big). \end{aligned}$$

Then we normalize $\mathbf{L}$ by the total number of links M and normalize time by the number of nodes N:

$$\frac{d}{dt}\mathbf{l} = \frac{N}{M}E[\Delta\mathbf{L}|\mathbf{L}] = \frac{N}{M}[D+(<k>-1)R]\,\mathbf{l}$$
$$= 2\left[\frac{1}{<k>}D+(\frac{<k>-1}{<k>})R\right]\mathbf{l}. \quad (1)$$

Now we get the ODE of $\mathbf{l}$ for the NG and $<k>$ is explicit in the formula. In the last line, the first term is linear and comes from the change of the link between the listener and the speaker. The second term is nonlinear and comes from the changes of all the related links. Under the mean field assumption, the first term does not exist, because there is no specific "speaker" and every one receives messages from the mean field. When $<k>\rightarrow 1$, the ODE becomes:

$$\frac{d}{dt}\mathbf{l} = 2D\mathbf{l}$$

which is a linear system. When $<k>\rightarrow\infty$, the ODE becomes:

$$\frac{d}{dt}\mathbf{l} = 2R\mathbf{l}$$

If in matrix R we further require $\overrightarrow{P(\cdot|A)} = \overrightarrow{P(\cdot|B)} = \overrightarrow{P(\cdot|AB)} = \mathbf{p}$ and transform the coordinates by $\mathbf{L}\rightarrow\mathbf{p}(\mathbf{L})$, the ODE just turns back to the one we have under the mean field assumption [6].

3 Numerical Results without committed agents

Next we show some numerical results. Fig.1 shows the comparison between our theoretical prediction (color lines) and the simulation on ER networks (black solid lines). The dotted lines are theoretical prediction by mean field approximation. We calculate the evolution of the fractions of nodes with A, B and AB opinions respectively and show that the prediction of mean field approximation deviates from the simulation significantly while that of homogeneous pair approximation matches simulations very well.

Fig.2 shows the trajectory of the macrostate mapped into two dimensional space (p_A,p_B), the black line is the trajectory predicted by the mean field approximation. We find that when $<k>$ is large enough, say 50, the homogeneous pair approximation is very close to the mean field approximation. When $<k>$ decreases, the trajectory tends to the line $p_{AB} = 1-p_A-p_B = 0$, which means there are fewer nodes with mixed opinions than predicted by the mean field. In this situation, opinions of neighbors are highly correlated forming the "opinion blocks", and mixed opinion nodes can only appear on the boundary.

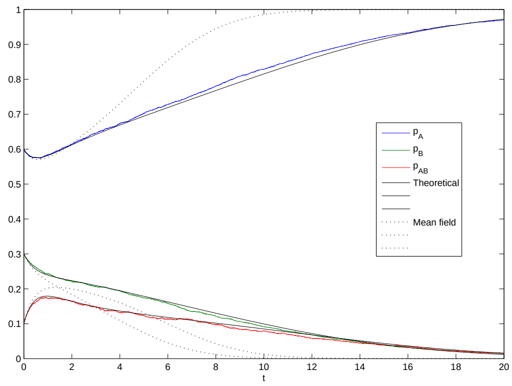

Fig. 1 Fractions of A, B and AB nodes as function of time. The three color lines are the averages of 50 runs of NG on ER network with $N = 500$ and $< k >= 5$. The black solid lines are solved from the ODE above with the same $< k >$. The black dotted lines are from the ODE using mean field assumption.

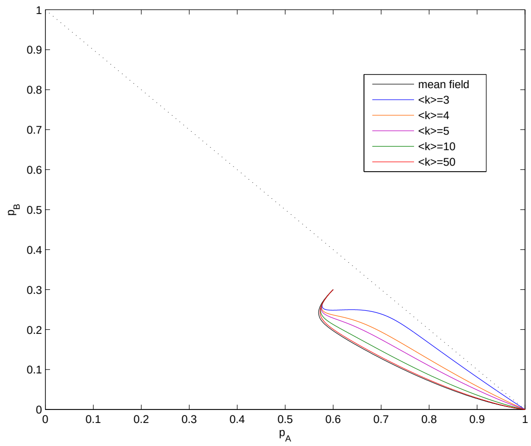

Fig. 2 The trajectories solved from the ODE with different $< k >$ mapped onto 2D macrostate space. When $< k > \rightarrow \infty$, the trajectory tends to that of the mean field equation. When $< k > \rightarrow 1$, the trajectory get close to the line $p_{AB} = 1 - p_A - p_B = 0$.

4 Committed Agents

Suppose we have p (fraction) committed agents (nodes that never change their opinions) of opinion A, and all the other nodes are initially of opinion B. Is it possible for the committed agents to persuade the others and achieve a global consensus? Previous studies found there is a critical value of p called tipping point. Above this value, it is possible and the persuasion takes a short time, while below this value, it is nearly impossible as it takes exponentially long time with respect to the system sizes.

Similar to what we did in the previous section. We derive the ODE for the macrostate, although the macrostate now contains three more dimensions. $\mathbf{L} = [L_{A-C}, L_{B-C}, L_{AB-C}, L_{A-A}, L_{A-B}, L_{A-AB}, L_{B-B}, L_{B-AB}, L_{AB-AB}]^T$, where C denotes the committed A opinion and A itself denotes the non-committed one. Hence we have a nine dimensional ODE which has the same form as equation 1, but with different details in D and R:

$$D = \begin{pmatrix} 0 & 0 & \frac{3}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{3}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{3}{4} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{4} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{4} & -1 \end{pmatrix}, Q_A = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \; Q_B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\begin{aligned} R = (&\mathbf{0}, \frac{1}{2} Q_B \overrightarrow{P(\cdot|B)}, -\frac{3}{4} Q_A \overrightarrow{P(\cdot|AB)}, \mathbf{0}, \frac{1}{2}[Q_A \overrightarrow{P(\cdot|A)} + Q_B \overrightarrow{P(\cdot|B)}], \\ &Q_A[\frac{1}{4}\overrightarrow{P(\cdot|A)} - \frac{3}{4}\overrightarrow{P(\cdot|AB)}], \mathbf{0}, Q_B[\frac{1}{4}\overrightarrow{P(\cdot|B)} - \frac{3}{4}\overrightarrow{P(\cdot|AB)}], -(Q_A + Q_B)\overrightarrow{P(\cdot|AB)}). \end{aligned}$$

Finally, we show the change of the tipping point with respect to the average degree $< k >$ in Fig.3. Starting from the state that $p_B = 1 - p$, the ODE system will go to a stable state for which $p_B = p_B^*$. p_B^* is 0 if the committed agents finally achieve the global consensus. The sharp drop of each curve indicates the tipping point transition with the corresponding $< k >$. According to the figure, the tipping point shifts left when the average degree $< k >$ decreases.

Acknowledgements This work was supported in part by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053, by the Army Research Office Grant No. W911NF-09-1-0254, and by the Office of Naval Research Grant No. N00014-09-1-0607. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.

References

1. A. Baronchelli, M. Felici, V. Loreto, E. Caglioti and L. Steels: Sharp transition towards shared vocabularies in multi-agent systems. *J. Stat. Mech.: Theory Exp.* **P06014**(2006).
2. A. Baronchelli: Role of feedback and broadcasting in the naming game. *Phys. Rev. E* **83**,046103 (2011).
3. E. Pugliese and C. Castellano: Heterogeneous pair approximation for voter models on networks. *Eur. Lett.* **88**, 5, pp. 58004 (2009).
4. F. Chung and L. Lu: The Average Distances in Random Graphs with Given Expected Degrees. *Proceeding of National Academy of Science* **99**,15879C15882 (2002).
5. F. Vazquez and V. M. Egułluz: Analytical solution of the voter model on uncorrelated networks. *New Journal of Physics* **10**, 063011 (2008).
6. J. Xie, S. Sreenivasan, G. Korniss, W. Zhang, C. Lim, B. K. Szymanski: Social Consensus through the Influence of Committed Minorities, *Phys. Rev. E* **84**, 011130 (2011).
7. W. Zhang, C. Lim, S. Sreenivasan, J. Xie, B. K. Szymanski, and G. Korniss: Social influencing and associated random walk models: Asymptotic consensus times on the complete graph *Chaos* **21**, 2, 025115 (2011).

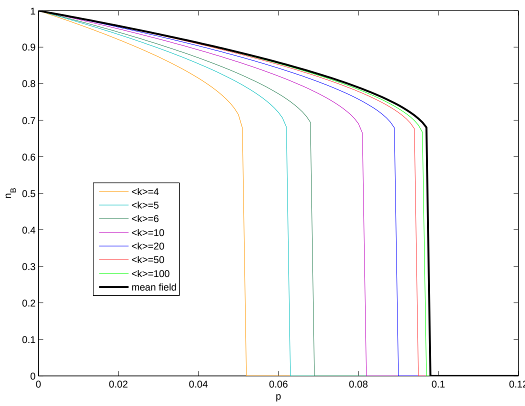

Fig. 3 Fraction of B nodes of the stable point (p_B^*) as a function of the fraction of nodes committed to A (p). The color lines consist of stable points obtained by tracking the ODE of NG on ER for a long enough time. The black lines are the stable points solved from the mean field ODE.