Generalized Mean-Field Approximation for Opinion Spreading in Social Networks

Sungsu Lim
KAIST
ssungssu@kaist.ac.kr

Kyomin Jung
KAIST
kyomin@kaist.edu

How political opinion, product adoption, rumors, and trends diffuse through social networks has been a fundamental question for many decades in a wide variety of research disciplines. The spread of opinion on networks can be thought of as a state dynamics, where each node decides its state based on interactions with its neighbors. In many such settings, the dynamics of the states of nodes are described by some *Markov process* on a graph $G = (V, E)$ with a finite state space S. I.e., the state of a node v at each time step t is determined solely by the states of v and its neighbors at a time period $\{t-j, \cdots, t-1\}$ for some nonnegative integer j. Examples of such Markov processes include information diffusion models [7], the voter model [15], the pricing model [6], the naming game [14], and the gossip algorithm [4], some of which we will present in more detail later.

To predict the behavior of opinion spreading under the Markov process, mean-field (MF) approximation with ordinary differential equations (ODEs) has been widely used [2, 3, 5, 11]. Such an approach describes the state change rule as a system of ODEs that reflect the inherent Markovian state dynamics. In essence, the actual state ratio dynamics converges to the solution of the ODEs as the number of nodes goes to infinity. However, the analyses of its convergence are known only for symmetric network structures such as complete or bipartite graphs [8]. In our work, we propose a generalized MF method that relaxes the condition on strong symmetry, and prove the convergence to the ODE solution on any *slightly dense* graphs.

More precisely, consider a network structure $G = (V, E)$ with $n = |V|$ nodes so that each node has degree $\omega(\log n)$ (hence, *slightly dense*). Initially each node determines its state at time 0 randomly according to the state ratios $(s_i)_{i \in S}$ in an i.i.d. manner. Then we prove that the solution of the actual state dynamics of the Markov process converges to the solution of the ODEs. In addition, we also show that the ratios of the states among the neighbors of any given node is close to the actual state ratios of the entire network. As a general framework, our analysis can be applied to many opinion spreading processes in social networks caused by public service advertising, group-targeted marketing, and external influence [1, 10]. Note that our result does not take any structural information of the network into account. Hence, surprisingly our results show that the MF approximation holds independently of the network structure, such as the community structure, as long as the initial states are drawn in an i.i.d. manner.

We adopt the standard (continuous) asynchronous time model [11] to express the state dynamics, where on average there are n state updates per unit time. Our result can also be applied to the slotted synchronous time model [4]. The following theorems formalize our main results.

Theorem 1 *Consider a Markov process with a finite state space* S *on a graph* $G = (V, E)$ *so that all nodes are of degrees* $\omega(\log n)$*, and with an initial state ratio* $(s_i)_{i \in S} \in [0,1]^{|S|}$*. For each time* $t \in \mathbb{R}_+$*, let* $\mathbf{a}(t) = (a_i(t))_{i \in S}$ *be the solution of the system of ODEs that corresponds to a given Markov process* [1]*. For* $t \in \mathbb{R}_+$*, let* $\mathbf{b}(t) = (b_i(t))_{i \in S}$ *be the random variable for the actual state ratio of* V*. Then, for any constants* $\epsilon > 0$*,* $\delta > 0$*, and* $T > 0$*,*

$$Pr\left(\sup_{0 \le t \le T} \|\mathbf{b}(t) - \mathbf{a}(t)\|_1 < \epsilon\right) = 1 - o(n^{-\delta}). \quad (1)$$

This means that the overall state ratio of V converges to $\mathbf{a}(t)$ uniformly over $t \in [0, T]$ as n goes to infinity. Furthermore, the state ratio of the neighbors of each node at time $t \in [0, T]$ is also very close to $\mathbf{a}(t)$, as stated in the following theorem.

Theorem 2 *For* $t \in \mathbb{R}_+$ *and* $v \in V$*, let* $\mathbf{b}_v(t) = (b_{v,i}(t))_{i \in S}$ *be the random variable for the actual state ratio of the neighbors of* v*. Under the same condition as in Theorem 1, for any constants* $\epsilon > 0$*,* $\delta > 0$*, and* $T > 0$*,*

$$Pr\left(\sup_{0 \le t \le T, v \in V} \|\mathbf{b}_v(t) - \mathbf{a}(t)\|_1 < \epsilon\right) = 1 - o(n^{-\delta}). \quad (2)$$

We conducted experiments on the network datasets consisting of two synthetic networks (Watts-Strogatz (WS) and the Barabási-Albert (BA)) of 10,000 nodes with average degree 100, and two real-world networks (ePinions [12] (75,879 nodes with average degree 13.4) and Slashdot [9] (77,360 nodes with average degree 23.4)) [2]. The opinion spreading models we used were the linear threshold model [7] and the ternary voter model [11]. Our empirical results confirmed that Theorem 1 and Theorem 2 are very accurate for WS and BA networks, and are validated quite well for real-world networks where the slightly dense condition is violated. Finally, we present the details of some of the popular Markovian opinion spreading models where our results can be applied.

- *Information diffusion models* [3,7,10]: In many information diffusion models, including the general threshold model and the independent cascade model, the state space can be expressed by $S = \{0(\text{inactive}), 1(\text{active})\}$. In the general threshold model, each node v becomes active if a function evaluated on the states of its neighbors exceeds the threshold of v. In the independent cascade model, every adoptor v has a single chance to influence its non-adopted neighbor u with a certain probability. Our work can also be applied to non-monotonic information diffusion models, such as SIS model and its recent variants [3], and the information diffusion model with external influence [10].

[1] Solve $\mathbf{a}(t)$ by a standard MF method with the system of ODEs.
[2] `http://snap.stanford.edu`

- *Voter model* [5, 11]: In the voter model, the state space corresponds to the set of candidates or items to vote, and a node v either updates its state by copying that of its neighbor chosen uniformly at random with some probability, or preserves its state with the remaining probability.
- *Naming game* [13, 14]: Naming game was originally intended to model language diffusion in a society. This model has been widely used for describing how a multi-agent system can converge towards a consensus state in a self-organized way. Similarly to the voter model, a node (listener) is selected, and it interacts with a randomly chosen neighbor (speaker). The speaker randomly selects a language from its language list and sends it to the listener. Then, the listener updates its language list.

References

[1] M.-F. Balcan, A. Blum, and Y. Mansour. Improved equilibria via public service advertising. In *Proceedings of SODA*, 2009.

[2] M. Benaïm and J. Weibull. Mean-field approximation of stochastic population processes in games. 2009.

[3] A. Beutel, B. A. Prakash, R. Rosenfeld, and C. Faloutsos. Interacting viruses in networks: Can both survive? In *Proceedings of ACM SIGKDD*, 2012.

[4] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Networking*, 14, 2005.

[5] K. Jung, B. Y. Kim, and M. Vojnovic. Distributed ranking in networks with limited memory and communication. In *Proceedings of ISIT*, 2012.

[6] M. L. Katz and C. Shapiro. Network Externalities, Competition, and Compatibility. *The American Economic Review*, 75, 1985.

[7] D. Kempe, J. Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of ACM SIGKDD*, 2003.

[8] T. Kurtz. *Approximation of Population Processes*. SIAM, 1981.

[9] J. Leskovec, K. Lang, A. Dasgupta, and M. Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 6, 2009.

[10] S. Myers, C. Zhu, and J. Leskovec. Information diffusion and external influence in networks. In *Proceedings ACM SIGKDD*, 2012.

[11] E. Perron, D. Vasudevan, and M. Vojnović. Using three states for binary consensus on complete graphs. In *Proceedings of INFOCOM*, 2009.

[12] M. Richardson, R. Agrawal, and P. Domingos. Trust management for the semantic web. In *Proceedings of ISWC*, 2003.

[13] L. Steels. A self-organizing spatial vocabulary. *Artificial Life*, 2, 1995.

[14] J. Xie, S. Sreenivasan, G. Korniss, W. Zhang, C. Lim, and B. K. Szymanski. Social consensus through the influence of committed minorities. *Phys. Rev. E*, 84, 2011.

[15] M. Yildiz, R. Pagliari, A. Ozdaglar, and A. Scaglione. Voting models in random networks. In *Proceedings of ITA*, 2010.