Parsimonious Algorithms for Decentralized Ranking in Social Networks

Kyomin Jung
KAIST
kyomin@kaist.edu

Boyoung Kim
KAIST
combicola@kaist.ac.kr

Milan Vojnović
Microsoft Research
milanv@microsoft.com

We consider decentralized algorithms where each node in a network aims at computing an aggregate quantity of all node states using only local information without any centralized agency. Recently, a large amount of interest on this type of algorithms has been arisen in various contexts such as social networks [1, 2], Internet [3, 4], and biological systems [5], because they are not only useful to explain phenomena observed in a network system, but also useful for the design of new computation protocols. Since many network systems inherently contain restrictions on memory and communications (i.e. parsimonious), designing a decentralized algorithm under such restrictions is an important problem. There have been many studies trying to account for these restrictions. For example, in the context of averaging algorithms, randomized gossip algorithms based on reversible Markov chains [6] have been considered as well as averaging algorithms based on non-reversible Markov chains [7]. For the averaging problem, also the effects of quantization of messages exchanged between nodes have been studied [8, 9].

In this paper, we study a *rank aggregation problem* where the goal is to rank a set of alternatives in decreasing order of users' preference in a decentralized manner. A specific example is voting over a set of alternatives, which frequently arises in social networks including surveys of consumer preferences. The goal is to identify a list of top k popular products in decreasing order of their popularity. Such cooperative decision making problem arises in a variety of applications such as in surveys of preference in social networks, decentralized database systems, and sensor networks.

The main contributions of our work are in (1) allowing for arbitrary number of alternatives $m \geq 2$, and (2) algorithms for ranking that are based on computing a generalized version of the *mode*, and (3) allowing for user preference across a set of alternatives. For computing the mode in network systems where each node prefers one out of two alternatives, the classical voter model [10–12] has been extensively studied. Algorithms for binary consensus were proposed to serve as an improvement of the voter model with respect to the error probability and the convergence speed to the correct consensus [13–15]. A quantized version of the gossip algorithm was suggested to identify the quantization interval containing the average value [8]. Using this algorithm, the majority voting problem can be solved with only four states per node. However, these works are restricted to the case of two alternatives.

The detailed setup that we consider is as follows: we consider a network system that consists of nodes $[n] = \{1, 2, \ldots, n\}$ where $n \geq 1$ and a finite set of alternatives $[m] = \{1, 2, \ldots, m\}$ where $m \geq 2$. The preference of each node $j \in [n]$ over alternatives is described by the vector of ranking scores $\vec{v_j} = (v_1, v_2, \ldots, v_m)$ where $v_i \geq 0$ and $\sum_{i=1}^{m} v_i = 1$. A vector of ranking scores $\vec{v}_j$ is such that the i-th coordinate of this vector represents preference of node j for alternative i. A *top-k ranking* is a tuple of alternatives $(a_1, a_2, \ldots, a_k)$, for $k \leq m$, such that $a_i \in [m]$ for every i, and $U(a_1) \geq U(a_2) \geq \ldots \geq U(a_m)$ where $U(a_i)$ is the sum of ranking scores for alternative a_i over all nodes. The ranking problem is to construct a decentralized algorithm which ensures that every node computes a top-k ranking correctly after finitely many number of iterations.

First, we propose an algorithm that computes the full ranking of alternatives for any connected network graph by generalizing the discretized averaging algorithm [8]. Our algorithm runs in a time equivalent to the mixing time of the corresponding random walk in the network. Our algorithm uses $2^{m(m-1)}$ states per node. Although this algorithm runs correctly on any connected network graph, it is not parsimonious in the required memory per node. Next, we present parsimonious algorithms for the mode computation and the top-k ranking computation for the case of small k.

Our mode computation algorithm is described as follows. As the behavior of bloggers, at each time step, a randomly chosen node observes another node chosen uniformly at random and updates its preference state. The main idea of the update rule is the introduction of two extra states (weak and strong) for each alternative $j \in [m]$. Based on this idea, we prove that this algorithm converges correctly with probability of error that diminishes exponentially with the total number of nodes; this result is established using mean field arguments along with a concentration inequality for random processes.

Finally, we propose an efficient algorithm for computing a top-k ranking. This algorithm starts with assigning to each node a random *k-ranking state* $(b_1, b_2, \ldots, b_k)$, where $b_i \in [m]$ for every i, according to a probability that depends on the ranking score vector of the node. At this step, $m(m-1)\ldots(m-k+1) = O(m^k)$ many k-ranking states are needed. We prove that the problem of computing the mode among a set of k-ranking states is equivalent to the problem of computing the top-k ranking on the set of the original alternatives. Using our mode computation algorithm, we prove that the error probability of our top-k ranking algorithm decays exponentially with the total number of nodes. We examine convergence of our algorithms using simulations.

References

[1] J. Kleinberg, "The small-world phenomenon and distributed search," in *SIAM News*, vol. 37, 2004.

[2] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," in *Review of Economic Studies*, 2010.

[3] J. Kleinberg, "Complex networks and decentralized search algorithms," in *Proc. of the International Congress of Mathematicians*, 2006.

[4] F. Vega-Redondo, A. Galeotti, S. Goyal, M. Jackson, and L. Yariv, and A. Steger, "Network games," in *Review of Economic Studies*, vol. 77, 2010.

[5] F. Kuhn, K. Panagiotou, J. H. Spencer, and A. Steger, "Synchrony and asynchrony in neural networks," in *Proc. of ACM-SIAM SODA*, 2010.

[6] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," in *IEEE Trans. on Information Theory*, vol. 52, pp. 2508-2530, 2006.

[7] K. Jung, D. Shah, and J. Shin, "Distributed averaging via lifted Markov chains," in *IEEE Trans. on Information Theory*, vol. 56, pp. 634-647, 2010.

[8] F. Benezit, P. Thiran, and M. Vetterli, "Interval consensus: from quantized gossip to voting," in *Proc. of IEEE ICASSP*, 2009.

[9] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "On distributed averaging algorithms and quantization effects," in *IEEE Trans. on Automatic Control*, vol. 54, pp. 2506-2517, 2009.

[10] P. Donnelly and D. Welsh, "Finite particle systems and infection models," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 94, pp. 167-182, 1993.

[11] D. Aldous and J. A. Fill, "Reversible Markov chains and random walks on graphs," in *Monograph in Preparations*, 1999.

[12] J. Cruise and A. Ganesh, "Probabilistic consensus via polling and majority rules," in *Proc. of Allerton Conference*, 2010.

[13] E. Perron, D. Vasudevan, and M. Vojnović, "Using three items for binary consensus on complete graphs," in *Proc. of INFOCOM*, 2009.

[14] M. Draief and M. Vojnović, "Convergence speed of binary interval consensus," in *Proc. of INFOCOM*, 2010.

[15] E. Mossel, G. Schoenebeck, "Reaching consensus on social networks," in *International Journal of Intelligent Systems - Decision Making in Social Networks*, vol. 25, 2010.