8-2016

Detecting communities using coordination games: A short paper

Radhika ARAVA

Pradeep VARAKANTHAM
Singapore Management University, pradeepv@smu.edu.sg

Citation
ARAVA, Radhika and Pradeep VARAKANTHAM. Detecting communities using coordination games: A short paper. (2016). *Proceedings on the 22nd European Conference on Artificial Intelligence: ECAI 2016, The Hague, Netherlands, 2016 August 29 - September 2*. 285, 1752-1753.
Available at: https://ink.library.smu.edu.sg/sis_research/3616

ECAI 2016
G.A. Kaminka et al. (Eds.)

doi:10.3233/978-1-61499-672-9-1752

Detecting Communities Using Coordination Games : A Short Paper

Radhika Arava [1] and **Pradeep Varakantham**[2]

Abstract. Communities typically capture homophily as people of the same community share many common features. This paper is motivated by the problem of community detection in social networks, as it can help improve our understanding of the network topology. Given the selfish nature of humans to align with like-minded people, we employ game theoretic models and algorithms to detect communities in this paper. Specifically, we employ coordination games to represent interactions between individuals in a social network. We provide a novel and scalable two phased algorithm *NashOverlap* to compute an accurate overlapping community structure in the given network. We evaluate our algorithm against the best existing methods for community detection and show that our algorithm improves significantly on benchmark networks with respect to standard *normalised mutual information* measure.

1 Introduction

In social networks like Facebook, Google+, there can be overlap in user's high-school friends' circle and his university friends' circle. It is important to identify the overlapping community structure of a social network, as it helps in understanding the network topology, the spread of information, rumor in that network.

In this paper, we provide a novel, scalable two-phase algorithm *NashOverlap* to compute the overlapping community structure of a network. We evaluate our algorithm against the current state of the art and we find that it works significantly better than the best existing methods on the standard LFR benchmark networks [2]. To the best of our knowledge, this is the first game theory based scalable overlapping community detection algorithm that can detect an accurate community structure for a given social network.

2 Community Detection Problem

In a social network $G = (V, E, w)$, V is the set of vertices and E is the set of undirected edges that represents relationships between vertices in the network and $w : E \mapsto \mathbb{R}$ is the weight function on edges. A community in a social network is a non-empty connected subset of vertices with denser connections within themselves than with the rest of the network. Formally, if $C_j \subseteq V$ denotes a community j, then our goal is to identify the community structure $\Gamma = \{C_1, C_2, \ldots, C_r\}$, where $\bigcup_j C_j = V$.

[1] Nanyang Technological University,Singapore,email:radhika.arava@gmail.com
[2] Singapore Management University,Singapore,email:pradeepv@smu.edu.sg

3 Algorithm *NashOverlap*

We design a novel two-phased algorithm *NashOverlap* to detect overlapping communities in a social network as illustrated in Figure (1).

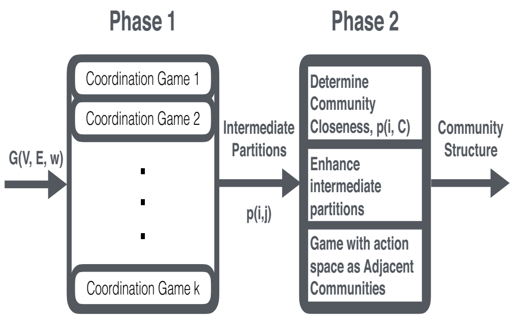

Figure 1. Algorithm: NashOverlap

3.1 First Phase

In this phase, we compute edge-closeness value $p(i, j)$ for each edge (i, j) and compute an intermediate partition of the network by solving k graph coordination games independently using local search.

Given a weighted network $G(V, E, w)$, we formulate each game, $\zeta^1 = (V, (S_i)_{i \in V}, (u_i)_{i \in V})$, where V is the set of players, S_i constitutes the set of r ($r \geq 2$) strategies for each player i and u_i is the utility function for each player i. Let $s_i \in \{1, \ldots, r\}$ denote the strategy that player i chooses; then $(s_i)_{i \in V}$ defines a strategy profile of the game. Let s_{-i} denote the strategies of all the other players other than i. Utility of a player i at a given strategy profile $s = (s_i, s_{-i})$ is given by

$$u_i(s_i, s_{-i}) = \sum_{\substack{j:(i,j) \in E \\ s_i = s_j}} t(i, j) \tag{1}$$

where tie-strength $t(i, j)$ of an edge (i, j) is defined as:

$$t(i, j) = w(i, j) + \sum_{\substack{k:(i,k) \in E \\ (k,j) \in E}} (w(i, k) + w(j, k)) \,. \tag{2}$$

We solve the above graph coordination game ζ^1 using local search in the following way. Initially, assign to each vertex $i \in V$, a strategy picked uniformly from S_i. Pick a uniform random vertex-ordering and each vertex i during its turn picks the strategy $s_i \in S_i$, whichever gives him the maximum utility.

This is repeated with the same vertex-ordering until game ζ^1 converges to a Nash equilibrium. We define edge-closeness value $p(i,j)$ of an edge as the proportion of k games in which its vertices choose the same strategy at Nash equilibrium.

We now show that ζ^1 is a potential game with the potential function $\Phi_1 : \{1, 2, \ldots, r\}^{|V|} \mapsto \mathbb{R}$ that is defined as:

$$\Phi_1 = \sum_{(i,j) \in E \setminus E_c} t(i,j),$$

where E_c is the set of cut-edges.

Theorem 1 Φ_1 *is a weighted potential function.*

We construct an intermediate partition which is the set of connected components containing only edges with $p(i,j) \geq 0.95$.

3.2 Second Phase

This phase takes care of the mistakenly identified edges in the intermediate partition of first phase using $p(i,j)$ values and outputs a stable overlapping community structure.

Definition 1 *Given a vertex i and a community C, define the community-closeness, $p(i,C)$ as the sum of edge-closeness values between i and C. That is,*

$$p(i,C) = \sum_{\substack{j:(i,j) \in E \\ j \in C}} p(i,j).$$

Intuitively, a vertex is said to be adjacent to a community, if it has any of its adjacent vertices in that community. If a vertex is not adjacent to a community, its community-closeness to that community is 0.

Consider a game $\zeta^2 = (V, (S_i)_{i \in V}, (u_i)_{i \in V})$ where S_i for each vertex i is the set of all communities, u_i is the sum of vertex $i's$ community-closeness values to all its adjacent communities.

We solve this graph coordination game using local search. Pick a random vertex-ordering and each vertex in its turn, computes its community-closeness to each of its adjacent communities. The vertex chooses to be a part of only those communities to which its community-closeness is at least α times its maximum community-closeness value, given that its utility increases with that decision. The vertex-ordering is repeated, until the game converges to a Nash equilibrium.

For a given overlapping community structure Γ, we shall duplicate every vertex in each of its overlapping communities. Let $\mathcal{C}_i$ be the set of communities of player i in a given overlapping community structure Γ. So, we put a copy of vertex i in each of its communities $C \in \mathcal{C}_i$. Observe that if the community structure is not overlapping, then $\mathcal{C}_i$ is of size 1 for each player i.

Let ζ^2 be a potential game with potential function $\Phi_2 : \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_{|V|}\} \mapsto \mathbb{R}$ that is given by

$$\Phi_2 = \frac{1}{2} \cdot \sum_{\substack{i \in V \\ C \in \mathcal{C}_i}} p(i,C) \tag{3}$$

Theorem 2 Φ_2 *is a weighted potential function.*

α is a overlap parameter and lies in $[0,1]$. If $\alpha = 1$, the resulting community structure has no overlap.

3.3 Discussion

We show that the games ζ^1 and ζ^2 always converges to a local optimum. Though, the problem of computing the local optimum for all these games is PLS-Complete, we can allow for minor changes in the game parameters and show that we can compute a stable overlapping community structure in linear time (linear in the number of edges).

For every network, given its community structure, mixing parameter μ is defined as the fraction of a vertex's links that connect to vertices sharing no communities. For fuzzier networks($\mu > 0.5$), our algorithm is likely to converge to global optimum. However for networks with $\mu < 0.5$, our algorithm detects an accurate overlapping community structure. We choose $k = 100$ and vary α in $[0,1]$ for our experiments.

We compared the performance of our algorithm against current best algorithms: CFinder [5], OSLOM [4], COPRA [1], SLPA [6] with respect to a standard measure *Normalized Mutual Information* [3] on networks with varying network sizes, community sizes, overlapping membership and mixing parameter. *NashOverlap* outperformed all the other algorithms, with respect to NMI measure in all experimental settings. We provide results for a setting that is representative of the comparison results observed over all the parameter settings in Table 1.

om	SLPA	COPRA	CFinder	OSLOM	NashOv
2	0.87444	0.95993	0.36046	0.97164	0.968639
3	0.79596	0.89308	0.3135	0.89679	0.94078
4	0.75394	0.80646	0.37866	0.82748	0.877891
5	0.692	0.74409	0.36277	0.76214	0.787768
6	0.65122	0.69011	0.32862	0.7129	0.727489
7	0.60913	0.64374	0.34251	0.66702	0.678984
8	0.56703	0.59485	0.34953	0.62428	0.63227

Table 1. Comparison of NMI measure for all algorithms on 5000 vertex networks varying overlapping membership (*om*) from 2 to 8 and community sizes in the range [20, 100] for $\mu = 0.5$ and $on = 10\%$.

We believe that these results along with the simplicity of our algorithm can help in understanding the topology of networks efficiently.

REFERENCES

[1] Steve Gregory, 'Finding overlapping communities in networks by label propagation', *New Journal of Physics*, **12**(10), 103018, (oct 2010).

[2] Andrea Lancichinetti and Santo Fortunato, 'Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities', *Physical Review E*, **80**(1), 016118, (2009).

[3] Andrea Lancichinetti, Santo Fortunato, and Jnos Kertsz, 'Detecting the overlapping and hierarchical community structure in complex networks', *New Journal of Physics*, **11**(3), 033015, (mar 2009).

[4] Andrea Lancichinetti, Filippo Radicchi, Jos J. Ramasco, and Santo Fortunato, 'Finding statistically significant communities in networks', *PLoS ONE*, **6**(4), e18961, (apr 2011).

[5] Gergely Palla, Imre Dernyi, Ills Farkas, and Tams Vicsek, 'Uncovering the overlapping community structure of complex networks in nature and society', *Nature*, **435**(7043), 814–818, (june 2005).

[6] Jierui Xie, Xiaoming Liu, and Boleslaw K. Szymanski, 'SLPA: Uncovering overlapping communities in social networks via a speaker-listener interaction dynamic process', *Proceedings of Data Mining technologies for computational collective intelligence workshop, at ICDM, Vancouver, CA*, 344–349, (dec 2011).