Supplemental Document #1

Community Detection Algorithms Description

In this supplemental document, the community detection algorithms used in this research are described.

To standardized the complexity notation, here is the list of notation used:

1. n : number of nodes.
2. m : number of edges.
3. k : number of iterations.
4. c : number of communities.
5. d : average node's degree.

No.	Algorithm	Description
1.	Belief	This algorithm seeks the consensus from subset of the network with high modularity value. Consensus is acquired with scalable message-passing algorithm, treating modularity as a Hamiltonian, and use the cavity method. Founder: Pan Zhang & Cristopher Moore Year: 2014 Publication: Zhang, Pan, and Cristopher Moore. "Scalable detection of statistically significant communities and hierarchies, using message passing for modularity." Proceedings of the National Academy of Sciences 111.51 (2014): 18144-18149. Method: Propagation-based algorithm, Statistical inference This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: O(kn)
2.	Constant Potts Model (CPM)	This algorithm use the constant potts model where the optimizing a quality function. Founder: Traag et al. Year: 2011 Publication:

		Traag, V. A., Van Dooren, P., & Nesterov, Y. (2011). Narrow scope for resolution-limit-free community detection. Physical Review E, 84(1), 016114. 10.1103/PhysRevE.84.016114. Method: Modularity maximization, Statistical inference This algorithm works for undirected, weighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: unknown or not applicable.
3.	Chinese Whispers	Clustering graph with propagation and include ambiguity with intermediate representation of initial network and hard clustering to get the crisp clusters in disambiguated intermediate network. Founder: Ustalov et al. Year: 2019 Publication: Ustalov, D., Panchenko, A., Biemann, C., Ponzetto, S.P.: `Watset: Local-Global Graph Clustering with Applications in Sense and Frame Induction.` Computational Linguistics 45(3), 423–479 (2019). Method: Propagation-based algorithm This algorithm works for undirected, weighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: O(km)
4.	Diffussion Entropy Reducer (DER)	This algorithm utilize random walks to represent the graph in measure's space, a modification of k-means in that space is applied. Walks is created, initialization is performed, and algorithm running to finally infer the communities. Founder: Kozdoba & Mannor Year: 2015 Publication: M. Kozdoba and S. Mannor, Community Detection via Measure Space Embedding, NIPS 2015. Method: Statistical inference, Information theory, Random walks This algorithm works for undirected, weighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities.

		Complexity: unknown or not applicable.
5.	Eigenvector	This algorithm use the basic concept of eigenvalue and eigenvector to optimize the modularity. Founder: Newman Year: 2006 Publication: Newman, Mark EJ. Finding community structure in networks using the eigenvectors of matrices. Physical review E 74.3 (2006): 036104. Method: Spectral clustering, Modularity maximization This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: $O(n^2)$
6.	Genetic Algorithm (GA)	Use Genetic Algorithm to optimize fitness function and identify densely connected community. Founder: Pizzuti Year: 2008 Publication: Pizzuti, C. (2008). Ga-net: A genetic algorithm for community detection in social networks. In Inter conf on parallel problem solving from nature, pages 1081–1090.Springer. Method: Modularity maximization This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: unknown or not applicable.
7.	Greedy Modularity	At every step of the algorithm, a pair of community that have a maximum positive value to global modularity is merged. Founder: Clauset et al. Year: 2004 Publication: Clauset, A., Newman, M. E., & Moore, C. Finding community structure in very large networks. Physical Review E 70(6), 2004.

		Method: Modularity maximization This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: unknown or not applicable.
8.	Label Propagation	Based on assumption that each node in the network is assigned to the similar community as its neighbours, this algorithm start labelling all the nodes with different unique label and randomized the order of node-picking. Node is then starting to take the majority label of its neighbours. The step will stop until each node has the same label as the majority of its neighbours. Founder: Cordasco & Gargano Year: 2010 Publication: Cordasco, G., & Gargano, L. (2010, December). Community detection via semi-synchronous label propagation algorithms. In 2010 IEEE international workshop on: business applications of social network analysis (BASNA) (pp. 1-8). IEEE. Method: Propagation-based algorithm This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: O(m)
9.	Leiden	Nodes is moved locally, partition is refined, and merging the network based on the refined partition are three of this algorithm principle to find communities in a network. Founder: Traag et al. Year: 2018 Publication: Traag, Vincent, Ludo Waltman, and Nees Jan van Eck. From Louvain to Leiden: guaranteeing well-connected communities. arXiv preprint arXiv:1810.08473 (2018). Method: Modularity maximization This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: unknown or not applicable.

10.	Louvain	This algorithm consists of two stages, local movement of the nodes and aggregation of the network. Founder: Blondel et al. Year: 2008 Publication: Blondel, Vincent D., et al. Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment 2008.10 (2008): P10008. Method: Modularity maximization This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: unknown or not applicable.
11.	Markov Clustering	This algorithm use stochastic approach, finding cluster/community with mathematical bootstrapping procedure. Founder: Enright et al. Year: 2002 Publication: Enright, Anton J., Stijn Van Dongen, and Christos A. Ouzounis. An efficient algorithm for large-scale detection of protein families. Nucleic acids research 30.7 (2002): 1575-1584. Method: Statistical inference, Random walks This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: $O(n^3)$
12.	RB Pots	RB Pots using a quality function and try to optimize it with adjustment in community detection. Founder: Reichardt & Bornholdt Year: 2006 Publication: Reichardt, J., & Bornholdt, S. (2006). Statistical mechanics of community detection. Physical Review E, 74(1), 016110. 10.1103/PhysRevE.74.016110. Method: Modularity maximization

		This algorithm works for directed, weighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: unknown or not applicable.
13.	RBER Pots	RBER Pots is an improvement of RB Pots with several attributes addition. Founder: Reichardt & Bornholdt Year: 2006 Publication: Reichardt, J., & Bornholdt, S. (2006). Statistical mechanics of community detection. Physical Review E, 74(1), 016110. 10.1103/PhysRevE.74.016110. Method: Modularity maximization This algorithm works for undirected, weighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: unknown or not applicable.
14.	Significance Community	Significance community algorithm has a quality function to optimize, in which its also calculate the significance, a statistic measurement. Founder: Traag et al. Year: 2013 Publication: Traag, V. A., Krings, G., & Van Dooren, P. (2013). Significant scales in community structure. Scientific Reports, 3, 2930. *10.1038/srep02930* *http://doi.org/10.1038/srep02930*. Method: Modularity maximization, Statistical inference This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: unknown or not applicable.
15.	Spinglass	This algorithm based on Potts model with the theory that edges should connect nodes of the same spin state (in the same community) and nodes of different state (community) should not be connected. This algorithm use Potts Hamiltonian and Simulated Annealing.

		Founder: Reichardt & Bornholdt Year: 2006 Publication: Reichardt, Jörg, and Stefan Bornholdt. Statistical mechanics of community detection. Physical Review E 74.1 (2006): 016110. Method: Statistical inference This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: $O(n^{3.2})$.
16.	Surprise Community	Surprise community has the quality function to optimize with the fraction of internal edges and binary Kullback-Leibler divergence. Founder: Traag et al. Year: 2015 Publication: Traag, V. A., Aldecoa, R., & Delvenne, J.-C. (2015). Detecting communities using asymptotical surprise. Physical Review E, 92(2), 022816. 10.1103/PhysRevE.92.022816. Method: Statistical inference, Modularity maximization This algorithm works for directed, weighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: O(kn)
17.	Walktrap	Short distance random walks should be stay in the same community. Starting from an empty non-clustered network subset, the distance between adjacent nodes are calculated and two adjacent nodes are merged into one community. Recalculate and continue merging all community until stopping condition. Foounder: Pons & Latapy. Year: 2006 Publication: Pons, Pascal, and Matthieu Latapy. Computing communities in large networks using random walks. J. Graph Algorithms Appl. 10.2 (2006): 191-218. Method: Random walks, Hierarchical clustering

		This algorithm works for undirected, unweighted, non-bipartite, non-feature-rich, and non-temporal networks. Result from this algorithm is crisp communities. Complexity: $O(n^2 \log n)$