Supplementary text S2

Stochastic block model for rich cores in single-layer networks

Suppose we have N nodes and we want to construct a single-layer network from which we can identify a partition into two sets: a core of size $N_c < N$ and a periphery of size $Np = N - N_c$. Here we test the performance of the single-layer algorithm to detect rich cores [1] on a simple stochastic block model.

Let us consider N nodes from which N_c drawn at random are chosen to be part of the network core, whereas the remaining N_p are part of the periphery. A network with core-periphery structure is such that its adjacency matrix can be decomposed into four different blocks: a dense diagonal block encoding information on core-core links, a sparser diagonal block describing links among peripheral nodes, and two off-diagonal blocks encoding core-periphery edges.

In our block model, we connect two nodes with probability ρ_1 if they both belong to the core, with probability ρ_2 if one of them belongs to the core and one to the periphery, and with probability ρ_3 if they both belong to the periphery, $\rho_1 \geq \rho_2 \geq \rho_3$. Given a stochastic realization of the block model, we can extract the rich core of the network and compare it with the ground-truth, i.e. the set of nodes originally labeled as core nodes. In particular, we can test the accuracy of the algorithm for different choice of the parameters ρ_1, ρ_2 and ρ_3.

Given the three probabilities, the expected total number of edges connecting two core nodes is $K_{cc} = \rho_1[(N_c - 1) * N_c/2]$, the expected total number of edges connecting two peripheral nodes is $K_{pp} = \rho_3[(N - N_c - 1) * (N - N_c)/2]$, and the expected total number of edges connecting a node in the core and a node in the periphery $K_{cp} = \rho_2[N_c * (N - N_c)]$. The total number of links is $K = K_{cc} + K_{cp} + K_{pp}$.

In the case $\rho_1 = \rho_2 = \rho_3 = \rho$ the nodes are statistically indistinguishable from a structural point of view, the network lacks a core-periphery structure and specifying the value of ρ simply sets the expected average degree of the network $\langle k \rangle = N\rho$. For instance, for $N = 250$ and $\rho = 0.04$ we obtain $\langle k \rangle = 10$ and $K = 1250$. Of the different blocks of the adjacency matrix, the exact value of the density of the block encoding links between core and periphery nodes does not play a significant role [2]. For such a reason here we set $\rho_2 = 0.04$, and study the core-periphery structure of the network as a function of ρ_1, with $\rho_1 > \rho_2$. The higher the value of ρ_1, the stronger the core-periphery structure of the system. In order to control for the density of the network, as we increases ρ_1 we have to opportunely decrease the value of ρ_3. The average degree $\langle k \rangle$ can be kept fixed by setting

$$\rho_3 = \frac{2}{(Np) * (Np - 1)} \bigg(K - K_{cc} - K_{cp} \bigg). \tag{1}$$

In our case with $N = 250$ and $\langle k \rangle = 10$, we have $K = 1250$ whereas K_{cc} and K_{cp} are set once we fix the core size N_c and the value of ρ_1. In Fig. 1 we show the average Jaccard index J computed for the ground-truth partition and the partition extracted by the algorithm on the stochastic realizations of the network as a function of different values of ρ_1 for different core size.

As shown, J increases quickly until $\rho_1 = 0.2$ and only mildly after this point. This indicates that $\rho_1 = 0.2$, corresponding to a value of $\rho_3 = 0.03$, can be considered as the smallest density of the core-core block at which the core-periphery structure of the network is sufficiently well-defined. For this reason, in the stochastic block model for multiplex networks with different values of core similarity S_c described in Fig. **??** of the main text, where we have $N = 250$ and $N_c = 50$ we set $\rho_1 = 0.2$.

Given the set of parameters ρ_1, ρ_2 and ρ_3 we can also compute the average degree $\langle k_c \rangle$ of core nodes

$$\langle k_c \rangle = \rho_1(N_c - 1) + \rho_2(N_p), \tag{2}$$

the average degree $\langle k_p \rangle$ of the peripheral nodes

$$\langle k_p \rangle = \rho_3(N_p - 1) + \rho_2(N_c). \tag{3}$$

so that we have

$$\langle k \rangle = \frac{N_c \langle k_c \rangle + N_p \langle k_p \rangle}{N}. \tag{4}$$

In Fig. 2 we show the average Jaccard index J computed for the ground-truth partition and the partition extracted by the algorithm as a function of $\langle k_c \rangle / \langle k_p \rangle$. The Jaccard index J is defined as

$$J = \frac{I_c^{[\alpha\beta]}}{N_c^{[\alpha]} + N_c^{[\beta]} - I_c^{[\alpha\beta]}}, \tag{5}$$

where $N_c^{[\alpha]}$ is the number of core nodes at layer α, $N_c^{[\beta]}$ is the number of core nodes at layer β and $I_c^{[\alpha\beta]}$ is the number of nodes that are part of the core at both layers α and β.

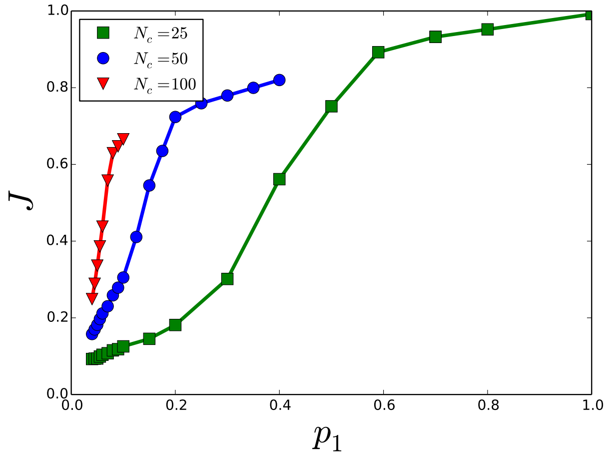

Figure 1: Jaccard index J for the groundtruth core-periphery partition and the partition obtained by the algorithm on realizations of the stochastic block model as a function of ρ_1 and for different core sizes N_c.

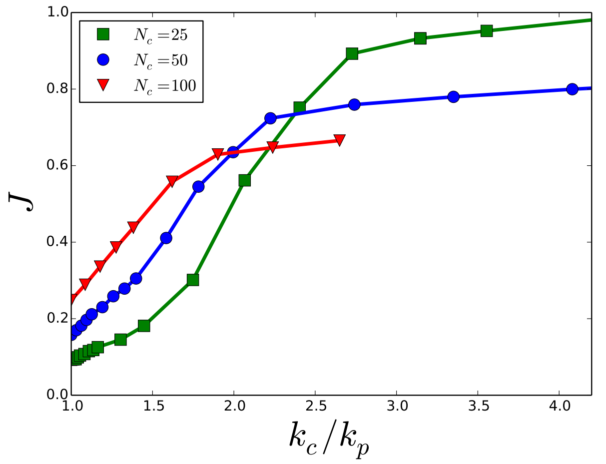

Figure 2: Jaccard index J for the ground-truth core-periphery partition and the partition obtained by the algorithm on realizations of the stochastic block model as a function of $\langle k_c \rangle / \langle k_p \rangle$ and for different core sizes N_c.

References

[1] A. Ma, R. J. Mondragn, Rich-Cores in Networks, PLOS ONE 10 (3) (2015) e0119678. doi:10.1371/journal.pone.0119678.

[2] S. P. Borgatti, M. G. Everett, Models of core/periphery structures, Social Networks 21 (4) (2000) 375–395. doi:10.1016/S0378-8733(99)00019-2.