

## 1 Một số tóm tắt tiền đề

- Giả thiết cơ bản (Fundamental Hypothesis): Các cộng đồng được mã hóa duy nhất trong sơ đồ liên kết của mạng. Chúng đại diện cho một sự thật tự nhiên vẫn đang được khám phá bằng các thuật toán thích hợp.
- Giả thiết liên thông và trù mật (Connectedness and Density Hypothesis): Một cộng đồng tương ứng với một đồ thị con được liên thông trù mật cục bộ.
- Giả thiết ngẫu nhiên (Random Hypothesis): Mạng có liên kết ngẫu nhiên thì không có cộng đồng.
- Giả thiết cực đại mô-đun hóa (Maximal Modularity Hypothesis): Phân vùng có tính mô-đun tối đa cung cấp cấu trúc cộng đồng tốt nhất, trong đó tính mô-đun được cung cấp bởi

$$M = \sum_{c=1}^{n_c} \left[ \frac{l_c}{L} - \left( \frac{k_c}{2L} \right)^2 \right] \quad (1)$$

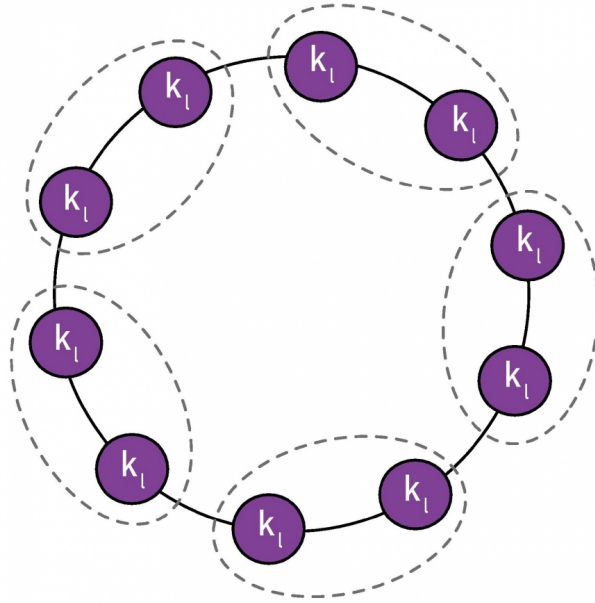
## 2 Các bài toán

A complex network graph with nodes colored purple, green, and orange, illustrating a hierarchical or clustered structure. The graph is composed of numerous nodes connected by edges, forming a dense, interconnected web. The nodes are arranged in a way that suggests a hierarchical or clustered organization, with a central core of nodes and many peripheral nodes branching out. The colors of the nodes (purple, green, and orange) likely represent different categories or classes within the network. The overall shape of the graph is roughly triangular, with a dense central region and more sparse, branching structures extending outwards.

**Problem 2.** Xem xét một lưới trong một chiều (1-dimensional) gồm  $N$  nút. Lưới này hình thành một đường tròn mà trong đó mỗi nút liên kết với hai láng giềng của nó. Phân hoạch đường tròn này thành  $n_c$  cụm liên tiếp nhau với kích thước  $N_c = \frac{N}{n_c}$ .

- (a) Tính toán modularity của phân hoạch thu được.
- (b) Dựa trên giả thiết cực đại modularity (Maximum Modularity Hypothesis), giá trị lớn nhất của  $M_c$  tương ứng với phân hoạch tốt nhất. Hãy đưa ra kích thước cộng đồng  $n_c$  tương ứng với phân hoạch tốt nhất.

**Problem 3.** Xem xét một mạng lưới bao gồm một vành  $n_c$  cliques (đồ thị con đầy đủ), mỗi clique có  $N_c$  nút và  $\frac{m(m-1)}{2}$  liên kết. Lân cận của các cliques được liên kết bởi một cạnh đơn. Hình bên dưới minh họa cho mô tả này. Mạng này có cấu trúc cộng đồng rõ ràng, mỗi



cộng đồng tương ứng với một clique.

- (a) Xác định modularity  $M_{single}$  của phân hoạch tự nhiên này, và modularity  $M_{pairs}$  của phân hoạch mà trong đó các cặp lân cận clique được gộp vào một cộng đồng đơn (như được mô tả trong hình trên bởi các đường đứt khúc).
- (b) Chứng minh rằng chỉ khi  $n_c < 2L$  thì môđun tối đa sẽ dự đoán phân vùng cộng đồng chính xác về mặt trực quan, trong đó

$$L = n_c \frac{m(m-1)}{2} + n_c \quad (2)$$

- (c) Đưa ra quan điểm khi vi phạm bất đẳng thức trên.

**Problem 4.** *Chứng minh rằng giá trị cực đại của modularity  $M$  mà được định nghĩa bởi*

$$M = \sum_{c=1}^{n_c} \left[ \frac{L_c}{L} - \left( \frac{k_c}{2L} \right)^2 \right] \quad (3)$$

*thì không vượt quá 1.*