NP - Hardness Proof

Viraj Parimi

August 2017

1 Problem Definition

Given a graph G(V, E) and a cost K, we need to identify a set of nodes S* in G such that :-

$$Cost(S^*) \ \leq \ K$$

[1]

$$\arg\max_{G} Modularity(G) - Modularity(G(V/S^*))$$

2 Reduction Problem

We will use *Maximum Vertex Coverage - Bipartite Graphs*(MVC-B) for reduction. This problem states the following,

Given a bipartite graph $G = (V, E), V = V_L \cup V_R, V_L \cap V_R = \emptyset$ and positive integers b, c, check if there exists a subset of vertices $S \subseteq V with \mid S \mid = b$ such that at least c edges are incident to nodes in S.

3 NP Completeness Proof

Given an instance of MVC-B with bipartite graph $G = (V, E)$, $V = V_L \cup V_R, V_L \cap V_R = \emptyset$ and positive integers b, c. Now we construct out problem instance $G' = (V', E')$ by connecting V_L, V_R to cliques K_L, K_R, respectively. We create an edge between all $\{(u, v)|u \in K_L, v \in V_L\}$ and all $\{(u, v)|u \in K_R, v \in V_R\}$. We choose the size of K_R, K_L in a way that when at least c edges are removed from E, $A(Algorithm)$ will detect two communities $K_L \cup V_L, K_R \cup V_R$ (and exclude the removed vertices) and one community otherwise. Let $k = b$ and,

$$a = \min_{S_L \in V_L, S_R \in V_R, |S_L| + |S_R| = b} Q(A(G'), Y')$$

where

$$Y' = \{K_L \cup V_L \setminus S_L, K_R \cup V_R \setminus S_R\}$$

[1]Current proof works on unit cost formulation

$$Q = Modularity(A) - Modularity(B)$$

Now assume we have a solution S, $\mid S \mid = k$ to our problem. As $Q(A(G'), A(G'[V' \setminus S]) \geq a, A(G'[V' \setminus S])$ must contain two communities. By construction, number of edges removed from E is at least c. If $S \subseteq V$ then we directly obtain the solution to MVC-B. If $\exists v \in S, v \notin V$, we can always find a vertex $u \in V, u \notin S$ and update S to S' $= S \cup \{u\} \setminus \{v\}$ while keeping the number of edges incident to S greater than c. Therefore, we have a solution $S', \mid S' \mid = k = b$ for MVC-B.

Now assume we have a solution $S, \mid S \mid = b$ to MVC-B. Then at least c edges in E are incident to vertices in S. If we remove all vertices in S, by construction, $A(G'[V' \setminus S)$ will output 2 communities, $K_L \cup V_L \setminus S'_L$ and $K_R \cup V_R \setminus S'_R$. Then we have,

$$Q(A(G'), A(G'[V' \setminus S]) \geq a$$

as a is the maximum Q value for communities in the form $\{K_L \cup V_L \setminus S_L, K_R \cup V_R \setminus S_R\}$. Therefore, we have a solution $S, \mid S \mid = b = k$ for our problem.

4 Conclusion

Since our problem statement has a solution if and only if *Maximum Vertex coverage - Bipartite Graph* has a solution, hence our problem statement is NP-Complete.