Proof of R2.O3

Anonymous

We prove that finding the optimal sampling strategy (for minimum expected τ-cover) costs unbearably when graph is large, by first proving that seeking a minimum expected t-cover is equivalent to a linear programming, and then showing the number of variables is exponential to the number of maximal cliques.

(a) (b)

Figure 1: Optimal solution

Firstly, we prove the problem equivalence with linear programming. We start with an example in Figure 1. Suppose that the entire graph consists of two maximal clique C_1 and C_2, which contains $m+x$ and $n+x$ vertices respectively. C_1 and C_2 intersects with x vertices. Now we calculate the optimal sampling probability for each vertex (in the three different areas) to produce an expected τ-cover with minimum size. The problem for finding the optimal probability is formalized as a linear programming:

$$\begin{aligned}
&\underset{p_m,p_n,p_x\in[0,1]}{\arg\min}\ (m\cdot p_m + n\cdot p_n + x\cdot p_x)\\
&\quad s.t.\ \begin{cases} m\cdot p_m + x\cdot p_x \geq \tau\cdot(m+x)\\ n\cdot p_n + x\cdot p_x \geq \tau\cdot(n+x)\end{cases}
\end{aligned}\tag{1}$$

W.l.o.g., we assume $m \geq n$.

Here we have used a fact that the m vertices obey an uniform sampling with probability p_m, similar for the x and n vertices. Now we use contradiction to prove such a fact. Given two vertices u and v among the m vertices, supposed that the optimal probability of sampling them are not equal:

$p_u(u) \neq p_v(v)$, and such a probability produces a minimum size cover. Now we exchange u with v. The sampling probability for them will be $p_v(u)$ and $p_u(v)$. However, since they are symmetry to each other, $p_u(u) = p_u(v), p_v(u) = p_v(v)$, the sampling algorithm produces a cover with the same size. Finally, we construct a new sampling probability for u and v to produce the cover with the same size again:

$$\begin{aligned} p(u) &= \frac{p_u(u) + p_v(u)}{2} \\ p(v) &= \frac{p_u(v) + p_v(v)}{2} \end{aligned} \tag{2}$$

One can verify that $p(u) = p(v)$, which is a uniform sampling for the two vertices. We can generalize such a case to m vertices (by replace the 2-vertex exchange with an m-vertex full permutation) and prove that if an optimal sampling exists, we can construct a uniform sampling which is also optimal. Now we assume that such a uniform sampling exists, and then see whether we can find it. The solution is shown as:

$$\begin{cases} p_x = \frac{m+x}{x} \cdot \tau,\ p_m = 0,\ p_n = 0, & \text{if } \ x \in [\frac{\tau}{1-\tau} \cdot m, +\infty) \\ p_x = 1,\ p_m = \tau - \frac{x}{m}(1-\tau),\ p_n = 0, & \text{if } \ x \in [\frac{\tau}{1-\tau} \cdot n, \frac{\tau}{1-\tau} \cdot m) \\ p_x = 1,\ p_m = \tau - \frac{x}{m}(1-\tau),\ p_n = \tau - \frac{x}{n}(1-\tau), & \text{if } \ x \in [0, \frac{\tau}{1-\tau} \cdot n) \end{cases} \tag{3}$$

It is inferred that when more cliques are added into the graph, then problem will transform into a linear programming with more variables (see Figure 1b).

Secondly, we show the number of variables when the graph contains more maximal cliques. If we add another clique to the graph (see Figure 1b), the solution to this linear programming will be much more complicated because there will be 7 variables in total. Given X to be the number of maximal cliques, one can verify that the number of variables can be $2^X - 1$. Such an exponential growth rate w.r.t. the clique number is unbearable, since the clique number itself is exponential to the number of vertices.

We finish this proof.