Algorithm Seminar: warm up session for art-gallery problem

Hirak Sarkar

April 3, 2015

1 March 13

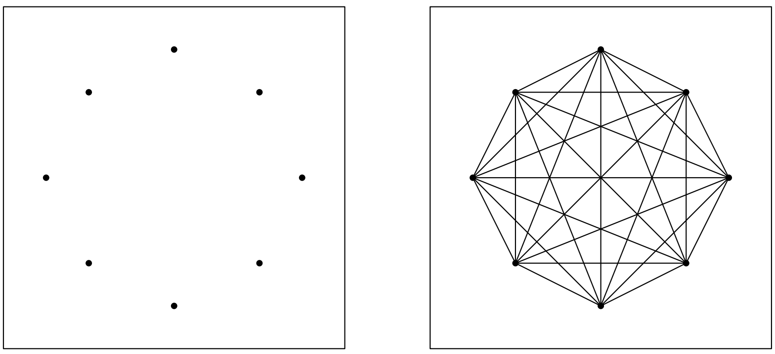
The question asked in the reading group was as follows,

Problem 1. Let G be a complete graph, and we are deleting edges from the graph uniformly, what is the expected number of deletions before we get a graph such that, maximum degree of a node is less than $n-1$.

Let us assume the number of such deletions is d. Now Prof. Skiena suggested, this problem can be approached from the reverse direction. For a graph G, we construct another graph $\overline{G}$ which the complement. That essentially means the edges present in G are absent in $\overline{G}$. It is shown in Figure 1

Now if we keep adding edges to $\overline{G}$, parallely the complete graph G is loosing edges. So adding an edge essentially means we are seeing an unseen vertex, and we would stop when we see the last unseen vertex. The situation is shown in Figure 2.

Let's assume the number of additions before seeing the last seeing vertex is a. We can claim that $a = d$, because when we see the last unseen vertex,



Empty Graph *Complete Graph*

Figure 1: Complement graph $\overline{G}$ and G

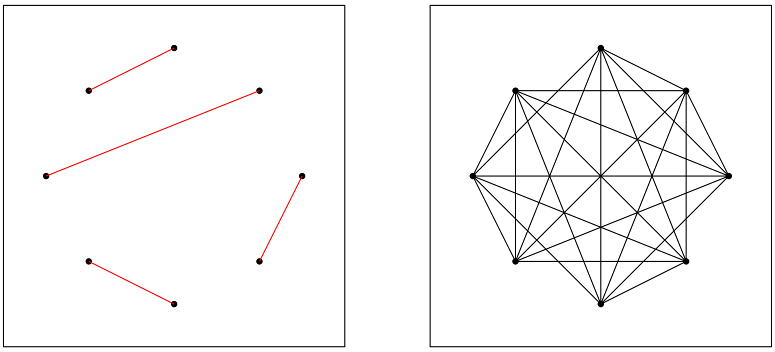


Figure 2: After addition of few edges to $\overline{G}$

then the original graph G will loose its last $(n-1)$-degree vertex. So the calculation of the d is now same as calculating a.

Now we can model 1 as follows,

Problem 2. Given an empty graph $\overline{G}$, what is the expected number of additions of edges after which we all nodes have degree at least 1.

Problem 2 can be viewed as a version of coupon collector's problem. The famous coupon collector's problem can be stated as follows,

Problem 3. Given n different coupons, how many coupons do you expect you need to draw with replacement before having drawn each coupon at least once?

The modified version could be formed by replacing the n different coupons with n vertices. Instead of "drawing" one coupon when we are adding edges, we are actually drawing two nodes. So the problem can be formally stated as,

Problem 4. Given n different coupons (vertices), and an urn containing packets of coupons where each packet contain two coupons, (edges) how many **packets** (edges) do you expect you need to draw with replacement before having drawn each coupon at least once?

Problem 4 has only one caveat, that is with the "replacement" policy, because when we add an edge in the empty graph then it is meaningless to add that edge again. But when we draw randomly the probability of taking a packet is low if it is already drawn, so for now we can ignore that gap.

The rest of the derivation is almost same as the coupon collector's problem. Although analysis would be a bit different. Let us recall the solution of Problem 3 first.

Let there be n different coupons. Now if we assume that to see the ith unseen coupon we need t_i trials. Then probability $P(t_i) = \frac{n-(i-1)}{n}$. Here distribution of t_i is geometric. Expectation, $E(t_i) = \frac{1}{P(t_i)}$. Expectation of seeing all the coupons is the summation over all i, from 1 to n. Which is a harmonic series and sums up to $n \log n$.

To generalize it to Problem 4, we will take a different approach. As the packets are of two now. There is always a possibility that we will see more than one unseen coupon in one trial. Let's assume if we have i distinct coupons, then we need $E(i)$ steps to see all the unseen coupons. At this step drawing a packet could mean one of three possibilities,

$$P(0 \text{ new coupon}|\text{i coupons are seen}) = \frac{\binom{i}{2}}{\binom{n}{2}}$$

,

$$P(1 \text{ new coupon}|\text{i coupons are seen}) = \frac{(n-i)i}{\binom{n}{2}}$$

and

$$P(2 \text{ new coupon}|\text{i coupons are seen}) = \frac{\binom{n-i}{2}}{\binom{n}{2}}$$

.

Now each trial would at 1 to the total number of trials. Total number of trials would be a recursive definition as follows,

$$E(i) = 1 + \frac{\binom{i}{2}}{\binom{n}{2}}E(i) + \frac{(n-i)i}{\binom{n}{2}}E(i+1) + \frac{\binom{n-i}{2}}{\binom{n}{2}}E(i+2)$$

From the definition $E(n) = 0$. We can solve it by back substitution, after few substitution steps, it converges to $n \log n$.

So for a random graph after $O(n \log n)$ expected deletions the dominating set would be more than 1.