Social and Technological Networks Edinburgh, 2018

Exercises 2. Cascades, approximations and submodular optimisations.

Rik Sarkar *Exercises*

Q 1. *Suppose $k = n/2$. Show that the number of possible subsets of size k is at least $2^{\Omega(n)}$.*

Answer. Divide $|V|$ into $n/2$ pairs. It is possible to create a set of size $n/2$ by choosing one element from each pair. There are $2^{n/2}$ possible ways to choose the set.

Q 2. *Suppose we want to find shortest paths. Is this is a maximization or minimization problem? What can you say about the approximation factor of an approximation algorithm for this problem?*

Q 3. *Write the proof that the greedy algorithm for submodular maximization produces a $(1 - 1/e)$ approximation.*

[Comment: The main elements of the proof are on the slides. You have to combine them into a formal proof.]

Q 4. *Suppose that we are doing social sensing. There are a set of social network accounts that publish news, which are reshared successively by others. Suppose we also have a probability on each edge that says how likely news is to propagate along that edge (i.e. shared by one node when it sees news shared by the other). How would you select your sensing nodes, or k nodes to follow, so as to maximize expected coverage of news?*

Answer. [Sketch of answer. Meaning you will be expected to answer more precisely using symbols and functions.] This can be framed as a submodular maximization problem, similar to what we did in class. Suppose $S \subseteq V$ is the set of nodes that are sources of news. Suppose $p : E \to [0, 1]$ the probability function of news being transmitted along each edge.

For each $v \in V$, we can compute the probability that the news from $s \in S$ will reach v. (How do you compute this?)

For a selected subset of k sensors, we can compute the Expected coverage of S provided by this set of k nodes. And we can now formulate this as a "Maximize expected coverage" by selecting sensing nodes. (Show that this problem is monotone, submodular.)