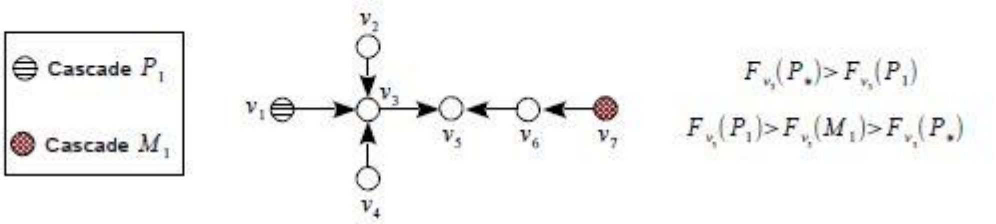
Homework 4

The Theory of social computing | Aadish Joshi

1. (2 points) Show that the objective function in the paper reading4-3 is neither submodular nor supermodular.



P1 is one positive cascade P1 and M1 is misinformation cascade

we deploy a new positive cascade P* and assume the candidate seed set V * is equal to V . Suppose that the probability on each edge is equal to 1, π(P1) = {v_1} and π (M1) = {v_7}, and the cascade priority at v_3 and v_5 is given as shown in the figure. We can observe that f({ϕ}) = 5, f({v_2}) = f({v_4}) = f({v_2; v_4}) = 4. Therefore, f({v_2}) < f(ϕ); and f({v_2}) + f({v_4}) < f({v_2} \ {v_4}) + f({v_2} ⌒ {v_4}): This illustrates that inappropriately selecting positive seed nodes may lead to a wider spread of misinformation.

2 (2 points) Show that the objective function in the paper reading4-2 has a monotone nondecreasing submodular upper bound and a monotone nondecreasing submodular lower bound.

For a seed set $\tau(P_*)$ of cascade P_*, the objective function is $f\left(\tau(P_*)\right)$ to denote the expected number of the $\bar{M}$-active nodes when the diffusion process terminates. Given a budget $k \in Z^+$ and a candidate set $V^* \subseteq V$, select a seed set $\tau(P_*) \subseteq V^*$ for P_* with $|\tau(P_*)| \leq k$ such that $f\left(\tau(P_*)\right)$ is maximized. The objective function is $f\left(\tau(P_*)\right)$ is not a submodular function. However, in some special cases, the objective function f is submodular.

The cascade priority is said to be homogeneous if each cascade has the same priority at each node. The cascade priority is said to be M-dominant if at each node, the priority of each misinformation cascade is higher than that of any positive cascade. The cascade priority is said to be P-dominant if at each node, the priority of each positive cascade is higher than that of any misinformation cascade. Then, we get two important results: (1) f is monotone nondecreasing and submodular if the cascade priority is M-dominant or P-dominant. (2) f is monotone nondecreasing and submodular if the cascade priority is homogeneous. Each cascade priority F_v induces another two cascade priorities, defined as follows:

$\bar{\mathrm{F}}_v$ is a cascade priority at node v induced by F_v, satisfying (a) for each $P_1, P_2 \in \mathrm{P} \cup \{P_*\}$, $\bar{\mathrm{F}}_v(P_1) < \bar{\mathrm{F}}_v(P_2)$ if $\mathrm{F}_v(P_1) < \mathrm{F}_v(P_2)$. (b) for each $M_1, M_2 \in \mathrm{M}$, $\bar{\mathrm{F}}_v(M_1) < \bar{\mathrm{F}}_v(M_2)$ if $\mathrm{F}_v(M_1) < \mathrm{F}_v(M_2)$. (c) for each $P \in \mathrm{P} \cup \{P_*\}$ and $M \in \mathrm{M}$, $\bar{\mathrm{F}}_v(M) < \bar{\mathrm{F}}_v(P)$. $\underline{\mathrm{F}}_v$ is a cascade priority at node v induced by F_v, satisfying (a) and (b) above, and for each $P \in \mathrm{P} \cup \{P_*\}$ and $M \in \mathrm{M}$, $\underline{\mathrm{F}}_v(M) > \underline{\mathrm{F}}_v(P)$. For a seed set $\tau(P_*) \subseteq V^*$ of cascade P_*, we use $\overline{f}(\tau(P_*))$ (resp. $\underline{f}(\tau(P_*))$) to denote the expected $\overline{M}$-active nodes when each node v replaces its cascade priority F_v by $\overline{\mathrm{F}}_v$ (resp. $\underline{\mathrm{F}}_v$). Because $\overline{\mathrm{F}}_v$ is P-dominant and $\underline{\mathrm{F}}_v$ is M-dominant, $\overline{f}$ and $\underline{f}$ are both monotone nondecreasing and submodular. Furthermore, $\overline{f}$ is an upper bound of f and $\underline{f}$ is a lower bound of f. For each $\tau(P_*) \subseteq V^*$, $\overline{f}(\tau(P_*)) \geq f(\tau(P_*)) \geq \underline{f}(\tau(P_*))$.

3 (2 points) Show that any set function has a monotone nonincreasing supermodular upper bound and a monotone nonincreasing supermodular lower bound.

Consider following definition of the set function.

A set function $f : 2^V \to R$ is called monotone increasing if $f(X) \leq f(Y)$ for any $X \subseteq Y$.

A set function $f : 2^V \to R$ is called submodular if $f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$ for any X and $Y \in 2^V$.

Lemma 2: γ () is monotone increasing and submodular.

By Lemma 2, $\gamma(A^{i1}(i_3, a)) - \gamma(A^{i1}) \geq \gamma(A^{i2}(i_3, a)) - \gamma(A^{i2})$, for any full-action A, $1 \leq i_i \leq i_2 \leq i_3 \leq k$ and $a \in A_{i3}$. Since $\gamma\text{-}(S) = \sum_A \Pr[A|S] \cdot \gamma(A)$.

A function $f : X \to R$ is supermodular if

$$(\forall x, y \in X) f(x) + f(y) \leq f(X \wedge Y) + f(X \cup Y)$$

lower bound is the worst case, expected error experienced when leader as a function.

$$F_c(S) = \max\{E_{\Pi s}(R(S|l_i)), c\}.$$

4 (2 points) Can you give a DS decomposition to the objective function in the paper reading4-2?

Partial set cover problem is defined as follows.

Problem 3 ($\pm$PSC problem). An instance of $\pm$PSC is a triplet (X, Y, Φ) where X and Y are two sets of elements with $X \cap Y = \emptyset$, and $\Phi = \{\phi_1, ..., \phi_m\} \subseteq 2^{X \cup Y}$ is collection of subsets over $X \cup Y$. For each $\Phi^* \subseteq \Phi$, its cost is defined as $|X \setminus (\cup_{\phi \in \Phi^*} \phi)| + |Y \cap (\cup_{\phi \in \Phi^*} \phi)|$. The $\pm$PSC problem seeks for a $\Phi^* \subseteq \Phi$ with the minimum cost.

Where objective function is defined as follows.

$$g(\Phi^*) = |X \setminus (\cup_{\phi \in \Phi^*} \phi)| + |Y \cap (\cup_{\phi \in \Phi^*} \phi)|.$$

Also according to read4-2,

$$f_{\mathrm{M}}(\Phi^*) = 3 + |X \setminus \cup_{\phi \in \Phi^*} \phi| + |Y \cap \cup_{\phi \in \Phi^*} \phi| = 3 + g(\Phi^*).$$

Suppose that ϕ* is an (alpha|V *|)-approximation to the Min-M problem for some (alpha|V*|) > 1. We have

$$f_{\mathrm{M}}(\Phi^*) \leq \alpha(|V^*|) \cdot f_{\mathrm{M}}(OPT) \iff 3 + g(\Phi^*) \leq \alpha(|V^*|) \cdot (3 + g(OPT))$$
$$\iff \frac{g(\Phi^*)}{g(OPT)} \leq \alpha(|V^*|) + \frac{3(\alpha(|V^*|) - 1)}{g(OPT)} \implies \frac{g(\Phi^*)}{g(OPT)} \leq 4\alpha(|V^*|) - 3.$$

Since |V*| = | ϕ | = m, ϕ* is (4.alpha(m) -3) approximation to the instance of PSC problem.

Algorithm 1.

Input: a function f over a ground set U and budget k:

$V_0 \leftarrow \phi$.

for $i = 1 : k$ do

$u \leftarrow \arg\max_{u \in Vh} (U_i + U\{u\}) - h(U_{i+1}$

$U_i \leftarrow U_{i-1} \cup \{u\}$

return $U \leftarrow \arg\max_{U_i} h(U_i)$.

Algorithm 2.

Input: $f \cdot \bar{f} \cdot \underline{f}$ V^*, k

$\bar{S}_A \leftarrow ALG1(\bar{f}, V^*, k)$:

$\underline{S_A} \leftarrow ALG1(\underline{f}, V^*, k)$.

$S_A \leftarrow ALG1(f, V^*, k)$

return $\acute{S}' = \arg\max_{S \in} (\bar{S}_A, \underline{S_A}, S) f(S)$

5 (2 points) Design a method to apply the submodular-supermodular procedure to the maximization problem in the paper reading4-2 in case that you cannot find a DS decomposition for the objective function.

Every set function f: 2X -> R can be expressed as difference of two monotone nondecreasing submodular function g and h. i.e. f = g – h. where x is a finite set

According to the definition, any set function can be expressed as a DS function. That is, for set function h, there exist two submodular functions f and g such that h = f – g.

We select a submodular function g such that alpha(g) > 0.

And f(x) = h(x) + |alpha(h)| / alpha(g) . g(X). Then alpha(f) > 0

H = f – h = f - |alpha(h)| / alpha(g). g

We g(x) = sqrt(|X|)

Also we define, if alpha(h) >= 0, then h is submodular.

$$\alpha(h) = \min_{X \subset Y \subseteq V \setminus j} \{\Delta_j h(X) - \Delta_j h(Y)\}$$