Proposal and Referee Biclustering for Panel Formation

Enver Kayaaslan, Cevdet Aykanat

April 19, 2011

We are given a set of referees and and a set of proposals to be coclustered such that the proficiency between the referees and the proposals within a cluster is maximized. Due to capacity of the referees, one should consider an additional constraint that balances the number of proposals and the number of referees within each cluster, as well. Current biclustering approaches do not capture this balance constraint. As a remedy to this deficiency, in this study, we redefined the problem and proposed a graph partitioning based model to solve.

We are given a set $\mathcal{R} = \{r_1, r_2, \ldots, r_n\}$ of n referees, a set $\mathcal{P} = \{p_1, p_2, \ldots, p_m\}$ of m proposals, and a function $f : \mathcal{R} \times \mathcal{P} \rightarrow [0 - 1]$ regarding to proficiency level between referees and proposals. That is, $f(r_i, p_j)$ refers to how referee r_i is proficient to evaluate proposal p_j. Let $\Pi_{\mathcal{R}} = \{R_1, R_2, \ldots, R_K\}$ and $\Pi_{\mathcal{P}} = \{P_1, P_2, \ldots, P_K\}$ be a K-way partition of referee set $\mathcal{R}$ and propsal set $\mathcal{P}$, respectively. We define the cutsize $(\Pi_{\mathcal{R}}, \Pi_{\mathcal{P}})$ of a given referee and proposal partition pair $(\Pi_{\mathcal{R}}, \Pi_{\mathcal{P}})$ as follows.

$$c(\Pi_{\mathcal{R}}, \Pi_{\mathcal{P}}) = \sum_{\substack{(r,p) \in \mathcal{R} \times \mathcal{P} \\ \pi_{\mathcal{R}}(r) \neq \pi_{\mathcal{P}}(p)}} f(r, p), \tag{1}$$

where $\pi_{\mathcal{R}}(r)$ ve $\pi_{\mathcal{P}}(p)$ refers to the index of the part that referee r and proposal p are assigned to, respectively. Given these definitions, the referee-proposal biclustering problem is redefined as follows.

Problem 1 (Referee-Proposal Biclustering Problem) *Given a referee set $\mathcal{R}$, a proposal set $\mathcal{P}$, a proficiency function f, a number R_{max} regarding to the maximum number of referees that a cluster can hold, and the maximum allowable referee-proposal imbalance ϵ, find a referee and proposal partition pair $(\Pi_{\mathcal{R}}, \Pi_{\mathcal{P}})$ minimizing the cutsize $c(\Pi_{\mathcal{R}}, \Pi_{\mathcal{P}})$ such that,*

1. $K = |\Pi_{\mathcal{R}}| = |\Pi_{\mathcal{P}}|$
2. $|R_k| \leq R_{max}$, *for* $1 \leq k \leq K$
3. $\frac{|\mathcal{P}|}{|\mathcal{R}|} \leq \max_k\{\frac{|P_k|}{|R_k|}\} \leq (1+\epsilon)(\frac{|\mathcal{P}|}{|\mathcal{R}|})$

The multi-capacity constrainted graph partitioning problem is defined as, for a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where each vertex is associated with a weight vector $\vec{w}(v_i)$ and each edge is associated with a cost $c(e_j)$, to find a K-way graph partition $\Pi_{\mathcal{G}} = \{V_1, V_2, \ldots, V_K\}$ maximizing the cutsize, i.e. the total costs of the edges in the cut, such that the total weights of each part does not exceed a given capacity vector $\vec{W}_{max}$, i.e., $\sum_{v_i \in V_k} \vec{w}(v_i) \leq \vec{W}_{max}$. We model the Referee-Proposal Biclustering Problem as a multi-capacity constrainted graph partitioning problem (with two-dimensional weights) as following algorithm.

Algorithm 1 Referee-Propoosal Biclustering

Require: $\mathcal{R}, \mathcal{P}, f, W_{max}, \epsilon$

$\vec{w}(r) \leftarrow (1, 0), \forall r \in \mathcal{V}_{\mathcal{R}}$

$\vec{w}(p) \leftarrow (0, 1), \forall p \in \mathcal{V}_{\mathcal{P}}$

$\mathcal{E} \leftarrow \{(r, p) \in V_{\mathcal{R}} \times \mathcal{V}_{\mathcal{P}} : f(r, p) > 0\}$

$\vec{W} \leftarrow (W_{max}, (1 + \epsilon)\frac{|\mathcal{P}|}{|\mathcal{H}|} W_{max})$

$\Pi_{\mathcal{G}} \leftarrow$ Graph-Partitioning$(G(\mathcal{V}_{\mathcal{R}} \cup \mathcal{V}_{\mathcal{P}}, \mathcal{E}), f, \vec{w}, \vec{W})$

$(\Pi_{\mathcal{R}}, \Pi_{\mathcal{P}}) \leftarrow$ Clustering-Refinement$(\Pi_{\mathcal{G}}, \mathcal{R}, \mathcal{P}, f, W_{max}, \epsilon)$

return $(\Pi_{\mathcal{R}}, \Pi_{\mathcal{P}})$