Thesis notes

30th March

The echo chamber problem - notation

- $G = (V, E^+, E^-)$ interaction graph
- $\mathcal{C}$ set of contents
- $C \in \mathcal{C}$ content, $\mathcal{T}_C$ set of threads associated with C. A thread $T \in \mathcal{T}_C$ is a subgraph of G
- $U \subseteq V$ subset of users, $T[U]$ subgraph of T induced by U. $|T(U)|$ is the number of edges of this subgraph

The echo chamber problem - notation

- $\eta(C)$ fraction of negative edges associated with C (analogous definition for a thread T). Content (or thread) controversial if $\eta \in [\alpha, 1]$
- $\hat{\mathcal{C}} \subseteq \mathcal{C}$ set of *controversial* contents
- $\mathcal{S}_C(U)$ set of *non controversial* threads induced by U, for *controversial* contents, i.e.

$$\mathcal{S}_C(U) = \{T[U] \; s.t. \; T[U] \; non \; controversial, T \in \mathcal{T}_C, C \in \hat{\mathcal{C}}, U \subseteq V\} \tag{1}$$

The echo chamber problem

Goal: given an interaction graph G, find $U \subseteq V$ maximing

$$\xi(U) = \sum_{C \in \hat{\mathcal{C}}} \sum_{T[U] \in S_C(U)} |T[U]| \tag{2}$$

The set of users maximing the expression is denoted as $\hat{U}$ and the corresponding score is $\xi(G)$

Computing exactly the score

$$\textit{maximize} \sum_{ij \in E(\hat{\mathcal{C}})} x_{ij} \tag{3}$$

$$x_{ij} \leq y_i \quad \forall ij \in E(\hat{\mathcal{C}}) \tag{4}$$

$$x_{ij} \leq y_j \quad \forall ij \in E(\hat{\mathcal{C}}) \tag{5}$$

$$x_{ij} \leq z_k \quad \forall ij \in E(T_k), T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}} \tag{6}$$

$$x_{ij} \geq -2 + y_i + y_j + z_k \quad \forall ij \in E(T_k), T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}} \tag{7}$$

$$\sum_{ij \in E^-(T_k)} x_{ij} - \alpha \sum_{ij \in E(T_k)} x_{ij} \leq 0 \quad \forall T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}} \tag{8}$$

$$y_i \in \{0, 1\} \quad \forall i \in V \tag{9}$$

$$0 \leq x_{ij} \leq 1 \quad \forall ij \in E(\hat{\mathcal{C}}) \tag{10}$$

$$0 \leq z_k \leq 1 \quad \forall T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}} \tag{11}$$

Computing exactly the score

$y_i = 1$ means that $y_i \in U$.
$z_k = 1$ means that thread k is non controversial.
$x_{ij} = 1$ means that the edge contributes to the score.

The following constraints enforce that only edges ij whose both vertices are in U and are associated to a non controversial thread can contribute.

$$x_{ij} \leq y_i \quad \forall ij \in E(\hat{\mathcal{C}})$$

$$x_{ij} \leq y_j \quad \forall ij \in E(\hat{\mathcal{C}})$$

$$x_{ij} \leq z_k \quad \forall ij \in E(T_k), T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}}$$

Computing exactly the score

Also, considering edges contributing to the score, they must not produce a controversial thread

$$\sum_{ij \in E^-(T_k)} x_{ij} - \alpha \sum_{ij \in E(T_k)} x_{ij} \leq 0 \quad \forall T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}}$$

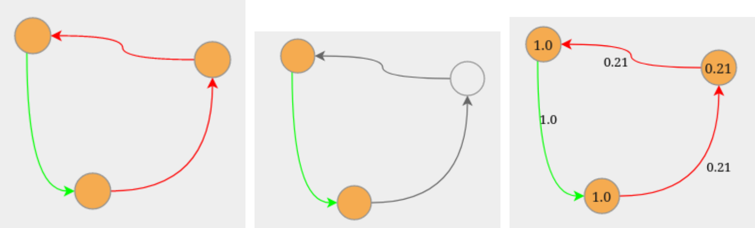
If $T_k[U]$ is controversial, then x_{ij} must be 0 $\forall ij \in E(T_k[U])$.

In general either all $ij \in E(T_k[U])$ are 0 or they are 1.

$$x_{ij} \leq z_k \quad \forall ij \in E(T_k), T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}}$$

$$x_{ij} \geq -2 + y_i + y_j + z_k \quad \forall ij \in E(T_k), T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}}$$

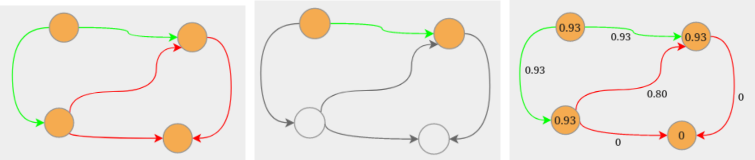
Computing exactly the score - relaxation example



(a) Original graph (b) Integer solution (c) Relaxed solution

Figure: Model solution for a graph with a single thread, for $\alpha = 0.3$

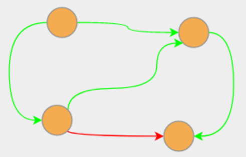
Computing exactly the score - relaxation example



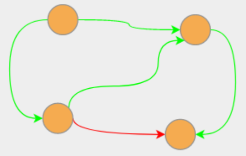
(a) Original graph (b) Integer solution (c) Relaxed solution

Figure: Model solution for a graph with a single thread, for $\alpha = 0.3$

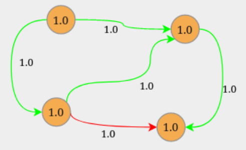
Computing exactly the score - relaxation example



(a) Original graph



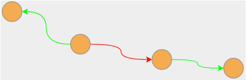
(b) Integer solution



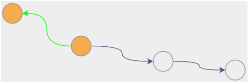
(c) Relaxed solution

Figure: Model solution for a graph with a single thread, for $\alpha = 0.3$

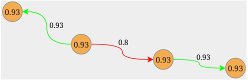
Computing exactly the score - relaxation example



(a) Original graph



(b) Integer solution



(c) Relaxed solution

Figure: Model solution for a graph with a single thread, for $\alpha = 0.3$

Reconstructing solutions from relaxation results

Let

$$U(r) := \{i \in V \ s.t. \ y_i \geq r\}$$

$$E(r) := \{ij \in E \ s.t. \ x_{ij} \geq r\}$$

$$\tilde{R} := \{x_{ij} \ \forall ij \in E\}$$

Algorithm 1: Relaxation results reconstruction

foreach $r \in \tilde{R}$ **do**
    while $(V(r), E(r)) \neq T[V(r)]$ **do**
        remove a vertex that misses one or more edges from $V(r)$ and its corresponding edges from $E(r)$;
    end
    Calculate $\xi(V(r))$
end

Evaluating echo chamber

Results of graph G can be evaluated against random shuffling of the graph G_r.

Similarly to [AW18], signs are allocated randomly on the same underlying structure while keeping the same fraction of negative edges. This is done separately for each thread.

Let $\xi(G_r)$ and $\sigma(G_r)$ be the average and standard deviation of ξ across the many G_r, respectively.
They use Z-score

$$Z = \frac{\xi(G) - \xi(G_r)}{\sigma(G_r)} \tag{12}$$

Bibliography

[AW18] Samin Aref and Mark C Wilson. "Balance and frustration in signed networks". In: *Journal of Complex Networks* 7.2 (Aug. 2018). Ed. by ErnestoEditor Estrada, pp. 163–189. ISSN: 2051-1329. DOI: 10.1093/comnet/cny015. URL: http://dx.doi.org/10.1093/comnet/cny015.