ROBUST INFLUENCE MAXIMIZATION FOR HYPERPARAMETRIC MODELS

Dimitris Kalimeris
Harvard University

Joint work with Gal Kaplan and Yaron Singer
ICML 2019
DIFFUSION IS MODELED AS INDEPENDENT CASCADE

Independent cascade (IC) [KKT 03]:
- Each active node tries to influence its neighbors independently
- Diffusion proceeds in discrete steps

Why is this model nice?
1. It is approximable: simple greedy alg obtains constant apx [KKT 03]
2. It is learnable: we can learn the diffusion probabilities [NPS 15]
INFLUENCE MAX IS NOT ROBUST TO LEARNING ERRORS

True model

\[ E[\#\text{nodes infected}] = n \]
INFLUENCE MAX IS NOT ROBUST TO LEARNING ERRORS

Learned model

\[
\mathbb{E}[\#\text{nodes infected}] = \text{const}
\]

\[
\frac{\mathbb{E}[\#\text{nodes estimated to be infected}]}{\mathbb{E}[\#\text{nodes that are really infected}]} \rightarrow 0 \quad n \rightarrow \infty
\]
ROBUST OPTIMIZATION TO THE RESCUE

Confidence intervals for each edge: \( P = \times_{e \in E} [l_e, u_e] \)

\[
\max_S \min_{\rho \in P} f_{\rho}(S)
\]

# nodes influenced by S when probs are p

- \( f \) is monotone in \( p \)
- The problem is trivial

What is the right formulation of the robust optimization problem?
FORMULATE ROBUST OPTIMIZATION VIA HYPERPARAMETER

Hyperparametric model: Every edge is associated with feature $x_e \in [-1,1]^d$ and $p_e$ is determined by hyperparameter $\theta \in \Theta \subseteq [-1,1]^d$

$$p_e = \sigma(\theta \cdot x_e) = \frac{1}{1 + e^{-\theta \cdot x_e}}$$

Example features: gender, age, location, degree, pagerank, etc.

Robust IM restated: $\max_S \min_{\theta \in \Theta} f_{\theta}(S)$
MAIN RESULT

Let $F = \{f_\theta : 2^V \to R \mid \theta \in \Theta \subseteq [-B, B]^d\}$ be a family of influence functions. There exists a randomized poly-time algorithm that produces a solution $\hat{S}$ s.t.

$$\min_{\theta \in \Theta} \mathbb{E}[f_\theta(\hat{S})] \geq (1 - 1/e) \max_{S:|S|\leq k} \min_{\theta \in \Theta} f_\theta(S) - \epsilon$$

Details about the algorithm and experiments in poster #268

Thanks!