Online Learning with Sleeping Experts and Feedback Graphs

Corinna Cortes[1], Giulia DeSalvo[1], Claudio Gentile[1], Mehryar Mohri[1,2], and Scott Yang[3]

1 Google Research, New York, NY
2 Courant Institute, New York, NY
3 D. E. Shaw & Co., New York, NY

Sequential Prediction

At round $t \in [T]$,

- A pair $(x_t, y_t) = z_t \sim \mathcal{D}$ is drawn i.i.d.
- Learner sees context x_t.
- Learner selects an expert out of set: $I_t \in \{\xi_j : j \in [K]\}$.
- Learner incurs loss : $L_t(\xi_{I_t}, z_t)$.

Sleeping Experts

Sleeping experts: only a subset of experts are available/awake at each round.

At round $t \in [T]$,

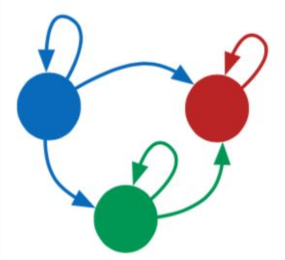
- A pair $(x_t, y_t) = z_t \sim \mathcal{D}$ is drawn i.i.d.
- Learner sees context x_t.
- Learner selects an expert out of an awake set : $I_t \in A_t \subseteq \{\xi_j : j \in [K]\}$.
- Learner incurs loss : $L_t(\xi_{I_t}, z_t)$.

Feedback Graphs

Feedback graph: losses observed by the learner modeled by a graph

At round $t \in [T]$,

- A pair $(x_t, y_t) = z_t \sim \mathcal{D}$ is drawn i.i.d.
- Learner sees context x_t.
- Learner selects an expert out of an awake set : $I_t \in A_t \subseteq \{\xi_j : j \in [K]\}$.
- Learner incurs loss : $L_t(\xi_{I_t}, z_t)$.
- Learner sees loss of chosen expert and others within its out-neighborhood as defined by a feedback graph.



Motivation

Web advertising:

- Feedback graph: related ads have similar rewards.
- Sleeping experts: ads availability changes.

Sensor networks:

- Feedback Graphs: sensor area can overlap.
- Sleeping experts: sensors may be broken.

Losses and awake sets can be dependent: can we design an algorithm with favorable guarantees that works well in practice?

Our Contribution for Two Settings

Independent awake sets and losses: feedback graph extension of AUER algorithm (Kleinberg et al. 2008); favorable guarantee with matching lower bound.

Dependent awake sets and losses

- General regret definition based on conditional expectations
  - Coincides with standard regret definition in the independent case
- Novel algorithm based on conditional expected losses of experts with favorable regret guarantees:

$$R_T^{\text{SLEEP}} = O\left(\sum_{k=1}^{p} p_k \min_{\mathcal{C}_k} \sum_{C \in \mathcal{C}_k} \frac{\max_{j \in C} \bar{\Delta}_{k,j}}{\min_{j \in C} \bar{\Delta}_{k,j}^2} \log(T)\right)$$

- Application to online abstention: novel algorithm outperforming state-of-the-art in an extensive suite of experiments.

Poster #152 in Pacific Ballroom