Bandits on Stochastic Blockmodel Graphs

Arthur Mensch, Michaël Weiss

Reinforcement Learning

January 2015

Table of contents

Problem definition

N arms, T steps.

1. The environment chooses losses for every arm noted $l_{t,i}$ for the arm i at the step t.

2. Following the algorithm we hope would minimize as much as possible the regret the player draws an arm I_t.

3. The player receives the loss l_{t,I_t}.

4. We define $(O_t)_{i \in [N]}$ as the indicative function of observed loss at step t. We have:

$$O_{t,I_t} = 1 \qquad \forall i \neq I_t, \ O_{t,i} \sim B(r)$$

$(O_t)_{i \in [N]}$ corresponds to the value of the logic expression i is neighbor of I_t in the Erdös-Rényi random graph drawn at step t.

5. For all i such that $O_{t,i} = 1$ the player can observe the loss $l_{t,i}$.

Duplex tricks

Definition geometrical random variables

$$M_t^* = \min\{1 \le i < N : O_{t-1,i}^{'} = 1\} \cup \{N\}$$
$$G_{t,i} = \min\left(K_{t,i}, M_t\right)$$

Independence

$$p_{t+2,i} \propto w_{t+2,i} = \frac{1}{N} \exp\left(-\eta_{t+2}\hat{L}_{t,i}\right)$$

Generalizing

ESTIMATE_R gives a safe lower bounding on r.

If $\underline{r} = 0$,   we run vanilla EXP3.
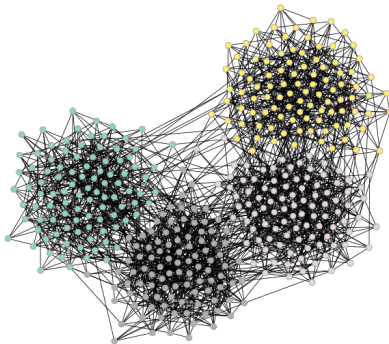
If $\underline{r} \geq \dfrac{\log T}{N}$,   we run vanilla DUPLEX.

If $0 < \underline{r} < \dfrac{\log T}{N}$, $A = \left\lceil \dfrac{\log T}{N\underline{r}} \right\rceil$ and run DUPLEX grouping A steps.

Table of contents

Random graph with Stochastic Block-Model

Neighbor communities

Probability of side information

- N clusters
- Labeled vertices
- $R = (r_{ij})_{1 \leq i,j \leq n}$ the matrix representing the probability of having side information.
- r_{ij} represents the probability that a vertex of the cluster j reveals his loss to a given vertex of the cluster i

Adapt M_t sampling

Unbiased estimator

$$\hat{l}_{t,i}^* = \frac{O_{t,i} l_{t,i}}{p_{t,i} + (1 - p_{t,i}) r_{C(I_t), C(i)}}$$

Using data...

$$\hat{l}_{t,i} = G_{t,i} O_{t,i} l_{t,i} \tag{1}$$

- Estimator $G_{i,t} \sim G(p_{t,i} + (1 - p_{t,i}) r_{I_t, i})$
- M_t^i truncated geometric law of parameter $p_{t,i}$
- $\rightarrow G_{t,i} = \min(K_{t,i}, M_t^i)$

Sampling M_t^i

Algorithm adaptation

- Using previous observations
- Ensure independence of M_t and O_t
- Find the last $O_{t'}$, with $t' \equiv t - 1[2]$ in which $C(I(t')) = C(I(t))$
- Average the distance between two observation within cluster $C(j)$.

Keeping track of last time we picked I in cluster i

Generalized algorithm

- We adapt ESTIMATE_R so that it returns *underlineR* where $\underline{r}_{ij} \leq r_{ij}$ with high probability.
- This is done at the relatively low cost
  $$\mathbb{E}(\tau) \leq \sum_{i,j \in [N]^2} \frac{4 \log T}{N_i} \mathbf{1}_{r_{ij} > \frac{1}{N_j}} + \sqrt{T} \mathbf{1}_{r_{ij} < \frac{1}{N_j}} + \mathrm{Cste}$$

If $\underline{R} = 0$,   we run vanilla EXP3.

If $\min\limits_{i,j,\ \underline{r}_{i,j} > 0} \left(\underline{r}_{ij} N_j \right) \geq \log T$,   we run adapted DUPLEX.

Else,
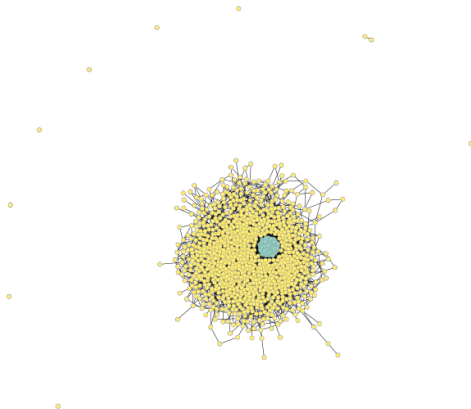$$A = \left\lceil \frac{\log T}{r_{N*}} \right\rceil \text{ where}$$
$$r_{N*} = \min\limits_{i,j,\ \underline{r}_{i,j} > 0} \left(\underline{r}_{ij} N_j \right)$$

and run DUPLEX grouping A steps.

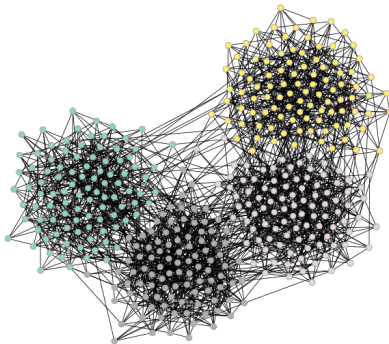# Table of contents

Heavily unbalanced graph

Neighbor communities

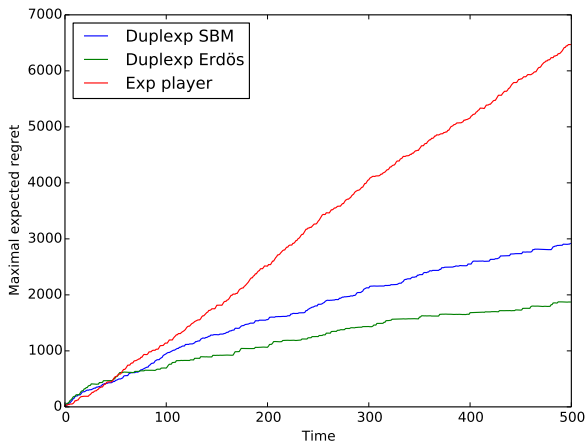# Graph 'close' to Erdös-Rényi
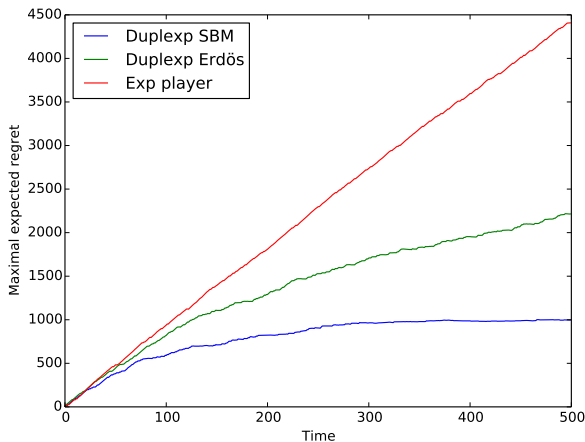
Orginal algorithm outperforms our algorithm

Graph 'close' to Erdös-Rényi

Neighbor communities

Graph 'close' to Erdös-Rényi

Orginal algorithm outperforms our algorithm