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

$$
\begin{aligned}
M_t^* &= \min\{1 \leq i < N : O'_{t-1,i} = 1\} \cup \{N\} \\
G_{t,i} &= \min\left(K_{t,i}, M_t\right)
\end{aligned}
$$

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.

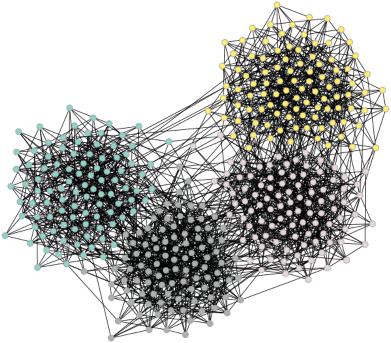
If $\underline{r} = 0$, we run vanilla Exp3.

If $\underline{r} \geq \frac{\log T}{N}$, we run vanilla Duplex.

If $0 < \underline{r} < \frac{\log T}{N}$, $\quad A = \left\lceil \frac{\log T}{N\underline{r}} \right\rceil$ 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 $p_{t,i}$

$\rightarrow$ $G_{t,i} = \min\left(K_{t,i}, M_t^i\right)$

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 $\textsc{Estimate\_R}$ so that it returns $\underline{R}$ 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 $\textsc{Exp3}$.

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

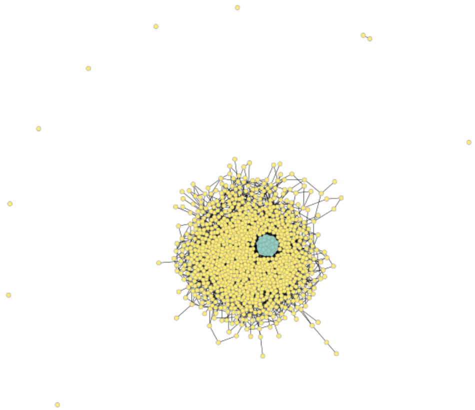
Else, $A = \left\lceil \frac{\log T}{r_{N*}} \right\rceil$ where

$$r_{N*} = \min_{i,j,\ \underline{r}_{i,j} > 0} \left(\underline{r}_{ij} N_j \right)$$

and run $\textsc{Duplex}$ grouping A steps.

Table of contents

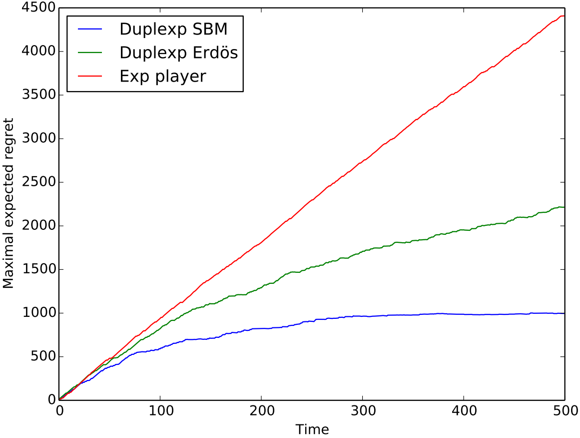
Heavily unbalanced graph



«««¡ HEAD

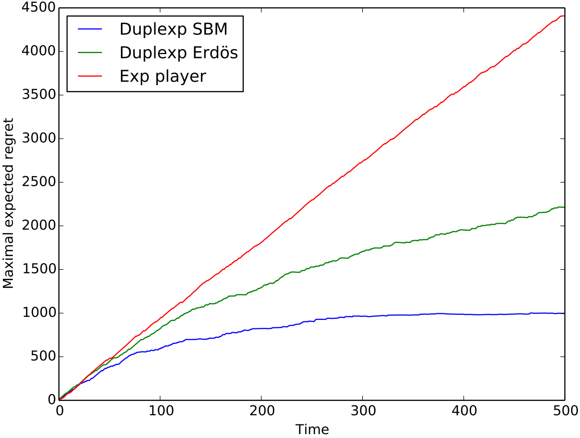
Graph (at one t)

Graph 'close' to Erdös-Rényi



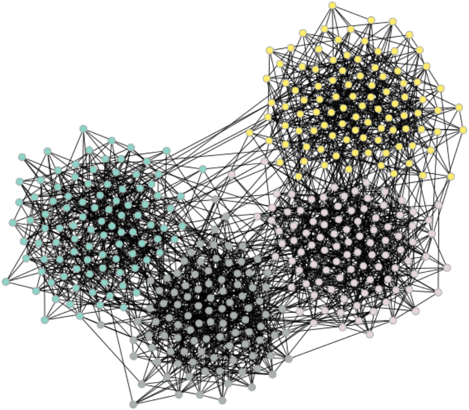
Easy problem

Graph 'close' to Erdös-Rényi



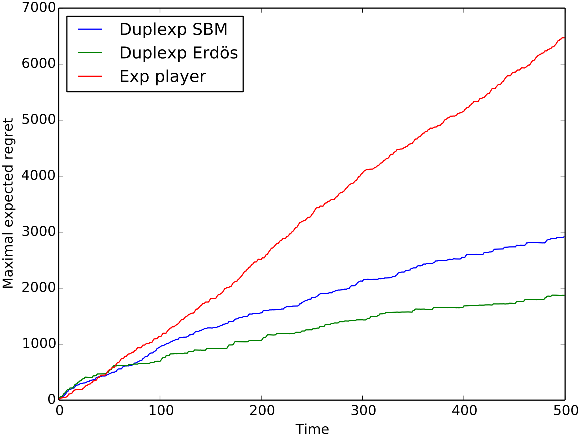
Hard problem

Graph 'close' to Erdös-Rényi



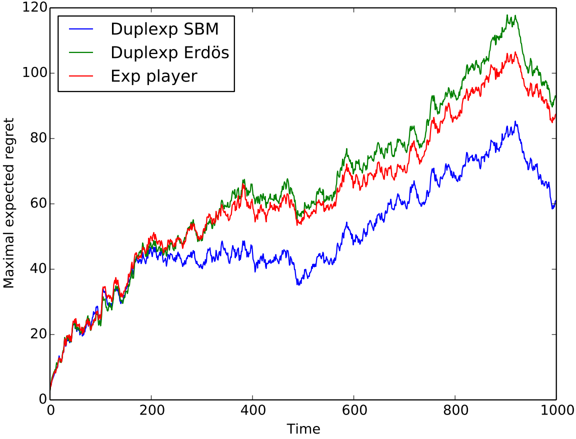
Graph (at one t)

Graph 'close' to Erdös-Rényi



Easy problem

Graph 'close' to Erdös-Rényi



Hard problem