Bandits on stochastic block-model graphs

Arthur Mensch, Michaël Weiss

January 14, 2015

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

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

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 $\underline{r} \geq \frac{log\ \ T}{N}$, 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.

Heavily unbalanced graph

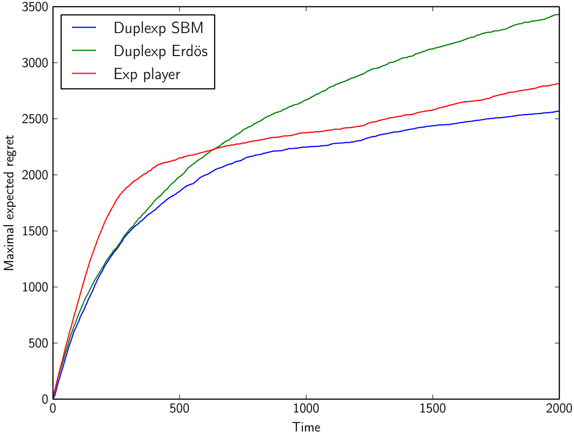

Figure : Outperformance of our algorithm

Graph 'close' to Erdös-Rényi

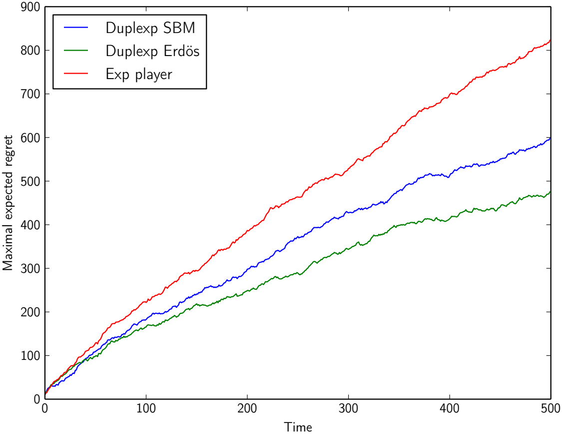

Figure : Orginal algorithm outperforms our algorithm