Optimal Discovery with Probabilistic Expert Advice [JMLR - arXiv:1110.5447]

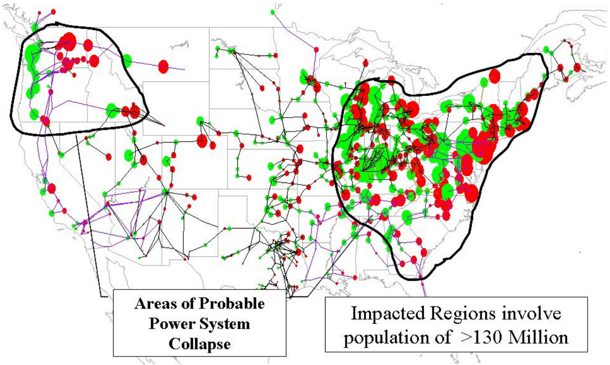
Sébastien Bubeck, Damien Ernst and Aurélien Garivier

January 12th, 2016

Outline

The problem

Power system security assessment

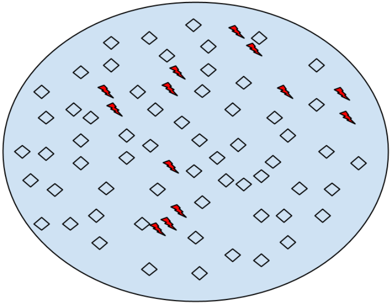


By Mark MacAlester, Federal Emergency Management Agency [Public domain], via Wikimedia Commons

Identifying contingencies/scenarios that could lead to unacceptable operating conditions (dangerous contingencies) if no preventive actions were taken.

The model

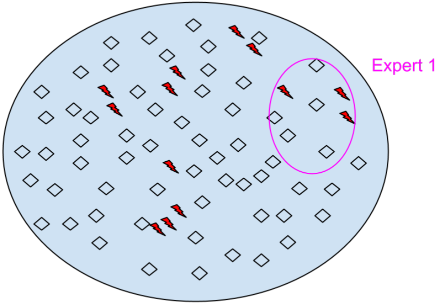
- Subset $A \subset \mathcal{X}$ of important items
- $|\mathcal{X}| \gg 1$, $|A| \ll |\mathcal{X}|$
- Access to $\mathcal{X}$ only by probabilistic experts $(P_i)_{1 \leq i \leq K}$: sequential independent draws



Goal : discover rapidly the elements of A

The model

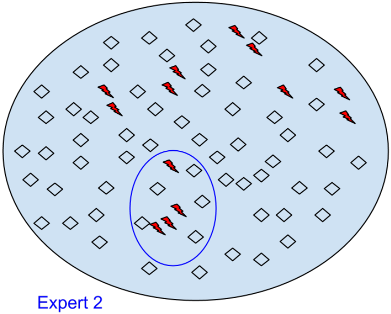
- Subset $A \subset \mathcal{X}$ of important items
- $|\mathcal{X}| \gg 1$, $|A| \ll |\mathcal{X}|$
- Access to $\mathcal{X}$ only by probabilistic experts $(P_i)_{1 \leq i \leq K}$: sequential independent draws



Goal : discover rapidly the elements of A

The model

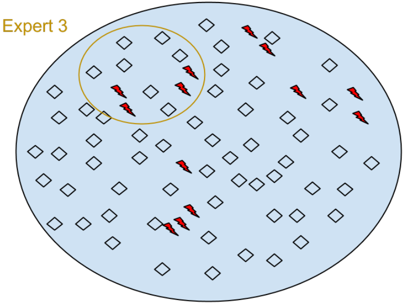
- Subset $A \subset \mathcal{X}$ of important items
- $|\mathcal{X}| \gg 1$, $|A| \ll |\mathcal{X}|$
- Access to $\mathcal{X}$ only by probabilistic experts $(P_i)_{1 \leq i \leq K}$: sequential independent draws



Goal : discover rapidly the elements of A

The model

- Subset $A \subset \mathcal{X}$ of important items
- $|\mathcal{X}| \gg 1$, $|A| \ll |\mathcal{X}|$
- Access to $\mathcal{X}$ only by probabilistic experts $(P_i)_{1 \leq i \leq K}$: sequential independent draws



Goal : discover rapidly the elements of A

Goal

At each time step $t = 1, 2, \dots$:

- pick an index $I_t = \pi_t\big(I_1, Y_1, \dots, I_{s-1}, Y_{s-1}\big) \in \{1, \dots, K\}$ according to past observations
- observe $Y_t = X_{I_t, n_{I_t,t}} \sim P_{I_t}$, where

$$n_{i,t} = \sum_{s \le t} \mathbb{1}\{I_s = i\}$$

Goal : design the strategy $\pi = (\pi_t)_t$ so as to maximize the number of important items found after t requests

$$F^\pi(t) = \Big| A \cap \{Y_1, \dots, Y_t\} \Big|$$

Assumption : non-intersecting supports

$$A \cap \mathrm{supp}(P_i) \cap \mathrm{supp}(P_j) = \emptyset \text{ for } i \neq j$$

Is it a Bandit Problem ?

It looks like a bandit problem. . .

- sequential choices among K options
- want to maximize cumulative rewards
- exploration vs exploitation dilemma

. . . but it is **not a bandit problem** !

- rewards are not i.i.d.
- destructive rewards : no interest to observe twice the same important item
- all strategies eventually equivalent

The oracle strategy

Proposition : Under the non-intersecting support hypothesis, the greedy oracle strategy

$$I_t^* \in \arg\max_{1\leq i\leq K} P_i\left(A \setminus \{Y_1, \ldots, Y_t\}\right)$$

is optimal : for every possible strategy π, $\mathbb{E}\left[F^\pi(t)\right] \leq \mathbb{E}\left[F^*(t)\right]$.

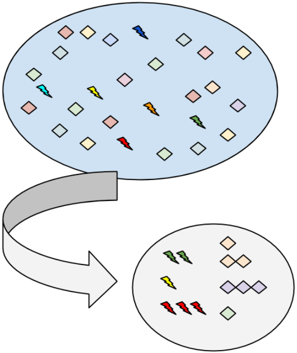
Remark : the proposition if false if the supports may intersect

$\Longrightarrow$ **estimate the "missing mass of important items" !**

Outline

Missing mass estimation

Let us first focus on one expert i : $P = P_i, X_n = X_{i,n}$



$X_1, \dots, X_n$ independent draws of P

$$O_n(x) = \sum_{m=1}^{n} \mathbb{1}\{X_m = x\}$$

How to 'estimate' the total mass of the *unseen* important items

$$R_n = \sum_{x \in A} P(x) \mathbb{1}\{O_n(x) = 0\} \text{ ?}$$

The Good-Turing Estimator

Idea : use the **hapaxes** = items seen only once (linguistic)

$$\hat{R}_n = \frac{U_n}{n}, \quad \text{where } U_n = \sum_{x \in A} \mathbb{1}\{O_n(x) = 1\}$$

Lemma [Good '53] : For *every* distribution P,

$$0 \leq \mathbb{E}\big[\hat{R}_n\big] - \mathbb{E}\big[R_n\big] \leq \frac{1}{n}$$

Proposition : With probability at least $1 - \delta$ for *every* P,

$$\hat{R}_n - \frac{1}{n} - (1 + \sqrt{2})\sqrt{\frac{\log(4/\delta)}{n}} \leq R_n \leq \hat{R}_n + (1 + \sqrt{2})\sqrt{\frac{\log(4/\delta)}{n}}$$

See [McAllester and Schapire '00, McAllester and Ortiz '03] :

- deviations of $\hat{R}_n$: McDiarmid's inequality
- deviations of R_n : negative association

The Good-UCB algorithm

Estimator of the missing important mass for expert i :

$$\hat{R}_{i,n_{i,t-1}} = \frac{1}{n_{i,t-1}} \sum_{x \in A} \mathbb{1}\left\{ \sum_{s=1}^{n_{i,t-1}} \mathbb{1}\{X_{i,s} = x\} = 1 \right.$$
$$\left. \text{and} \sum_{j=1}^{K} \sum_{s=1}^{n_{j,t-1}} \mathbb{1}\{X_{j,s} = x\} = 1 \right\}$$

Good-UCB algorithm :

1: For $1 \leq t \leq K$ choose $I_t = t$.
2: **for** $t \geq K + 1$ **do**
3: Choose $I_t = \arg\max_{1 \leq i \leq K} \left\{ \hat{R}_{i,n_{i,t-1}} + C\sqrt{\frac{\log(4t)}{n_{i,t-1}}} \right\}$
4: Observe Y_t distributed as P_{I_t}
5: Update the missing mass estimates accordingly
6: **end for**

Outline

Classical analysis

Theorem : For any $t \geq 1$, under the non-intersecting support assumption, Good-UCB (with constant $C = (1+\sqrt{2})\sqrt{3}$) satisfies

$$\mathbb{E}\left[F^*(t) - F^{UCB}(t)\right] \leq 17\sqrt{Kt\log(t)} + 20\sqrt{Kt} + K + K\log(t/K)$$

Remark : Usual result for bandit problem, but not-so-simple analysis

Sketch of proof

1. On a set $\tilde{\Omega}$ of probability at least $1-\sqrt{\frac{K}{t}}$, the "confidence intervals" hold true simultaneously all $u \geq \sqrt{Kt}$
2. Let $\bar{I}_u = \arg\max_{1\leq i\leq K} R_{i,n_{i,u-1}}$. On $\tilde{\Omega}$,

$$R_{I_u,n_{I_u,u-1}} \geq R_{\bar{I}_u,n_{\bar{I}_u,u-1}} - \frac{1}{n_{I_u,u-1}} - 2(1+\sqrt{2})\sqrt{\frac{3\log(4u)}{n_{I_u,u-1}}}$$

3. But one shows that $\mathbb{E}F^*(t) \leq \sum_{u=1}^{t} \mathbb{E}R_{\bar{I}_u,n^{\pi}_{\bar{I}_u,u-1}}$
4. Thus

$$\begin{aligned}
&\mathbb{E}\left[F^*(t) - F^{UCB}(t)\right] \\
&\leq \sqrt{Kt} + \mathbb{E}\left[\sum_{u=1}^{t} \frac{1}{n_{I_u,u-1}} + 2(1+\sqrt{2})\sqrt{\frac{3\log(4t)}{n_{I_u,u-1}}}\right] \\
&\leq \sqrt{Kt} + K + K\log(t/K) + 4(1+\sqrt{2})\sqrt{3Kt\log(4t)}
\end{aligned}$$

Experiment : restoring property

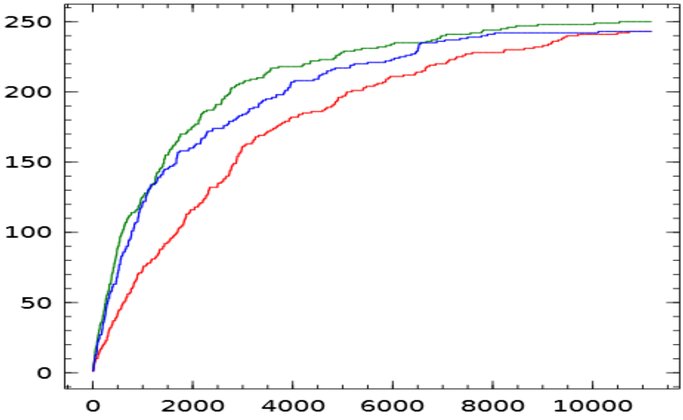


FIGURE – green : oracle, blue : Good-UCB, red : uniform sampling

Another analysis of Good-UCB

For $\lambda \in (0,1)$, $T(\lambda)$ = time at which missing mass of important items is smaller than λ on all experts :

$$T(\lambda) = \inf \left\{ t : \forall i \in \{1,\ldots,K\}, P_i(A \setminus \{Y_1,\ldots,Y_t\}) \leq \lambda \right\}$$

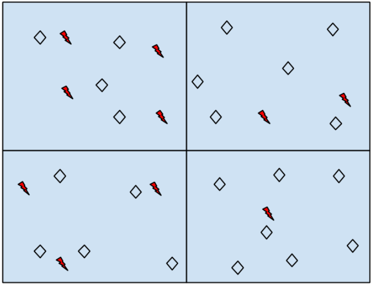
Theorem : Let $c > 0$ and $S \geq 1$. Under the non-intersecting support assumption, for Good-UCB with $C = (1+\sqrt{2})\sqrt{c+2}$, with probability at least $1 - \frac{K}{cS^c}$, for any $\lambda \in (0,1)$,

$$T_{UCB}(\lambda) \leq T^* + KS\log\left(8T^* + 16KS\log(KS)\right),$$

$$\text{where} \quad T^* = T^*\left(\lambda - \frac{3}{S} - 2(1+\sqrt{2})\sqrt{\frac{c+2}{S}}\right)$$

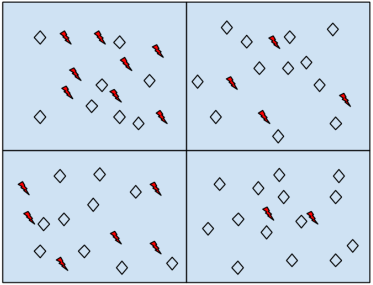
The macroscopic limit

- Restricted framework : $P_i = \mathcal{U}\{1, \ldots, N\}$
- $N \to \infty$
- $|A \cap \mathrm{supp}(P_i)|/N \to q_i \in (0,1)$, $q = \sum_i q_i$



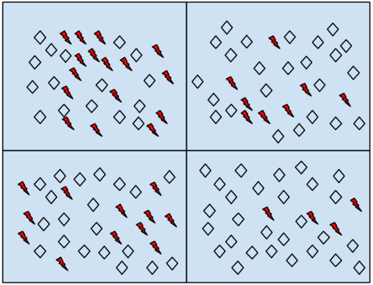
The macroscopic limit

- Restricted framework : $P_i = \mathcal{U}\{1, \ldots, N\}$
- $N \to \infty$
- $|A \cap \mathrm{supp}(P_i)|/N \to q_i \in (0,1)$, $q = \sum_i q_i$



The macroscopic limit

- Restricted framework : $P_i = \mathcal{U}\{1, \ldots, N\}$
- $N \to \infty$
- $|A \cap \mathrm{supp}(P_i)|/N \to q_i \in (0,1)$, $q = \sum_i q_i$



The Oracle behaviour

The limiting discovery process of the Oracle strategy is *deterministic*

Proposition : For every $\lambda \in (0, q_1)$, for every sequence $(\lambda^N)_N$ converging to λ as N goes to infinity, almost surely

$$\lim_{N \to \infty} \frac{T_*^N(\lambda^N)}{N} = \sum_i \left(\log \frac{q_i}{\lambda}\right)_+$$

Oracle vs. uniform sampling

Oracle : The proportion of important items not found after Nt draws tends to

$$q - F^*(t) = I(t)\underline{q}_{I(t)} \exp\left(-t/I(t)\right) \leq K\underline{q}_K \exp(-t/K)$$

with $\underline{q}_K = \left(\prod_{i=1}^K q_i\right)^{1/K}$ the geometric mean of the $(q_i)_i$.

Uniform : The proportion of important items not found after Nt draws tends to $K\bar{q}_K \exp(-t/K)$

$\Longrightarrow$ Asymptotic ratio of efficiency

$$\rho(q) = \frac{\bar{q}_K}{\underline{q}_K} = \frac{\frac{1}{K}\sum_{i=1}^k q_i}{\left(\prod_{i=1}^k q_i\right)^{1/K}} \geq 1$$

larger if the $(q_i)_i$ are unbalanced

Macroscopic optimality

Theorem : Take $C = (1+\sqrt{2})\sqrt{c+2}$ with $c > 3/2$ in the Good-UCB algorithm.

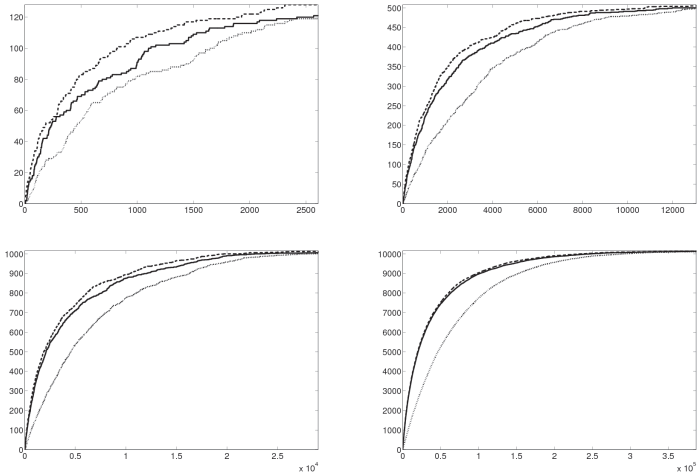
- For every sequence $(\lambda^N)_N$ converging to λ as N goes to infinity, almost surely

$$\limsup_{N \to +\infty} \frac{T^N_{UCB}(\lambda^N)}{N} \leq \sum_i \left(\log \frac{q_i}{\lambda}\right)_+$$

- The proportion of items found after Nt steps F^{GUCB} converges *uniformly* to F^* as N goes to infinity

Experiment

Number of items found by Good-UCB (solid), the OCL (dashed), and uniform sampling (dotted) as a function of time for sizes $N = 128, N = 500, N = 1000$ and $N = 10000$ in a 7-experts setting.



Conclusion and perspectives

- We propose an algorithm for the optimal discovery with probabilistic expert advice
- We give a standard regret analysis under the only assumption that the supports of the experts are non-overlapping
- We propose a different optimality result, which permits a macroscopic analysis in the uniform case

- Another interesting limit to consider is when the number of important items to find is fixed, but the total number of items tends to infinity (Poisson regime)
- Then, the behavior of the algorithm is not very good : too large confidence bonus because no tight deviations bounds for the Good-Turing estimator when the proportion of important items tends to 0