Supplementary Material for paper Bandit Algorithms boost motor-task selection for Brain Computer Interfaces

A The $UCB - classif$ algorithm

A.1 Some intuition on bandit algorithms

Figure 3 illustrates how the *UCB-classif* algorithm works.

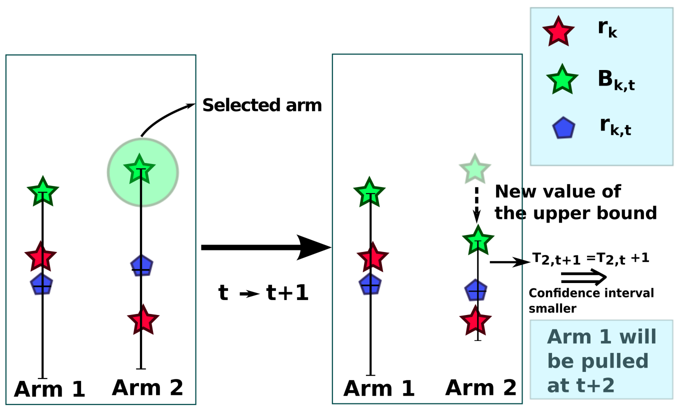

Figure 3: This figure represents two snapshots, a time t and $t+1$, of a bandit with 2 arms. Although arm 1 is the best arm ($r_1^* > r_2^*$, represented by the red stars), at time t, $B_{1,t} < B_{2,t}$, therefore the arm 2 is selected. Pulling the arm 2 gives a better estimate $\hat{r}_{2,t+1}$ of r_2^* and reduces the confidence interval. At time $t+1$, $B_{1,t+1}$ will be greater than $B_{2,t+1}$, so arm 1 will be selected.

A.2 Proof of Theorem 1

Reminder of Vapnik-Chervonenkis's bound in classification Let $\mathcal{D}$ be a probability distribution in $\mathbb{R}^d \times \{0,1\}$. Let $\mathcal{H}$ be the set of binary linear classifiers in $\mathbb{R}^d$, i.e. if $(X,Y) \sim \mathcal{D}$, (i.e. are drawn from $\mathcal{D}$) then $h(X)$ is the inferred class of the sample while the true class is Y.

We define the $\{0,1\}$ loss of a classifier h (which is not always equal to the loss $l(.,.)$ of the SVM classifier) as

$$L_{\mathcal{D}}(h) = \mathbb{E}_{(X,Y)\sim\mathcal{D}}[\mathbf{1}\{h(X) \neq Y\}].$$

Let h^* be the best linear classifier on $\mathcal{D}$ for the $\{0,1\}$ loss, i.e.

$$h^* = \arg\min_{h\in\mathcal{H}} L_{\mathcal{D}}(h).$$

Let now $\mathcal{X} = \{(X_1,Y_1),\ldots,(X_T,Y_T)\}$ be T i.i.d. points in $\mathbb{R}^d \times \{0,1\}$, sampled from $\mathcal{D}$.

We define the $\{0,1\}$ empirical loss of a classifier h as

$$\hat{L}_{\mathcal{X}}(h) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{h(X_t) \neq Y_t\}.$$

Let $\widehat{h} \in \mathcal{H}$ be the linear SVM classifier on $\mathcal{X}$ in $\mathcal{H}$. We have the following Theorem (see [15] for a survey on this).

Theorem 2 (Vapnik, 1982) *We have with probability* $1-2\delta$ *a bounded error on the* $(0,1)$ *loss in generalization, and a bounded error in the estimate of the* $(0,1)$ *loss, that is to say*

$$L_{\mathcal{D}}(\widehat{h}) - L_{\mathcal{D}}(h^*) \leq \sqrt{\frac{d(\log(2T/d)+1)+\log(4/\delta)}{T}},$$

and

$$|L_{\mathcal{D}}(h^*) - \hat{L}_{\mathcal{X}}(\widehat{h})| \leq 2\sqrt{\frac{d(\log(2T/d)+1)+\log(4/\delta)}{T}}.$$

Adaptation of Vapnik-Chervonenkis's bound in our context Write $\hat{R}_{k,t}$ the empirical estimate of the linear SVM classifier's classification error on the t first samples of task k (and with any samples of idle condition).

Define the following event

$$\xi = \bigcap_{k \leq K} \bigcap_{t \leq n} \left\{ |r_k^* - \hat{R}_{k,t}| \leq 2\sqrt{\frac{d(\log(2t/d)+1)+\log(8NK/\delta)}{t}} \right\}. \tag{3}$$

The previous Theorem states that this event is of probability at least $1-\delta$ (by an union bound).

In our setting and for task k, we have $1-r_k^*$ which is the $\{0,1\}$ loss of the best classifier for task k and $1-\hat{r}_{k,t}$ which is the empirical $\{0,1\}$ loss of the linear SVM classifier for task k with $T_{k,t}$ samples. As a corollary, we obtain that with probability $1-\delta$, for any task k and any time t,

$$|r_k^* - \hat{r}_{k,t}| \leq 2\sqrt{\frac{d(\log(2T_{k,t}/d)+1)+\log(8NK/\delta)}{T_{k,t}}},$$

where d is the number of features.

Overview of the way the algorithm works As $T_{k,t} < N$, we have on ξ that for any task k and any time t,

$$|r_k^* - \hat{r}_{k,t}| \leq 2\sqrt{\frac{d(\log(2N/d)+1)+\log(8NK/\delta)}{T_{k,t}}} \leq \sqrt{\frac{a\log(8NK/\delta)}{T_{k,t}}},$$

where $a = 5(d+1)$.

We thus have on ξ

$$r_k^* \leq \hat{r}_{k,t} + \sqrt{\frac{a\log(8NK/\delta)}{T_{k,t}}} \leq r_k^* + 2\sqrt{\frac{a\log(8NK/\delta)}{T_{k,t}}}.$$

Note here that $B_{k,t} = \hat{r}_{k,t} + \sqrt{\frac{a\log(8NK/\delta)}{T_{k,t}}}$ is an upper bound on ξ on r_k^*.

In the event ξ of large probability such that this is true for any k and any N, we know that we pull at time t a sub-optimal arm k if for the best arm $*$ with reward r^*, $B_{*,t} \leq B_{k,t}$, which according to the last equation leads to:

$$r^* \leq B_{*,t} \leq B_{k,t} \leq r_k^* + 2\sqrt{\frac{a\log(8NK/\delta)}{T_{k,t}}},$$

This means by a simple computation that on ξ we pull a sub-optimal arm k only if

$$T_{k,t} \leq 4\frac{a\log(8NK/\delta)}{(r^*-r_k^*)^2}.$$

We then pull with probability $1-\delta$ the suboptimal arms only a number of times in $O(\log(8NK/\delta))$, as $T_{k,N} \leq 4\frac{a\log(8NK/\delta)}{(r^*-r_k^*)^2}$ and thus pull the optimal arm $N - O(\log(8NK/\delta))$ times, more precisely at least $N - \sum_{k\neq *} 4\frac{a\log(8NK/\delta)}{(r^*-r_k^*)^2}$.

Finally, the error of the empirical classifier on the best arm is such that with probability $1 - \delta$

$$|r^* - \hat{r}^*| \leq \sqrt{\frac{a \log(8NK/\delta)}{N - \sum_{k \neq *} 4 \frac{a \log(8NK/\delta)}{(r^* - r_k^*)^2}}}.$$

If for instance we choose $\delta = 1/N$, we have that with probability at least $1 - /N$, the best arm is pulled at least $N - \sum_{k \neq *} 8 \frac{a \log(8NK)}{(r^* - r_k^*)^2}$