Package ‘CBT’

October 12, 2022

Type Package

Title Confidence Bound Target Algorithm

Version 1.0

Date 2018-05-16

Author Hock Peng Chan and Shouri Hu

Maintainer Shouri Hu <e0054325@u.nus.edu>

Description The Confidence Bound Target (CBT) algorithm is designed for infinite arms bandit problem. It is shown that CBT algorithm achieves the regret lower bound for general reward distributions. Reference: Hock Peng Chan and Shouri Hu (2018) <arXiv:1805.11793>.

License GPL-2

RoxygenNote 6.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2018-05-31 20:38:43 UTC

R topics documented:

`CBT` *Confidence Bound Target (CBT) Algorithm*

Description

`CBT` and `EMp_CBT` provide simution to infinite arms with Bernoulli Rewards. `CBT` assumes prior ditribution in known whereas `EMp_CBT` does not. `Ana_CBT` performs analysis to real data.

Usage

```
CBT(n, prior, bn = log(log(n)), cn = log(log(n)))
Emp_CBT(n, prior, bn = log(log(n)), cn = log(log(n)))
Ana_CBT(n, data,  bn = log(log(n)), cn = log(log(n)))
```

Arguments

`n`	total number of rewards.
`prior`	prior distribution on mean of the rewards. Currently avaiable priors: "Uniform", "Sine" and "Cosine".
`bn`	bn should increse slowly to infinity with n.
`cn`	cn should increse slowly to infinity with n.
`data`	A matrix or dataframe. Each column is a population.

Details

If `bn` or `cn` are not specified they assume the default value of `log(log(n))`.
The confidence bound for an arm with t observations is

$$L = max(xbar/bn, xbar - cn * sigma/sqrt(t)),$$

where xbar and sigma are the mean and standatd deviation of the rewards from that paticular arm. CBT is a non-recalling algorithm. An arm is played until its confidence bound L drops below the target mean μ_*, and it is not played after that.
If the prior distribution is unknown, we shall apply empirical CBT, in which the target mean μ_* is replaced by S/n, with S the sum of rewards among all arms played at current stage. Unlike CBT howerver empirical CBT is a recalling algorithm which decides from among all arms which to play further, rather than to consider only the current arm.

Value

A list including elements

`regret`	cumulative regret generated by n rewards.
`K`	total number of experimented arms.

Author(s)

Hock Peng Chan and Shouri Hu

References

H.P. Chan and S. Hu (2018) Infinite Arms Bandit: Optimality via Confidence Bounds <arXiv:1805.11793>

Examples

```
R = 1000

cum_regret = numeric(R)
arms = numeric(R)

for(i in 1:R){
  result = CBT(n = 10000, prior = "Sine")
  cum_regret[i] = result$regret
  arms[i] = result$K
}

mean(cum_regret)
sd(cum_regret)/sqrt(R)
mean(arms)
sd(arms)/sqrt(R)
```

Index