result overviews

Wenjia Wang

June 2018

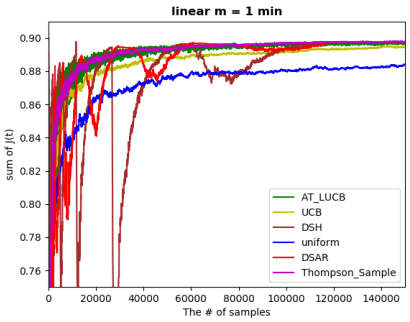
1 Introduction

The bandits used in the article include:
Linear and Polynomial bandits with $m = 1$,10 and 50.
Caption bandits with $m = 50$.
Preventive bandits with $m = 3$.

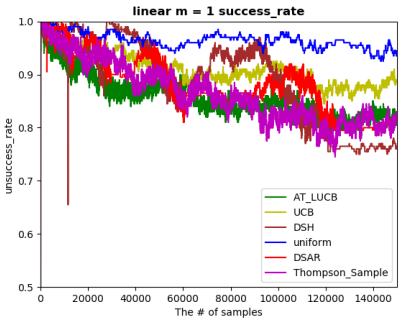
the results will show the smallest true mean of the best m bandits,sum of true mean of the best m bandits and the success rate of the best m bandits.

2 Plot

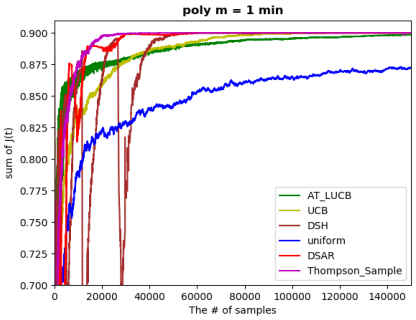
2.1 m = 1



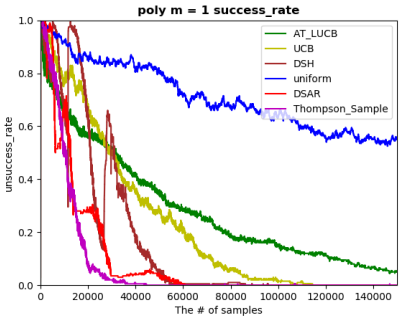
(a) Linear min



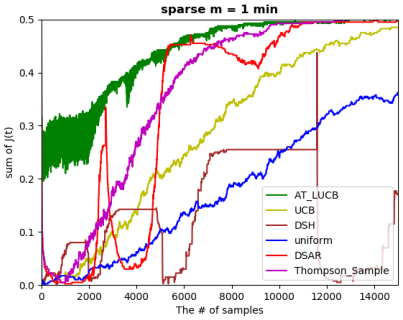
(b) Linear success rate



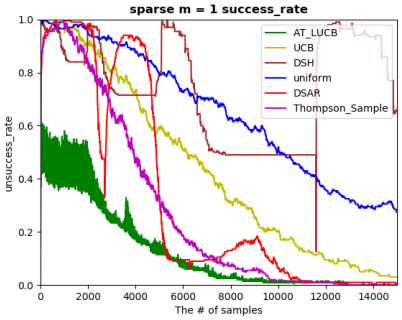
(c) Polynomial min



(d) Polynomial success rate

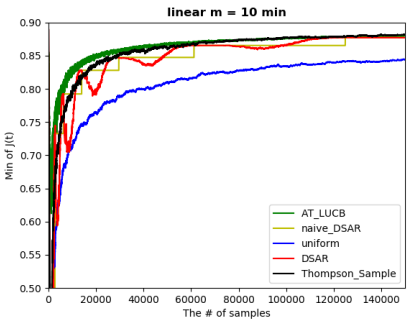


(e) Sparse min

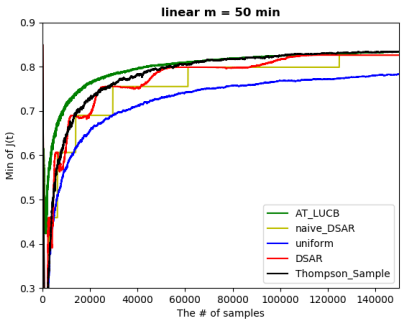


(f) Sparse success rate

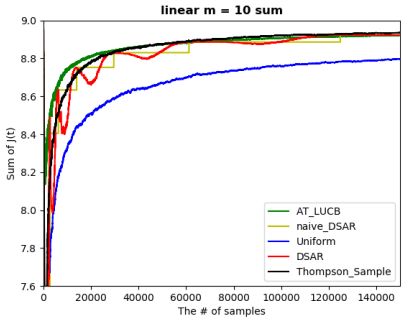
Figure 1: Anytime Explore-1 results:this figure shows the true mean of best arm ($m = 1$) for different types of bandits with different algorithms. The three group figures are the result of linear, polynomial and sparse bandit respectively and in each group, the first figure show the true mean of best arm of sample steps, the second figure shows the average mismatching rate(result is not the true best mean) over 200 times running



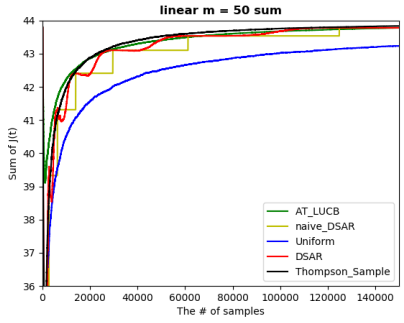
(a) Linear m = 10 min



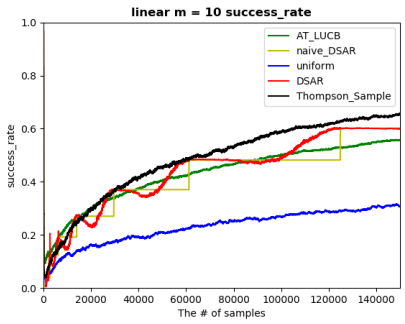
(b) Linear m = 50 min



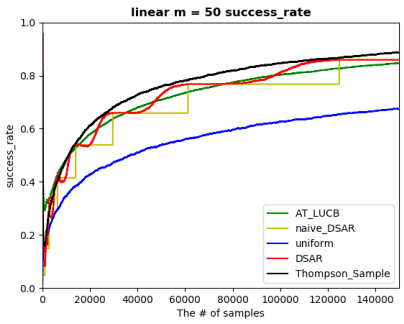
(c) Linear m = 50 min



(d) Linear m = 50 sum

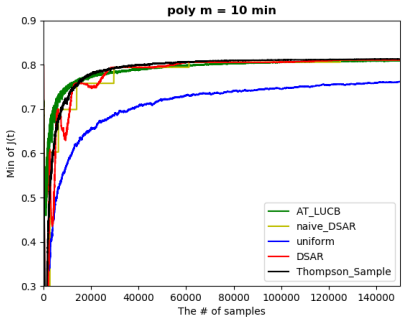


(e) linear m = 10 success rate

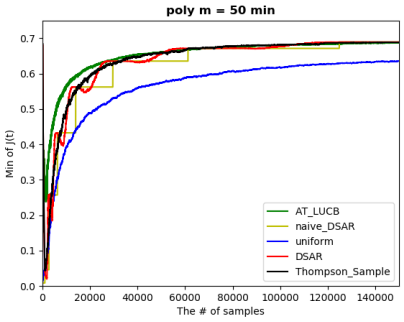


(f) linear m = 50 success rate

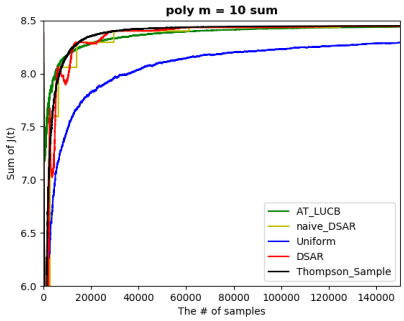
Figure 2: Anytime Explore-Linear results:this figure shows the linear bandit results. The left is for $m = 10$ and the right is $m = 50$ respectively. The first one shows the smallest true mean in the top m best result, the second figure shows the sum of the top m mean, the third one is the success rate. All the results are averaged from 200 times running.



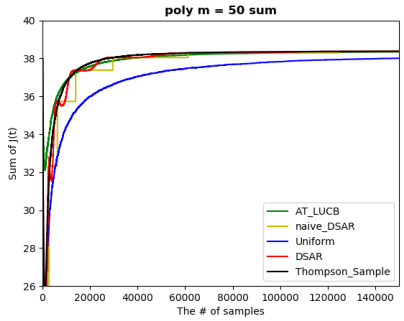
(a) Polynomial m = 10 min



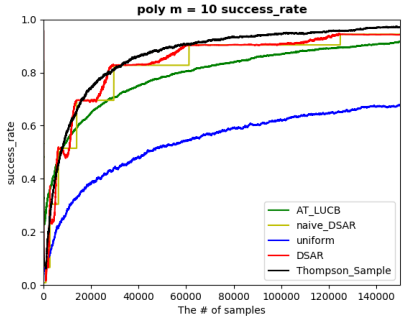
(b) Polynomial m = 50 min



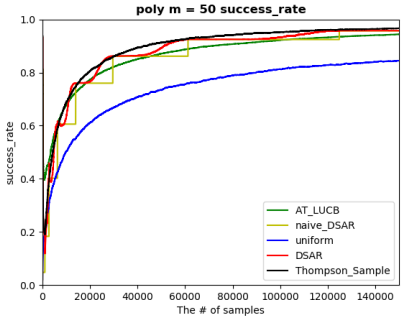
(c) Polynomial m = 50 min



(d) Polynomial m = 50 sum

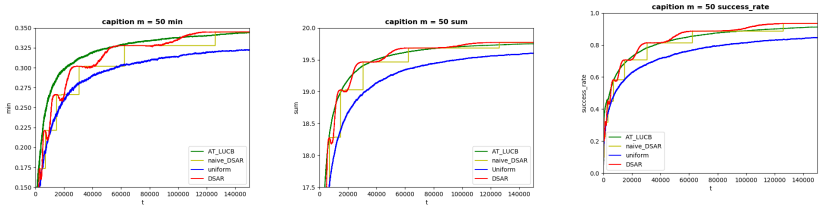


(e) Polynomial m = 10 success rate



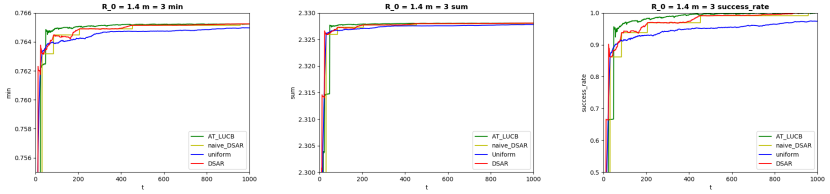
(f) Polynomial m = 50 success rate

Figure 3: Anytime Explore-Polynomial results:this figure shows the Polynomial bandit results. The left is for $m = 10$ and the right is $m = 50$ respectively. The first one shows the smallest true mean in the top m best result, the second figure shows the sum of the top m mean, the third one is the success rate. All the results are averaged from 200 times running.



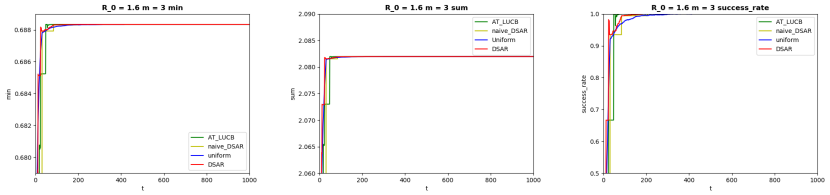
(a) Caption m = 50 min (b) Caption m = 50 sum (c) Caption m = 50 success rate

Figure 4: Anytime Explore-Caption($m = 50$) results:this figure shows the Caption bandit results. All the results are averaged from 200 times running.



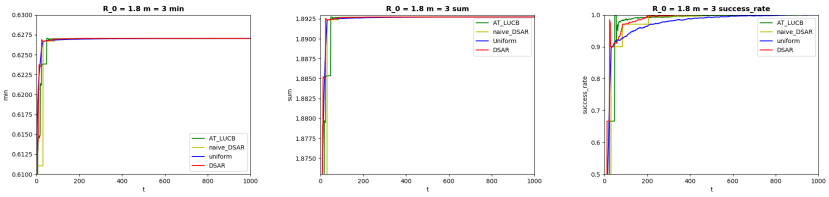
(a) $R_0 = 1.4$ min (b) $R_0 = 1.4$ sum (c) $R_0 = 1.4$ success rate

Figure 5: Anytime Explore-Preventive results with $R0_0 = 1.4$:this figure shows the Preventive bandit results with $m = 3$. All the results are averaged from 200 times running.



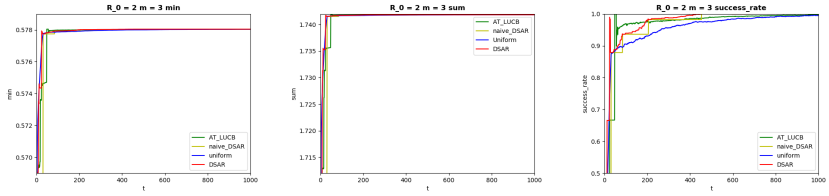
(a) $R_0 = 1.6$ min (b) $R_0 = 1.6$ sum (c) $R_0 = 1.6$ success rate

Figure 6: Anytime Explore-Preventive results with $R0_0 = 1.6$:this figure shows the Preventive bandit results with $m = 3$. All the results are averaged from 200 times running.



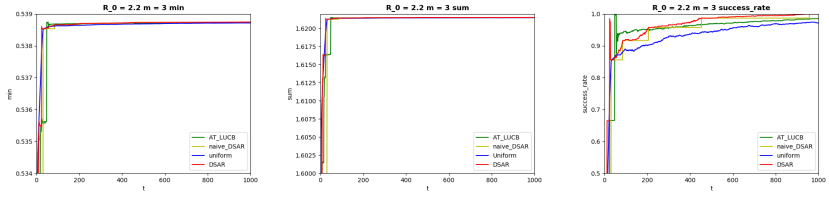
(a) $R_0 = 1.8$ min (b) $R_0 = 1.8$ sum (c) $R_0 = 1.8$ success rate

Figure 7: Anytime Explore-Preventive results with $R0_0 = 1.8$:this figure shows the Preventive bandit results with $m = 3$. All the results are averaged from 200 times running.



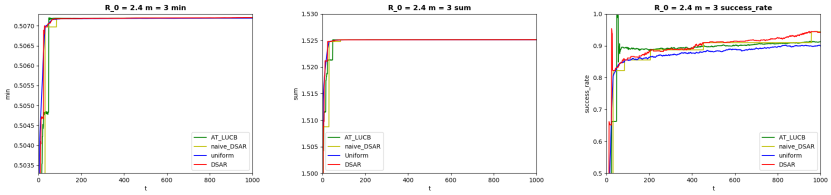
(a) $R_0 = 2.0$ min (b) $R_0 = 2.0$ sum (c) $R_0 = 2.0$ success rate

Figure 8: Anytime Explore-Preventive results with $R0_0 = 2.0$:this figure shows the Preventive bandit results with $m = 3$. All the results are averaged from 200 times running.



(a) $R_0 = 2.2$ min (b) $R_0 = 2.2$ sum (c) $R_0 = 2.2$ success rate

Figure 9: Anytime Explore-Preventive results with $R0_0 = 2.2$:this figure shows the Preventive bandit results with $m = 3$. All the results are averaged from 200 times running.



(a) $R_0 = 2.4$ min (b) $R_0 = 2.4$ sum (c) $R_0 = 2.4$ success rate

Figure 10: Anytime Explore-Preventive results with $R0_0 = 2.4$:this figure shows the Preventive bandit results with $m = 3$. All the results are averaged from 200 times running.