Non-stationary Dueling Bandits for Online Learning to Rank

[illegible] L.[1], [illegible] M.[2], P. [illegible][2], [illegible] H.[2], [illegible] L. [illegible][1(✉)]

[1] [illegible] Lab [illegible], Na[illegible] U[illegible], Na[illegible] 210023, C[illegible]a

{lusy,zhanglj}@lamda.nju.edu.cn

[2] A[illegible]baba G[illegible], Ha[illegible] 311121, C[illegible]a

{miaoyuan.my,yangping.yangping,yaoohu}@alibaba-inc.com

Abstract. [illegible] (OL2R), [illegible] OL2R [illegible] OL2R [illegible] OL2R [illegible] (DBGD) [illegible] DBGD [illegible] DBGD- [illegible]

Keywords: [illegible] · D[illegible] ba[illegible] · N[illegible]

1 Introduction

A [illegible] 7,15. I[illegible] fff [illegible]. A [illegible] fff [illegible] 4. F[illegible],

B. L[illegible]. (E[illegible].): AP[illegible]-AIM 2022, LNC[illegible] 13422, [illegible]. 166–174, 2023.

[illegible]://[illegible]/10.1007/978-3-031-25198-6_13

, 6 .

T , (OL2R), ž 3 . C ﬀ -

, OL2R .

A OL2R - ﬀ

. T ,

2 Problem Setup

W - , - . L $\mathcal{W} \subseteq$ T . F 8,11,12 ,

I $t \in [T] = \{1, \ldots, T\}$, $\mathbf{w} \in \mathcal{W}$ $\mathbf{w}' \in \mathcal{W}$, . T , 5,9 . T . S , . F , .

W $\mathbf{w} \succ \mathbf{w}'$ $\mathbf{w}$ $\mathbf{w}'$. W , - . S , t, $\mathbf{w} \succ \mathbf{w}'$

$$\Pr(\mathbf{w} \succ \mathbf{w}'|t) = f(\mathbf{w}, \mathbf{w}') = \sigma(v(\mathbf{w}) - v(\mathbf{w}')) \tag{1}$$

σ , v t. F 11,14 , :

T $\mathcal{W}$

$$\max_{\mathbf{w} \in \mathcal{W}} \|\mathbf{w}\|_2 \leq R. \tag{2}$$

T σ -

$$\sigma(x) = 1 - \sigma(-x). \tag{3}$$

T σ ,

$$\sigma(-\infty) = 0, \quad \sigma(0) = 1/2, \quad \sigma(\infty) = 1.$$

T σ L -L , v , $t \in [T]$ L -L . F , σ L_2-L .[2]

D , $L = L\ L$, f , $t \in [T]$ L-L .

L $\mathbf{w}^* = \arg\max_{\mathbf{w} \in \mathcal{W}} v(\mathbf{w})$ - t. W ,

$$\mathrm{DR}(T) = \sum_{=1} \big(f(\mathbf{w}^*, \mathbf{w}) + f(\mathbf{w}^*, \mathbf{w}') - 2f(\mathbf{w}^*, \mathbf{w}^*)\big).$$

O , , .

[2] I OL2R, a d sd c s s d c $\sigma(x) = 1/\big(1+\exp(-x)\big)$, c a s sa a ss s

3 Method

I , (DBGD) , DBGD- .

3.1 a a

A A , 1, DBGD δ γ; - . I t, DBGD $\mathbf{u}$ $\triangleq \{\mathbf{x} \in : \|\mathbf{x}\|_2 = 1\}$. T ,

$$\mathbf{w}' = \Pi_{\mathcal{W}}[\mathbf{w} + \delta \mathbf{u}] \quad (4)$$

$\mathbf{w}$ $\Pi_{\mathcal{W}}[\cdot]$ $\mathcal{W}$. N , $\mathbf{w}$ $\mathbf{w}'$ 5 , . F , , DBGD . S , $\mathbf{w}'$, , γ: $\mathbf{w}_{+1} = \Pi_{\mathcal{W}}[\mathbf{w} + \gamma \mathbf{u}]$. O , . W DBGD - .

1. *Let C be the path length of the optimal rankers over T rounds, defined as*

$$C = \sum_{=2} \|\mathbf{w}^* - \mathbf{w}^*_{-1}\|_2. \quad (5)$$

By setting $\delta = \sqrt{\frac{2}{(11+2\ )\ \sqrt{\ }}}$ *and* $\gamma = \sqrt{\frac{5\ \ ^2+2}{\ \ \ T}}$*, the dynamic regret of DBGD satisfies*

$$[\mathrm{DR}(T)] \leq \sqrt{2(11+2\lambda)\lambda d L}\left(1 + \sqrt{5R^2 + 2RC}\right) T^{\frac{3}{4}}.$$

3.2 a a

W DBGD - $C = o(\sqrt{T})$, - C , γ, C

1. DBGD

Re e: s s a δ a d a γ

1: I a a a $\mathbf{w}_1 \in \mathcal{W}$ a b a

2: **f** $t = 1, 2, \ldots, T$ **d**

3: D a a c $\mathbf{u}_t$ a a d

4: C a a a a $\mathbf{w}'_t = \Pi_{\mathcal{W}}[\mathbf{w}_t + \delta \mathbf{u}_t]$

5: C a $\mathbf{w}_t$ a d $\mathbf{w}'_t$ b bab sc a

6: **f** $\mathbf{w}'_t \succ \mathbf{w}_t$ **e**

7: S $\mathbf{w}_{t+1} = \Pi_{\mathcal{W}}[\mathbf{w}_t + \gamma \mathbf{u}_t]$

8: **e e**

9: S $\mathbf{w}_{t+1} = \mathbf{w}_t$

10: **e d f**

11: **e d f**

$\mathbf{w}_1^*, \ldots, \mathbf{w}^*$. T , γ, 2,16,17 . T DBGD , ff γ . W DBGD , DBGD .

W DBGD M M L (DM^{2}L)

Meta Algorithm A A 2, ff γ. A ,

$$N = \left\lceil \log_2 \sqrt{1 + 4T/5} \right\rceil + 1 \tag{6}$$

γ *i*-

$$\gamma \ = 2^{\ -1} R \sqrt{5/T}, \quad i = 1, \ldots, N. \tag{7}$$

E $i \in [N]$ - π , j *i*. F :

$$\pi_1 = \frac{N+1}{i(i+1)N}, \quad i = 1, \ldots, N. \tag{8}$$

I *t*, **w** $i \in [N]$ π , $i \in [N]$ **w** =

$\sum_{=1} \pi \mathbf{w}$. T , u $\mathbf{w}$ $\mathbf{w}' = \Pi_{\mathcal{W}}[\mathbf{w} + \delta \mathbf{u}]$, $\{\mathbf{w}'_t \succ \mathbf{w}_t\}$. N ,

$$\pi_{+1} = \frac{\pi \exp(-\alpha\ell(\mathbf{w}))}{\sum_{=1} \pi \exp(-\alpha\ell(\mathbf{w}))}, \quad i = 1, \ldots, N \tag{9}$$

$\ell(\mathbf{w})$,

$$\ell(\mathbf{w}) = -\frac{d}{\delta}\langle_{\{\mathbf{w}'_t \succ \mathbf{w}_t\}} \mathbf{u}, \mathbf{w} - \mathbf{w}\rangle$$

. F , $\{\mathbf{w}'_t \succ \mathbf{w}_t\}$ u .

Expert Algorithm. A A 3, DBGD. I t, $i \in [N]$ $\mathbf{w}$. T , $\{\mathbf{w}'_t \succ \mathbf{w}_t\}$ u . F ,

$$\mathbf{w}_{+1} = \Pi_{\mathcal{W}}[\mathbf{w} + \gamma_{\{\mathbf{w}'_t \succ \mathbf{w}}$$

2. DM^{2}L: M A

Re e: b N, δ, $\gamma_1, \ldots, \gamma_N$, a a α

I A 3 γ_i ac $i \in [N]$

I a $\pi_1^i, i \in [N]$ b (8)

f $t = 1, 2, \ldots, T$ **d**

R c a w_t^i ac $i \in [N]$

A a a a $\mathsf{w}_t = \sum_{i=1}^{N} \pi_t^i \mathsf{w}_t^i$

D a a c u_t a a d

C a a a a $\mathsf{w}_t' = \Pi_{\mathcal{W}}[\mathsf{w}_t + \delta \mathsf{u}_t]$

C a w_t a d w_t' b bab c a

U da ac $\pi_t^i, i \in [N]$ b (9)

S d $\{\mathbf{w}_t' \succ \mathbf{w}_t\}$ a d u_t ac $i \in [N]$

e d f

3. DM^{2}L: E A

Re e: a γ_i

I a a a $\mathsf{w}_1^i \in \mathcal{W}$ a b a

f $t = 1, 2, \ldots, T$ **d**

S d a w_t^i A 2

R c $\{\mathbf{w}_t' \succ \mathbf{w}_t\}$ a d u_t A 2

U da a w+11

Ac ed e e . T a a a d b NSFC (61976112) a d J a SF (BK20200064). W a a c

References

1. C -B a c , N., L , G.: P d c , L a , a d Ga Ca b d U P (2006)
2. a E , T., K , W.M.: M a ad: M a a a . I : Ad a c N a I a P c S . 29, . 3666 3674 (2016)
3. G , A., R , M.: O a a a . I : P c d 39 I a a ACM SIGIR C c R a c a d D I a R a , . 1215 1218 (2016)
4. H a , K.: Fa a d ab a a a . P .D. D a (2013)
5. H a , K., W , S., R , M.: A bab c d c c I : P c d 20 ACM I a a C c I a a d K d Ma a , . 249 258 (2011)
6. H a , K., W , S., R , M.: Ba a c a a d a a d a a a a . I . R . **16**(1), 63 90 (2013)
7. L , T.Y.: L a a a a . F d. T d I . R . **3**(3), 225 331 (2009)
8. O H., R , M.: D ff ab b d a a . I : P c d 27 ACM I a a C c I a a d K d Ma a , . 1293 1302 (2018)
9. Rad F., K , M., J ac T.: H d da a c a a ? I : P c d 17 ACM C c I a a d K d Ma a , . 43 52 (2008)
10. Sc , A., O H., W , S., R , M.: M a ad d a a a . I : P c d 9 ACM I a a C c W b S a c a d Da a M , . 457 466 (2016)
11. Wa , H., K , S., McC d-S , E., W , Q., Wa , H.: Va a c d c ad a a . I : P c d 42 d I a a ACM SIGIR C c R a c a d D I a R a , . 835 844 (2019)
12. Wa , H., La , R., K , S., McC d-S , E., Wa , H.: Effic a ad ac a a . I : P c d 41 I a a ACM SIGIR C c R a c a d D I a R a , . 145 154 (2018)
13. Y , Y., B d , J., K b , R., J ac T.: T -a d d ba d b . J. C . S . Sc **78**(5), 1538 1556 (2012)
14. Y , Y., J ac T.: I ac a a a d ba d b . I : P c d 26 I a a C c Mac L a , . 1201 1208 (2009)
15. Za , Y., a .: GISDCN: A a -ba d a a d d ab a . I : I a a C c Da aba S Ad a c d A ca . 289 297. S , C a (2022). https://doi.org/10.1007/978-3-031-00126-0_21

16. Z a , L., L , S., Z , Z.H.: Ada a d a c s I : Ad a c s N a I a P c s S s s . 31, . 1323 1333 (2018)
17. Z a , P., Wa , G., Z a , L., Z , Z.H.: Ba d c a - s a a s I : P c d s 23 d I a a C c A c a I c a d S a s c s . 1508 1518 (2020)