Online Appendix to: Directly Optimize Diversity Evaluation Measures: A New Approach to Search Result Diversification

JUN XU, LONG XIA, YANYAN LAN, JIAFENG GUO, and XUEQI CHENG,
CAS Key Lab of Network Data Science and Technology, Institute of Computing Technology, Chinese Academy of Sciences

In this Appendix, as supplementary material to the results shown in Section 7.3, we show the experimental results of SGDMM-Log and SGDMM-Exp on the dataset of WT2009.

A. ABILITY TO IMPROVE THE EVALUATION MEASURES

Similar to the experiments conducted in Section 7.3.2, we also conducted experiments to see whether SGDMM-Log and SGDMM-Exp have the ability to improve diverse ranking quality in terms of a measure by using the measure in training. Figures 10, 11, and 12 show the results in terms of ERR-IA@20, α-NDCG@20, and D#-NDCG@20, respectively. From the results, we can conclude that the algorithms derived under the proposed framework can indeed enhance the diverse ranking quality in terms of a measure by using the measure in training.

B. EFFECTS OF POSITIVE AND NEGATIVE RANKINGS

Similar to the experiments conducted in Section 7.3.3, we also examined the effects of the number of positive rankings generated per query (parameter τ^+) and the effects of the number of negative rankings per query (parameter τ^-) in SGDMM-Log and SGDMM-Exp. Figure 13 shows the performance curves in terms of ERR-IA@20 and α-NDCG@20. From the results, we can see that the curves do not change much with different τ^+ values, and the performance increases steadily with the increasing τ^- values until $\tau^- = 15$.

C. CONVERGENCE

Similar to the experiments conducted in Section 7.3.5, we also conducted experiments to show whether SGDMM-Log and SGDMM-Exp can converge in terms of the diversity evaluation measures. Figures 14 and 15 show the learning curve of SGDMM-Log and SGDMM-Exp in terms of α-NDCG@20, ERR-IA@20, and D#-NDCG@20 with respect to the number of training iterations. At each training iteration, the model parameters are outputted and evaluated on the test data. The results indicate that SGDMM-Log and SGDMM-Exp converge fast and conduct the training efficiently.

© 2017 ACM 2157-6904/2017/01-ART41 $15.00
DOI: http://dx.doi.org/10.1145/2983921

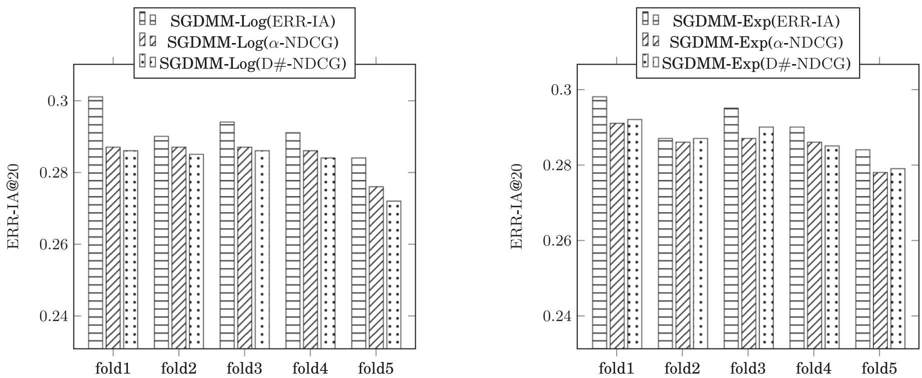

Fig. 10. Performance in terms of ERR-IA@20 when model is trained with ERR-IA@20, α-NDCG@20, or D#-NDCG@20.

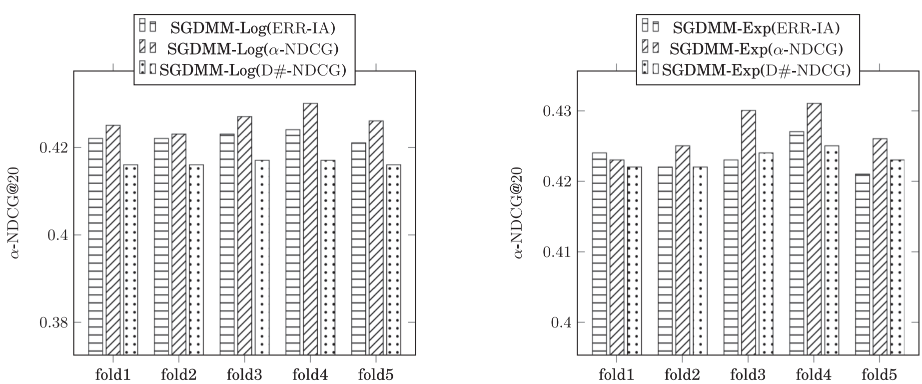

Fig. 11. Performance in terms of α-NDCG@20 when model is trained with ERR-IA@20, α-NDCG@20 or D#-NDCG@20.

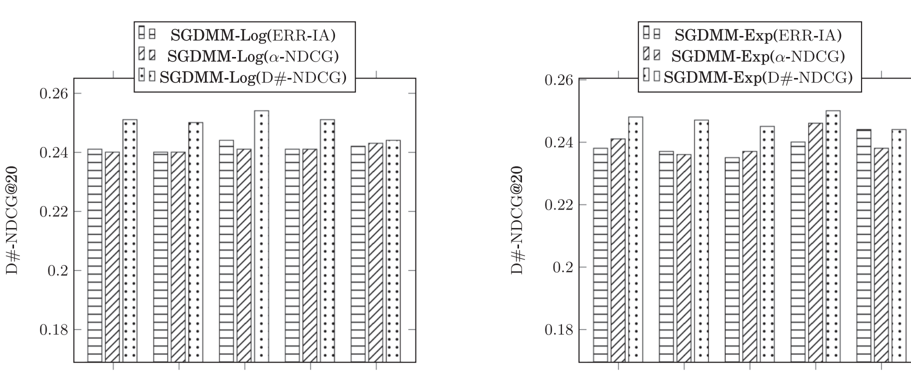

Fig. 12. Performance in terms of D#-NDCG@20 when model is trained with ERR-IA@20, α-NDCG@20, or D#-NDCG@20.

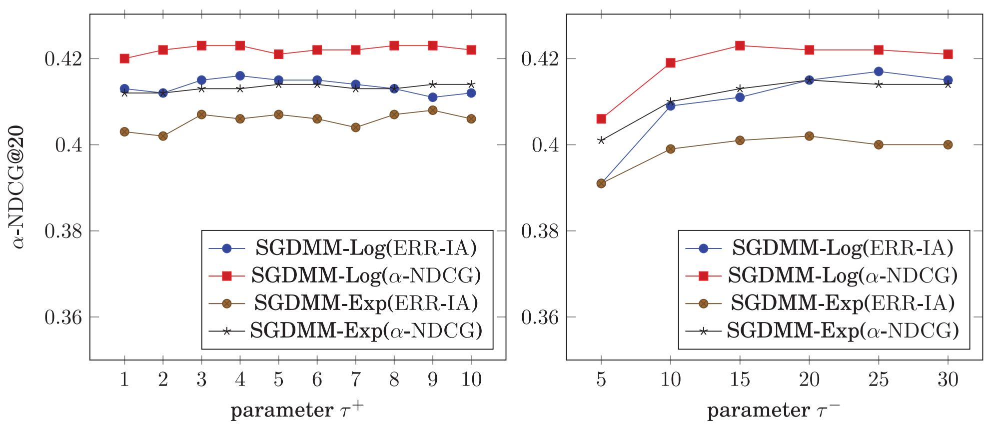

Fig. 13. Ranking accuracies and training time with respect to τ^+ (left) and τ^- (right).

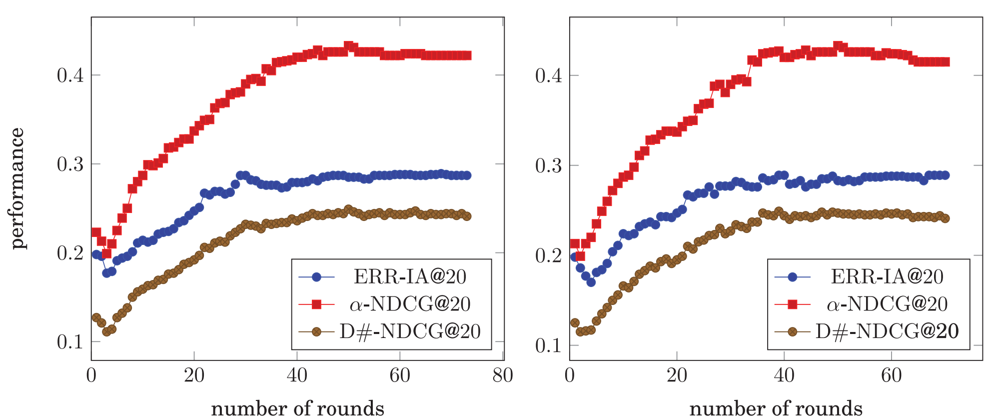

Fig. 14. Learning curve of SGDMM-Log(α-NDCG) (left) and SGDMM-Log(ERR-IA) (right).

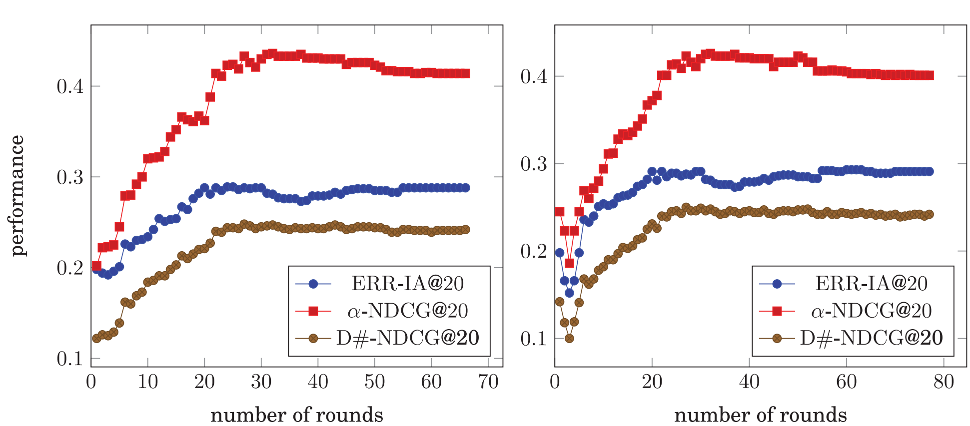

Fig. 15. Learning curve of SGDMM-Exp(α-NDCG) (left) and SGDMM-Exp(ERR-IA) (right).