

Contents

Symbols	xxiii
1 Introduction	1
1.1 Reinforcement learning	2
1.1.1 Temporal Difference (TD) learning	5
1.2 Policy Gradient Algorithms	6
1.2.1 Actor-critic Methods	7
1.3 Ranking in Search Systems and Recommender Systems	8
1.3.1 Recommender systems	11
1.3.2 Search systems in Information retrieval	12
1.3.3 Learning to Rank and Top-N Recommender systems	13
1.4 Applications of Recommender systems and Search systems	15
1.5 Major Challenges	18
1.6 Motivation	19
1.7 Contribution	21
2 Background	23
2.1 Literature Review	23
2.1.1 Literature Review for Recommender systems	23
2.1.2 Literature Review for Ranking in search systems	26
2.2 Deep Reinforcement Learning	28
2.3 Multi Agent Reinforcement Learning	30

2.4	Ranking metrics	33
3	A Deep Actor Critic Reinforcement Learning framework for Learning to rank	37
3.1	Formulation of Learning to Rank as MDP	39
3.2	Algorithm	41
3.2.1	Significance of Each Component in Ranking Tasks	46
3.3	Experiments	49
3.3.1	Evaluation datasets and experimental settings	49
3.4	Experimental Results	50
4	A Multi Agent Deep Reinforcement Learning based approach for large scale Learning to Rank	57
4.1	Problem Statement	63
4.2	Multi Agent MDP Formulation	64
4.2.1	Actor Critic Module	66
4.3	Algorithm	69
4.4	Experiments	71
4.5	Experimental Results	72
5	Proximal Policy Optimization based Hybrid Recommender Systems for large scale recommendations	77
5.1	Proximal Policy Optimization	80
5.2	Problem statement	83
5.3	MDP based Collaborative filtering	83
5.4	Autoencoder based content recommender system	85
5.5	MDP based PPO Algorithm	86
5.6	Experiments	87
5.6.1	Experimental Settings	87
5.7	Evaluation baselines and criterion	89
5.8	Results	91

CONTENTS

6 Discussions 97

7 Conclusion and Future Work 104

 7.0.1 Future Work 106

8 List of Publications 108