

# Contents

## Part I Theory

<b>1 Introduction</b>	3
1.1 The Gambler's Dilemma	3
1.2 A Taxonomy of Multi-armed Bandit Problems	4
1.3 Organization	6
References	6
<b>2 Stochastic Multi-armed Bandit</b>	9
2.1 Problem Formulation	9
2.2 Theoretical Lower Bound	11
2.3 Algorithms	12
2.3.1 Upper Confidence Bound (UCB) Strategies	13
2.3.2 $\epsilon$ -Greedy Policy	17
2.3.3 Thompson Sampling Policy	18
2.4 Variants of Stochastic Multi-armed Bandit	19
2.4.1 Multiplay MAB	19
2.4.2 MAB with Switching Costs	20
2.4.3 Pure Exploration MAB	21
2.5 Summary	24
References	24
<b>3 Markov Multi-armed Bandit</b>	27
3.1 Problem Formulation	27
3.1.1 Markov MAB and Markov Decision Process	28
3.1.2 Optimal Policies for Restless Markov MABs with Complete Information	29

3.2	Algorithms . . . . .	30
3.2.1	Rested Markov MAB . . . . .	31
3.2.2	Restless Markov MAB . . . . .	32
3.3	Summary . . . . .	39
	References. . . . .	39
<b>4</b>	<b>Adversarial Multi-armed Bandit . . . . .</b>	<b>41</b>
4.1	Problem Formulation . . . . .	41
4.2	Algorithms . . . . .	45
4.2.1	Weighted Average Prediction Algorithm. . . . .	45
4.2.2	Following-the-Perturbed-Leader (FPL) Algorithm. . . . .	48
4.2.3	Implicitly Normalized Forecaster (INF) Algorithm . . . . .	51
4.2.4	Internal-Regret Minimizing Algorithm . . . . .	52
4.3	Game Theoretical Results for Multiplayer Adversarial Multi-armed Bandit . . . . .	54
4.4	Summary . . . . .	57
	References. . . . .	57
 <b>Part II Applications</b>		
<b>5</b>	<b>Spectrum Sensing and Access in Cognitive Radio Networks . . . . .</b>	<b>61</b>
5.1	Introduction . . . . .	61
5.2	Problem Formulation . . . . .	63
5.2.1	Single SU with IID Rewards. . . . .	63
5.2.2	Single SU with Markov Reward Processes . . . . .	64
5.2.3	Multiple SUs. . . . .	64
5.3	Solution Approaches. . . . .	65
5.3.1	Cooperative Spectrum Access. . . . .	65
5.3.2	Distributed Learning and Allocation . . . . .	67
5.4	Summary . . . . .	68
	References. . . . .	68
<b>6</b>	<b>Sniffer-Channel Assignment in Multichannel Wireless Networks. . .</b>	<b>71</b>
6.1	Introduction . . . . .	71
6.2	Problem Formulation . . . . .	73
6.2.1	Optimal Channel Assignment in the Nominal Form . . . . .	74
6.2.2	Linear Bandit for Optimal Channel Assignment with Uncertainty . . . . .	75
6.2.3	Extensions. . . . .	77
6.3	Solution Approaches. . . . .	79
6.3.1	Spanners . . . . .	79
6.3.2	An Upper Confidence Bound (UCB)-Based Policy. . . . .	79
6.3.3	An $\varepsilon$ -Greedy Algorithm with Spanner. . . . .	80

6.3.4	An Upper Confidence Bound (UCB)-Based Policy for Switching Cost Consideration . . . . .	81
6.3.5	Approximate Learning Algorithms . . . . .	82
6.4	Numerical Results. . . . .	86
6.5	Summary . . . . .	88
	References. . . . .	89
<b>7</b>	<b>Online Routing in Multi-hop Wireless Networks . . . . .</b>	<b>91</b>
7.1	Introduction . . . . .	91
7.2	System Model and Problem Statement . . . . .	93
7.2.1	Network Model. . . . .	93
7.2.2	Online Routing Problem . . . . .	94
7.3	Algorithm. . . . .	95
7.4	Numerical Results. . . . .	96
7.5	Summary . . . . .	98
	Appendix . . . . .	99
	References. . . . .	101
<b>8</b>	<b>Channel Selection and User Association in WiFi Networks . . . . .</b>	<b>103</b>
8.1	Introduction . . . . .	103
8.2	System Models and Problem Statement . . . . .	105
8.2.1	Network Model. . . . .	105
8.2.2	Airtime Cost . . . . .	106
8.2.3	Channel Selection and User Association (CSUA) Problem. . . . .	106
8.3	Adversarial MAB Framework for CSUA Problem and Algorithms. . . . .	107
8.3.1	Adversarial MAB Formulation . . . . .	107
8.3.2	Algorithm . . . . .	109
8.4	Performance Evaluation . . . . .	111
8.4.1	Simulation Setup. . . . .	111
8.4.2	Simulation Results . . . . .	111
8.5	Summary . . . . .	114
	References. . . . .	115
	<b>Index . . . . .</b>	<b>117</b>

Sequential Learning and Decision-Making in Wireless  
Resource Management

Zheng, R.; Hua, C.

2016, XIII, 118 p. 22 illus., Hardcover

ISBN: 978-3-319-50501-5