THE OHIO STATE UNIVERSITY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

UNDERGRADUATE RESEARCH THESIS

Optimal Adaptive Coordinated Cyber-Attacks on Power Grids using ϵ -Greedy Method

Author

PEIMENG(PG) GUAN

Advisor

Dr. JIANKANG WANG

Abstract

The future power grid is supported by Information and Communication Technology, which also exposes it to cyber-attacks. In particular, Coordinated Cyber-Attacks (CCAs) are highly threatening and difficult to defend against. In this research, we propose a stochastic game model to capture the interaction between attackers and the grid operator. In particular, we consider the most vicious CCAs, which intend to cause cascading power blackouts, through a non-cooperative zero-sum game. The CCA attack vector is derived with the Multi-Armed Bandit ϵ -Greedy method. Distinct from few existing studies on CCAs, the attack model is more realistic in twofold: (i) it does not assume attackers with prior knowledge of the power grid, and (ii) attackers could adapt their strategies in response to defense actions. The result of this research provides important implications in defense resource allocation and cybersecurity infrastructure reinforcement in the power grid. The proposed model and the attack vector is validated using the New England 39 bus power system model.

April 16, 2020

Contents

1	Introduction	2
1.1	Previous Works	2
1.2	Our Work	2
2	Background	3
2.1	Coordinated Control Cyber-Attacks	3
2.2	Assumptions	3
2.3	Defense Implications	3
3	Modeling Attack-Defender Interaction	4
3.1	The Defender	4
3.2	The Attacker	4
3.3	Game Payoff	4
3.4	Attack Vector	5
4	Power System Implementation	6
5	Numerical Experiments	7
5.1	Single Attack Validation	7
5.2	Single Attack on a Fixed Defense	7
5.3	Coordinated Attack	7
6	Conclusions & Future Works	8
7	Acknowledgement	8

List of Figures

1	The power grid evolution after cascading failure	5
2	IEEE 39-bus system	6
3	Game Value with Single Attack over Iteration	7
4	Strategies for Single Attack	8
5	Strategies for CCAs	8
6	Attack Strategy for Fixed Defense	9
7	Game Value for CCAs over Iteration	9

1 Introduction

The future Power Grid will be characterized by the pervasive use of heterogeneous and non-proprietary Information and Communication Technology (ICT), which exposes the power grid to a wide scope of cyber-attacks. In particular, opposed to single random attacks, Coordinated Cyber-Attacks (CCAs) can cause disastrous consequences (e.g., cascading blackouts) by attacking multiple targets, so are highly threatening and difficult to defend against. The famous cyber-attacks against the Ukrainian power grid (2015) affected three distribution companies and thirty substations, leaving 230,000 consumers without power [1].

1.1 Previous Works

To protect the power grid against CCAs, many defense strategies have been proposed in [2–5]. However, two deficiencies are observed in those works. First, the CCAs are oversimplified by modeling the attackers with fixed strategies. Realistically, adversaries learn from the past perceptions by adapting attack strategies to defense reaction. Secondly, those works assume the defender has enough time to *react* against CCAs. For example, in [6] the defender reacts against CCAs by adapting the control actions in the Security Constrained Economic Dispatch, which happens every 5 minutes in the real-time electricity markets. This assumption ignores the instantaneity of cyber-attacks. In fact, transmission circuit breakers usually act within $15ms$ to $200ms$ once receiving the tripping signal, which will leave no time for the operator to take control response under cyber-attacks. Therefore, the defense strategies based on this assumption would be ineffective due to the slow response.

To address these deficiencies, game-theoretic approaches have been adopted to derive autonomous defense response against adaptive attackers [6–9]. Most of those works, nevertheless, require the attackers with prior knowledge of the system, including the network configuration, grid operating status, and victim applications’ model. Such a strong assumption would allow attackers’ free willing of damaging the power grid, deviating from reality.

1.2 Our Work

In this research, we derive a CCA vector with a multi-armed bandit method. This method was initially derived to solve for limited resources allocation between competing choices. For example, a gambler decides which arm of a K-slot machine to pull to maximize his/her reward in a series of trials. During the course, each choice’s properties are only partially known at the time of resource allocation (*i.e.*, with limited prior knowledge), and may become better understood by allocating resources to the choice (*i.e.*, through learning) [10]. We derive the CCA vector under the most vicious attack goal, maximizing the power loss, by modeling the physical mechanism that triggers cascading blackouts. In comparison to existing studies, the primary contribution of the research is twofold:

1. The CCA vector does not require attackers’ prior knowledge of the power grid.
2. The CCA vector considers the behavioral stochasticity of attackers with a Markov-game model, and derive the “optimal” attack strategy with the ϵ -Greedy method using the New England 39-bus power system model.

The rest of the research is organized as follows. Section 2 provides the background, study scope, and assumptions. Section 3 introduces the attack model in a general context. Section 4 derives the CCA vector with the proposed attack model. Section 5 validates the proposed approach using IEEE-39 bus system. Finally, Section 6 concludes the paper.

It should be noticed that the thesis uses control attacks, which falsify/hijack control commands to the power grid transducers, to introduce the proposed approach. We remark, however, that the proposed approach can be extended to measurement-attacks [11]. In measurement

attacks, the attacker corrupts sensors' information to (i) manipulate power application (*e.g.*, Economic Dispatch) or (ii) masquerade control attacks.

2 Background

2.1 Coordinated Control Cyber-Attacks

Control attacks is a primary form of cyber-attacks. They directly falsify/hijack control commands sent to the power grid transducers. For example, the attacker can manipulate the power network configuration by attacking circuit breakers and phase-shifting transformer, leading to significant power flow change and power supply disruption.

In the worst case, control-attacks, if coordinately launched across the grid simultaneously, could lead to cascading power failure even blackouts. This is due to the fact that tripping several heavily loaded lines will further overload other lines and power delivery assets, such as transformers. The tripping-chain will propagate until sufficient demand is disconnected by the operator or the power system protection scheme, remaining the electricity supply and demand in balance. While the device tripping in the first stage is caused by the control attacks, it is the power grid physics that drive the grid's evolution through the rest of stages [2].

2.2 Assumptions

The scope of this paper is limited to several assumptions on the power system's accessibility to attackers. We assume the attackers can access either the control center or remote substations. To hack into these units, the attacker can (i) exploit vulnerabilities in the open communication network (*i.e.*, WAN) connected from the enterprise end, (ii) directly acquire access to the Ethernet (*i.e.*, LAN), for example, through phishing emails. Once get inside these units, the attacker can take control of the Human Machine Interface (HMI), Remote Terminal Units, etc. to execute control cyber-attacks. In addition, we assume that attackers in a CCA are cooperative and do not take competitive actions.

On the contrary, we relax several assumptions in the prior works and profile attackers with the following characteristics:

1. The attackers are adaptive, *i.e.*, they will modify the attack actions after learning the defense strategies.
2. The attackers have limited resources; they can only compromise a limited set of cyber-assets.
3. The attackers are aware of the extent to which their common goal is accomplished, *i.e.*, s/he can measure the attack payoff.

2.3 Defense Implications

We define the entities, including human operators and automatic programs, which detect, intercept, and mitigate the attacks, as defenders of the power grid. A defender could install and reinforce defense-in-depth tools at the cyber-layer, such as, Intrusion Detection Systems (IDS), authentication, firewalls, etc.

During real-time interaction, the defender, with automatic algorithm setups, can distribute the defense resources once observing the attack actions or the power response. Over the long term, the defender could reinforce the critical cyber-assets, which could lead to the maximum expected losses, after learning the CCA vector.

3 Modeling Attack-Defender Interaction

This section introduces the proposed approach in a general context. We model the interaction between the adaptive attacker and the defender with a stochastic repeated game. A stochastic game describes the players' interactions, which returns a payoff (either a positive reward or a negative cost) to the players. In a two-player stochastic game, we define the action sets of the attacker and the defender. The attacks are one-shot, while the players can learn each other's strategies over time.

Notations: Let $\mathbb{R}$ denote the set of real numbers and $\mathbb{R}_+$ be the set of positive real numbers. For a set $\mathcal{V}$, we let $2^{\mathcal{V}}$ denote its power set, *i.e.*, the set of all subsets of $\mathcal{V}$.

3.1 The Defender

We assume all the defense measures are unified under a single entity, which is *The Defender*. Let $\mathcal{D} := \{d_1, \dots, d_N\}$ denote the set of N primary defense strategies, and let a defense configuration for the power grid be a set of defense strategies. In addition, we define the j th defense configuration as $\mathcal{D}_j \in 2^{\mathcal{D}}$, where $2^{\mathcal{D}}$ is the defender's action set.

To each defense configuration, we associate the defense cost map $C_d : 2^{\mathcal{D}} \rightarrow \mathbb{R}_+$ defined as $C_d(\mathcal{D}_j) = \sum_{d \in \mathcal{D}_j} C_d(d)$, where we assume independence among distinct defense strategies. The j th defense strategy $d_j \in \mathcal{D}$ can only *protect* a small set of assets $P_{d_j} \subset \mathcal{A}$. In other words, d_j can protect the asset a_i if and only if $a_i \in P_{d_j}$.

3.2 The Attacker

Under the cooperative attackers' assumption, we aggregate all the attack actions under a single attack entity, defining it as *The Attacker*. In a CCA, the attacker can apply the actions on multiple elements. Let $\mathcal{A} := \{a_1, \dots, a_M\}$ denote the set of M target assets, *i.e.*, grid transducers. Thus, we define the i th coordinated attack as $\mathcal{A}_i \in 2^{\mathcal{A}}$, where $2^{\mathcal{A}}$ is the attacker's action set.

To each coordinated attack, we associate the attack cost map $C_a : 2^{\mathcal{A}} \rightarrow \mathbb{R}_+$ where $C_a(\mathcal{A}_i) = \sum_{a \in \mathcal{A}_i} C_a(a)$. If $\mathcal{A}_i$ is successful, the attacker will disrupt the power grid. The reward for the attacker from this disruption is captured by the map $C_r : 2^{\mathcal{A}} \rightarrow \mathbb{R}_+$. Subsections III-C and III-D introduce the construction of the cost map C_a and the reward map C_r .

3.3 Game Payoff

With the above definitions, the interaction between the Attacker and the Defender can be modeled as a zero-sum game between, in which the two players have oppositely equal objectives. At any instant, it is expressed as

$$r(\mathcal{A}_i, \mathcal{D}_j) = r_a(\mathcal{A}_i, \mathcal{D}_j) = -r_d(\mathcal{A}_i, \mathcal{D}_j), \quad (1)$$

where $r : 2^{\mathcal{A}} \times 2^{\mathcal{D}} \rightarrow \mathbb{R}$ denotes the game cost, r_a is the cost function for the attacker and r_d is the cost function for the defender. The game cost r is computed as the amount of power outage. Given the defense configuration $\mathcal{D}_j$, the game cost is expressed as:

$$r(\mathcal{A}_i, \mathcal{D}_j) = \begin{cases} 0, & \text{if } a \in \cup_{d_k \in \mathcal{D}_j} P_{d_k}, \forall a \in \mathcal{A}_i \\ \gamma, & \text{otherwise} \end{cases} \quad (2)$$

where $\gamma \in \mathbb{R}_+$ quantifies the damage to the grid and is computed in Subsection 4.

It should be noticed that (2) is the definition but cannot be used by either the attacker nor the defender to compute the game cost. This is because they do not know the opponent's strategies; they are unable to map the action ($\mathcal{A}_i$ or $\mathcal{D}_j$) to the cost $r(\mathcal{A}_i, \mathcal{D}_j)$. Thus, in a repeated game,

Figure 1: The power grid evolution after cascading failure

the attacker (*resp.* the defender) estimates the average payoff associated with an action $\mathcal{A}_i$ (*resp.* $\mathcal{D}_j$) as:

$$\begin{aligned}\bar{r}_a(\mathcal{A}_i) &= \frac{\sum (\text{past payoffs with } \mathcal{A}_i)}{\text{frequency of } \mathcal{A}_i} \\ \bar{r}_d(\mathcal{D}_j) &= \frac{\sum (\text{past payoffs with } \mathcal{D}_j)}{\text{frequency of } \mathcal{D}_j}\end{aligned}\tag{3}$$

The cost set of all the actions for the attacker and the defender are $\bar{\mathbf{r}}_a = \{\bar{r}_a(\mathcal{A}_1), \dots, \bar{r}_a(\mathcal{A}_{2^{|\mathcal{A}|}})\}$ and $\bar{\mathbf{r}}_d = \{\bar{r}_d(\mathcal{D}_1), \dots, \bar{r}_d(\mathcal{D}_{2^{|\mathcal{D}|}})\}$, respectively.

3.4 Attack Vector

We consider the attacker maximizing the damage to the grid, while the defender minimizes it, *i.e.*, $\mathcal{A}_i = \arg \max_{\mathcal{A}_i} \bar{r}_a(\mathcal{A}_i)$ and $\mathcal{D}_j = \arg \min_{\mathcal{D}_j} -\bar{r}_d(\mathcal{D}_j)$. Furthermore, the attacker and defender rationalize their actions through exploitation (trying strategies that win in the past) and exploration (trying other strategies) [12]. This rationalization procedure allow us to derive the attack vector with a greed algorithm detailed as below.

Let the greedy factor $\epsilon \in [0, 1]$ measure the level of exploration. For a probability of ϵ , the attacker explores strategies by random selection; For $(1 - \epsilon)$, s/he selects the strategy with the highest reward evaluation.

Denote the probability distribution of the strategies for the defender and the attacker as $\mathbf{f}(\mathcal{D}_j)$ and $\mathbf{g}(\mathcal{A}_i)$, where $\mathcal{D}_j \in 2^{\mathcal{D}}$ and $\mathcal{A}_i \in 2^{\mathcal{A}}$. Clearly, the maps $\mathbf{f} : 2^{\mathcal{D}} \rightarrow [0, 1]$ and $\mathbf{g} : 2^{\mathcal{A}} \rightarrow [0, 1]$ must satisfy $\sum_{j=1}^{2^{|\mathcal{D}|}} \mathbf{f}(\mathcal{D}_j) = 1$ and $\sum_{i=1}^{2^{|\mathcal{A}|}} \mathbf{g}(\mathcal{A}_i) = 1$. Then, the average cost over all possible strategies $\mathbb{E}[C_a]$ for attacker and $\mathbb{E}[C_d]$ for defender are computed as:

$$\begin{aligned}\mathbb{E}[C_a] &= \mathbf{g} \cdot \bar{\mathbf{r}}_a \\ \mathbb{E}[C_d] &= \mathbf{f} \cdot \bar{\mathbf{r}}_d\end{aligned}\tag{4}$$

The algorithm for deriving the CCA vector is summarized in Algorithm 1.

Algorithm 1 ϵ -Greedy multi-armed bandit algorithm

- 1: Initialize the action sets $\mathcal{A}$
 - 2: $n(\mathcal{A}) \leftarrow 0$, $r_a \leftarrow 0$, $\bar{\mathbf{r}} \leftarrow 0$
 - 3: **repeat**
 - 4: $\mathcal{A}_i \leftarrow \begin{cases} \mathcal{A}_i = \arg \max_{\mathcal{A}_i} \bar{r}_a(\mathcal{A}_i), & \text{with probability } 1 - \epsilon \\ \text{random action selection,} & \text{with probability } \epsilon \end{cases}$
 - 5: $n(\mathcal{A}) \leftarrow n(\mathcal{A}) + 1$
 - 6: update $\mathbf{f}(\mathcal{D}_j)$
 - 7: measure the reward r_a from the environment
 - 8: update $\bar{r}_a(\mathcal{A}_i) = \frac{1}{n(\mathcal{A})} (\text{Sum of award taking action } \mathcal{A}_i)$
 - 9: **until** The average cost stabilize
-

4 Power System Implementation

This section derives the game payoff, $\gamma \in \mathbb{R}_+$ in (2), under a specific physical mechanism, the cascading blackouts.

Given $\mathcal{A}_i \in 2^{\mathcal{A}}$ and $\mathcal{D}_j \in 2^{\mathcal{D}}$ denote the attack strategy and defense configuration. To describe the set of assets $a_k \in \mathcal{A}_i$ (*i.e.*, power transducers) not protected by defense configuration $\mathcal{D}_j$, we introduce the following set:

$$\mathcal{A}_{\text{n-pr}} := \{a \in \mathcal{A}_i \mid a \notin \cup_{d_k \in \mathcal{D}_j} P_{d_k}\} \subset \mathcal{A}_i.$$

This set $\mathcal{A}_{\text{n-p}}$ describes the power transducers disconnected during the attack $\mathcal{A}_i$.

After disconnecting the power transducers described in $\mathcal{A}_{\text{n-p}}$, the power flow in another line may exceed their predefined limits. Thus, the line is disconnected, *i.e.*, a single failure occurs. This single failure can trigger other failures. This cascade of failures continues until all parameters stay within their limits. The whole procedure can be simulated through the *Cascading Failure Evolution - Pseudo-Inverse Based (CFE-PB)* algorithm introduced in [2].

The following example illustrates how a CCA triggers cascading blackouts. We consider the IEEE 39-bus New England Power Grid Model, depicted in Fig. 2. In addition, we assume that the attacker and the defender have limited 2 strategies. Specifically, the attack strategies are to disconnect the line (31,6) (*i.e.*, the line connecting buses 31 and 6) or line (8,9), *i.e.*, $\mathcal{A}_i = \{(31,6), (8,9)\}$, and the defender protects the circuit breaker on the line (8,9). Thus, line (31,6) is not protected, *i.e.*, $\mathcal{A}_{\text{n-p}} = \{(31,6)\}$. By disconnecting line (31,6), the Attacker triggers cascading failures. Fig. 1 illustrates the evolution of the cascading initiated by the attack. The defective (*resp.* unaffected) lines are depicted as dotted red (*resp.* solid blue) lines. Fig. 1 also illustrates that the system reaches a new steady state (with reduced load) after four cascading steps.

In this example, the game payoff of a CCA is the amount of load shed $P_{\text{dis}\%}$, which is computed after the CFE-PB algorithm converges, expressed as

$$\gamma = C P_{\text{dis}\%},$$

where $C \in \mathbb{R}_+$ is a normalizing constant, with

$$P_{\text{dis}\%} = \frac{P_{\text{dis}}}{P_t} \in [0, 1],$$

where P_{dis} is the amount of load lost and P_t the total demand in the grid.

Figure 2: IEEE 39-bus system

5 Numerical Experiments

We tested our method using the IEEE New England 39-bus system depicted in Fig. 2. This system has 10 generators, 46 transmission lines (with circuit breakers), and 12 transformers. For simplicity, we assume generators and transformers are well protected, leaving the lines as the potential victims (*i.e.*, $M = 46$ potential targets). The Attack/Defender will know their Reward/Cost after every round of the game.

The greedy factor ϵ is set using the ϵ -Decreasing Strategy [10]: $\epsilon = 1/\log(k)$ when $\epsilon < 1$, and $\epsilon = 1$ elsewhere. In comparison to the ϵ -Greedy method, it rationalize the behavior of players by allowing them to explore more at the beginning and slowly selecting his/her optimal strategy.

5.1 Single Attack Validation

Figure 3: Game Value with Single Attack over Iteration

To illustrate how to obtain the game solution, we provide the following example. First, we considered no Coordinated attack. Every attack aims at one transmission line $\{\mathcal{A}_1, \dots, \mathcal{A}_{46}\}$, thus, 46 attack strategies in total. To validate the game algorithm, we assume the defender having same defense configurations $\{\mathcal{D}_1, \dots, \mathcal{D}_{46}\}$. $\mathcal{D}_i \in 2^{\mathcal{D}}$ can stop $\mathcal{A}_i$ (*i.e.*, $\mathcal{A}_i \subset \cup_{d \in \mathcal{D}_i} P_d$ for $\forall i \in M$).

Fig. 3 shows average cost over iteration. The expected cost reaches its peak value in the first 50 iterations. After ensuring one optimal strategy, the curve then levels down and stabilizes. Fig. 4 shows the frequency of strategies used over 200 iterations, where strategy 14 is the line 6-31, by switching among strategies, the attacker and defender both found the most vulnerable line.

Rational attackers aim to disrupt the power transducers that will cause the most damage to the grid. Thus, to validate the critical targets, we rank all the branches (*i.e.*, lines and transformers) using the Most Vulnerable Edges Selection – Resistance distance Based (MVES-RB) algorithm described in [2]. This algorithm determines the branches that have a greater impact on the power flow, where Line 6-31, 23-36 and 10-32 are ranked at top three which match with the result from the simulation.

5.2 Single Attack on a Fixed Defense

In this experiment, we validate our algorithm using the same action space for the attacker however, fixing the defense strategy by leaving only the line 10-32 and 23-24 unprotected. Fig. 6 shows that the attacker choose to attack line 23-24 by giving only a few explorations. From the MVES-RB algorithm, we know that disconnecting line 23-24 causes higher power shed than disconnecting line 10-32.

5.3 Coordinated Attack

We allow the attacker to hack into two targets at once, therefore, with $M = 46$, he has 1081 possible choices of attack target ($C_{46}^1 + C_{46}^2$). The total number of attacks increases exponentially as allowing more targets in a coordinated attack. For coordinated Attack, we chose ϵ to be

Figure 4: Strategies for Single Attack

Figure 5: Strategies for CCAs

$3/\log(k)$ when $\epsilon < 1$ and 1 elsewhere, allowing more exploration at higher time step. Fig. 7 shows the average cost over iterations, and Fig. 5 shows the strategy frequencies. The strategies are marked with numbers from 1 to 1081 for simplification. The distribution is more spread with more explorations on strategies, but both the defender and the attacker choose their targets on line 6-31 and 16-21. In fact, the line 6-31 single ranked the first in the MVES-RB algorithm, and together with line 16-21 give a greater power loss.

6 Conclusions & Future Works

Coordinated Cyber-Attacks (CCAs) at the power grid are highly threatening and difficult to defend against. In this research, we derive the attack vector for CCAs of most vicious intent, causing the cascading blackouts. The proposed attack vector relaxes two common assumptions in the prior works: (i) the attackers with deterministic strategies, and (ii) the attacker has the prior knowledge of the power grid. In particular, the interaction between the attacker and the defender is modeled with a Markov-game, and the attacker’s strategies are rationalized with the ϵ -Greedy method. This approach provides a much more realistic and reliable estimation of CCA’s action paths. The results of this research provides important implications in defense resource allocation as well as cybersecurity infrastructure reinforcement in the power grid.

7 Acknowledgement

I would like to express my deepest gratitude to my research advisor Jiankang Wang, research mentors Gonzalo Constante, Christian Moya, Ziran Gao and Jack Mao for all the help and guidance in my research!

References

- [1] Y. Tang, Q. Chen, M. Li, Q. Wang, M. Ni, and X. Fu, “Challenge and evolution of cyber attacks in cyber physical power system,” in *2016 IEEE PES Asia-Pacific Power and Energy Engineering*

Figure 6: Attack Strategy for Fixed Defense

Figure 7: Game Value for CCAs over Iteration

Conference (APPEEC), Oct 2016, pp. 857–862.

- [2] S. Soltan, D. Mazauric, and G. Zussman, “Cascading failures in power grids: Analysis and algorithms,” in *Proceedings of the 5th International Conference on Future Energy Systems*, ser. e-Energy ’14. New York, NY, USA: ACM, 2014, pp. 195–206. [Online]. Available: <http://doi.acm.org/10.1145/2602044.2602066>
- [3] F. Pasqualetti, F. Dörfler, and F. Bullo, “Attack detection and identification in cyber-physical systems,” *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, Nov 2013.
- [4] J. Zhang and L. Sankar, “Physical system consequences of unobservable state-and-topology cyber-physical attacks,” *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 2016–2025, July 2016.
- [5] Y. Yuan, Z. Li, and K. Ren, “Quantitative analysis of load redistribution attacks in power systems,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1731–1738, Sept 2012.
- [6] L. Wei, A. I. Sarwat, W. Saad, and S. Biswas, “Stochastic games for power grid protection against coordinated cyber-physical attacks,” *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 684–694, March 2018.
- [7] Y. Xiang and L. Wang, “A robust optimization-based strategy for optimal power system protection considering uncertainties,” in *2017 IEEE Power Energy Society General Meeting*, July 2017, pp. 1–5.
- [8] F. Wu, Z. Hu, and K. Chen, “Game theory based power system security analysis,” in *2011 International Conference on Control, Automation and Systems Engineering (CASE)*, July 2011, pp. 1–3.
- [9] S. Hasan, A. Ghafouri, A. Dubey, G. Karsai, and X. Koutsoukos, “Vulnerability analysis of power systems based on cyber-attack and defense models,” in *2018 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Feb 2018, pp. 1–5.
- [10] J. Vermorel and M. Mohri, “Multi-armed bandit algorithms and empirical evaluation,” in *Machine Learning: ECML 2005*, J. Gama, R. Camacho, P. B. Brazdil, A. M. Jorge, and L. Torgo, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 437–448.
- [11] H. Song, R. Srinivasan, T. Sookoor, and S. Jeschke, *Smart cities: foundations, principles, and applications*. John Wiley & Sons, 2017.
- [12] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, “Gambling in a rigged casino: The adversarial multi-armed bandit problem,” in *Proceedings of IEEE 36th Annual Foundations of Computer Science*, Oct 1995, pp. 322–331.