RGC Ref. No.:
UGC / FDS14 / E01 / 19
(please insert ref. above)

RESEARCH GRANTS COUNCIL
COMPETITIVE RESEARCH FUNDING SCHEMES FOR
THE LOCAL SELF-FINANCING DEGREE SECTOR

FACULTY DEVELOPMENT SCHEME (FDS)

Completion Report
(for completed projects only)

Submission Deadlines:
1. *Auditor's report with unspent balance, if any: within **six** months of the approved project completion date.*
2. *Completion report: within **12** months of the approved project completion date.*

Part A: The Project and Investigator(s)

1. Project Title

Securing Smart-City Infrastructures using Markov Game

2. Investigator(s) and Academic Department(s) / Unit(s) Involved

Research Team	Name / Post	Unit / Department / Institution
Principal Investigator	MA Yu Tak	The Department of Computer Science, The Hang Seng University of Hong Kong
Co-Investigator(s)	N/A	N/A
Others		

3. Project Duration

	Original	Revised	Date of RGC / Institution Approval *(must be quoted)*
Project Start Date	1 Jan 2020		
Project Completion Date	31 Dec 2022	30 June 2023	22 July 2022 (HSUHK)
Duration *(in month)*	31 Dec 2022	30 June 2023	22 July 2022 (HSUHK)

Deadline for Submission of Completion Report	31 Dec 2023	30 June 2024	22 July 2022 (HSUHK)

4.4 Please attach photo(s) of acknowledgement of RGC-funded facilities / equipment.
N/A

Part B: The Final Report

5. Project Objectives

5.1 Objectives as per original application

1. Develop a Markov Game framework, which has the scalability issue resolved, to model the interactions between the provider and the attacker of different critical infrastructures in a smart city under different interaction scenarios, such as different knowledge possessed by the players, and determine the best strategy of the provider.

2. Develop a simulation platform that will interface with other existing simulation tools, which have been developed to model the running of critical infrastructures, to simulate the interactions between the provider and the attacker of the modeled system.

3. Prove the feasibility of the framework model with empirical analysis using the developed simulation platform and widely-used datasets of critical infrastructures.

5.2 Revised objectives

Date of approval from the RGC: N/A

Reasons for the change: N/A

5.3 Realisation of the objectives

(Maximum 1 page; please state how and to what extent the project objectives have been achieved; give reasons for under-achievements and outline attempts to overcome problems, if any)

Among the three objectives of the project, the components involving the Markov Game framework have seen the most substantial progress, while the aspects related to the development of the simulation platform and the empirical analysis have faced the greatest setbacks and remain the least completed.

For instance, a thorough literature review has been conducted, and various approaches to address the "curse of dimensionality" issue in MDPs and Markov Games have been identified. One of such approaches, *lazy evaluation with pruning*, has been implemented. Meanwhile, two simulators, GridLabD for power grid simulation and CloudSim for cloud computing infrastructures, were investigated and were readily integrated with the simulation platform with the help of the research assistant.

The primary reason for this uneven level of achievement is because of the team's inability to recruit a research assistant with the necessary experience in Markov Decision Processes (MDPs) and game theory to assist with the implementation. Despite continuous efforts to post job advertisements and conducted more than ten interviews, as well as increasing the research assistant headcount to two in order to catch up with the progress, the team has been unable to find a suitable candidate to fill this crucial role after the departure of the first one. Without this crucial support, the team has struggled to fully realize the ambitious goals set out for the simulation platform and the empirical evaluation.

From the literature review conducted, it is evident that many of the approaches proposed to address the "curse of dimensionality" issue in MDPs and Markov Games often lack rigorous analytical proofs regarding their performance guarantees or error bounds. Techniques such as the use of basis functions or deep neural networks to approximate the action-value function (Q-function), approximation of the Markov chain with a factored one, or other approximate dynamic programming methods, are frequently only proven to converge under highly stringent conditions that are often violated in practice.

Recognizing the limitations of these theoretical approaches, the research team made the decision to pursue a more empirical, experimental path to study and compare the performance of these various approaches. The goal was to assess these approaches against metrics such as the payoff of the identified "optimal" policy of the player against that of the opponent computed using other approaches, the computational complexity (in terms of time using the same hardware), memory usage, and their sensitivity to different parameter choices (both within the approach and the transition probabilities of the system).

However, without the dedicated support of a research assistant with the appropriate expertise in this domain, the team has been unable to conduct thorough experimentation and comparative analysis as envisioned. The lack of this crucial role has significantly hampered the progress on the simulation platform development and the planned empirical investigations.

5.4 Summary of objectives addressed to date

Objectives *(as per 5.1/5.2 above)*	**Addressed** *(please tick)*	**Percentage Achieved** *(please estimate)*
1. Markov Game framework	✓ Study existing approaches in Markov Decision Process and investigate their applicability to Markov Game ✓ Develop the Markov Game framework with the computation complexity issue addressed	75%
2. Simulation platform	✓ Study existing simulators ✓ Interface with the existing simulators to simulate attack scenarios and obtain the payoff value ✓ Implement the Markov Game framework	70%
3. Empirical study	✓ Perform empirical study using the developed platform with existing simulators	20%

6. Research Outcome

6.1 Major findings and research outcome
(Maximum 1 page; please make reference to Part C where necessary)

An extensive literature review has been conducted, and one promising direction is to use reinforcement learning to find the optimal policy. Reinforcement learning is suitable because even though we have the simulator(s) to provide the rewards, and hence, we do have a "reward model" of the environment, the problem remains intractable due to the huge state space and action space involved. To handle the "curse of dimensionality" in reinforcement learning for Markov Decision Processes and Markov Games, different approaches have been proposed in the literature, such as Approximate Dynamic Programming and Deep Q-Networks.

One possible approach in Approximate Dynamic Programming is to approximate the Q-function (action-value function) by a linear combination of basis functions. These basis functions can be either handcrafted using domain knowledge or found by performing spectral analysis of the Markov chain (through random walk in the underlying topology). Another Approximate Dynamic Programming approach is to exploit the underlying features of the Markov chain and approximate it with a Factored Markov Decision Process, where the state transition model can be represented by a set of Dynamic Bayesian Networks, and the reward function and value function can also be approximately factored.

However, these approaches primarily focus on solving the problem of handling the intractable Q-function. The large number of states and actions in many practical problems remain a challenge that these methods do not fully address. Moreover, the performance (error bound) of these approaches is often lacking in analytical proof. As pointed out by [1], when extending approximate Q-learning algorithms from Markov Decision Processes (MDPs) to Markov Games (MGs), "This approach would yield the same stability and performance guarantees of ordinary Q-learning with function approximation, that is, essentially none." Alternatively, the convergence of these methods may only be proven under stringent conditions that are often violated in practice. For instance, some approaches require that each state-action tuple be visited infinitely often, which may not be practical in many real-world Scenarios (because of cost or time constraints).

Another direction is to perform state aggregation. However, this approach requires knowing the model of the environment, such as the transition probabilities and the reward, or even relying on domain-specific knowledge. While state aggregation can be effective in certain domains, such domain-specific groupings may not be easily generalizable to other types of infrastructures. In those cases, states with similar rewards and transition probabilities, as provided by the model, may be aggregated together. As a preliminary work, we studied a one-shot game between an attacker and a provider under a hierarchical formulation of infrastructures (as a kind of system-aggregation) [First publication in Part C] and data transfer infrastructure [Second publication in Part C]. Nash Equilibrium is derived in these formulations.

Due to the lack of strong theoretical guarantees for the performance of these various approaches, we decided to empirically study and compare their performance through experimentation, which failed to materialize as intended because of the departure of the research assistant and the difficulty in recruiting a suitable replacement, which impacted the progress of the simulation platform development.

In summary, the key research outcomes include the exploration of various approximation and simplification methods to handle the curse of dimensionality in the Markov Game framework, the recognition of the limitations of existing approaches, and the plan to conduct a rigorous empirical study to compare the performance of different approximation techniques. The inability to fully implement the planned empirical study is also acknowledged as a limitation of the current research progress.

[1] Lagoudakis, M., & Parr, R. (2012). Value Function Approximation in Zero-Sum Markov Games. arXiv preprint arXiv:1301.0580.

6.2 Potential for further development of the research and the proposed course of action *(Maximum half a page)*

As discussed in Section 6.1, because of the lack of strong theoretical guarantees for the performance of various approximation and simplification approaches, it is important to evaluate their performance empirically. Moreover, such an empirical study needs to be performed comprehensively. It is because although simulators can be utilized to provide the rewards associated with different states, state transitions are often not modeled by these simulators. It is also difficult for us to model such state transitions exactly because these transition probabilities are not readily disclosed by the infrastructure operators, and there may not be a closed-form analytical solution to derive them. Hence, using only one set of transition probabilities in the simulation may not model the system in practice.

Consequently, the research can be further developed by comparing and evaluating more accurately the true performance of various approximation approaches, such as the use of basis functions, neural networks, or other approximate dynamic programming techniques such as entropy-regularized soft Q-learning, through rigorous empirical studies using different possible and reasonable parameter configurations.

7. Layman's Summary

(Describe in layman's language the nature, significance and value of the research project, in no more than 200 words)

The research project focused on protecting critical cyber-physical infrastructure in smart cities using the mathematical model of Markov Games. This is an important issue because the increasing integration of cyber components in these infrastructures significantly expands the attack surface, making it crucial for infrastructure providers to determine the most effective approach to secure their systems.

Markov Games are an appropriate model for this problem, as they represent situations where multiple decision-makers, such as different organizations or agents, interact with each other and their environment over time, and must make strategic choices that balance their own interests with the interests of others. However, solving Markov Games exactly is computationally intensive and often intractable.

To overcome this challenge, the project explored various approximation and simplification methods proposed in the literature. By leveraging these techniques, the research aimed to enhance the decision-making capabilities of infrastructure providers, enabling them to better mitigate risks and vulnerabilities.

The value of this research lies in its ability to equip infrastructure providers with tools to make more informed and strategic decisions regarding resource allocation and system resilience, even in the face of the strongest potential attackers.

Part C: Research Output

8. Peer-Reviewed Journal Publication(s) Arising Directly From This Research Project

(Please attach a copy of the publication and/or the letter of acceptance if not yet submitted in the previous progress report(s). All listed publications must acknowledge RGC's funding support by quoting the specific grant reference.)

The Latest Status of Publications				**Author(s)** *(denote the corresponding author with an asterisk*)*	**Title and Journal / Book** *(with the volume, pages and other necessary publishing details specified)*	**Submitted to RGC** *(indicate the year ending of the relevant progress report)*	**Attached to this Report** *(Yes or No)*	**Acknowledged the Support of RGC** *(Yes or No)*	**Accessible from the Institutional Repository** *(Yes or No)*
Year of Publication	**Year of Acceptance** *(For paper accepted but not yet published)*	**Under Review**	**Under Preparation** *(optional)*						
2023				**Nageswara S. V. Rao, Chris Y. T. Ma, and Fei He**	**Game-Theoretic Strategies for Cyber-Physical Infrastructures Under Component Disruptions IEEE Transactions on Reliability, 72(2), June 2023**	**N/A**	**Yes (Annex I)**	**Yes**	**Yes** https://researchdb.hsu.edu.hk/view/publication/202300100
	2024			**Nageswara S. V. Rao, Chris Y. T. Ma, and Fei He**	**Game Strategies for Data Transfer Infrastructures Against ML-Profile Exploits IEEE Transactions on Machine Learning in Communications and Networking**	**N/A**	**Yes (Annex II)**	**Yes**	**No**

9. Recognized International Conference(s) In Which Paper(s) Related To This Research Project Was / Were Delivered

(Please attach a copy of each conference abstract)

Month / Year / Place	**Title**	**Conference Name**	**Submitted to RGC** *(indicate the year ending of the relevant progress report)*	**Attached to this Report** *(Yes or No)*	**Acknowledged the Support of RGC** *(Yes or No)*	**Accessible from the Institutional Repository** *(Yes or No)*
N/A						

10. Whether Research Experience And New Knowledge Has Been Transferred / Has Contributed To Teaching And Learning
(Please elaborate)

The research experience and new knowledge gained from this work on security protection of realistic infrastructure has contributed to teaching and learning at a high level, but not in low-level details.

At a high level, this research has highlighted the significant complexity and difficulty involved in providing robust security protections for realistic, large-scale infrastructure systems. The work has demonstrated that these problems often involve huge state and action spaces, making them intractable to solve exactly. This emphasizes to students the inherent challenges in developing effective security measures for real-world, mission-critical systems.

However, the specific technical details of the approaches evaluated, such as Markov Decision Processes, Markov Games, reinforcement learning, and approximate dynamic programming techniques, are not directly covered in the courses offered by the department. These topics tend to be at a more advanced, graduate-level research focus. Hence, while the high-level insights have been transferred, the low-level technical knowledge has not been directly integrated into teaching materials and courses at this time.

In summary, this research experience has meaningfully informed teaching and learning by highlighting the complexity of real-world security challenges, although the specific technical content is not directly covered in existing coursework.

11. Student(s) Trained
(Please attach a copy of the title page of the thesis)

Name	Degree Registered for	Date of Registration	Date of Thesis Submission / Graduation
N/A			

12. Other Impact
(e.g. award of patents or prizes, collaboration with other research institutions, technology transfer, teaching enhancement, etc.)

At a high level, the idea of challenges in developing effective security measures for real-world mission-critical systems has been used to improve the Security chapter of the

course COM3301 Professionalism and Ethics in Computing.

13. Statistics on Research Outputs

	Peer-reviewed Journal Publications	Conference Papers	Scholarly Books, Monographs and Chapters	Patents Awarded	Other Research Outputs (please specify)	
No. of outputs arising directly from this research project	1 published and 1 accepted	0	0	0	Type	No.
					N/A	

14. Public Access Of Completion Report

(Please specify the information, if any, that cannot be provided for public access and give the reasons.)

Information that Cannot Be Provided for Public Access	Reasons
N/A	