REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
05-02-2023	Final Report	20-Jun-2019 - 19-Dec-2022

4. TITLE AND SUBTITLE
Final Report: 5.3.2: Graph Theoretic Approaches for Cyber Physical Security in Networks

5a. CONTRACT NUMBER
W911NF-19-1-0362

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Southern California
Contracts & Grants
3720 S. Flower St.
Los Angeles, CA 90089 -0701

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
72973-NC.15

12. DISTRIBUTION AVAILIBILITY STATEMENT
2 Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Viktor Prasanna
UU	UU	UU	UU		19b. TELEPHONE NUMBER +12-137-4044

RPPR

as of 14-Feb-2023

Agency Code:

Proposal Number: **Agreement Number:**

Organization:
Address: , ,
Country:
DUNS Number: EIN:
Report Date: Date Received:
for Period Beginning and Ending
Title:
Begin Performance Period: **End Performance Period:**
Report Term: -
Submitted By: Email:
Phone:

Distribution Statement: -

STEM Degrees: **STEM Participants:**

Major Goals:

Accomplishments:

Training Opportunities:

Results Dissemination:

Plans Next Period:

Honors and Awards:

Protocol Activity Status:

Technology Transfer:

I certify that the information in the report is complete and accurate:
Signature:
Signature Date:

Major Goals

The major goal of the project is to advance the techniques for cyber security in smart grids in the following ways:

* Develop fast, accurate and robust graph based data-driven models and learning techniques of the smart grids.

* Develop identification techniques for critical/influential nodes of the smart grid using only easily available data such as load, generation, electricity price data, etc.

* Develop optimal protection scheme and robust decision making techniques for safe, smooth and secure grid operations.

*Develop a framework for quantitative evaluation of risks associated with software used in critical domains such as smart grids.

Accomplished under Goals

In this reporting period, we developed advanced techniques for robust learning, fast decision making for safe and secure operations, and a theoretical framework for cyber-insurance. Specifically, we accomplished the following during the reporting period:

(a) We improved the behavior regularized offline reinforcement learning, a state of the art Reinforcement Learning algorithm and proposed BRAC+. First, we proposed quantification of the out-of-distribution actions and conducted comparisons between using Kullback–Leibler divergence versus using Maximum Mean Discrepancy as the regularization protocol. We proposed an analytical upper bound on the KL divergence as the behavior regularizer to reduce variance associated with sample based estimations. Second, we mathematically showed that the learned Q values can diverge even using behavior regularized policy update under mild assumptions. This leads to large overestimations of the Q values and performance deterioration of the learned policy. To mitigate this issue, we added a gradient penalty term to the policy evaluation objective. By doing so, the Q values are guaranteed to converge. On challenging offline RL benchmarks, we demonstrated that BRAC+ outperforms the baseline behavior regularized approaches by 40% ~ 87% and the state-of-the-art approach by 6%. This result was published in the Asian Conference on Machine Learning, 2021.

(b) We developed the use of membership vectors from soft clustering of electricity consumption time series as features for prediction of socio-demographic characteristics. The membership vector indicates how similar each customer is to every consumption pattern cluster. By discovering the correlation between clusters and characteristics, more accurate prediction of characteristics can be performed. Our experiments on a real survey dataset showed that the combination of using statistical features of consumption and clustering membership vector as input features gives better classification accuracy than solely using either type of features. This result was published in the Future Technologies Conference, 2021.

(c) We developed a framework for generating scalable reinforcement learning implementations on multi-core systems. Replay Buffer is a key component of RL algorithms which facilitates storage of samples obtained from environmental interactions and data sampling for the learning process. We defined a new data structure for Prioritized Replay Buffer based on K-ary sum tree that supports asynchronous parallel insertions, sampling, and priority updates. To address the challenge of irregular memory accesses, we proposed a novel data layout to store the nodes of the sum tree that

reduces the number of cache misses. Additionally, we proposed lazy writing mechanism to reduce thread-level synchronization overheads of the Replay Buffer operations. Our framework employs parallel actors to concurrently collect data via environmental interactions, and parallel learners to perform stochastic gradient descent using the collected data. Our framework supports a wide range of reinforcement learning algorithms including DQN, DDPG, etc. We demonstrated the effectiveness of our framework in accelerating RL algorithms by performing experiments on CPU+ GPU platform using OpenAI benchmarks. Our results showed that the performance of our K-ary sum tree based Prioritized Replay Buffer improves the baseline implementations by around 4x-100x. Our proposed synchronization optimizations improved the performance by around 2x-4.4x compared with using a global lock. By plugging our Replay Buffer implementation into existing open source reinforcement learning frameworks, we achieved 1.19x-1.75x speedup for various algorithms. This result was published in International Conference on High Performance Computing, Data, and Analytics (HiPC), 2021.

(d) We studied safe building HVAC control via batch reinforcement learning. Random exploration in building HVAC control is infeasible due to safety considerations. However, diverse states are necessary for RL algorithms to learn useful policies. To enable safety during exploration, we proposed guided exploration by adding a Gaussian noise to a hand-crafted rule-based controller. Adjusting the variance of the noise provides a tradeoff between the diversity of the dataset and the safety. We applied Conservative Q Learning (CQL) to learn a policy. CQL ensures that the trained policy stays within the policy distribution used to collect the dataset, thereby guarantees safety at deployment. To select the optimal policy during the offline training, we applied model-based performance evaluation. We used the widely adopted CityLearn testbed to evaluate the performance of our proposed method. Compared with a rule-based controller, our approach obtained 12%∼35% reduction in ramping, 3%∼10% reduction in 1-load factor, 3%∼8% reduction in daily peak at deployment with less than 10% performance degradation during the exploration. This result was published in the IEEE Transactions on Sustainable Computing, 2022.

(e) We put forward a new method of approximate quantum circuit reconstruction. Current and imminent quantum hardware lack reliability and applicability due to noise and limited qubit counts. Quantum circuit cutting --- a technique dividing large quantum circuits into smaller subcircuits with sizes appropriate for the limited quantum resource at hand --- is used to mitigate these problems. However, classical postprocessing involved in circuit cutting generally grows exponentially with the number of cuts and quantum counts. This article introduces the notion of approximate circuit reconstruction. Using a sampling-based method like Markov Chain Monte Carlo (MCMC), we probabilistically select bit strings of high probability upon reconstruction. This avoids excessive calculations when reconstructing the full probability distribution. Our results show that such a sampling-based postprocessing method holds great potential for fast and reliable circuit reconstruction in the NISQ era and beyond. This research was published in QCE 2022.

(f) We developed a new algorithm that finds optimal non-datapath caching strategies via network flow algorithms. Flash and non-volatile memory (NVM) devices have only a limited number of write-erase cycles. Consequently, when employed as caches, cache management policies may choose not to cache certain requested items in order to extend device lifespan. In this work, we propose a simple single-parameter utility function to model the trade-off between maximizing hit-rate and minimizing write-erase cycles for such caches, and study the problem of developing an off-line strategy for deciding whether to write a new item to cache, and if so which item already in the cache to replace. Our main result is mOPT, an efficient, network flow based algorithm which finds optimal cache management policy under this new setting. This research has been submitted for journal publication.