

INTERNATIONAL
CONFERENCE ON
PARALLEL
PROCESSING

ICPP / 2021 / CHICAGO / USA

AUGUST 9-12, 2021



Multi-Agent Reinforcement Learning based Distributed Renewable Energy Matching for Datacenters

Haoyu Wang, Haiying Shen, Jiechao Gao, Kevin Zheng, Xiaoying Li

University of Virginia



UNIVERSITY
of VIRGINIA



1%
(total electricity used worldwide) [1]

18%
(carbon emission worldwide) [2]

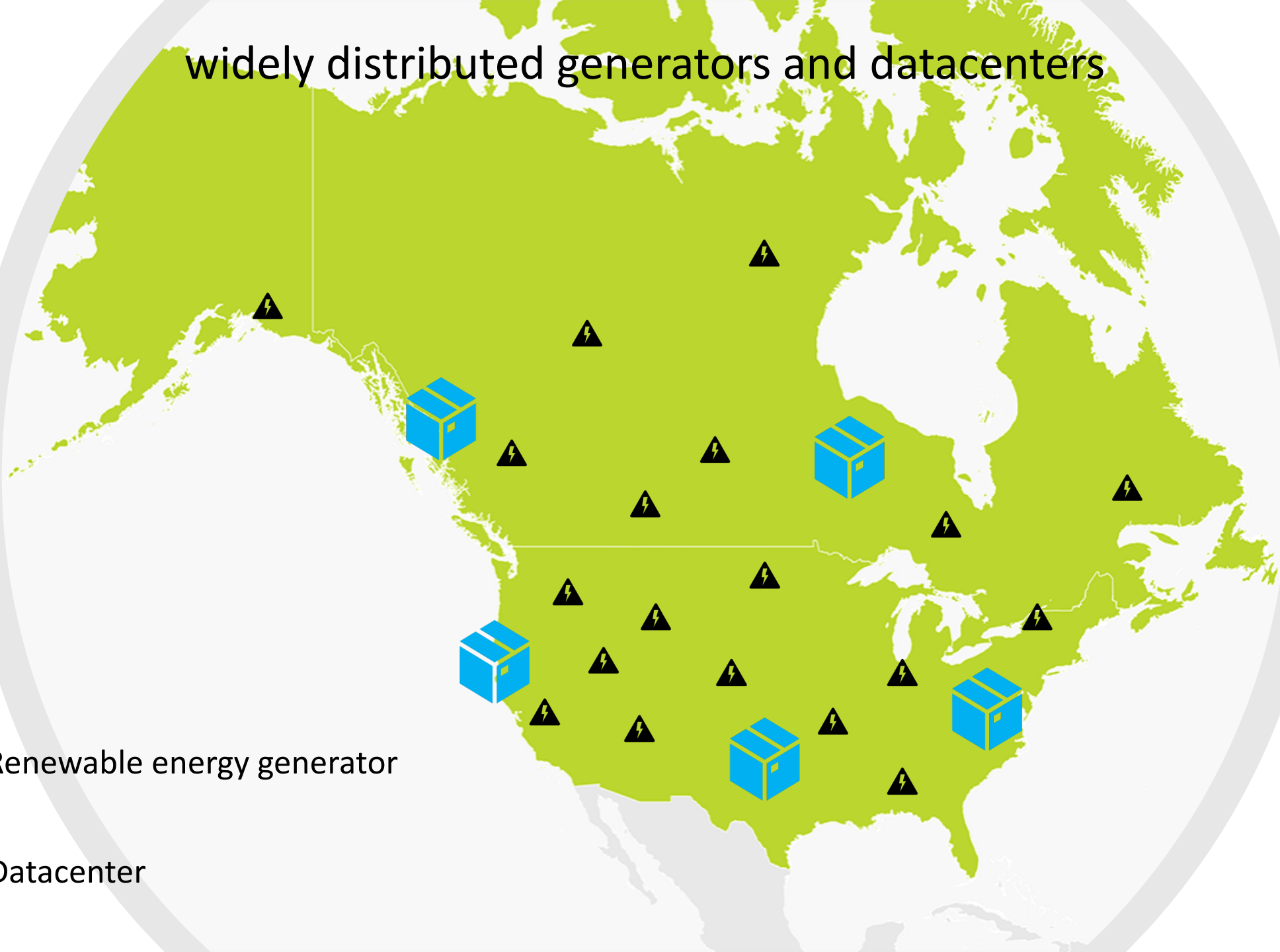


[1] Study: Data Centers Responsible for 1 Percent of All Electricity Consumed Worldwide. In <https://www.datacenterknowledge.com/energy/study-data-centers-responsible-1-percent-all-electricity-consumed-worldwide>.

[2] Consequences of Carbon Emissions for Humans. In <https://sciencing.com/consequences-of-carbon-emissions-for-humans-12730960.html>.



widely distributed generators and datacenters



Renewable energy generator



Datacenter

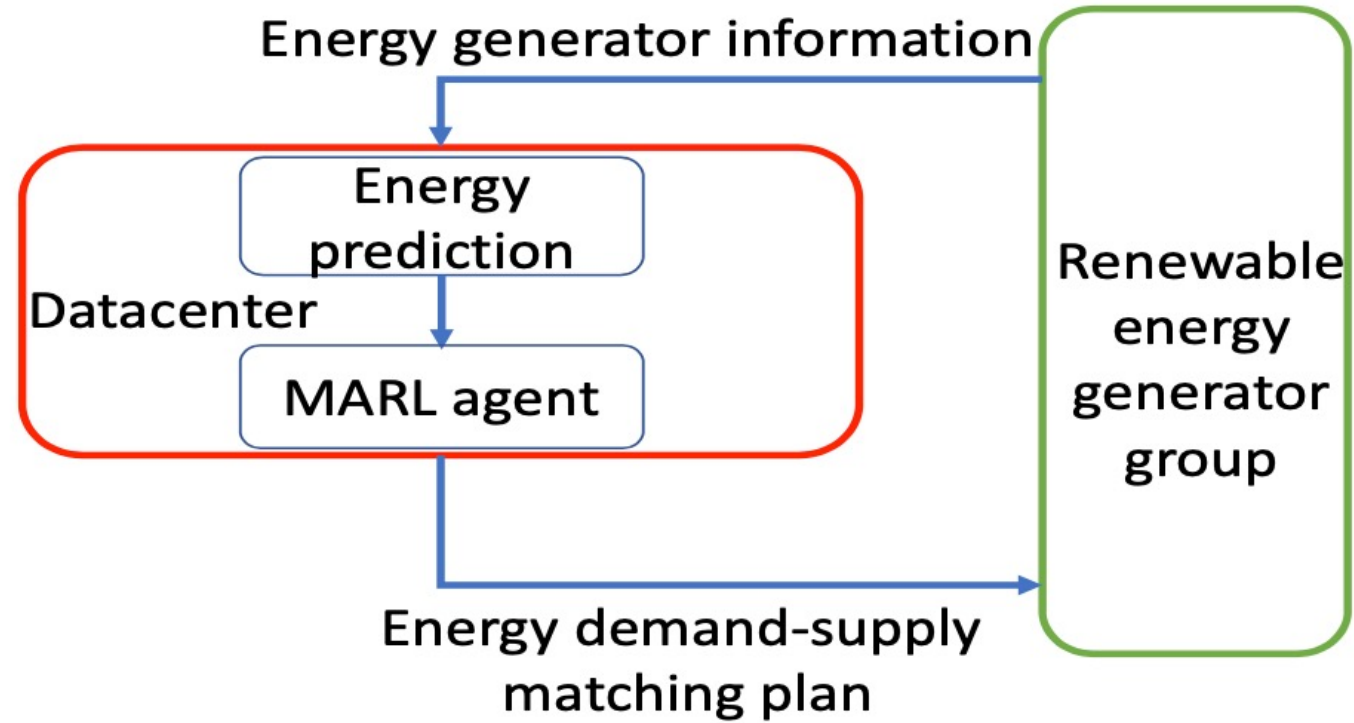
For multiple datacenters and multiple renewable energy generators

How should the renewable energy generators be matched to different datacenters in order to fulfill their energy demands in a long term (e.g., one month) to minimize application SLO violation rate (due to interruption from insufficient renewable energy), total carbon emission and monetary cost of each datacenter?

Related Work

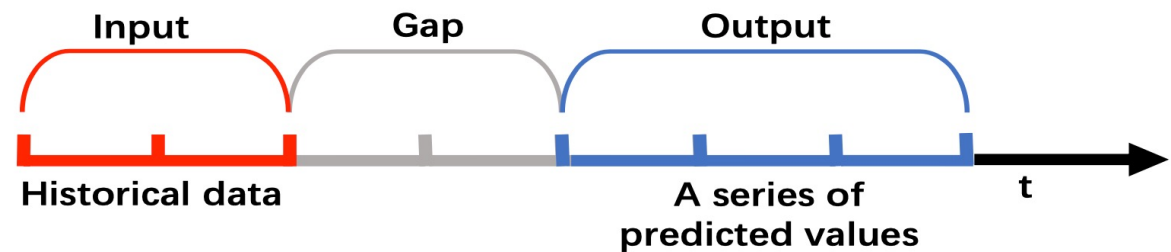
Work	Multi datacenters	Constrained by fixed matchings	Carbon emission	Monetary cost	SLO	Multi CSPs
Cplex[IMSA'2010]	✓	X	✓	X	✓	X
REA [NSE'2018]	✓	X	✓	X	X	X
WST [IOP'2018]	✓	X	✓	X	X	X
TM [CST'2015]	X	X	✓	X	X	X
REM [TSG'2021]	X	X	✓	✓	✓	X
GS [SMTP'2019]	X	X	✓	X	✓	X
FF_LPT [IPDPS'2020]	X	X	✓	✓	X	X
Linear [Access'2019]	✓	✓	X	✓	✓	X
OPT [INFOCOM'2010]	✓	✓	✓	✓	X	X
SRL [ICPP'2020]	X	✓	✓	✓	✓	X
Our work	✓	✓	✓	✓	✓	✓

MARL based System

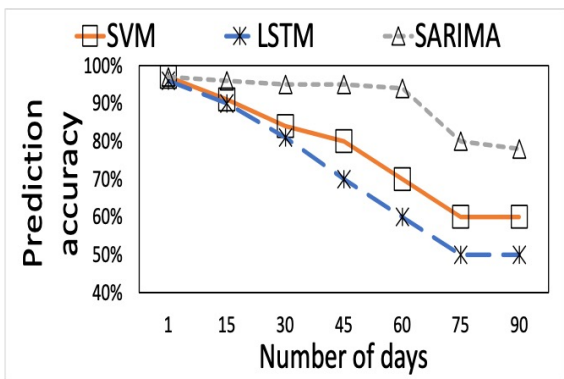
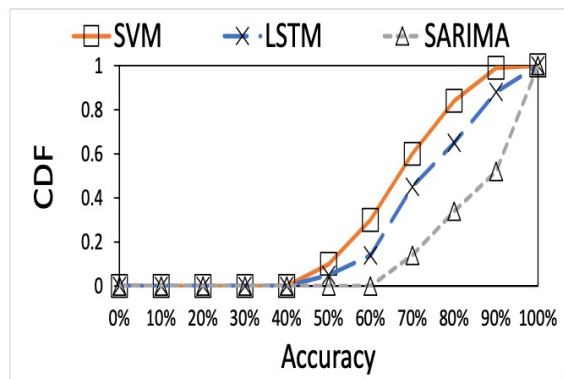
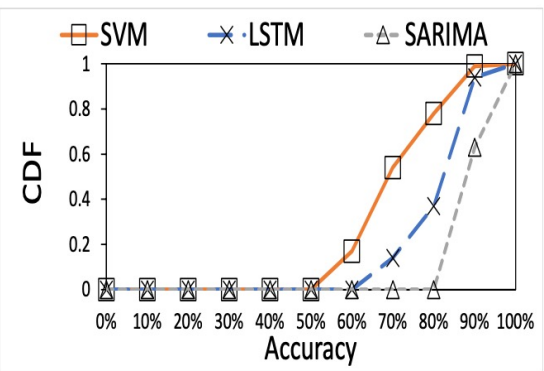
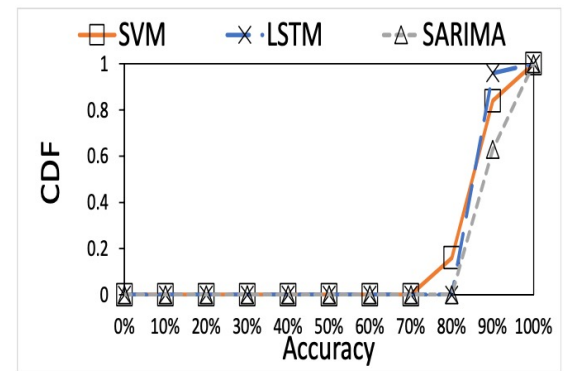


Deadline-Guaranteed Job Postponement is introduced in Section 3.4

MARL based System



Energy prediction



MARL based System

State space: $\mathcal{S}^i = \{\mathcal{D}^i, (\mathcal{G}_{G_1}, u_{G_1}), \dots, (\mathcal{G}_{G_k}, u_{G_k}), \dots, (\mathcal{G}_{G_K}, u_{G_K})\}$ (6)

$$\mathcal{A}^i = \{\mathcal{E}_{G_1}, \mathcal{E}_{G_2}, \dots, \mathcal{E}_{G_k}, \dots, \mathcal{E}_{G_K}\} \quad (7)$$

Action space: where

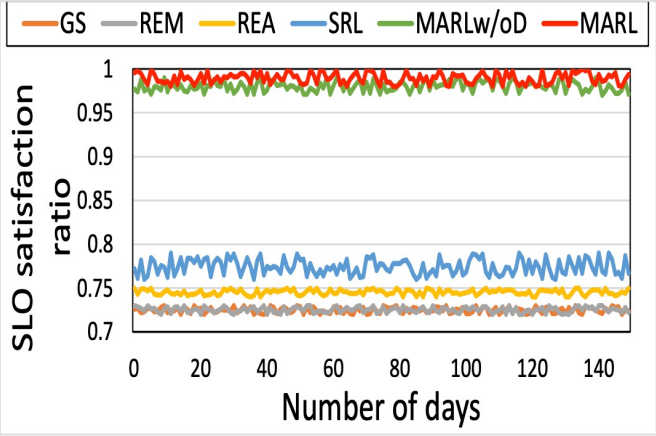
$$\mathcal{E}_{G_k} = \{\mathcal{E}_{G_k, t_1}, \mathcal{E}_{G_k, t_2}, \dots, \mathcal{E}_{G_k, t_z}, \dots, \mathcal{E}_{G_k, t_Z}\} \quad (8)$$

Reward function:
$$\mathcal{R}^i = \sum_{t_z \in Z} \sum_{k \in K} \frac{1}{\alpha_1 * \mathcal{C}_{G_k, t_z}^i + \alpha_2 * W_{G_k, t_z}^i + \alpha_3 * V_{a^i, t_z}}, \quad (11)$$

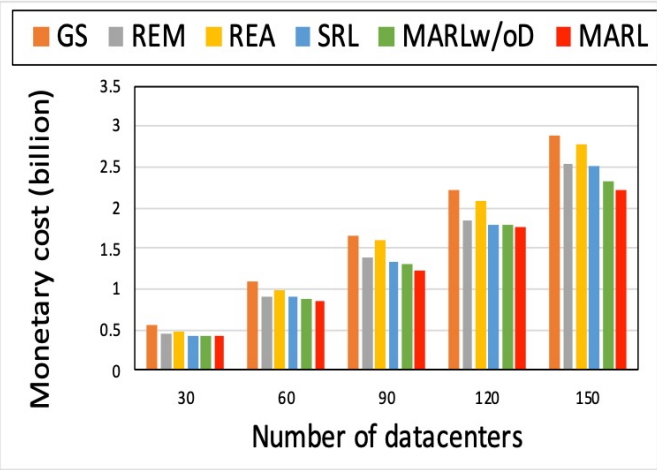
Performance Evaluation

- Workload generation: 30 million Wikipedia web pages into one datacenter and consider one request as one job.
- Experiment scale: 30 to 150 datacenters. 60 renewable energy generators.
- Energy trace: Energy generation trace is from Virginia, Arizona, and California.
- Comparison methods: GS [SMTP'2019], REM [TSG'2021], REA [NSE'2018], SRL [IPDPS'2020]

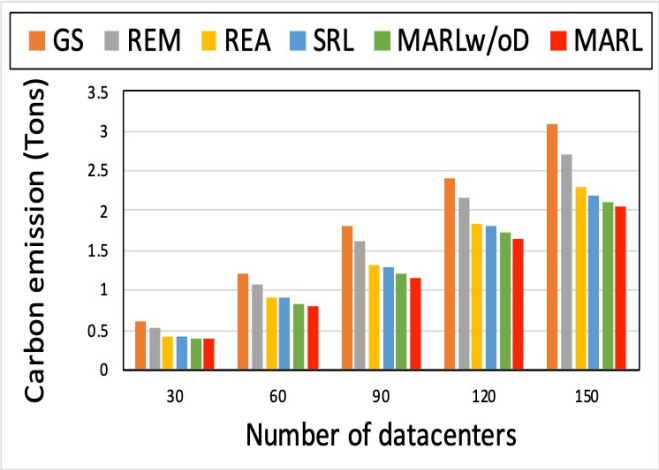
Performance Evaluation



SLO satisfaction ratio for each day



Money cost



Carbon emission

Conclusion

- (1) We compared the prediction accuracies of several ML techniques using real datasets and chose SARIMA that can achieve the highest accuracy.
- (2) Based on the predicted renewable energy supply and energy demand, each datacenter uses MARL to determine how much renewable energy to request from each generator to achieve the goals in the problem above.
- (3) We conduct comprehensive real trace-driven experiments to compare our method with other three methods and the experimental results show that our method can achieve a much lower SLO violation ratio, total energy monetary cost, and total carbon emission compared to the other methods.
- (4) In the future, we will investigate how to jointly conduct workload balance considering the job computing resource competition and how to distribute the generated energy to datacenter requesters.

Thank you!



Questions?

Haoyu Wang: hw8c@virginia.edu