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

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

University of Virginia
1% (total electricity used worldwide) [1]

18% (carbon emission worldwide) [2]


widely distributed generators and datacenters
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

<table>
<thead>
<tr>
<th>Work</th>
<th>Multi datacenters</th>
<th>Constrained by fixed matching</th>
<th>Carbon emission</th>
<th>Monetary cost</th>
<th>SLO</th>
<th>Multi CSPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cplex[IMSA’2010]</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>REA [NSE’2018]</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
<td>X</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WST [IOP’2018]</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
<td>X</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>TM [CST’2015]</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>X</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>REM [TSG’2021]</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>GS [SMTP’2019]</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>FF_LPT [IPDPS’2020]</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>x</td>
</tr>
<tr>
<td>Linear [Access’2019]</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>OPT [INFOCOM’2010]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>SRL [ICPP’2020]</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Our work</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
MARL based System

Energy generator information

Datacenter

Energy prediction

MARL agent

Energy demand-supply matching plan

Renewable energy generator group

Deadline-Guaranteed Job Postponement is introduced in Section 3.4
MARL based System

Energy prediction
MARL based System

State space: 

\[ S^i = \{ D^i, (G_{G_1}, u_{G_1}), ..., (G_{G_k}, u_{G_k}), ..., (G_{G_K}, u_{G_K}) \} \]  

(6)

Action space: 

\[ A^i = \{ \varepsilon_{G_1}, \varepsilon_{G_2}, ..., \varepsilon_{G_k}, ..., \varepsilon_{G_K} \} \]  

(7)

where

\[ \varepsilon_{G_k} = \{ \varepsilon_{G_k,t_1}, \varepsilon_{G_k,t_2}, ..., \varepsilon_{G_k,t_z}, ..., \varepsilon_{G_k,t_{Z}} \} \]  

(8)

Reward function: 

\[ R^i = \sum_{t_z \in Z} \sum_{k \in K} \frac{1}{\alpha_1 \cdot \varnothing^i_{G_k,t_z} + \alpha_2 \cdot W^i_{G_k,t_z} + \alpha_3 \cdot V^i_{a,t_z}} \]  

(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