CHEDDAR

UNIVERSITY OF LEEDS

SPRINT LAB
Smart Pervasive Connectivity Lab

CARES: Carbon-Risk-Aware Reinforcement Learning for Edge Sustainability

David Naseh, Syed Ali Raza Zaidi, Maryam Hafeez

School of Electronic and Electrical Engineering, University of Leeds, UK

MOTIVATION & CRITICAL GAPS

The rapid growth of edge computing enables ultra-low-latency IoT applications but dramatically increases carbon emissions. Existing approaches suffer from **three critical limitations:**

1. Static Trade-off Weights: Unable to adapt to real-time grid carbon intensity changes.

2. Deterministic Renewables: Ignore the inherent uncertainty in solar/wind generation forecasts.

3. Limited Action Spaces: Restrict actions to offloading only, neglecting critical energy levers like DVFS and server sleep modes.

Approach	Dynamic Weights	Uncertainty-Aware	Comprehensive Actions
Heterogeneous MADRL [4]	No	No	No (offloading only)
REP-DRL [1]	No	No	No (offloading only)
CAROF [16]	No	No	Partial (VM placement)
CADDTO-PPO [10]	No	No	Partial (power-control)
CECO [3]	No	No	Partial (CPU+offloading)
TTMADRL [6]	No	No	No (trading only)
CCFRL [11]	No	No	No (client selection)
CARES (Proposed)	Yes	Yes	Yes (DVFS, sleep, migration, offloading)

PROPOSED SOLUTION: CARES FRAMEWORK

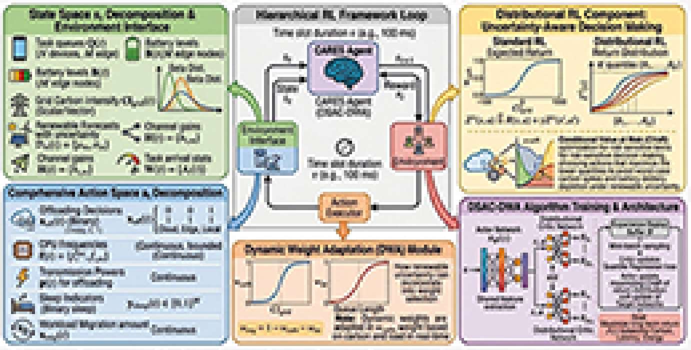
To address these gaps, CARES introduces three innovations:

1. Dynamic Reward Shaping: Dynamically adjusts reward weights using softmax normalization based on real-time grid carbon intensity and renewable availability.

2. Uncertainty-Aware Orchestration: Employs Distributional Reinforcement Learning to explicitly model renewable forecast uncertainty, enabling risk-averse decisions via Conditional Value at Risk (CVaR) optimization.

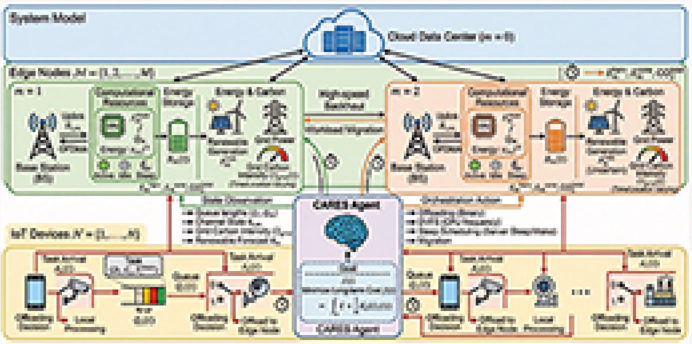
3. Comprehensive Action Space: Utilizes a full range of actions including DVFS (Dynamic Voltage/Frequency Scaling), deep-sleep scheduling, workload migration, and offloading.

DSAC-DWA ALGORITHM ARCHITECTURE



RL agent interacts with the edge orchestration environment. It observes system state, selects actions, and receives a reward shaped by dynamic weights. The distributional critic captures renewable uncertainty, while the actor learns a policy that balances multiple objectives.

CARES SYSTEM ARCHITECTURE



IoT devices generate tasks that can be processed locally or offloaded to edge nodes. Edge nodes are equipped with renewable energy harvesters, batteries, and can communicate with the cloud. The CARES agent observes system state (queue lengths, carbon intensity, renewable forecasts, etc.) and selects actions (offloading, DVFS, sleep scheduling, migration) to minimize a dynamic weighted cost.

EXPERIMENTAL SETUP

Simulator: Custom Python/PyTorch simulator run on NVIDIA RTX 3090 GPUs.

Environment: 10 Edge Nodes (2.0 GHz max), 100 IoT devices.

Traces: Bursty IoT workloads (Healthcare/Autonomous/Video), Sinusoidal Grid Carbon (0.05 - 0.30 kgCO_2/kWh), Beta-distributed Solar generation.

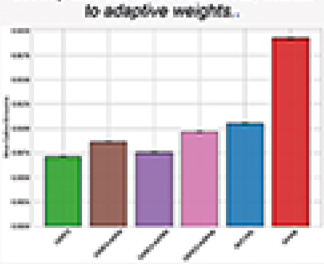
Baselines: REP-DRL (SOTA deterministic), Greedy (Latency-first), CARES-noDWA (Ablation), CARES-noSleep (Ablation).

CORE PERFORMANCE SUMMARY (Mean ± Std over 10 runs)

Algorithm	Carbon (kgCO_2)	Energy Eff. (bits/J)	Latency (ms)	p (vs CARES)
CARES	0.0071 ± 0.0000	9.8 ± 0.0	2.1 ± 0.7	—
CARES-noDist	0.0086 ± 0.0000	8.0 ± 0.0	0.3 ± 0.0	= 0.015
CARES-noDWA	0.0075 ± 0.0000	8.9 ± 0.0	2.1 ± 0.9	= 0.008
CARES-noSleep	0.0097 ± 0.0000	6.9 ± 0.0	20.3 ± 1.7	= 0.012
REP-DRL	0.0105 ± 0.0001	5.8 ± 0.0	60.8 ± 3.4	< 0.001
Greedy	0.0192 ± 0.0001	2.7 ± 0.0	0.1 ± 0.0	< 0.001

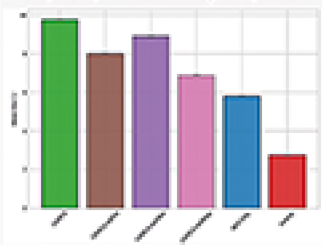
TOTAL CARBON EMISSIONS

CARES reduces emissions by 63.0% vs Greedy and 32.4% vs REP-DRL, thanks to adaptive weights.



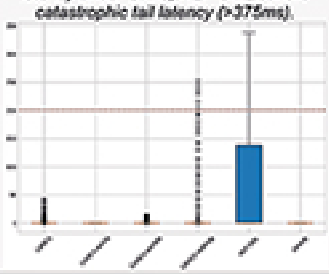
ENERGY EFFICIENCY

CARES achieves 1.7x higher bits/Joule than REP-DRL by leveraging DVFS and deep-sleep to eliminate "ghost power".



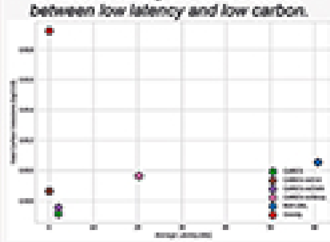
LATENCY DISTRIBUTION

CARES maintains 99th percentile latency under 20ms; REP-DRL suffers catastrophic tail latency (>375ms).



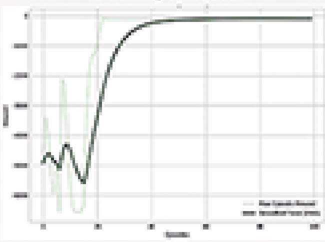
MULTI-OBJECTIVE TRADE-OFF: CARBON VERSUS LATENCY

CARES forms a dominant Pareto frontier, breaking the strict trade-off between low latency and low carbon.



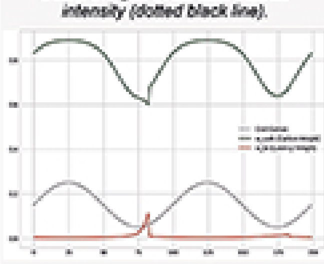
TRAINING CONVERGENCE

DSAC-DWA shows rapid convergence, achieving a near-optimal policy within 40 training episodes.



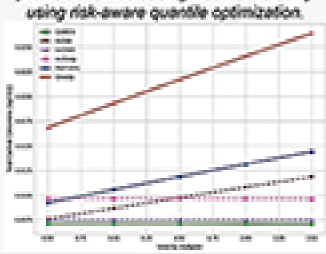
DYNAMIC WEIGHT ADAPTATION

The carbon penalty weight (green) tracks the grid's real-time carbon intensity (dotted black line).



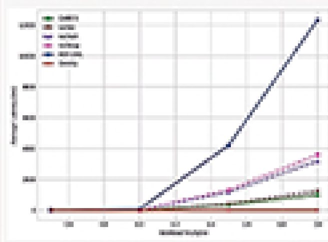
ROBUSTNESS TO RENEWABLE UNCERTAINTY

CARES maintains a flat, low carbon profile under increasing weather volatility using risk-aware quantile optimization.



SCALABILITY

CARES demonstrates graceful degradation (under 1000ms latency) at 2.0x workload, while REP-DRL collapses to 12,000ms.



CONCLUSION & PRACTICAL IMPACT

CARES demonstrates that **sustainability and high performance can coexist** in edge computing. By combining dynamic reward shaping, distributional risk-awareness, and a comprehensive action space, the framework achieves state-of-the-art carbon reductions without violating QoS (latency < 100 ms). The agent's inference time is under 5 ms, making it immediately deployable for edge operators using commercial carbon APIs (WattTime, ElectricityMap).

Acknowledgment: This work was supported by EPSRC CHEDDAR under Grant UK EP/X040518/1.