Traffic-Aware Joint Clustering and Routing in Energy-Harvesting Industrial Internet of Things

Hongyi Zheng, Tingting Zhang, Dongxuan He, *Member, IEEE,* Nan Wu, *Senior Member, IEEE,* Chuhong Yang

***Abstract*—The energy-harvesting industrial IoT networks exhibit intertwined quasi-periodic dynamics in traffic patterns and energy budgeting. However, conventional myopic approaches optimize only instantaneous states, resulting in a mismatch between traffic demand and energy supply. To address this, we present a traffic-aware joint clustering and routing (TJCR) approach. Firstly, we develop a lightweight forecaster that utilizes learnable basis decomposition to predict future traffic and energy profiles. Then, we establish a collaborative optimization scheme that jointly adapts clustering and routing, iteratively refining the clusters based on routing-induced feedback. Simulations demonstrate that TJCR improves network lifetime over representative baselines while maintaining high throughput.**

***Index Terms*—Industrial Internet of Things (IIoT), Traffic prediction, Routing, Clustering, Deep Reinforcement Learning (DRL).**

I. Introduction

The Industrial Internet of Things (IIoT) is a key enabler for smart manufacturing and automation, where wireless sensor networks (WSNs) are widely deployed at the network edge to aggregate and forward mission-critical sensing traffic under stringent energy budgets [1]. To support long-term unattended operation, energy harvesting (EH) is increasingly integrated into such deployments [2]. In EH-IIoT, traffic is often strongly time-correlated and quasi-periodic due to production cycles and control loops [3], while harvested energy (e.g., solar) is time-varying yet typically quasi-periodic with stochastic fluctuations. Consequently, neglecting these time-varying yet predictable dynamics often yields a critical misalignment between energy availability and traffic demands, accelerating energy depletion and precipitating premature network failure.

Existing cluster-based routing protocols predominantly rely on instantaneous state observations to optimize [4]. LEACH [5] employs probabilistic cluster head (CH) rotation, whereas distance-aware CH selection has also been explored to reduce communication cost [6]. Recent learning-enhanced schemes further exploit high-dimensional state features for routing and clustering. WOAD3QN [7] and graph neural network (GNN) based methods [8] respectively exploit deep reinforcement learning (DRL) enhanced optimization and topology-aware representation learning. Nevertheless, clustering and routing are usually addressed separately, without feedback-coupled joint optimization. Moreover, they mainly rely on current observations and lack forecast-aware control, limiting their ability to proactively address future traffic and energy-harvesting variations.

Prior studies have explored IIoT traffic and harvested-energy forecasting, including spatio-temporal graph models for traffic prediction [9], DRL-based predictors for non-stationary traffic patterns [10], [11], and DQN-based traffic-prediction-assisted load-balancing routing [12]. Harvested-energy prediction has also been investigated for EH-enabled WSNs [13]. However, these methods mainly optimize prediction accuracy, while the forecasts are not fully exploited for downstream network control.

To address these issues, we propose a traffic-aware joint clustering and routing (TJCR) approach that moves beyond prediction-only or decoupled clustering-routing designs by using long-horizon traffic and harvested-energy forecasts to guide network reconfiguration. We first develop a lightweight forecaster utilizing period-aware learnable basis decomposition to predict traffic load and energy profiles with low overhead. Based on these forecasts, a collaborative optimization scheme is proposed to jointly optimize clustering and routing. This scheme dynamically shapes the clustering structure to align traffic loads with the energy budget, establishing an optimized topology for routing policies while leveraging routing-induced feedback to iteratively refine the configuration.

II. System Model and Problem Formulation

A. System Model

We consider an IIoT network consisting of a resource-rich data sink $\mathcal{S}$ and N energy-harvesting devices $\mathcal{N}_\mathrm{s} = n_1, \ldots, n_N$, each equipped with a rechargeable battery and a photovoltaic energy harvester. Each device delivers traffic generated by the underlying sensors to $\mathcal{S}$. To avoid prohibitive long-range direct transmissions and improve scalability, we adopt a clustering architecture that partitions $\mathcal{N}_\mathrm{s}$ into disjoint clusters $C^{(\tau_\mathrm{c})} = \{c_1^{(\tau_\mathrm{c})}, \ldots, c_K^{(\tau_\mathrm{c})}\}$. Each cluster is coordinated by a CH $h_j^{(\tau_\mathrm{c})} \in \mathcal{H}^{(\tau_\mathrm{c})}$, which collects intra-cluster traffic via single-hop uplinks and routes the resulting traffic to $\mathcal{S}$ over inter-cluster multi-hop links. In the architecture, sensor nodes periodically report local states to the sink through their corresponding CHs. The sink then constructs state summaries from the received reports for forecasting, clustering, and routing decisions. Since static clustering can cause uneven energy depletion under stringent energy constraints, the system updates the clustering and CH set at the beginning of each decision window τ_c with duration T_c, while keeping the cluster structure fixed within the window to reduce signaling overhead. Nevertheless, traffic

Hongyi Zheng is with the School of Integrated Circuits and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: 3120231368@bit.edu.cn).

Tingting Zhang, Dongxuan He, Nan Wu and Chuhong Yang are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: zhangtt@bit.edu.cn, dongxuan_he@bit.edu.cn, wunan@bit.edu.cn, 3120230733@bit.edu.cn).

Corresponding Author is Tingting Zhang.

and harvested energy fluctuate with finer granularity than T_c. Therefore, each window is divided into multiple slots of equal-duration t (indexed by τ), in which inter-cluster routing is updated to adapt to the time-varying network state. In light of such temporal dynamics, the effectiveness of clustering and routing decisions benefits from the foresight of future traffic patterns and energy availability. To provide long-term foresight, we leverage the quasi-periodic of industrial traffic [3] and solar harvesting by introducing a lightweight Unified Forecasting Module. It extracts temporal features from historical L slots to predict profiles over a long-term horizon of H slots. As illustrated in Fig. 1, these forecasts drive the proposed TJCR approach, which couples clustering and routing in a collaborative optimization loop. The routing agent optimizes paths over the current clusters, while routing-induced feedback updates the cluster structure to align the network configuration with long-term energy budgets and traffic loads. Following this offline joint optimization, the learned policies are deployed to support real-time inference.

Then, we specify the per-slot dynamics. In slot τ, device n_i generates traffic demand $B_i^{(\tau)}$ and harvests energy $H_i^{(\tau)} = \eta_i A_i I_i^{(\tau)} t$, where $A_i, \eta_i, I_i^{(\tau)}$ denote the panel area, efficiency, and irradiance, respectively [13]. On the consumption side, we adopt the standard first-order radio model [5], where the energy to transmit and receive b bits for distance d is $E_{\text{tx}}(b,d) = b(E_{\text{elec}}+\varepsilon_{\text{amp}} d^{\alpha})$ and $E_{\text{rx}}(b) = bE_{\text{elec}}$, respectively. This model adopts a network-layer abstraction, where channel effects are captured by the distance-dependent term d^{α} and practical link impairments are reflected through link feasibility, packet/link loss, and routing-layer link-quality indicators. The total energy consumption $E_{i,\text{cons}}^{(\tau)}$ comprises the detection and aggregation cost $E_{i,\text{c}}^{(\tau)}$ and the communication overhead accumulation:

$$E_{i,\text{cons}}^{(\tau)} = E_{i,\text{c}}^{(\tau)} + \sum_{(b,d)\in\mathcal{P}_i^{(\tau)}} E_{\text{tx}}(b,d) + \sum_{b\in\mathcal{R}_i^{(\tau)}} E_{\text{rx}}(b), \quad (1)$$

where $\mathcal{P}_i^{(\tau)}$ and $\mathcal{R}_i^{(\tau)}$ denote the sets of transmitted and received packets characterized by bit-length b by node n_i in slot τ, respectively. Finally, the battery state evolves according to the harvesting-consumption balance:

$$E_i^{(\tau+1)} = \min\Big(E_{i,\max},\ \max\left(0,\ E_i^{(\tau)} - E_{i,\text{cons}}^{(\tau)} + \eta_{\text{b}} H_i^{(\tau)}\right)\Big), \quad (2)$$

where η_{b} is the charging efficiency. Crucially, specific nodes face a resource misalignment: periods of high traffic demand may coincide with harvesting deficits, necessitating the proposed predictive load-balancing strategy.

B. Problem Formulation

We formulate the joint clustering and routing problem to maximize network lifetime and total throughput over a maximum operational period $T_{\max}$. Let $a_i^{(\tau)} \in \{0,1\}$ denote the alive status of node i ($E_i^{(\tau)} \geq E_{\min}$) and $\Gamma^{(\tau)} \in \{0,1\}$ denote the network functional status. The network is considered alive ($\Gamma^{(\tau)} = 1$) if and only if at least half of the nodes are alive. For clustering, we define binary variable $z_{ik}^{(\tau)}$ for node membership and $y_i^{(k,\tau)}$ for CH selection, where $s_k^{(\tau)} \in \{0,1\}$ indicates

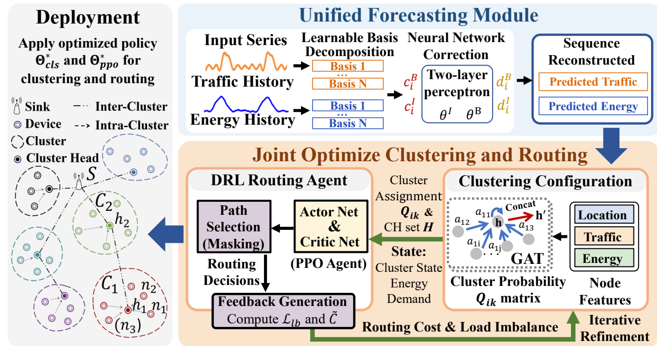

Fig. 1. Overview of the proposed traffic-aware joint clustering and routing (TJCR) approach.

if cluster k is active. The aggregated traffic of cluster k is $L_k^{(\tau)} \triangleq \sum_i z_{ik}^{(\tau)} B_i^{(\tau)}$, which is routed on a directed CH-overlay graph $G_{\text{H}}^{(\tau)}$ via binary link variables $x_{uv}^{(k,\tau)}$. Let $\Lambda_k^{(\tau)} \in \{0,1\}$ indicate whether $L_k^{(\tau)}$ is successfully delivered to $\mathcal{S}$, and denote the neighbor sets on $G_{\text{H}}^{(\tau)}$ by $\mathcal{N}^{+}(u)$ and $\mathcal{N}^{-}(u)$, respectively. The optimization problem (3) is formulated as:

$$\max_{\mathbf{z},\mathbf{y},\mathbf{x}} \quad w_1 \sum_{\tau=1}^{T_{\max}} \Gamma^{(\tau)} + w_2 \sum_{\tau=1}^{T_{\max}} \sum_{k=1}^{K} L_k^{(\tau)} \Lambda_k^{(\tau)} \quad (3a)$$

$$\text{s.t.} \quad E_{\min} - M(1-a_i^{(\tau)}) \leq E_i^{(\tau)} \leq E_{\min} - \varepsilon + M a_i^{(\tau)}, \quad \forall i,\tau, \quad (3b)$$

$$\lceil N/2 \rceil \Gamma^{(\tau)} \leq \sum_{i=1}^{N} a_i^{(\tau)} \leq (\lceil N/2 \rceil - 1) + N\Gamma^{(\tau)}, \forall \tau, \quad (3c)$$

$$\Gamma^{(1)} = 1, \quad \Gamma^{(\tau+1)} \leq \Gamma^{(\tau)}, \quad \forall \tau < T_{\max}, \quad (3d)$$

$$\textstyle\sum_k z_{ik}^{(\tau)} = a_i^{(\tau)}, \quad z_{ik}^{(\tau)} \leq s_k^{(\tau)}, \quad \forall i,k,\tau, \quad (3e)$$

$$s_k^{(\tau)} \leq \textstyle\sum_i z_{ik}^{(\tau)}, \quad \forall k,\tau, \quad (3f)$$

$$\textstyle\sum_i y_i^{(k,\tau)} = s_k^{(\tau)}, \quad y_i^{(k,\tau)} \leq z_{ik}^{(\tau)}, \quad \forall i,k,\tau, \quad (3g)$$

$$\sum_{v\in\mathcal{N}^{+}(u)} x_{uv}^{(k,\tau)} - \sum_{v\in\mathcal{N}^{-}(u)} x_{vu}^{(k,\tau)} = \begin{cases} \Lambda_k^{(\tau)} y_u^{(k,\tau)}, & u \in \mathcal{H}^{(\tau)}, \\ -\Lambda_k^{(\tau)}, & u = \mathcal{S}, \end{cases} \quad \forall u,k,\tau, \quad (3h)$$

$$x_{uv}^{(k,\tau)} \leq \Lambda_k^{(\tau)} \leq s_k^{(\tau)}, \ \Lambda_k^{(\tau)} \leq \Gamma^{(\tau)}, \ \forall (u,v),k,\tau. \quad (3i)$$

where w_1, w_2 are weights balancing lifetime and throughput. The battery state evolves according to (2). Constraints (3b)–(3d) enforce the network lifetime definition via Big-M formulation. Constraints (3e)–(3g) ensure that every alive node belongs to an active cluster and each active cluster selects exactly one CH. Finally, (3h) enforces flow conservation for inter-cluster routing, while (3i) ensures that data is only routed for active clusters while the network is alive. The formulated problem is NP-hard because it generalizes classical location-allocation and constrained routing problems [14].

III. The Proposed Algorithm

A. Unified Traffic and Energy Forecasting via Decomposition

To enable proactive energy management and load balancing, the proposed TJCR approach necessitates accurate forecasts for two distinct time series: traffic demand $B_i^{(\tau)}$ and solar energy

Algorithm 1 Windowed Alternating Co-Training of TJCR

Input: $\hat{\mathbf{B}}, \hat{\mathbf{H}}$, T_c, cycles C, PPO updates N_{ppo}, alive threshold $N_{\min}$
Output: Θ_{cls} (clustering), Θ_{ppo} (routing)

Warm-up: Train Θ_{cls} with $\hat{\mathbf{B}}, \hat{\mathbf{H}}$; fix clustering and warm-start Θ_{ppo}.
for cycle $c = 1, \dots, C$ **do**
    % Phase 1: Global Feedback Collection
    Roll out: Execute current policy $\pi_{\Theta_{\text{ppo}}}$ over the entire horizon H.
    Collect: Record snapshots of network states and routing costs for each window $w \in \{1, \dots, \lceil H/T_c \rceil\}$.
    % Phase 2: Batch Clustering Adaptation
    for window $w = 1, \dots, \lceil H/T_c \rceil$ **do**
        if active nodes in window $w \geq N_{\min}$ **then**
            Build graph using forecasts and feedback from Phase 1.
            Optimize Θ_{cls} via Eq. (9) to generate cluster configurations.
        end if
    end for
    % Phase 3: Routing Policy Improvement
    Environment Update: Apply new cluster configurations to the environment and update Θ_{ppo} for N_{ppo} time steps using PPO.
end for

harvested $H_i^{(\tau)}$. Since industrial operations and solar irradiance both exhibit pronounced periodic patterns (e.g., production cycles and diurnal rhythms), we employ a unified predictive architecture based on periodic basis decomposition for both tasks. Let $\mathbf{x}^{(L)} \in \mathbb{R}^L$ denote the historical sequence of a generic target variable (representing either traffic bits or harvested joules) over the last L slots, and $\mathbf{x}^{\text{gt}} \in \mathbb{R}^H$ denote the ground truth for the future horizon H.

First, a Hann-windowed Fast Fourier Transform (FFT) is applied to $\mathbf{x}^{(L)}$ to identify the set $\mathcal{P} = \{P_1, \dots, P_F\}$ containing the F most significant periods. For each period $P_f \in \mathcal{P}$, we initialize a trainable complex weight vector $\mathbf{w}_f \in \mathbb{C}^{P_f}$ and construct a periodic template $\boldsymbol{\phi}_f \in \mathbb{C}^{L+H}$ by repeating $\mathbf{w}_f$. The basis dictionary $\mathbf{D} \in \mathbb{C}^{F \times (L+H)}$ is formed by stacking these templates, partitioned into look-back $\mathbf{D}^{(L)}$ and prediction $\mathbf{D}^{(H)}$ segments. The projection of the history $\mathbf{x}^{(L)}$ onto the basis is computed via complex inner products. We use the Hermitian inner product in $\mathbb{C}^L$, i.e., $\langle \mathbf{a}, \mathbf{b} \rangle \triangleq \mathbf{a}^{\text{H}}\mathbf{b}$.

$$c_f = \big(\boldsymbol{\phi}_f^{(L)}\big)^{\text{H}} \mathbf{x}^{(L)}, \quad f = 1, \dots, F. \tag{4}$$

The resulting coefficient vector $\mathbf{c} \in \mathbb{C}^F$ effectively encodes the temporal composition. However, since the learned dictionary $\mathbf{D}$ is not strictly orthonormal, the raw coefficients $\mathbf{c}$ suffer from inter-basis crosstalk. Calculating the exact inverse matrix for compensation is computationally prohibitive for IoT nodes. Therefore, we employ a lightweight generic neural network Φ_{cmp} (parameterized by Θ_{cmp}) to approximate the linear compensation, yielding corrected coefficients $\mathbf{d} = \Phi_{\text{cmp}}(\mathbf{c}; \Theta_{\text{cmp}})$. Finally, the horizon-$H$ forecast $\hat{\mathbf{x}}$ is reconstructed via a weighted superposition

$$\hat{\mathbf{x}} = \text{Re}\big(\mathbf{d}^\top \mathbf{D}^{(H)}\big). \tag{5}$$

The parameters $\{\mathbf{w}_f\} \cup \Theta_{\text{cmp}}$ are optimized by minimizing a unified objective function $\mathcal{L}_{\text{pre}}$, which combines the prediction error and a soft orthogonality regularization, given by

$$\mathcal{L}_{\text{pre}} = \frac{1}{H} \|\hat{\mathbf{x}} - \mathbf{x}^{\text{gt}}\|_2^2 + \lambda \big\|\mathbf{D}^{(L)} \mathbf{D}^{(L)\text{H}} - \mathbf{I}_F\big\|_F^2, \tag{6}$$

By training separate instances of this architecture for traffic and energy data, the system provides reliable inputs for the subsequent clustering and routing modules.

B. Learning-Based Joint Clustering and Routing

Leveraging the forecasted profiles, the proposed TJCR approach addresses problem (3) by coordinating a Graph Attention Network (GAT) for periodic cluster adaptation and a Proximal Policy Optimization (PPO) agent to facilitate stable policy learning for masked discrete inter-cluster routing decisions amid dynamic traffic, energy, and link conditions. However, training the two models independently is generally suboptimal due to their cyclic coupling, where the clustering configuration limits feasible routing paths while the routing strategy influences the energy consumption of CHs, modifying the residual energy states that constrain subsequent clustering. To resolve this dependency and stabilize learning, we employ a co-training strategy which incorporates a decoupled warm-up phase followed by an alternating fine-tuning phase. Specifically, the process begins with the warm-up phase, where models are pre-trained individually to establish energy-aware compact clusters and stable initial routing policy. Subsequently, the system transitions to the fine-tuning phase, where both modules are iteratively updated using mutual feedback to enhance coordination, as summarized in Algorithm 1.

During clustering warm-up, we pre-train the GAT to generate window-level clusters from the forecasts, yielding a stable CH overlay for initializing the routing policy. Within each window τ_c, the GAT operates on the geometric graph $\mathcal{G}_n = (\mathcal{N}_s, \mathcal{E}_n)$ constructed over $\mathcal{N}_s$ via k-nearest neighbors, where $|\mathcal{E}_n|$ is the number of edges, using the normalized input vector $\mathbf{h}_i^{(0)} = [x_i, y_i, E_i^{(\tau)}, \bar{b}_i, \bar{h}_i]$. After L_a attention layers, node i has embedding $\mathbf{h}_i^{(L_a)}$. Its soft membership is computed by $Q_{ik} = \text{softmax}_k\big(\mathbf{h}_i^{(L_a)} \mathbf{W}_{\text{cls}}\big)$, $\forall k \in \{1, \dots, K\}$. To adapt cluster boundaries to the heterogeneous energy landscape, we use a hybrid semi-supervised objective that integrates spatial pseudo-supervision with an energy sustainability constraint. The objective seeks to equalize the Traffic-to-Energy Ratio (TER), rather than cluster sizes, thereby aligning traffic demand with energy availability. To achieve this, we formulate a Log-Ratio TER Consistency Loss that robustly penalizes deviations from the network-wide average efficiency $\mathcal{L}_{\text{TER}} = \frac{1}{K} \sum_{k=1}^{K} \text{Huber}\left(\log(\mathcal{D}_k + \epsilon) - \log(\mathcal{E}_k + \epsilon) - R_{\text{ref}}\right)$, where $\mathcal{D}_k = \sum_i Q_{ik} \bar{b}_i$ and $\mathcal{E}_k = \sum_i Q_{ik} (E_i^{(\tau)} + \bar{h}_i)$ represent the aggregated soft demand and potential energy reserves, respectively. R_{ref} is a target TER derived from network-wide statistics. To stabilize the initial learning phase and maintain topological coherence, we incorporate spatial priors via pseudo-labels generated by K-Means. The total composite loss $\mathcal{L}_{\text{cluster}}$ balances this energy-aware alignment with spatial regularization: $\mathcal{L}_{\text{cluster}} = \mathcal{L}_{\text{pseudo}} + \lambda_{\text{b}} \mathcal{L}_{\text{TER}}$, where $\mathcal{L}_{\text{pseudo}}$ denotes the cross-entropy loss against the spatial pseudo-labels, and λ_{b} governs the trade-off. CH election is determined by a suitability score S_i designed to balance energy sustainability with topological centrality: $S_i = \frac{E_{\text{res},i} + \beta \bar{h}_i}{E_{\text{i,max}}} - \text{Cent}_i - \alpha L_{\text{relay},i}$, where $\beta \in [0, 1]$ weights the confidence in harvesting forecasts, Cent_i denotes the normalized distance to the cluster centroid, and the term $\alpha L_{\text{relay},i}$ penalizes normalized historical relay loads to prevent premature node depletion. The node with the highest S_i is identified as the CH candidate.

Moving to routing warm-up, we fix the window-level

This article has been accepted for publication in IEEE Communications Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCOMM.2026.3723351

clustering and the resulting CH overlay from the previous phase. For each active cluster k in slot τ, we pre-compute a compact candidate route set $\mathcal{R}_{\mathrm{f}}^{(\tau)}(k)$ using shortest-path-based enumeration on the overlay, keeping at most R routes per cluster, and let PPO select one route as its action. A route $r \in \mathcal{R}_{\mathrm{f}}^{(\tau)}(k)$ is feasible if it is loop-free, every relay cluster on r has at least one alive node, and the hop count does not exceed $H_{\max}$. To guide PPO over this compact action space, each route r is assigned a per-bit surrogate cost

$$\tilde{E}_{kr}^{(\tau)} = \sum_{(c\to c')\in r} \left(e_{\mathrm{tx}}(d_{cc'}) + \lambda_{\mathrm{e}}\big(1-\bar{\rho}_{c'}^{(\tau)}\big) + \lambda_\phi \frac{\phi_{\mathrm{fut}}^{(\tau)}(c')}{\bar{\rho}_{c'}^{(\tau)}+\epsilon} \right), \quad (7)$$

where $d_{cc'}$ denotes the overlay-hop distance, the per-bit transmit energy is $e_{\mathrm{tx}}(d) \triangleq E_{\mathrm{tx}}(1,d)$, $\bar{\rho}_c^{(\tau)}$ is the residual energy ratio of cluster c, and $\phi_{\mathrm{fut}}^{(\tau)}(c)$ is the normalized look-ahead traffic intensity. The term $\lambda_\phi \phi_{\mathrm{fut}}^{(\tau)}(c')/(\bar{\rho}_{c'}^{(\tau)}+\epsilon)$ penalizes routes traversing relay clusters with high predicted future load. Here $\tilde{E}_{kr}^{(\tau)}$ is a normalized surrogate cost, and $\lambda_{\mathrm{e}}, \lambda_\phi$ control the energy-scarcity and look-ahead penalties. Within each slot τ, PPO operates in an accumulate-commit manner. It assigns route indices to active clusters over several micro-steps, while the environment only records the selected indices and returns a lightweight planning reward; the slot index τ is not advanced. At micro-step t_c of slot τ, the policy selects a feasible route for the current cluster based on an augmented observation $\mathbf{o}_\tau = \big[\tilde{\mathbf{B}}^{(\tau)},\ \mathbf{g}_\tau,\ \tilde{\mathbf{B}}_{\mathrm{fut}},\ \tilde{\mathbf{H}}_{\mathrm{fut}},\ \tilde{\mathbf{E}}_{\mathcal{H}}^{(\tau)},\ \mathbf{c}_\tau\big]$, where $\tilde{\mathbf{B}}^{(\tau)}$ is the normalized current demand, $\mathbf{g}_\tau$ summarizes global energy statistics, including the mean and variance of $\{E_i^{(\tau)}\}$, $\tilde{\mathbf{B}}_{\mathrm{fut}}$ and $\tilde{\mathbf{H}}_{\mathrm{fut}}$ are window-level look-ahead forecasts, $\tilde{\mathbf{E}}_{\mathcal{H}}^{(\tau)}$ collects CH energy states, and $\mathbf{c}_\tau$ encodes normalized surrogate-cost features of $\{\tilde{E}_{kr}^{(\tau)}\}_{r\in\mathcal{R}_{\mathrm{f}}^{(\tau)}(k)}$ relative to $\tilde{E}_{k,\min}^{(\tau)} = \min_{r\in\mathcal{R}_{\mathrm{f}}^{(\tau)}(k)} \tilde{E}_{kr}^{(\tau)}$. The environment commits after a sufficient set of active clusters has been assigned or a micro-step budget is reached, and then advances to $\tau+1$. The reward includes a micro-step shaping term and a slot-level commit reward. At each micro-step for cluster k, we penalize the relative suboptimality of the selected route by comparing its surrogate cost to the best feasible one, scaled by λ_{p}: $r_{\mathrm{plan}}(k,r) = -\lambda_{\mathrm{p}} \max\left(0, \frac{\tilde{E}_{kr}^{(\tau)}}{\tilde{E}_{k,\min}^{(\tau)}+\epsilon} - 1\right)$. A slot-level reward is returned upon commit at the end of slot τ, given by

$$\begin{aligned} r_{\mathrm{commit}}^{(\tau)} = {} & \beta_S \tanh\big(\kappa_S S^{(\tau)}\big) - \beta_E \tanh\big(\kappa_E E^{(\tau)}\big) \\ & - \beta_V \tanh\big(\kappa_V V^{(\tau)}\big) - \beta_A \mathcal{P}^{(\tau)}. \end{aligned} \quad (8)$$

Here, $S^{(\tau)}$ is the delivered-bit ratio in slot τ, $E^{(\tau)}$ is the normalized energy expenditure, and $V^{(\tau)}$ is the variance of normalized residual energies over alive nodes. The auxiliary penalty $\mathcal{P}^{(\tau)}$, with a small weight β_A, aggregates undesirable events, including newly depleted nodes, direct-to-sink fallback transmissions due to routing infeasibility, and high energy per delivered bit. The $\tanh(\kappa\,\cdot)$ function bounds each reward component and smooths scale variations, preventing large-magnitude terms from dominating the update.

To close the co-training loop, routing statistics are collected under the current clustering configuration. For each cluster k, let $\bar{C}_k$ denote the traffic-weighted mean per-bit cost of its selected routes, and let $\bar{p}_{kj}$ denote the corresponding fraction of traffic passing through CH h_j. These statistics are detached

TABLE I

SIMULATION SETTINGS AND HYPERPARAMETERS.

Item	Setting
Deploy.	N devices around industrial-zone centroids; sink $\mathcal{S}$
Energy	$E_0 = 100$ J; 1 cm^2 solar panel; 10% efficiency
Time	$t = 10$ min; $T_c/t = 40$
Radio	$E_{\mathrm{elec}} = 50$ nJ/bit; $E_{\mathrm{fs}} = 100$ pJ/bit/m^2; $E_{\mathrm{mp}} = 1.3\times10^{-15}$ J/bit/m^4; $E_{\mathrm{agg}} = 50$ nJ/bit
Scenarios	S1: $(200, 400\times500)$; S2: $(200, 500\times600)$; S3: $(400, 400\times500)$; S4: $(400, 500\times600)$
Forecast	$(L, H, F, \lambda) = (144, 144, 5, 10^{-3})$
GAT	$L_a = 2$; heads = 3; $d_a = 32$; dropout = 0.4; lr = 5×10^{-4}
PPO	MLP $[128, 64]$; lr = 2×10^{-4}; rollout = 1024; batch = 512; $\gamma = 0.995$; $\epsilon = 0.2$; entropy = 0.001
Reward	$(\beta_S, \beta_E, \beta_V, \beta_A) = (1, 0.3, 0.6, 0.3)$; $(\kappa_S, \kappa_E, \kappa_V) = (1, 2, 2)$

during clustering updates and coupled with the current soft cluster demand D_k as $\widetilde{C} = \frac{\sum_{k=1}^K D_k \bar{C}_k}{\sum_{k=1}^K D_k + \epsilon}$, $u_j = \sum_{k=1}^K D_k \bar{p}_{kj}$. Here, $\bar{C}_k$ is the traffic-weighted mean per-bit cost of the routes selected for cluster k, and $\bar{p}_{kj}$ is the corresponding fraction of traffic passing through CH h_j. We calculate Jain's index $J(\mathbf{u})$ and set $\mathcal{L}_{\mathrm{lb}} = 1 - J(\mathbf{u})$. The clustering model is then fine-tuned using

$$\mathcal{L}_{\mathrm{cluster}}^{+} = \mathcal{L}_{\mathrm{cluster}} + \eta \widetilde{C} + \mu \mathcal{L}_{\mathrm{lb}}. \quad (9)$$

Since D_k depends on Q, both feedback terms provide gradients for clustering updates.

For analyzing the computational complexity, we focus on the inference cost at deployment. Accordingly, the amortized per-slot complexity can is $O\big(\mathcal{T}_{\mathrm{pre}}/H + \mathcal{T}_{\mathrm{cls}}/T_{\mathrm{c}} + K\mathcal{T}_\pi\big)$, where $\mathcal{T}_{\mathrm{pre}} = O(N\,(L\log L + F(L{+}H) + FH))$ denotes the per-run costs of forecasting, $\mathcal{T}_{\mathrm{cls}} = O\big(L_a\left(|\mathcal{E}_n| d_a + N d_a^2\right) + N d_a K\big)$ denotes clustering with hidden embedding dimension d_a, and $\mathcal{T}_\pi = O(N(o)\, d_\pi + d_\pi R)$, where $N(o)$ is the observation dimension with hidden width d_π. In the default setting in Table I, the proposed forecasting module contains only 1.3K parameters and requires only 6.5K MACs per inference.

IV. Simulation Results and Discussion

We evaluate the effectiveness of TJCR using real-world **Telecom Milan traffic traces** [11], [15] to emulate bursty industrial data generation. These traces are time-synchronized with **NREL solar irradiance profiles** [13] to simulate realistic energy harvesting dynamics. The main simulation settings and model hyperparameters are summarized in Table I. All experiments are conducted on a 12th Gen Intel(R) Core(TM) i7-12700K CPU platform. The baselines include LEACH [5], FC-CRA [16], NVE-C [8], LBR-DQN [12], and HCEH-UC [17], covering classical clustering, adaptive clustering/routing, learning-based clustering, prediction-aware routing, and EH-aware clustering/routing.

Fig. 2 illustrates the number of alive nodes (NAN), network residual energy (NRE), and cumulative network throughput (CNT) for scenario S1. As shown in Fig. 2(a), TJCR maintains the largest NAN in S1, with most nodes remaining alive after around 150 slots. Fig. 2(b) shows that TJCR preserves higher residual energy throughout 300 slots, whereas several baselines nearly deplete their energy. As a result, TJCR achieves the highest CNT in Fig. 2(c), exceeding 4×10^9 bits at 300 slots. These gains mainly come from long-horizon forecasting and routing-feedback-guided clustering, which align future load

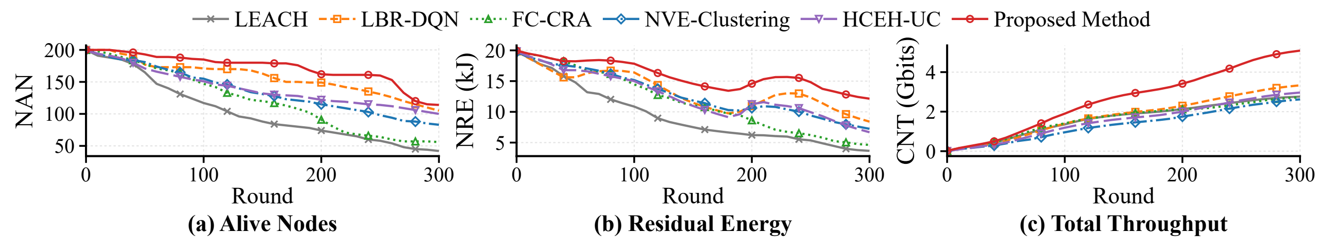

Fig. 2. Performance comparison in terms of (a) the number of alive nodes, (b) the network residual energy, and (c) the cumulative delivered data.

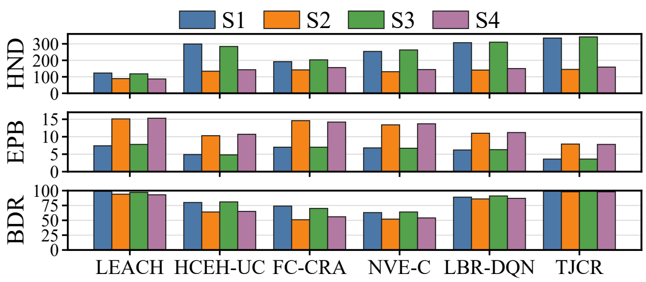

Fig. 3. Key performance metrics under different scenarios.

TABLE II
Expanded ablation results on S1 and S2.

Variant	S1			S2		
	HND	EPB	BDR	HND	EPB	BDR
w/o Prediction	275	4.6	95	118	9.6	95
w/o Clustering	251	4.8	94	104	11.2	93
w/o PPO Routing	238	5.6	92	100	12.0	87
w/o Co-training	302	4.0	97	127	8.8	96
TJCR (Full)	**335**	**3.6**	**99**	**145**	**7.9**	**98**

with harvested-energy budgets and mitigate relay hotspots. To further assess scalability, Fig. 3 summarizes the half-node depletion (HND), average energy per delivered bit (EPB, in μJ/bit), and bit delivery ratio (BDR) for S1–S4, where HND corresponds to the network lifetime metric defined in Sec. II-B. Overall, TJCR achieves longer HND and the lowest EPB across the evaluated scenarios, demonstrating improved lifetime and energy efficiency while maintaining competitive BDR.

To quantify the contribution of each module, ablation studies are conducted: w/o Prediction removes long-horizon traffic and energy forecasts, w/o Clustering disables GAT-based adaptive clustering, w/o PPO Routing replaces PPO with shortest-path routing, and w/o Co-training removes routing-feedback-guided refinement. Table II shows that the full TJCR achieves the best overall performance. The results show that prediction supports HND improvement by enabling lifetime-oriented decisions, while GAT clustering and PPO routing improve energy efficiency through better load organization and dynamic path selection. The gap between TJCR and w/o Co-training further confirms that feeding routing outcomes back to clustering strengthens clustering–routing coordination.

V. Conclusion

This letter presented TJCR, a traffic-aware joint clustering and routing approach for energy-harvesting IIoT WSNs, which integrated long-horizon traffic and energy forecasting, routing-feedback-guided clustering refinement, and PPO-based hierarchical routing. Simulations showed that TJCR alleviated relay hotspots and improved energy sustainability, achieving longer network lifetime than the considered baselines while maintaining high throughput. Future work will extend TJCR to more complex and realistic EH-IIoT scenarios.

References

[1] T. Z. H. Ernest and A. S. Madhukumar, "Age of information outage probability analysis for computation offloading in IIoT networks," *IEEE Commun. Lett.*, vol. 27, no. 9, pp. 2471–2475, 2023.

[2] A. Barat, P. K. J., and S. Bhatnagar, "Energy management in a cooperative energy harvesting wireless sensor network," *IEEE Commun. Lett.*, vol. 28, no. 1, pp. 243–247, 2024.

[3] D. E. Ruíz-Guirola, O. L. A. López, S. Montejo-Sánchez, R. D. Souza, and M. Bennis, "Performance analysis of ML-based MTC traffic pattern predictors," *IEEE Wireless Commun. Lett.*, vol. 12, no. 7, pp. 1144–1148, 2023.

[4] J. Leu, T. Chiang, M. Yu, and K. Su, "Energy efficient clustering scheme for prolonging the lifetime of wireless sensor network with isolated nodes," *IEEE Commun. Lett.*, vol. 19, no. 2, pp. 259–262, 2015.

[5] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, 2002.

[6] S. H. Kang and T. Nguyen, "Distance based thresholds for cluster head selection in wireless sensor networks," *IEEE Commun. Lett.*, vol. 16, no. 9, pp. 1396–1399, 2012.

[7] X. Yang, J. Yan, D. Wang, Y. Xu, and G. Hua, "WOAD3QN-RP: An intelligent routing protocol in wireless sensor networks—a swarm intelligence and deep reinforcement learning based approach," *Expert Syst. Appl.*, vol. 246, p. 123089, 2024.

[8] M. Saadati, S. M. Mazinani, A. A. Khazaei, and S. J. Seyyed Mahdavi Chabok, "Energy efficient clustering for dense wireless sensor network by applying graph neural networks with coverage metrics," *Ad Hoc Netw.*, vol. 156, p. 103432, 2024.

[9] R. Wang, Y. Zhang, L. Peng, G. Fortino, and P.-H. Ho, "Time-varying-aware network traffic prediction via deep learning in IIoT," *IEEE Trans. Ind. Informat.*, vol. 18, no. 11, pp. 8129–8137, 2022.

[10] L. Nie *et al.*, "A reinforcement learning-based network traffic prediction mechanism in intelligent Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 2169–2180, 2021.

[11] Y. Wang, C. Pan, F. Gao, H. Shi, and J. Shi, "Reinforcement learning empowered grey system theory: Advancing intelligence in traffic prediction for wireless sensor networks," *IEEE Sensors J.*, vol. 24, no. 14, pp. 22 023–22 036, 2024.

[12] Q. Ding *et al.*, "Traffic prediction and load balancing routing algorithm based on deep q-network for sd-iot," *Adv. Eng. Informat.*, vol. 68, p. 103596, 2025.

[13] D. K. Sah, A. Hazra, R. Kumar, and T. Amgoth, "Harvested energy prediction technique for solar-powered wireless sensor networks," *IEEE Sensors J.*, vol. 23, no. 8, pp. 8933–8940, 2023.

[14] O. Kariv and S. L. Hakimi, "An algorithmic approach to network location problems. part i: The p-centers," *SIAM Journal on Applied Mathematics*, vol. 37, no. 3, pp. 513–538, 1979.

[15] G. Barlacchi *et al.*, "A multi-source dataset of urban life in the city of milan and the province of trentino," *Sci. Data*, vol. 2, p. 150055, 2015.

[16] B. Fan and Y. Xin, "A clustering and routing algorithm for fast changes of large-scale WSN in IoT," *IEEE Internet Things J.*, vol. 11, no. 3, pp. 5036–5049, 2024.

[17] B. Han, F. Ran, J. Li, L. Yan, H. Shen, and A. Li, "A novel adaptive cluster based routing protocol for energy-harvesting wireless sensor networks," *Sensors*, vol. 22, no. 4, p. 1564, 2022.