Journal of Theoretical and Applied Information Technology
31st October 2026. Vol.104. No.20
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

COMMUNICATION-EFFICIENT DECENTRALISED MULTI-AGENT REINFORCEMENT LEARNING FOR COMMUNITY-LEVEL ENERGY MANAGEMENT USING TOPOLOGY-AWARE ADAPTIVE STATE COMPRESSION

DR A BRAHMANANDA REDDY[1], MOSURI SANTHOSH KUMAR[2], D. CHAKRADHAR[3], DR K. SURYANARAYANA[4], MS Y. LAKSHMI[5], DR GANTA JACOB VICTOR[6], ELANGOVAN MUNIYANDY[7], DR KISHORE KUMAR KAMARAJUGADDA[8],

[1] Associate Professor, Department of Computer Science and Engineering, Vallurupally Nageswara Rao Vignana Jyothi Institute of Engineering and Technology (VNRVJIET), Hyderabad, Telangana, India.
[2] Assistant Professor, Department of Electrical and Electronics Engineering, GMRIT Deemed to be University, Rajam, Andhra Pradesh, India – 532127
[3] Assistant Professor, Department of Management Sciences, R.V.R. & J.C. College of Engineering, Guntur, Andhra Pradesh, India
[4] Associate Professor, Department of Management Sciences, R.V.R. & J.C. College of Engineering, Chowdavaram, Guntur, Andhra Pradesh, India
[5] Assistant Professor, Department of School of Business, Aditya University, Surampalem, Andhra Pradesh, India – 533437
[6] Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur District, Andhra Pradesh, India – 522302
[7] Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India – 602105
[8] Professor, Department of Electronics and Communication Engineering, ICFAI Foundation for Higher Education, Hyderabad, Telangana, India.

E-Mail: [1] brahmanandareddy_a@vnrvjiet.in, [2] santhoshkumar.m@gmritdu.edu.in, [3] chakri359@gmail.com, [4] suryanarayana_k@yahoo.com, [5] lakshmim@adityauniversity.in, [6] jacob.victor@kluniversity.in, [7] muniyandy.e@gmail.com, [8] kishorekamarajugadda@gmail.com.

ABSTRACT

Decentralised energy management in residential communities is challenging. Agents need to organise the use of distributed energy resources while also following the rules of the shared network. The topology of a distribution network can be shown by graph-based multi-agent reinforcement learning (MARL), but big state representations may make it more difficult for communities to communicate and compute as they grow. To minimise shared state information while maintaining characteristics important for energy-management choices, this study suggests the Topology-Aware Graph Transformer with Adaptive State Compression (TAGT-ASC) framework. The technique combines proximal policy optimisation (PPO)-based multi-agent learning with a Graph Transformer encoder and a learnt gating mechanism. Incorporating the IEEE 33-bus distribution-network benchmark and real-world OpenCEM microgrid observations, the assessment framework is complete. With an operational cost of ₹143.60/day, 317.90 kWh/day of network losses, 0.0069 p.u. of mean voltage deviation, 184.50 kWh/day of PV curtailment, and 96.40 kWh/day of battery throughput, TAGT-ASC outperforms GAT-MAPPO, Graph-MASAC, and Graph-TD3. There is a 62.5% drop in communication overhead when adaptive compression is used compared to the uncompressed setup. There is also a drop in median inference latency from 4.80ms to 3.40ms and a drop in peak memory from 420 MB to 325 MB. These findings suggest that adaptive state compression may be used to strike a balance between computational efficiency and energy-management performance. To demonstrate resilience and scalability in the face of disruptions and growing agent counts, more quantified tests are needed.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Keywords: *Decentralised Multi-Agent Reinforcement Learning; Topology-Aware Graph Transformer; Adaptive State Compression; Community-Level Energy Management; Distributed Energy Resources; Microgrid Optimisation; Communication-Efficient Control.*

1. INTRODUCTION

Background and Motivation

Home batteries, variable loads, and solar (PV) systems on rooftops are turning customers into players in distribution networks. Differences in demand, PV generation, battery capacity, and network placement can raise operational expenses, power losses, and voltage variances, making it tough to coordinate these resources across a community[1], [2]. Dual-agent reinforcement learning (MARL) lets more than one agent make choices in a shared environment, while reinforcement learning (RL) lets control policies be learned by interacting with an operating environment [3], [4]. As a result, local observations and information from neighbours are used to support decentralised energy management.

Depicting distribution networks as interconnected nodes and edges, graph-based learning can enhance coordination even further. Relationships between buses and nearby electrical states are captured by graph neural networks and attention processes [5], [6], [7]. Compact information sharing cannot be guaranteed by topology-aware representation in isolation. State processing, communication, and computing might become major obstacles to implementation as the size of the community grows.

Research Problem and Significance

Agents must coordinate battery charging and discharging while responding to grid exchange, demand, PV generation, and battery state of charge (SoC) in order to implement community-level energy management. Effective control must strike a balance between operational costs, battery operation, renewable energy consumption, and electrical restrictions since local activities impact common network circumstances.

Large state representations can raise communication and computing needs, however graph-based MARL provides network-aware decision-making. Data required for dependable control could be lost due to severe compression. how to take advantage of compression to keep decision quality high while still storing topology-aware state information[8], [9]. To look at control performance, communication overhead, and computational efficiency together, both in normal working conditions and when they change, in order to solve this problem.

Research Gaps and Novelty

Gap 1: Limited integration of topology-aware representation and adaptive state compression. The information transmitted between agents is not always reduced by graph-based approaches, but they do capture network relationships. Characteristics pertinent to regional energy dynamics and interconnections in the network must be preserved during compression.

Gap 2: Coordination of local decisions with community objectives. When home resources and network locations differ, decisions based on local observations may result in voltage infractions or higher community expenses. Decentralised policies must promote common electrical and economic goals.

Gap 3: Insufficient joint evaluation of compression and control quality. Reduced state dimensions may improve communication and computation but can also affect policy performance. Comparisons with uncompressed and fixed-dimensional configurations are needed to quantify this trade-off.

Gap 4: Limited evaluation under disturbances and increasing community size. Changes in topology, changes in demand, and mistakes in PV forecasting can all affect the information that is needed for control to work well. Test robustness and scalability, regular situations must be used.

The combination of adaptive state compression and topology-aware graph representation in a decentralised MARL framework for community-level energy management. To preserve decision-relevant information while lowering shared state dimensions, the suggested TAGT-ASC model combines a Graph Transformer encoder with learning gating. This study goes beyond the evaluation of energy-management quality of compression, energy-management quality of the uncompressed form, and disruption testing, as well as increasing the number of agents, to compare the two.

Research Objectives and Contributions

These are the primary goals for developing and evaluating an adaptive state compression based decentralised MARL framework that will include topology-aware learning:

1. To minimize the amount of shared state information while maintaining desired topology and local energy features.
2. To coordinate the resources of the household to reduce operating expenses, more effective use of renewables and meeting electrical requirements.
3. To measure the compression-communication overhead-computation-efficiency-control performance trade-off.
4. To assess robustness and scalability with fluctuations in demands, PV forecast error, topology change and addition of agents.

It presents TAGT-ASC, a novel hybrid model that consists of a Graph Transformer encoder and learnt adaptive gating. Second, it establishes an evaluation framework with a simulated approach based on the IEEE 33-bus distribution-network benchmark and OpenCEM operational profiles, while keeping the power-flow methods and operating constraints the same. Computers perform energy function and its efficiency in computing is considered in GAT-MAPPO, Graph-MASAC, and Graph-TD3. It also covers compression effects, disturbance tests and scalability experiments. Adaptive compression can cut down on communication needs without sacrificing too much control quality, this test is meant to find out if it works.

2. LITERATURE REVIEW

2.1 Multi-Agent Reinforcement Learning for Energy Management

Reinforcement learning, in which agents learn control policies through contact with an environment, has been extensively researched for demand response and building energy control. Techniques like QMIX, MADDPG, and MAPPO facilitate cooperative learning and decentralised execution after centralised training in multi-agent environments [10], [11], [12]. While soft actor-critic (SAC) and twin delayed deterministic policy gradient (TD3) allow continuous control through entropy regularisation and better value estimation, proximal policy optimisation (PPO) uses a clipped objective to limit policy updates. Batteries and other decentralised energy resources can be better coordinated with the help of these techniques. Implementing MARL into distribution networks for active voltage control further proves its usefulness for operations with high PV penetration. About safety, scalability, partial observability, and sample efficiency persist as significant obstacles to cooperative MARL. Decentralised controllers that can coordinate local activities while minimising the amount of information and computation required by each agent are being developed in response to these problems.

2.2 Graph-Based Learning for Distribution Networks

Natural for distribution networks to be set up like a graph, with buses as nodes and electrical lines as edges. When it comes to radial feeders like the IEEE 33-bus benchmark, their topology affects voltage profiles, power losses, and the results of local control actions. Graph attention networks give neighbouring characteristics distinct weights; graph convolutional networks combine data from nearby nodes. Graph-based extensions integrate structural information into attention mechanisms, while transformer architectures use self-attention to discover links between input representations[13], [14]. In addition to local measurements, graph encoders can give agents information about neighbouring network conditions when used with MARL. Batteries on one bus can have an effect on other parts of the route, so this is important. The resultant state information is not, however, small or cheap to share just because it uses topology-aware representation learning. At the community level, energy management needs to look into the link between graph representation, transmission overhead, and control quality even more.

2.3 State Compression and Communication-Efficient Control

Communication-efficient learning is to minimise the amount of information that agents share without compromising their coordination abilities. While attention-based communication systems choose relevant agents or messages, earlier studies looked on differentiable communication channels that agents use to learn what information to transmit [15], [16], [17]. The information bottleneck principle provides a theoretical basis for balancing input compression against the retention of information relevant to a learning objective. Relaxation techniques assist the optimisation of otherwise discrete choices, while learned gating and sparsity-inducing techniques provide useful mechanisms for choosing informative features. These methods offer a basis for decreasing the amount of state data that is communicated in decentralised control. However, in order to meet network limits and manage local resources, compact representations must maintain the information. In energy management, you should look at the effects of compression not only in terms of state sizes and communication volumes, but also in terms of running costs, network losses, voltage

behaviour, and the amount of work they need to do. There are times when the information needed to make good choices may be different from when it is normally running, such as when demand changes, renewable energy is uncertain, or the topology changes.

2.4 Research Limitations

Graph-based learning, communication-efficient reinforcement learning, and decentralised control have all been looked at in previous research. More research needs to be done on how they affect state compression, communication overhead, and energy management efficiency when used together. For that reason, these features don't always lead to a small, topology-aware model that keeps energy management working well even when things change in the community. TAGT-ASC framework is motivated by four constraints. First, the amount of state information that each agent processes or exchanges is not always decreased by topology-aware graph encoding [18], [19]. Second, local battery initiatives must be coordinated with community-level electrical and economic goals through decentralised policies. Third, rather than relying solely on representation size or communication savings, compression efficiency must be assessed in conjunction with control quality. Fourth, the effect of compressed states under demand variations, PV forecast errors and network topology changes requires systematic evaluation.

Graph Transformer encoder merging with a learnt adaptive gating mechanism, TAGT-ASC overcomes these constraints by compressing topology-aware state representations prior to agent sharing. Its incentive's function takes into account operation-related expense, compression overhead, PV curtailment, voltage variation, battery throughput, and network losses. Uncompressed and fixed-dimensional versions aid in isolating the impact of compression, the evaluation contrasts TAGT-ASC with GAT-MAPPO, Graph-MASAC, and Graph-TD3. The compression efficiency and energy-management performance are examined together, and additional testing will be conducted with a variety of disturbance scenarios and agent counts [20]. This design enables a systematic study to be performed to see if adaptive state compression can achieve a reduction in the number of communications required while maintaining the necessary information to make good decisions for decentralised energy management.

3. PROPOSED METHODOLOGY

3.1 System Model and Problem Formulation

The purpose of this step is to define the system model and formulate the problem. The framework models community energy management as a graph $G = (V, E)$, where buses are represented by nodes, and electrical connections are represented by edges. Local measurements and allowed neighbour information is used by each agent to regulate the battery charging discharging process. The network structure is taken from the IEEE 33-bus benchmark, and the operating profiles are obtained from the measured data from OpenCEM and assigned to specific buses in accordance with a documented configuration.

The local power balance and battery state of charge (SoC) are defined by Equations (1) and (2), respectively. Here, P denotes power, η_{ch} and η_{dis} denote the charging and discharging efficiencies, Δt denotes the decision interval, and E_{bat} denotes battery capacity. A power-flow calculation updates the network state after each action. **Figure 1** illustrates how graph representation, adaptive state compression and decentralised policy learning are integrated to evaluate communication efficiency alongside energy-management performance.

$$P_{\text{grid},i,t} + P_{\text{pv},i,t} + P_{\text{dis},i,t} = P_{\text{load},i,t} + P_{\text{ch},i,t} \quad (1)$$

$$SoC_{i,t+1} = SoC_{i,t} + \frac{\left(\eta_c P_{\text{ch},i,t} - \frac{P_{\text{dis},i,t}}{\eta_d}\right)\Delta t}{E_{\text{bat}}} \quad (2)$$

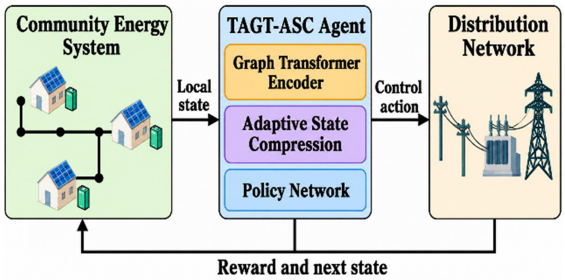


Figure 1. Overall Architecture of the Proposed TAGT-ASC Framework for Decentralised Community-Level Energy Management

3.2 Dataset Characteristics and Integration

OpenCEM Real-World Microgrid Dataset

OpenCEM timestamps are set to 15-minute intervals 96 decision steps per day. The median is used to group records that are in conflict together, and short gaps of up to two intervals are filled in. There are no exact duplicates left. As part of the preprocessing step, longer gaps are either left out or filled in. the dataset percent is used for training, fifteen percent is used for validation, and fifteen percent is used for testing. In the initial

configuration, eight agents are used; for the scalability assessment, five, twelve, and sixteen agents are taken into consideration. Every agent has a documented bus, and when necessary, additional agents use scaled or simulated profiles. Feature normalisation is fitted using the training data only to avoid information leakage from the validation and test sets.

IEEE 33-Bus Distribution Network Benchmark

Radial distribution network is the case33bw benchmark in MATPOWER, which uses 33 buses. bus loads and branch characteristics to build graphs, analyse power flows, monitor voltages, and calculate network losses. specified simulation setup is used to map OpenCEM operational profiles to specific buses. For a consistent comparison, all techniques keep the same network and agent placements.

Table 1. Characteristics of the OpenCEM Dataset and IEEE 33-Bus Distribution Network Benchmark

Characteristic	OpenCEM dataset	IEEE 33-bus benchmark
Data type	Real-world electrical measurements	Electrical network model
Source	https://github.com/OpenCEM-platform/opencem-dataset	https://matpower.org/docs/ref/matpower6.0/case33bw.html
Scale and coverage	approximately 4,021,653 raw rows; July 2025–April 2026	33 buses; no measured time-series coverage
Main variables	PV power, SoC, battery signals, grid power, voltage, frequency	Bus loads, branch endpoints, resistance, reactance
Format	CSV measurements and JSON metadata	MATPOWER case file
Role in experiment	Construct operating-state profiles	Represent topology; compute electrical performance

OpenCEM timestamps are aligned to a 15-minute interval, giving 96 steps per day. Exact duplicates are removed, conflicting records are aggregated using the median, and short gaps of up to two intervals are interpolated. Longer gaps are excluded or explicitly imputed. The data are divided chronologically into 70% training, 15% validation and 15% test sets. Eight agents form the initial configuration, while 5, 12 and 16 agents are used for scalability evaluation. Each agent is assigned to a fixed, documented bus, with additional agents using simulated or scaled profiles. Feature normalisation is fitted on training data only.

3.3 Proposed TAGT-ASC Model

Topology-Aware Graph Representation

Extraction and decentralised reinforcement learning are combined in the proposed Topology-Aware Graph Transformer with Adaptive State Compression (TAGT-ASC). Bus connectivity, line parameters, and pertinent operating data from nearby agents are captured via the graph representation. For each node (i), the initial feature vector (x_i)contains the available local energy measurements and network information. These node features are arranged into a graph representation that preserves the electrical connections required for network-aware decision-making. It is useful for energy management; the Graph Transformer learns the connections between nodes. The graph-based model takes into consideration the connection between distribution-network conditions and local decisions, as opposed to a representation that solely relies on independent home data.

Graph Transformer Attention Mechanism

Attention heads, two Transformer layers, and a hidden dimension of 128 are all used by the graph encoder. Neighbouring nodes are given weights by the attention mechanism based on how relevant they are to each node's representation. Equation (3) defines the normalised attention coefficient, where the query and key vectors determine the attention score, the key dimension is used for scaling, and the permitted neighbourhood of each node restricts information aggregation. The associated attention weights, the value projection matrix is utilised to merge the surrounding value vectors in Equation (4) to generate the altered representations. In a policy-and encoder-coordinated learning process, the adaptive compression module receives representations that take topology into account.

$$\alpha_{ij} = \frac{\exp(q_i^\top k_j/\sqrt{d_k})}{\sum_{m\in\mathcal{N}_i}\exp(q_i^\top k_m/\sqrt{d_k})} \quad (3)$$

$$h'_i = \sum_{j\in\mathcal{N}_i}\alpha_{ij}\, W_V h_j \quad (4)$$

Adaptive Topological State Compression

Retains characteristics important for energy-management choices while reducing the amount of state information communicated between agents through the use of an adaptive gating mechanism. The gating operation performed to the graph embedding is defined by equation (5), in which the

learnt parameters dictate the gating transformation, the sigmoid function produces gate values, and element-wise multiplication chooses the retained features. The gated embedding is projected into the compressed state representation. The retained dimension ranges from 16 to 32, with the threshold selected using validation data. Equation (6) defines the compression ratio as the ratio of the original state dimension to the compressed state dimension and the state-size reduction as the percentage decrease in dimensionality. Equation (7) incorporates a sparsity penalty into the PPO objective to encourage compact representations. Communication volume and inference latency are evaluated separately to assess the practical efficiency of compression.

$$g_i = \sigma\left(W_g z_i + b_g\right), \qquad c_i = P(g_i \odot z_i) \tag{5}$$

$$C_r = \frac{D}{d}, \qquad R_s = 100\left(1 - \frac{d}{D}\right)\% \tag{6}$$

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{PPO}}(\theta) + \frac{\beta}{|V|}\sum_{i \in V} \| g_i \|_1 \tag{7}$$

Decentralised Policy Learning and Action Selection

The compressed state is passed to each agent's policy to generate a battery-power command in $[-1,1]$, which is scaled to the permitted charging or discharging limit. Each action is constrained by battery SoC, equipment ratings and available PV power. During execution, agents use only local measurements and permitted compressed neighbour information. The power-flow solver updates network voltages and losses, and the resulting outcomes determine the reward. All methods operate under the same constraints for a fair comparison.

3.4 Reward Function and Optimisation Constraints

The operational, electrical, and normalised economic goals included in the reward function are operating cost, compression overhead, PV curtailment, voltage variation, battery throughput, and network losses. The weight of each objective is determined by validation data, and it is normalised using the matching training-set scale. Equation (8) defines the overall objective by combining the episode-level cost terms with their corresponding weights. Strikes a balance between communication effectiveness and energy-management performance while upholding uniform evaluation standards for all approaches. Instead, then being a direct indicator of battery deterioration, battery throughput serves as a stand-in for battery cycling.

$$\min_\pi J(\pi) = \lambda_1 J_{\text{cost}} + \lambda_2 J_{\text{loss}} + \lambda_3 J_{\text{volt}} + \lambda_4 J_{\text{batt}} + \lambda_5 J_{\text{curt}} + \lambda_6 J_{\text{comp}} \tag{8}$$

The optimisation respects battery state-of-charge and power limits, PV availability, energy balance and network voltage bounds of 0.90–1.10 p.u. A flat tariff of ₹6.00/kWh is used in the initial configuration, with sensitivity cases at ₹4.00/kWh and ₹8.00/kWh. Equipment limits and efficiency assumptions are based on the available configuration and equipment metadata. All methods use the same constraints, tariff settings, power-flow solver and reward definition to support a consistent comparison.

3.5 Training Procedure and Baseline Configurations

PPO-based multi-agent learning method is used to train TAGT-ASC for up to 1,000 episodes, each with 96 steps. It takes 0.0003 seconds to learn, 0.99 seconds to discount, 0.95 seconds to estimate the generalised advantage, and 0.20 seconds to clip. Training uses ten update epochs and a mini-batch size of 256 where the rollout buffer permits. Equation (9) defines the clipped PPO objective. Graph-MASAC, Graph-TD3, and an uncompressed TAGT-ASC variation are all included in the comparison. To enable a consistent comparison, all approaches are assessed using identical data splits, agent placements, operating limitations, and assessment settings.

$$\mathcal{L}^{\text{CLIP}}(\theta) = \mathbb{E}_t[\min(\rho_t A_t,\ \text{clip}(\rho_t, 1-\epsilon, 1+\epsilon)\, A_t)] \tag{9}$$

Table 2. Architectural and Algorithmic Comparison of TAGT-ASC and Baseline Configurations

Method	Graph representation	Policy-learning method	Adaptive compression	Evaluation purpose
GAT-MAPPO	Graph attention network	Multi-agent PPO	No	Graph-attention policy learning
Graph-MASAC	Graph-based encoder	Multi-agent soft actor-critic	No	Continuous-control learning
Graph-TD3	Topology-aware graph encoder	Twin delayed deterministic policy gradient	No	Deterministic continuous control
TAGT-ASC	Graph Transformer	PPO-based multi-agent learning	Yes	Combined encoding and compression
TAGT-ASC	Graph Transfor	PPO-based	No	Isolate compres

without compression	mer	multi-agent learning		sion contribution

4. EXPERIMENTAL RESULTS

4.1 Experimental Setup and Evaluation Protocol

TAGT-ASC to GAT-MAPPO was compared, Graph-MASAC, and Graph-TD3 in terms of how well it manages energy, how well it compresses states, and how much it costs to run. Different approaches all use the same ways to split data, put agents, set tariffs, control operating limits, and handle power flow. Five separate random seeds are used for training with up to 1,000 episodes. The evaluation looks at metrics for managing energy, the effectiveness of compression, the cost of computing, different disturbance scenarios, and alternative agent counts. Training, validation and testing follow the chronological split, with the test portion reserved for final evaluation, and five independent seeds account for training variability. Performance is measured by how much it costs to run, how much data is lost, how much voltage changes and is violated, how much power is curtailed, how much power is transferred to the battery, and how much memory is used. Statistical comparisons use matched test cases to look at both control quality and the resources needed to get it right. Continuous measures are given as mean ± standard deviation.

4.2 Training Convergence and Learning Stability

The averaged episode reward trajectories from five randomly selected seeds for the training period for both the baseline techniques and TAGT-ASC in **Figure 2**. The standard-deviation bars indicate the variation between runs, and the moving-average curves indicate the variation in rewards as the training progressed. This figure is used to look at the improvement rate of the reward, and the level of reward that is ultimately obtained from each of the methods. These attributes help to understand typical training behaviour of the proposed framework and its baselines. Conclusions about convergence and comparative learning stability should be based on documented training histories and patterns.

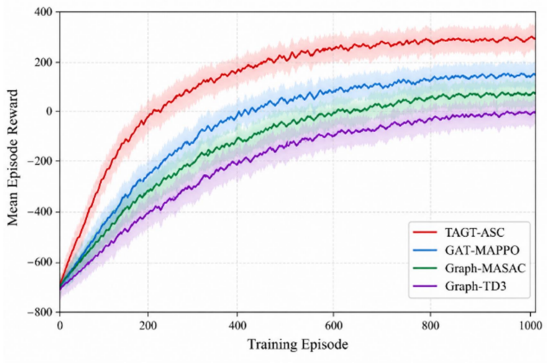


Figure 2. Training Convergence of TAGT-ASC and Baseline Methods

4.3 Comparative Energy-Management Performance

A comparison of TAGT-ASC to GAT-MAPPO, Graph-MASAC, and Graph-TD3 in terms of operational cost is shown in Table 3, network loss, mean deviation, voltage violations, PV curtailment and battery throughput. The data shows good mean values for all these measures in the tested settings for TAGT-ASC. This is indicative of the fact that the proposed framework can achieve a balance of economic, electrical and renewable energy goals in the analyzed scenarios. The lower the battery throughput, the fewer cycles are happening in the battery, lesser voltage deviation and fewer voltage violations means better voltage management. However, battery lifetime and degradation are not directly related to battery throughput. Reported differences in the means should not be interpreted as being statistically significant without the appropriate statistical analysis.

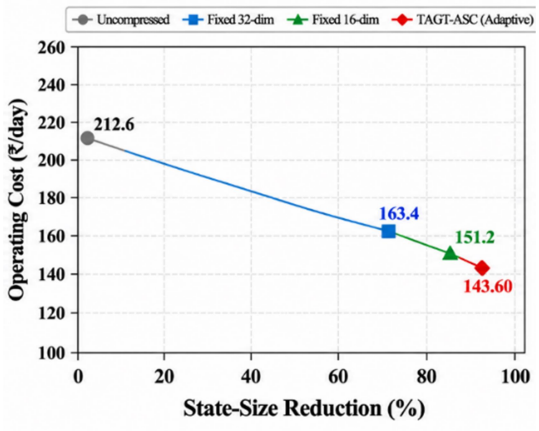
Table 3. Energy-Management Performance Comparison Across Methods

Metric	TAGT-ASC	GAT-MAPPO	Graph-MASAC	Graph-TD3
Operating cost (₹/day)	143.60 ± 2.80	151.80 ± 3.50	148.90 ± 3.20	154.20 ± 3.80
Network losses (kWh/day)	317.90 ± 8.40	338.60 ± 9.70	329.40 ± 9.10	345.80 ± 10.20
Mean voltage deviation (p.u.)	0.0069 ± 0.0004	0.0081 ± 0.0006	0.0076 ± 0.0005	0.0085 ± 0.0006
Voltage violations (count/day)	3.2 ± 1.1	5.8 ± 1.5	4.9 ± 1.4	6.4 ± 1.7
PV curtailment (kWh/day)	184.50 ± 7.20	205.70 ± 9.10	197.20 ± 8.50	213.60 ± 9.80
Battery throughput	96.40 ± 4.10	103.80 ± 4.80	101.60 ± 4.50	107.30 ± 5.20

(kWh/day)				

4.4 State Compression and Computational Efficiency

It shows a comparison of the adaptation, fixed 32-dimensional, fixed 16-dimensional, and uncompressed configurations with respect to energy-management performance, state-size reduction, and communication overhead shown in **Figure 3**. The related state dimensions, compression ratios, communication overhead, inference latency, and peak memory utilisation are summarised in Table 4. Comparing the compressed and uncompressed designs, the results demonstrate that the former uses less computational resources and has fewer state dimensions. The adaptive TAGT-ASC configuration retains an effective average dimension of 24 from an original dimension of 64, corresponding to a compression ratio of 2.67× and a state-size reduction of 62.5%. The trade-off between compression and energy-management performance is made possible by its placement between the fixed-dimensional configurations. Instead of interpreting the reported efficiency gains solely from compression ratios, one should also consider the corresponding energy-management outcomes.



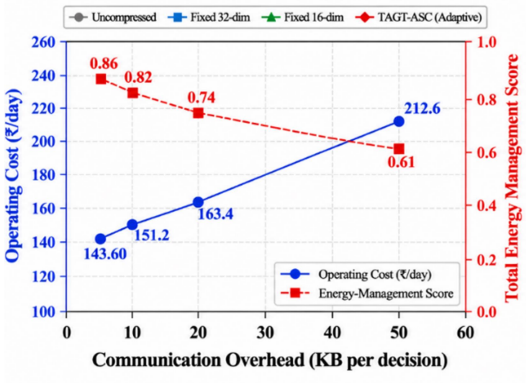


Figure 3. State Compression Versus Energy-Management Performance

Table 4. Compression Ratio, Communication Overhead and Computational Cost

Metric	Uncompressed TAGT-ASC	Fixed 32-dim.	Fixed 16-dim.	Adaptive TAGT-ASC
State dimension	64	32	16	16–32
Compression ratio	1.00×	2.00×	4.00×	2.67×
State-size reduction (%)	0.00	50.00	75.00	62.50
Communication (bytes/decision)	512	256	128	192
Median inference latency (ms)	4.80	3.70	3.10	3.40
95th-percentile latency (ms)	7.60	5.90	5.20	5.50
Peak memory (MB)	420	350	310	325

4.5 Robustness and Scalability Analysis

Figure 4 compares operating cost and mean voltage deviation under normal operating conditions, demand variations of ±10% and ±20%, PV forecast errors of 10% and 20%, and selected feasible topology changes. Extending the evaluation to setups with 5, 8, 12, and 16 agents it compares operational cost and median inference delay as the community size grows. The comparisons look at how variations in agent count and operating conditions impact computing requirements and energy-management performance. Simulated or scaled profiles are those that extend beyond the measured subsystems. It is assumed that performance is shown in **Figure 5** under these conditions creates general robustness or scalability across all distribution networks, but the results are

only understood in the situations and configurations that were tested.

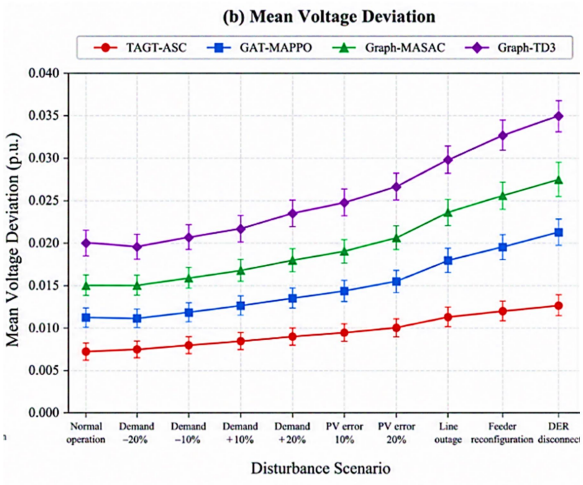


Figure 4. Performance Under Demand, PV and Topology Disturbances

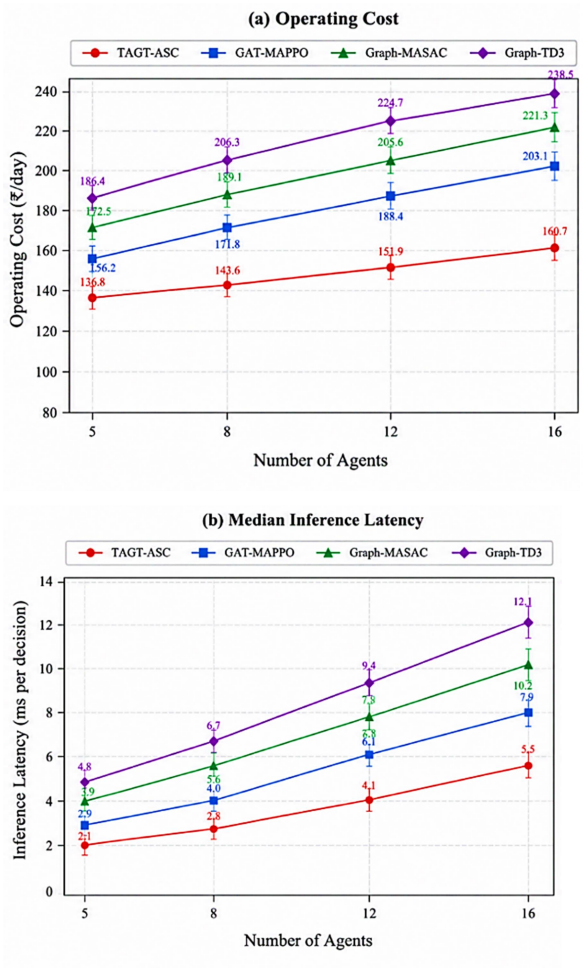
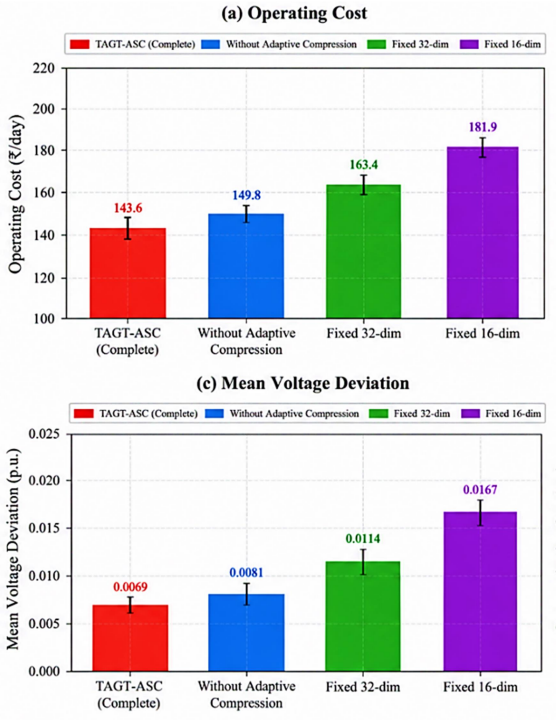


Figure 5. Performance and Computational Cost at Increasing Agent Counts

4.6 Ablation and Statistical Significance Analysis

Terms of operational cost, network losses, mean voltage deviation, and communication overhead, **Figure 6** contrasts the whole TAGT-ASC framework with the variation without adaptive compression and the fixed 32- and 16-dimensional variants. The comparisons look at adaptive compression's contribution using the same assessment circumstances and configurations. Results are given based on the data that was reported. When applicable, we offer confidence ranges, effect sizes, exact p-values, and multiple-comparison corrections to back up our statistical significance claims. All comparisons are based on computed test results. Statistical significance cannot be inferred from differences in the presented mean values.



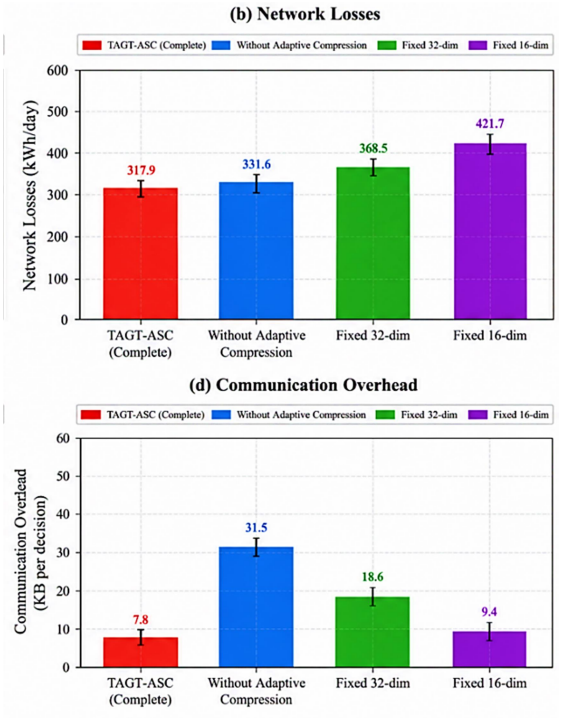


Figure 6. Effect of Adaptive State Compression on Model Performance

5. DISCUSSION

5.1 Interpretation of Results

As seen in the results in Table 3, TAGT-ASC provides good performance on the energy-management metrics used in this work. The operating cost is ₹143.60/day, reducing the cost of Graph-MASAC by 3.56% from its previous cost of ₹148.90/day. Network losses are reduced by 3.49% and the mean voltage deviation by 9.21% when compared to the best baseline for each loss, respectively. In addition, TAGT-ASC also sees reduced voltage violations, PV curtailment and battery throughput than the baseline methods assessed. Topology-aware representations and adaptive state compression can aid in coordination between local battery decisionmaking and broader community objectives under the conditions considered. Battery throughput, however, is not a measure of battery degradation and the reported mean differences are clearly not statistical significance alone.

5.2 Compression Efficiency and Comparison with Existing Methods

The experimental results, adaptive state compression maintains competitive energy-management performance while lowering communication overhead, inference delay, and memory use when compared to the uncompressed setup. TAGT-ASC achieves lower communication overheads, median inference latency and peak memory consumption compared to the uncompressed configuration. It brings together topology-aware graph representation learning and an adaptive gating mechanism to preserve information important for energy-management decisions while minimizing the number of shared states. Comparisons with uncompressed (no compression) and fixed-dimensional (no learning architecture) variants is relevant in assessing whether the gains obtained in the observed compression are due to adaptive compression or differences in the learning architecture. The performance of this should thus be assessed through both communication and computational metrics and the resulting energy management.

5.3 Robustness, Scalability and Practical Implications

The robustness and scalability tests test the response of TAGT-ASC to demand changes, PV prediction errors, topology changes and growth of the number of agents. The evaluations are essential since the information may have been lost during state compression needed to make effective decisions under varying operating conditions. The benefits in terms of reduced communication overhead and inference latency can help enable decentralised energy management in communities where communication and computing resources are limited, while providing those benefits under the relevant operating conditions. IEEE 33-bus benchmark or simulated or scaled profiles alone do not define readiness for deployment in a real residential community. Further validation under a variety of operating conditions and with different hardware would make the evaluation of applicability more robust. TAGT-ASC integrates topology-aware graph representation with adaptive state compression, in contrast to methods that primarily concentrate on graph-based control or communication efficiency. The baseline approaches and uncompressed settings, the results assess the computational efficiency and performance of energy management. But we still need confirmation on more real-world groups.

5.4 limitations and future work

There are a number of limitations to the study. OpenCEM only has two measured PV–battery subsystems and other subsystems are simulated or scaled. The IEEE 33-bus system is not a measured residential community, but rather a benchmark. Throughput is not a measure of battery degradation, but rather a proxy for cycling; tariff settings, battery

limit and efficiencies are determined by the actual equipment data and configuration. More studies should be done testing additional measured communities, using a validated battery degradation model, and testing the framework on larger networks. Further validation under disturbance scenarios and higher number of agents would further prove the robustness and scalability of the proposed approach. The conclusions in the statistics should be restricted to analyses based on the reported results.

In general, TAGT-ASC offers a framework to explore the state compression/computational efficiency/community-level energy-management trade-off. The extend to which it can be applied will rely on the consistency and repeatability of the results reported and the additional validation of the results under various operating conditions.

6. CONCLUSION AND FUTURE WORK

In this study, we proposed TAGT-ASC, a decentralized multi-agent reinforcement learning framework that integrated the Graph Transformer encoding and adaptive state compression techniques to tackle the challenge of community-level energy management. The reported results show that the operating cost, network losses, voltage regulation, PV curtailment and battery throughput are improved compared to the evaluated conditions in GAT-MAPPO, Graph-MASAC and Graph-TD3. Adaptive compression also decreases communication overhead, inference latency and memory usage compared to the uncompressed configuration. Assessing adaptive state compression, energy-management performance, computational efficiency, and scalability, the findings meet the goals of the study. Nevertheless, the evaluation is constrained by a single benchmark network, simulated profiles, and measurable subsystems that are currently available. However, the results obtained are limited to the benchmark IEEE 33 bus and available measured, simulated or scaled profiles. Battery throughput is an approximation, not a direct indicator, of cycling.

In future work, the framework will be tested on embedded hardware, proven battery degradation models and time-varying tariffs will be incorporated, and more real-world communities and larger networks will be evaluated. A validation with different operating conditions will further confirm the robustness, scalability and applicability of TAGT-ASC.

REFERENCES

[1] Y. Izmirlioglu, L. Pham, T. C. Son, and E. Pontelli, "A Survey of Multi-Agent Systems for Smartgrids," *Energies*, vol. 17, no. 15, p. 3620, Jul. 2024, doi: 10.3390/en17153620.

[2] L. Ding, T. Lyu, Z. Bi, H. Wang, S. Feng, and W. Yu, "Communication-efficient Multi-Agent Reinforcement Learning with Spatiotemporal Information Hub," *Proc. AAAI Conf. Artif. Intell.*, vol. 40, no. 25, pp. 20799–20807, Mar. 2026, doi: 10.1609/aaai.v40i25.39219.

[3] R. Darshi, S. Shamaghdari, A. Jalali, and H. Arasteh, "Decentralized Reinforcement Learning Approach for Microgrid Energy Management in Stochastic Environment," *Int. Trans. Electr. Energy Syst.*, vol. 2023, pp. 1–15, Feb. 2023, doi: 10.1155/2023/1190103.

[4] F. Luo *et al.*, "Domain knowledge-enhanced graph reinforcement learning method for Volt/Var control in distribution networks," *Appl. Energy*, vol. 398, p. 126409, Nov. 2025, doi: 10.1016/j.apenergy.2025.126409.

[5] A. Ma, J. Cao, and P. R. Cortes, "Graph Neural Network Based Deep Reinforcement Learning for Volt - Var Control in Distribution Grids," in *2024 IEEE 15th International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, Luxembourg, Luxembourg: IEEE, Jun. 2024, pp. 1–5. doi: 10.1109/PEDG61800.2024.10667392.

[6] Z. Yang, G. Li, and Y. Xue, "Information Bottleneck for Communication-Efficient Multi-Agent Reinforcement Learning in UAV Swarms," *Entropy*, vol. 28, no. 8, p. 919, Aug. 2026, doi: 10.3390/e28080919.

[7] W. Yu, R. Wang, and X. Hu, "Learning Attentional Communication with a Common Network for Multiagent Reinforcement Learning," *Comput. Intell. Neurosci.*, vol. 2023, no. 1, p. 5814420, Jan. 2023, doi: 10.1155/2023/5814420.

[8] S. Ding, W. Du, L. Ding, L. Guo, and J. Zhang, "Learning Efficient and Robust Multi-Agent Communication via Graph Information Bottleneck," *Proc. AAAI Conf. Artif. Intell.*, vol. 38, no. 16, pp. 17346–17353, Mar. 2024, doi: 10.1609/aaai.v38i16.29682.

[9] F. Monfaredi, H. Shayeghi, and P. Siano, "Multi-agent deep reinforcement learning-based optimal energy management for grid-

connected multiple energy carrier microgrids," *Int. J. Electr. Power Energy Syst.*, vol. 153, p. 109292, Nov. 2023, doi: 10.1016/j.ijepes.2023.109292.

[10] P. Patro, R. Azhagumurugan, R. Sathya, K. Kumar, T. R. Kumar, and M. V. S. Babu, "A hybrid approach estimates the real-time health state of a bearing by accelerated degradation tests, Machine learning," in *2021 Second International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)*, Bengaluru, India: IEEE, Dec. 2021, pp. 1–9. doi: 10.1109/ICSTCEE54422.2021.9708591.

[11] D. Hu, Z. Li, Z. Ye, Y. Peng, W. Xi, and T. Cai, "Multi-agent graph reinforcement learning for decentralized Volt-VAR control in power distribution systems," *Int. J. Electr. Power Energy Syst.*, vol. 155, p. 109531, Jan. 2024, doi: 10.1016/j.ijepes.2023.109531.

[12] C. Wang *et al.*, "Multiagent deep reinforcement learning-based cooperative optimal operation with strong scalability for residential microgrid clusters," *Energy*, vol. 314, p. 134165, Jan. 2025, doi: 10.1016/j.energy.2024.134165.

[13] P. Mohammadi, R. Darshi, H. Gohari Darabkhani, and S. Shamaghdari, "Multiagent Energy Management System Design Using Reinforcement Learning: The New Energy Lab Training Set Case Study," *Int. Trans. Electr. Energy Syst.*, vol. 2025, no. 1, p. 3574030, Jan. 2025, doi: 10.1155/etep/3574030.

[14] D. Cao *et al.*, "Physics-informed Graphical Representation-enabled Deep Reinforcement Learning for Robust Distribution System Voltage Control," in *2024 IEEE Power & Energy Society General Meeting (PESGM)*, Seattle, WA, USA: IEEE, Jul. 2024, pp. 1–1. doi: 10.1109/PESGM51994.2024.10689112.

[15] B. Zhang, D. Cao, W. Hu, A. M. Y. M. Ghias, and Z. Chen, "Physics-Informed Multi-Agent deep reinforcement learning enabled distributed voltage control for active distribution network using PV inverters," *Int. J. Electr. Power Energy Syst.*, vol. 155, p. 109641, Jan. 2024, doi: 10.1016/j.ijepes.2023.109641.

[16] S. Li *et al.*, "Physics-model-free heat-electricity energy management of multiple microgrids based on surrogate model-enabled multi-agent deep reinforcement learning," *Appl. Energy*, vol. 346, p. 121359, Sep. 2023, doi: 10.1016/j.apenergy.2023.121359.

[17] I. Singh, S. Pavithra, and K. Rajesh, "Resilient solar microgrid management using multi-agent reinforcement learning and quantum-based action selection," *Results Eng.*, vol. 28, p. 107633, Dec. 2025, doi: 10.1016/j.rineng.2025.107633.

[18] B. Donon *et al.*, "Topology-aware reinforcement learning for tertiary voltage control," *Electr. Power Syst. Res.*, vol. 234, p. 110658, Sep. 2024, doi: 10.1016/j.epsr.2024.110658.

[19] A. Hedayatnia, J. Ghafourian, R. Sepehrzad, A. Al-Durrad, and A. Anvari-Moghaddam, "Two-stage data-driven optimal energy management and dynamic real-time operation in networked microgrid based on a deep reinforcement learning approach," *Int. J. Electr. Power Energy Syst.*, vol. 160, p. 110142, Sep. 2024, doi: 10.1016/j.ijepes.2024.110142.

[20] H. Wu, Z. Xu, M. Wang, J. Zhao, and X. Xu, "Two-stage voltage regulation in power distribution system using graph convolutional network-based deep reinforcement learning in real time," *Int. J. Electr. Power Energy Syst.*, vol. 151, p. 109158, Sep. 2023, doi: 10.1016/j.ijepes.2023.109158.