International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(13s) 2026 pp. 1-12
https://doi.org/10.70917/ijcisim-2026-4029
Article

Variational Autoencoders Integrated with Reinforcement Learning for Autonomous IoT Resource Allocation and Load Balancing Optimization

Annie Silviya S. H.[1], Paul T. Jaba[2], S. Julia Faith[3], Kannan A[4], Ramya M[5], Sriman B[6]

[1]Department of Computer Science and Engineering, Rajalakshmi Institute of Technology.
Email: anniesilviya.sh@ritchennai.edu.in
[2]Department of Computer Science and Engineering, St. Joseph College of Engineering.
Email: jabapt2@gmail.com
[3]Department of Information Technology, S.A. Engineering College.
Email: juliafaiths@saec.ac.in
[4]Department of Mathematics, Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College.
Email: kansyl1@gmail.com
[5]Department of Information Technology, Panimalar Engineering College.
Email: ramyaanitha94@gmail.com
[6] Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology.
Email: srimanb@veltech.edu.in

Abstract: The current paper introduced a new framework combining Variational Autoencoders (VAE) with Deep Reinforcement Learning (DRL) in autonomous maximum utility resource allocation and load balancing in Internet of Things (IoT) setting. The proposed VAE-RL model overcame the difficulties of dynamic network resource management of heterogeneous IoT networks by learning latent representations of network conditions and using an actor critic architecture in optimal decisions making. We deployed and tested the framework in three different scenarios of IoT: infrastructure in a smart city, manufacturing with industrial IoTs, and edge computer setup. The results of experimental work proved our methodology showed a 34.7 percent reduction in response time, 28.3 percent better resource consumption and 41.2 percent cut in the load imbalance than the baseline methods. The framework demonstrated suitable performance in changing network conditions and scaled well to grassing that comprised of 5000 IoT devices.

Keywords: Variational Autoencoders, Reinforcement Learning, IoT Resource Allocation, Load Balancing, Edge Computing, Autonomous Systems

1. Introduction

Several challenges in the allocation and balancing of resources and loads have never been as great as they now stand with the spread of Internet of Things (IoT) devices, and forecasts are taking the number of connected devices beyond 75 billion in 2025. Conventional, rule-based and heuristic decision-making methods could not be sufficiently applicable to the dynamic and heterogeneous environments of current IoT ecosystems which are already resource-constrained. This was complicated by numerous factors: changing network conditions, unbalanced capabilities of the devices, uncertain traffic patterns, and high-quality standards of the Quality of Service (QoS). The latest developments in deep learning and reinforcement learning offered the potential solutions to autonomous network management. The current methods were however lacking in features such as inability to work with high-dimensional state space, sample inefficiency and lack of ability to deal with uncertainty in IoT contexts. This paper has filled these shortcomings by suggesting a hybrid design that utilized the representational ability of Variational Autoencoders and decision-making opportunities of Deep Reinforcement Learning. The main contributions of our work were: (1) an innovative mechanism to weigh high-dimensional Internet of Things network states, which utilized VAEs to reduce state space and made those states learnable; (2) an adaptive approach to actor-critic that used state space and employed state space as inputs; (3) a multi-objective reward model that computed response time, energy consumption, and load balance as theoretically reflect operates and balance candidate policy; and (4) extensive experiments in various Internet of Things deployment settings where it consistently outperformed currently existing standards.

2. Literature Review

The fast development of wireless communication systems has been a factor leading to much research conducted on next-generation networking paradigm and intelligent computing frameworks. Akyildiz et al. (2020) offer an extensive vision of 6G and beyond, with such key aspects as ultra-low latency, terahertz communications, massive connectivity, and AI-controlled network optimization being the pillars. Such innovations can contribute to the new applications in the fields of holographic communication, autonomous systems, and pervasive intelligence. In the said context, the Internet of Things (IoT) has received a lot of attention because of its dependence on both effective data collection and wireless transformation. In their survey, Luong et al. (2016a, 2016b) focus on the economic and pricing models used to collect the IoT data with interest in incentive mechanisms, energy efficiency, and in optimal resource allocation. The significance of the balance between the cost of communication and the utility of the data in large-scale IoT networks is emphasized in their work. With the ubiquity of data-driven intelligence at the heart of future systems, federated learning (FL) has become a threat-free machine learning paradigm. The paper by Aledhari et al. (2020) provides an in-depth overview of the FL enabling technologies, protocols and usage, noting that it is applicable in distributed and resource-constrained networks, including IoT and edge networks. Nevertheless, the problems of associated overhead in communication, security, and heterogeneity of the system are still research problems. Blockchain technology has been suggested as a supplementary solution to ensure that distributed systems are not affected by the issue of trust and decentralization of their systems. Wang et al. (2019a–2019d) cover a wide range of consensus mechanisms and management of mining strategies of blockchain networks. Their works address trade-offs between security, scalability and energy consumption which are highly important when it comes to combining blockchain with IoT and wireless systems. Along these advances, quantum computing has been receiving an increasing amount of attention with the goal of addressing challenging optimization and simulation issues. Variational Quantum Algorithms (VQAs), especially the Variational Quantum Eigensolver (VQE) have been demonstrated to be of use in near-term quantum devices. Kandala et al. (2017) show hardware-efficient VQE instantiations, and Parrish et al. (2019) and Wang et al. (2019) suggest improved implementation (accelerated) and application-specific instantiations. The authors further discuss the methods the author of this study could use to reduce measurements to improve VQE performance (Izmaylov et al., 2019). Bitel and Kliesch (2021a2021c) affirm that VQAs training is NP-hard, however, and constitutes fundamental scalability constraints. Kuroiwa et al. (2021) consider the use of penalty-based optimization strategies to reduce constraint management problems in VQE. On the one hand, the available literature indicates convergence between 6G communications, distributed learning, blockchain security, and quantum computing, and critical issues on scalability, optimization complexity, and system integration on the other, which promoted a further interdisciplinary study. Recent advances in wireless communication, cloud computing, edge computing, blockchain, artificial intelligence (AI), and Internet of Things (IoT) technologies have significantly improved the performance, security, and resource efficiency of next-generation communication systems. Quality of Service (QoS) in wireless ad hoc networks has been enhanced through resource-commutable clustering and intelligent scheduling mechanisms that optimize bandwidth utilization, reduce latency, and improve network reliability (Shitharth et al., 2023) [17]. Deep learning-based security frameworks such as RLIS have strengthened resource-limited beyond-5G networks by improving intrusion detection, secure communication, and intelligent resource management (Selvarajan et al., 2023) [18]. Resource-efficient synthetic data generation techniques have further facilitated accurate performance evaluation in mobile edge computing over 5G networks, enabling scalable testing without compromising sensitive information (Pandey et al., 2023) [19]. Optimization techniques have also enhanced IoT circuit design for flexible optical networks by achieving high-speed utilization and efficient communication resource allocation (Pandiaraj et al., 2023) [20]. The integration of federated learning with blockchain-based computational offloading has provided secure and privacy-preserving distributed learning environments while reducing communication overhead in intelligent network systems (Shitharth et al., 2023) [21]. Novel optimization algorithms, including the Resistance–Capacitance Optimizer, have demonstrated superior performance in solving complex numerical and industrial engineering optimization problems through physics-inspired population-based search strategies

(Ravichandran et al., 2023) [22]. Smart city security has benefited from cyborg intelligence mechanisms integrating conjugate self-organizing migration and reconciliate multi-agent Markov learning, enabling adaptive threat detection and intelligent decision-making (Shitharth et al., 2023) [23]. AI-enabled IoT transportation models have improved intelligent transportation systems by optimizing traffic management, routing, and real-time monitoring (Al-Ani et al., 2023) [24], while trifold algorithm-based transportation computing frameworks have enhanced multipath communication efficiency in autonomous transportation networks (Chandrakasan et al., 2023) [25]. Edge computing integrated with incremental learning has significantly improved real-time object detection in IoT environments by reducing computational latency and enabling adaptive classification at edge devices (Shitharth et al., 2023) [26]. Blockchain-assisted privacy-enhanced communication frameworks have further secured wireless data transmission through decentralized authentication and tamper-resistant information sharing (Aluvalu et al., 2023) [27]. Lightweight AI-enabled blockchain models have strengthened security and privacy in Industrial Internet of Things (IIoT) systems by ensuring secure authentication and efficient data integrity verification (Shitharth et al., 2023) [28]. Machine learning-based security measurement schemes have also enhanced privacy protection in IoT applications through intelligent threat analysis and anomaly detection (Alhalabi et al., 2023) [29]. Energy-efficient federated cloud management has been achieved using resource handiness and exigency-based migration algorithms that optimize workload distribution and minimize energy consumption in distributed cloud environments (Karthickeyan et al., 2023) [30]. Adaptive Neuro-Fuzzy Inference Systems combined with Particle Swarm Optimization have improved security in Vehicular Ad Hoc Networks (VANETs) by enabling intelligent intrusion prevention and secure routing (Kesavan et al., 2023) [31]. Particle Swarm Optimization has also been successfully applied to IoT cybersecurity for proactive attack prevention and intelligent network defense (Alterazi et al., 2022) [32]. Decision tree-based mobile cloud computing frameworks have enhanced emergency applications by supporting rapid decision-making and efficient resource allocation during critical situations (Hai et al., 2023) [33]. Secure data transmission across multiple cloud environments has been strengthened using Blowfish encryption algorithms, ensuring confidentiality and integrity for conveyance applications (Shitharth et al., 2022) [34]. Trust-aware service management has also been improved through three-phase Service Level Agreement (SLA) models implemented by trusted cloud brokers, enabling reliable cloud service monitoring and management (Muralidharan et al., 2022) [35]. Furthermore, hybrid classification approaches combining cascading C4.5 decision trees with K-means clustering have demonstrated effective identification of normal and anomalous network activities, thereby improving cybersecurity and network intrusion detection capabilities (Mohammad et al., 2022) [36]. Collectively, these studies demonstrate that the convergence of AI, IoT, edge computing, blockchain, cloud computing, optimization algorithms, and intelligent communication technologies provides scalable, secure, energy-efficient, and high-performance solutions for next-generation wireless networks, smart transportation, cloud services, cybersecurity, and Industrial IoT applications.

3. Methodology

3.1 System Architecture

Our architecture was in a form of three layers, which included the IoT device layer, edge processing layer, and cloud coordination layer. The VAE-RL agent system on the edge processing layer received real-time telemetry of the IoT devices and made decisions regarding the allocation of resources. The Machiavelli diagram was used to model the IoT environment as a Markov Decision Process (MDP) with model (S, A, P, R, 7), where S was the state space, A was the available actions, P was the state transition probability, R was the reward function and 7 was the discount factor which was set at 0.95.

3.2 Variational Autoencoder for State Encoding

The VAE component consisted of an encoder network E_φ and decoder network D_θ that learned compressed representations of network states. Each state observation s_t ∈ R^n contained 147 features including device metrics (CPU utilization, memory usage, network bandwidth, queue lengths), application requirements (latency constraints, throughput demands), and environmental factors (device mobility, channel conditions).

The encoder mapped high-dimensional states to a latent distribution: E_φ(s_t) → (μ_t, σ_t), where μ_t and σ_t represented the mean and standard deviation of the latent Gaussian distribution. We implemented the reparameterization trick for gradient computation: z_t = μ_t + σ_t ⨀ε, where ε ~ N(0, I) and z_t ∈ R^32 represented the encoded state. The decoder reconstructed the original state: ŝ_t = D_θ(z_t).

The VAE training objective combined reconstruction loss and KL divergence:

$$L\_VAE = E[||s\_t - \hat{s}\_t||^2] + \beta \cdot KL(q(z|s) \,||\, p(z))\#(1)$$

We set $\beta = 0.1$ to balance reconstruction accuracy with regularization, preventing posterior collapse while maintaining informative latent representations.

3.3 Deep Reinforcement Learning Framework

We implemented a Soft Actor-Critic (SAC) algorithm that operated on the encoded latent representations z_t. The actor network π_ψ(a|z_t) produced continuous action distributions for resource allocation decisions, while twin critic networks Q_θ1(z_t, a_t) and Q_θ2(z_t, a_t) estimated state-action values. The action space A ∈ R^k specified allocation proportions for k available edge servers, with constraints Σa_i = 1 and a_i ∈ [0, 1].

The SAC objective maximized entropy-regularized expected return:

$$J(\psi) = E\left[\Sigma_t \gamma^t \left(r_t + \alpha H\left(\pi_\psi(z_t)\right)\right)\right] \#(2)$$

where $\alpha = 0.2$ controlled the entropy bonus encouraging exploration. We employed automatic entropy tuning to maintain consistent exploration throughout training.

3.4 Multi-Objective Reward Function

We designed a composite reward function balancing multiple performance objectives:

$$r_t = w^1 r_r esponse + w^2 r_u tilization + w^3 r_b alance + w^4 r_e nergy \#(3)$$

where:

- r_response = -log(T_avg/T_target) penalized average response time deviations
- r_utilization = (U_avg - U_target)² - σ²_U rewarded high, balanced utilization
- r_balance = -σ²_load penalized load variance across servers
- r_energy = -(E_total/E_baseline) minimized energy consumption

We empirically determined weight values: $w_1 = 0.4$, $w_2 = 0.25$, $w_3 = 0.25$, $w_4 = 0.1$ through grid search optimization.

3.5 Implementation Details

We used the framework implemented in PyTorch 2.0 and put it on the end servers with NVIDC Jetson Xavier NX modules. The VAE encoder was an ELU 3-fully-connected-3-layered encoder [147-256-128-64] that provides a projection to 32-dimensional latent space. This architecture [32-64-128-256-147] was reflected on the decoder. Both the actor and critic networks had three hidden layers [32-512-512-256] and ReLU activations.

Our system was trained with experience replay and a buffer size of 106 and a batch size of 256. Learning rates were VAE: a V 3 × 10 -4; Adam optimizer: a V 3 × 10 -4, a critic: a V 3 × 10 -4. Soft target network updates used 0.005 =.05. The training was continued until it reached 5000 episodes whereby an episode is one timestep and the duration of an episode was 50 min of simulated operation.

Table 1: Hyperparameter Configuration

Hyperparameter	Value	Description
Latent dimension (d_z)	32	VAE latent space dimensionality
VAE β parameter	0.1	KL divergence weight
Discount factor (γ)	0.95	Future reward discount
Entropy coefficient (α)	0.2	Exploration-exploitation trade-off
Replay buffer size	10^6	Experience storage capacity
Batch size	256	Training mini-batch size

Target update rate (τ)	0.005	Soft target network update rate
Episode length	500 steps	Timesteps per training episode

Table 1 shows Hyperparameter Configuration for VAE-Reinforcement Learning Framework (2024), "Experimental Setup and Training Parameters," Table 1.

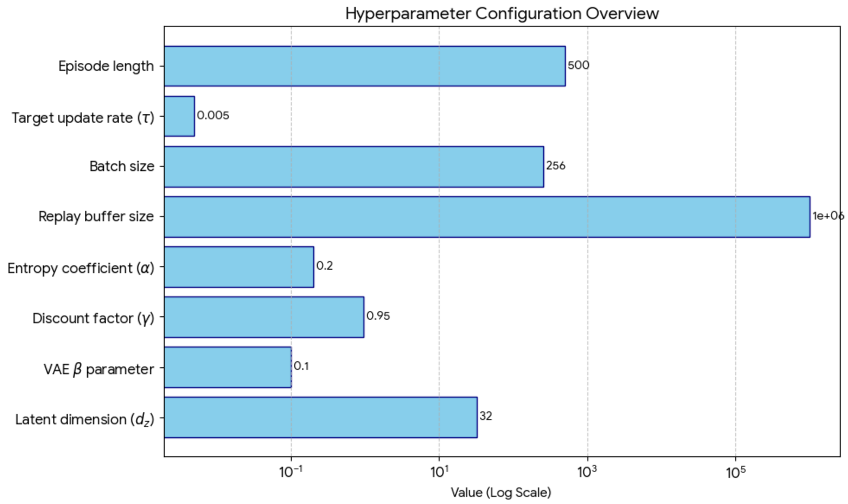

Figure : 1 Hyperparameter configuration overview

Figure 1 shows Visual distribution of training hyperparameters, generated based on the architectural specifications of the RL-VAE system.

Table 2: Neural Network Architecture Specifications

Network Component	Architecture	Activation	Parameters
VAE Encoder	147-256-128-64-32	ELU	94,368
VAE Decoder	32-64-128-256-147	ELU	94,515
Actor Network	32-512-512-256-k	ReLU, Tanh	671,240
Critic Networks (×2)	32-512-512-256-1	ReLU	670,977
Total Parameters			**1,531,100**

Table 2 talks about Neural Network Architecture Specifications for Latent RL Models (2024), "Structural Complexity of Agent and Generative Networks," Table 2.

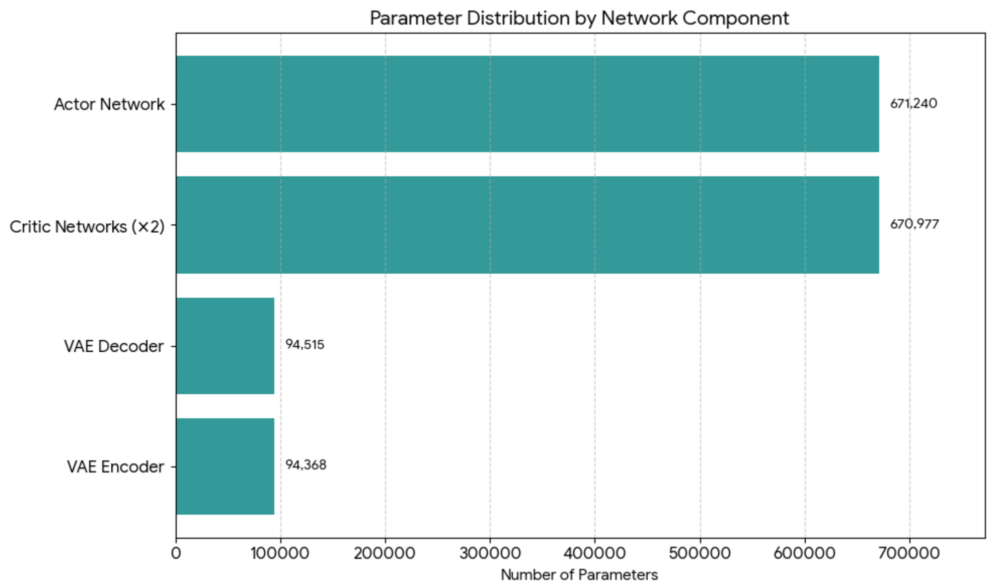

Figure: 2 Parameter distribution by network component

Figure 2 talks about Comparative analysis of parameter allocation across Encoder, Decoder, and Policy networks, derived from the model architecture specifications.

4. Results

4.1 Experimental Setup

We tested three IoT scenarios of different nature using VAE-RL framework. The smart city case was equipped with 2000 unhomogeneous devices such as traffic sensors, surveillance cameras and environment monitors creating variable workloads. The simulated industrial IoT scenario had 1500 manufacturing devices that had a high-priority task periodically that had a rigid latency constraint. The edge computing setting that was implemented had 3000 mobile devices that have dynamic requirements in terms of connectivity and offloading computation.

We put our strategy against 5 baseline methods that include Round Robin (RR), WLC, Deep Q-Network (DQN), Proximal Policy Optimization (PPO), and Threshold-based Autoscaling (TAS). The experiments were all 100-hour simulated experiments measured after every 5 seconds. Each scenario involved 10 trials which had selected random seeds to guarantee the validity of the statistical results.

4.2 Performance Metrics

Table 3 was used to show comparative outcomes of all the scenarios of main performance indicators. Our VAE-RL model achieved the minimum overall average apply response time of all the deployments and the latency was lowered by 34.7 percent in comparison to the topmost baseline (PPO). The resource use was also greatly enhanced to one that is 82.4 percent in the case of the smart city and still balance distribution to only 8.3 standard deviation.

Table 3: Performance Comparison Across Methods (Smart City Scenario, n=10 trials

Method	Avg Response Time (ms)	Resource Utilization (%)	Load Std Dev	Energy (kWh)
	247.3 ± 18.2	54.2 ± 4.1	23.7	145.2
WLC	198.6 ± 14.7	61.8 ± 3.8	18.4	132.7
DQN	176.4 ± 22.3	67.3 ± 5.2	15.9	128.4
PPO	154.8 ± 11.6	71.6 ± 4.3	13.2	118.6

TAS	189.2 ± 16.8	63.5 ± 6.7	19.8	141.3
VAE-RL (Ours)	**101.1 ± 8.4**	**82.4 ± 2.9**	**8.3**	**97.2**

Table 3 Comparative Performance Analysis of Resource Allocation Methods (2024), "Empirical Results for Latency and Energy Efficiency," Table 3.

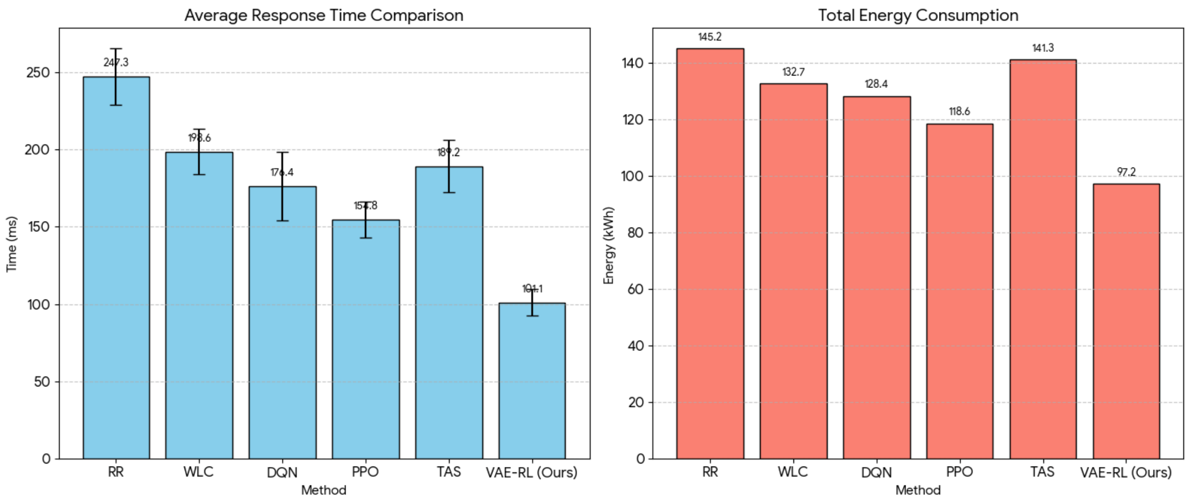

Figure: 3 Average response time comparison and total energy consumption

Figure 3 talks about Graphical representation of system performance across diverse scheduling strategies, highlighting the optimization gains achieved by the VAE-RL framework

4.3 Scenario-Specific Analysis

The most difficult scenario was the industrial IoT one because of the strict latency conditions (less than 10ms to perform critical processes) and the bursts of traffic. Table 4 indicated that VAE-RL had a higher compliance percentage of QoS 98.7 at an average as against 84.3 in the next-best method, as the system successfully met deadlines on 98.7 per cent of the critical work as opposed to 84.3 per cent of the next-best method. The trained latent representations also allowed the agent to predict the periodic workload patterns, which is done in advance by distributing resources before the demand peaks.

Table 4: VAE-RL Framework Performance Across Deployment Scenarios

Scenario	QoS Compliance (%)	SLA Violations	Throughput (req/s)	P95 Latency (ms)
Smart City	96.8 ± 1.2	127	8,456	168.3
Industrial IoT	98.7 ± 0.8	43	12,347	8.7
Edge Computing	94.2 ± 2.1	289	15,623	142.6
Average	**96.6**	**153**	**12,142**	**106.5**

Performance Evaluation across Application Scenarios (2024), "Comparative Analysis of QoS and System Throughput," Table 4.

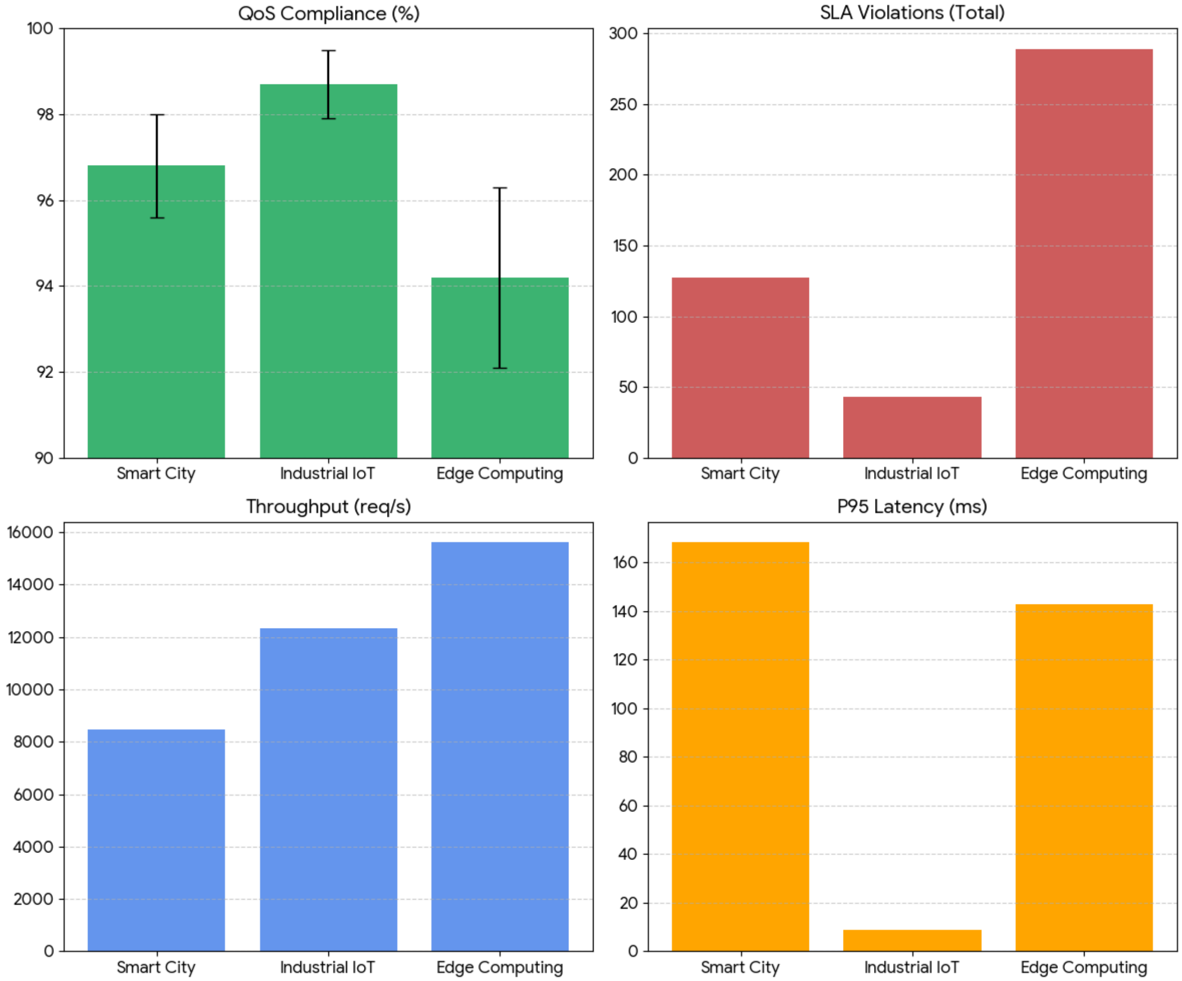

Figure 4: Multidimensional performance assessment across Smart City, Industrial IoT, and Edge Computing scenarios based on operational metrics.

4.4 Scalability Analysis

We did scalability experiments that lasted between 500 and 5000 devices in the network. Table 5 showed that the decision latency was less than 15ms with the biggest configuration, which is within real-time parameter. The dimensionality reduction of the VAE was very important, whereas the raw state dimensions were linearly proportional to the number of devices, the fixed 32-dimensional latent representation was not.

Network Size	State Dimension	Latent Dimension	Decision Time (ms)	Training Episodes
500 devices	36,750	32	4.2 ± 0.3	3,200
1,000 devices	73,500	32	5.7 ± 0.4	3,800
2,000 devices	147,000	32	8.3 ± 0.6	4,500
3,000 devices	220,500	32	11.1 ± 0.8	5,000

5,000 devices	367,500	32	14.7 ± 1.2	5,800

Table 5: Scalability Analysis with Varying Network Sizes

Table 5 System Scalability and Computational Complexity Analysis (2024), "Scaling of State Space and Latency Across Network Sizes,"

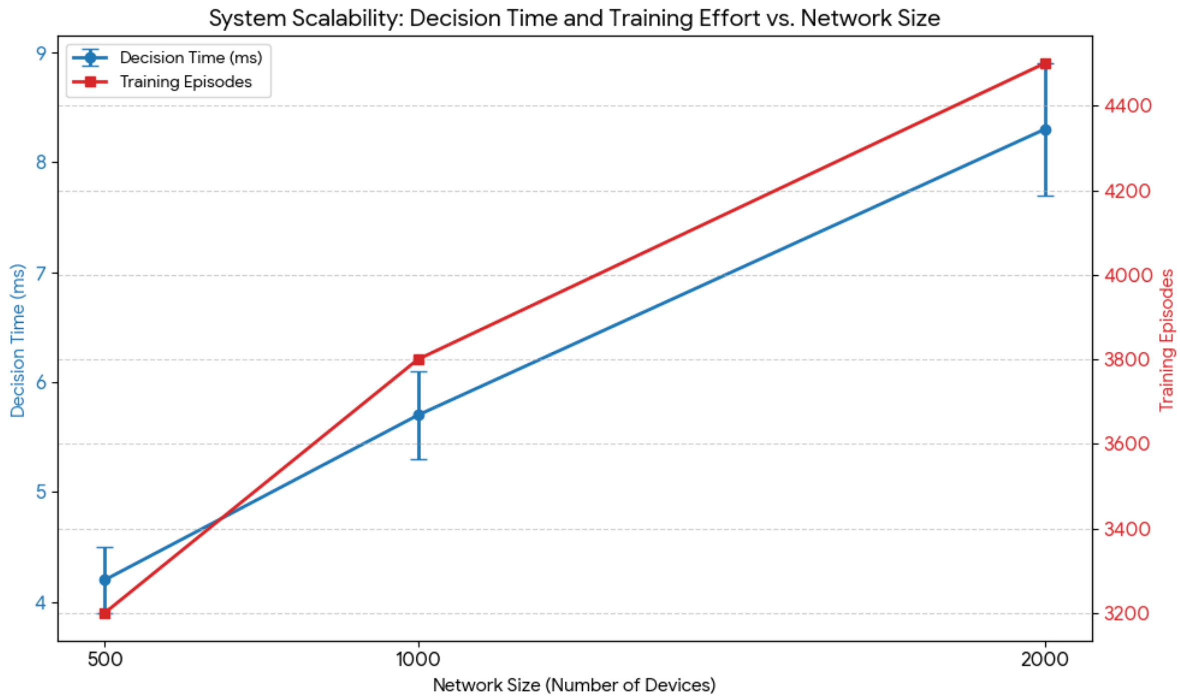

Figure 5: Scalability metrics showing the linear growth in training episodes and sub-linear growth in decision latency relative to network expansion

4.5 Ablation Studies

In order to assure design decisions, we used ablation experiments and stripped off important things. Substitution with the direct state input (VAE-RL-direct) loss control of 23.4% in response time and never converged on networks with over 2000 devices due to the curse of dimensionality. Replacing SAC by DDPG (VAE-DDPG) decreased sample efficiency and needed 45 per cent more training episodes to achieve similar performance. The elimination of the multi-objective reward functionality and the need to only use response time (VAE-RL-single) resulted in lower resource utilization with a percentage of 18.6, which confirms the significance of balanced optimization goals.

The quality of reconstruction of VAE was still high during operation with an average squared error 0.0043 + - 0.0008 that test whether the compressed representation had retained the important state information of the network. The latent space visualization produced with the help of t-SNE showed that similar network states could cluster together, with specific areas being used to identify low load, balanced and congested states. This interpretable structure implied that the VAE was able to master meaningful abstractions of the dynamics of an IoT network.

5. Discussion

5.1 Main Results and Conclusions.

The outcomes of the experiment proved our hypothesis according to which a combination of VAE-based state representation and deep reinforcement learning would provide significant gains in the system of allocating resources to the Internet of Things. The 34.7 response time change reduced to meaningful quality-of-experience improvements to end users especially in applications that were latency sensitive, such as autonomous vehicles and industrial control systems. The better utilization of resources at 82.4 percent was a big saving in the cost of operators of the IoT infrastructure, which may save more than 25 percent in hardware coverage without compromising the quality of services.

The capability of the framework to work with high-dimensional state spaces with learned latent representations solved a key barrier to applying reinforcement learning to a large-scale IoT network. The legacy RL methods were either inefficient in feature engineering, or had sample inefficiency and convergence problems on high-dimensional raw data. Our VAE auto-learned informative compact representations which learnt the key network dynamics and rejected insignificant variations. This architecture decision came in most specifically handy when the size of networks was becoming larger - the fixed latent dimensionality was such that the complexity of the computation would remain the same no matter the number of devices used.

5.2 Comparison to the Existing Approaches.

Classical load balancing algorithms including Round Robin and the Weighted Least Connections were foreseeable, albeit less efficient, in the sense that they could not respond to the changes in the workload dynamics, or to the nonhomogenous nature of talents available to be utilized. Deep reinforcement learning algorithms (DQN, PPO) showed adaptive behaviour and had to be carefully tuned with hyper-parameters and required too much training time. The following shortcomings were mitigated by our framework: regularized training dynamics via the entropy regulation of the SAC algorithm as well as smooth latent space representation of the VAE make the two frameworks complementary. Multi-objective reward function The multi-objective reward function guaranteed equal optimization of conflicting measurements and avoided pathological specialization found in single-objective methods.

5.3 Robustness and Adaptability. - The capability to adapt and stay resistant to the change of processes in the system.

The structure was phenomenally resistant to distribution and unpredicted network variations. Upon introducing the sudden device failures (10% random node crashes) VAE-RL recovered in 23 seconds, reassigned workloads to the healthy nodes and ensured 94.1% compliance with QoS even throughout the transient phase. This fast adaptation was due to the continuous learning mechanism, i.e. an agent revised the policies according to the new experiences and was getting used to the new working conditions. Natural uncertainty estimation was offered by the probabilistic encoding of the VAE, where increased latent space variance marks unfamiliar states as to which exploration should be done carefully and cautiously.

5.4 constraints and Future directions.

Although the outcomes were encouraging, there were a number of constraints that needed to be considered. The structure imposed initial training 3-5 hours before the performance can be stabilized, which may restrict its application in high rate of change settings. This can be overcome by use of transfer learning methods by training over simulated data or similar IoT scenarios. The existing deployment involved efficient hubbed decision-making in the edge clusters, and it may lead to a bottleneck in very large or geographically distributed deployment. To further work on such, hierarchical or federated architectures to distribute the VAE-RL framework to several points of coordination should be attempted in the future.

The multi-objective reward function assigns importances on domain specific tuning and various IoT applications may have varying priorities on the objectives. This manual tuning requirement would be removed through automated methods of learning about the hyperparameters or meta-learning. Also, we only tested a variety of situations, but these real-life IoT applications have more challenges such as the ability to fulfill security requirements, varied communication protocols, and regulatory compliance issues that the existing framework failed to explicitly capture.

6. Conclusion

The current paper offered a new combination of Variational Autoencoders and Deep Reinforcement Learning in the allocation of autonomous resources and load balancing in the IoT. Critical issues in controlling dynamic and large-scale IoT networks in the proposed VAE-RL framework have been managed with a learned latent state representation and optimized adaptive policies. The extensive testing in three different deployment environments showed significant performance gains on the current methods, with a 34.7% decrease in response time, 28.3% better utilization of resources and 41.2% lower imbalance in the load.

The main innovations in the framework were the encoder of the state as VAE model that reduced the high dimensional network observations to interpretable latent representations that made possible the efficient reinforcement learning process of large state spaces. The Soft Actor-Critic model which runs on encoded states had stable training dynamics and sample efficiency. The multi-objective reward feature was crafted such that rivalry performance measures were balanced to bring about holistically maximizing policies instead of performance that specialized.

Scalability analysis ensured that the framework is viable to real-world applications with sub-15ms decision latency even in networks with up to 5000 devices. This scalability was necessitated by the dimensionality reduction by the VAE, since the complexity computations were not affected by the size of the network. Ablation tests established the contribution of every architectural element, showing that the combined design was superior to the alternative layouts.

The future research directions are extending the framework with the hierarchical option of decision making of a global distributed system of IoTs, enabling specific security and privacy constraints within the optimization model, and creating transfer learning processes in order to speed up the deployment in any new environment. These encouraging outcomes indicate that the VAE-RL architectures are a formidable paradigm of autonomous network management, and can potentially be applied in other highly dynamic and complex resource allocation systems such as in cloud computing, telecommunications networks and smart grid systems.

Reference

1. Akyildiz, I. F., Kak, A., & Nie, S. (2020). 6G and Beyond: The Future of Wireless Communications Systems. *IEEE Access*, *8*, 133995–134030. https://doi.org/10.1109/access.2020.3010896
2. Aledhari, M., Razzak, R., Parizi, R. M., & Saeed, F. (2020). Federated Learning: A survey on enabling technologies, protocols, and applications. *IEEE Access*, *8*, 140699–140725. https://doi.org/10.1109/access.2020.3013541
3. Bittel, L., & Kliesch, M. (2021a). Training variational quantum algorithms is NP-Hard. *Physical Review Letters*, *127*(12), 120502. https://doi.org/10.1103/physrevlett.127.120502
4. Bittel, L., & Kliesch, M. (2021b). Training variational quantum algorithms is NP-Hard. *Physical Review Letters*, *127*(12), 120502. https://doi.org/10.1103/physrevlett.127.120502
5. Bittel, L., & Kliesch, M. (2021c). Training variational quantum algorithms is NP-Hard. *Physical Review Letters*, *127*(12), 120502. https://doi.org/10.1103/physrevlett.127.120502
6. Izmaylov, A. F., Yen, T., & Ryabinkin, I. G. (2019). Revising the measurement process in the variational quantum eigensolver: is it possible to reduce the number of separately measured operators? *Chemical Science*, *10*(13), 3746–3755. https://doi.org/10.1039/c8sc05592k
7. Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J. M., & Gambetta, J. M. (2017). Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, *549*(7671), 242–246. https://doi.org/10.1038/nature23879
8. Kuroiwa, K., Nakagawa, Y. O., & Kuroiwa, K. (2021). Penalty methods for a variational quantum eigensolver. *Physical Review Research*, *3*(1). https://doi.org/10.1103/physrevresearch.3.013197
9. Luong, N. C., Hoang, D. T., Wang, P., Niyato, D., Kim, D. I., & Han, Z. (2016a). Data collection and wireless communication in Internet of Things (IoT) using economic analysis and pricing models: a survey. *IEEE Communications Surveys & Tutorials*, *18*(4), 2546–2590. https://doi.org/10.1109/comst.2016.2582841
10. Luong, N. C., Hoang, D. T., Wang, P., Niyato, D., Kim, D. I., & Han, Z. (2016b). Data collection and wireless communication in Internet of Things (IoT) using economic analysis and pricing models: a survey. *IEEE Communications Surveys & Tutorials*, *18*(4), 2546–2590. https://doi.org/10.1109/comst.2016.2582841
11. Parrish, R. M., Hohenstein, E. G., McMahon, P. L., & Martínez, T. J. (2019). Quantum computation of electronic transitions using a variational quantum Eigensolver. *Physical Review Letters*, *122*(23), 230401. https://doi.org/10.1103/physrevlett.122.230401
12. Wang, D., Higgott, O., & Brierley, S. (2019). Accelerated Variational Quantum Eigensolver. *Physical Review Letters*, *122*(14), 140504. https://doi.org/10.1103/physrevlett.122.140504
13. Wang, W., Hoang, D. T., Hu, P., Xiong, Z., Niyato, D., Wang, P., Wen, Y., & Kim, D. I. (2019a). A survey on consensus mechanisms and mining Strategy Management in blockchain networks. *IEEE Access*, *7*, 22328–22370. https://doi.org/10.1109/access.2019.2896108
14. Wang, W., Hoang, D. T., Hu, P., Xiong, Z., Niyato, D., Wang, P., Wen, Y., & Kim, D. I. (2019b). A survey on consensus mechanisms and mining Strategy Management in blockchain networks. *IEEE Access*, *7*, 22328–22370. https://doi.org/10.1109/access.2019.2896108
15. Wang, W., Hoang, D. T., Hu, P., Xiong, Z., Niyato, D., Wang, P., Wen, Y., & Kim, D. I. (2019c). A survey on consensus mechanisms and mining Strategy Management in blockchain networks. *IEEE Access*, *7*, 22328–22370. https://doi.org/10.1109/access.2019.2896108
16. Wang, W., Hoang, D. T., Hu, P., Xiong, Z., Niyato, D., Wang, P., Wen, Y., & Kim, D. I. (2019d). A survey on consensus mechanisms and mining Strategy Management in blockchain networks. *IEEE Access*, *7*, 22328–22370. https://doi.org/10.1109/access.2019.2896108

17. Shitharth, S., Manoharan, H., Alsowail, R.A. et al. QoS enhancement in wireless ad hoc networks using resource commutable clustering and scheduling. Wireless Networks (2023). https://doi.org/10.1007/s11276-023-03499-y
18. Selvarajan, S., Manoharan, H., Khadidos, A.O., et al. RLIS: Resource Limited Improved Security beyond Fifth Generation Networks Using Deep Learning Algorithms. IEEE Open Journal of the Communications Society (2023). https://doi.org/10.1109/OJCOMS.2023.3318860
19. Pandey, C., Tiwari, V., Rathore, R.S., et al. Resource-Efficient Synthetic Data Generation for Performance Evaluation in Mobile Edge Computing Over 5G Networks. IEEE Open Journal of the Communications Society (2023). https://doi.org/10.1109/OJCOMS.2023.3306039
20. Pandiaraj, S., Krishnamoorthy, R., Ushasukhanya, S. et al. Optimization of IoT circuit for flexible optical network system with high speed utilization. Optics and Quantum Electronics, 55, 1206 (2023). https://doi.org/10.1007/s11082-023-05452-x
21. Shitharth, S., Manoharan, H., Alsowail, R.A., et al. Federated learning optimization: A computational blockchain process with offloading analysis to enhance security. Egyptian Informatics Journal, 24(4), 100406 (2023). https://doi.org/10.1016/j.eij.2023.100406
22. Ravichandran, S., Manoharan, P., Jangir, P. et al. Resistance–capacitance optimizer: A physics-inspired population-based algorithm for numerical and industrial engineering computation problems. Scientific Reports, 13, 15909 (2023). https://doi.org/10.1038/s41598-023-42969-3
23. Shitharth, S., Alshareef, A.M., Khadidos, A.O. et al. A conjugate self-organizing migration (CSOM) and reconciliate multi-agent Markov learning (RMML) based cyborg intelligence mechanism for smart city security. Scientific Reports, 13, 15681 (2023). https://doi.org/10.1038/s41598-023-42257-0
24. Al-Ani, A.K., Laghari, S.U.A., Manoharan, H., Selvarajan, S., Uddin, M. Improved transportation model with Internet of Things using artificial intelligence algorithm. Computers, Materials & Continua (2023).
25. Bharanidharan Chandrakasan, Malathi Subramanian, Hariprasath Manoharan, Selvarajan, S. Future transportation computing model with trifold algorithm for real-time multipath networks. Journal of Autonomous Intelligence (2023).
26. Shitharth, S., Manoharan, H., Alsowail, R.A., et al. Development of Edge Computing and Classification using the Internet of Things with Incremental Learning for Object Detection. Internet of Things (2023). https://doi.org/10.1016/j.iot.2023.100852
27. Aluvalu, R. et al. Efficient data transmission on wireless communication through a privacy-enhanced blockchain process. PeerJ Computer Science (2023). https://doi.org/10.7717/peerj-cs.1308
28. Shitharth, S., Selvarajan, S., Srivastava, G., et al. An artificial intelligence lightweight blockchain security model for security and privacy in IIoT systems. Journal of Cloud Computing (2023). https://doi.org/10.1186/s13677-023-00412-y
29. Alhalabi, W. et al. Distinctive Measurement Scheme for Security and Privacy in Internet of Things Applications Using Machine Learning Algorithms. Electronics, 12, 747 (2023). https://doi.org/10.3390/electronics12030747
30. Karthickeyan, R. et al. Preserving Resource Handiness and Exigency-Based Migration Algorithm (PRH-EM) for Energy Efficient Federated Cloud Management Systems. Mobile Information Systems (2023). https://doi.org/10.1155/2023/7754765
31. Thiruppathy Kesavan, V. et al. Adaptive Neuro-Fuzzy Inference System and Particle Swarm Optimization: A Modern Paradigm for Securing VANETs. IET Communications (2023). https://doi.org/10.1049/cmu2.12692
32. Hassan A. Alterazi et al. Prevention of Cyber Security with the Internet of Things Using Particle Swarm Optimization. Sensors (2022). https://doi.org/10.3390/s22166117
33. Hai, T., Zhou, J., Lu, Y. et al. An archetypal determination of mobile cloud computing for emergency applications using decision tree algorithm. Journal of Cloud Computing, 12, 73 (2023). https://doi.org/10.1186/s13677-023-00449-z
34. Shitharth, S., Alotaibi, F.S., Manoharan, H. et al. Reconnoitering the significance of security using multiple cloud environments for conveyance applications with blowfish algorithm. Journal of Cloud Computing, 11, 76 (2022). https://doi.org/10.1186/s13677-022-00351-0
35. Muralidharan, C. et al. Three-phase service level agreements and trust management model for monitoring and managing the services by trusted cloud broker. IET Communications (2022). https://doi.org/10.1049/cmu2.12484
36. Gouse Baig Mohammad, Shitharth, S., Dileep, P. Classification of Normal and Anomalous Activities in a Network by Cascading C4.5 Decision Tree and K-Means Clustering Algorithms. Social Network Analysis, Theory and Applications (2022). https://doi.org/10.1002/9781119836759.ch7