A Lightweight Deep Learning Framework for Intelligent and Secure IoT Network Management in Smart City Environments

CH.Nagaraju[1,2], Sai Kiran Oruganti[3]

[1]Post-Doctoral Researcher, Lincoln University college, 4730, petaling Jaya, selangor Darul Ehsan, Malaysisa. ***(Student ID: LUCPDF202604060093) (Application Number: LUC/CPGS/PDF/2025301002).***

[2]Professor, Department of Electronics and communication Engineering, Annamacharya University , Rajampet, Andhra Pradesh , India

[1,2]chrajuaits@gmail.com, [1,2]pdf.Choudavaram@lincoln.edu.my, [3]saisharma@lincoln.edu.my, [3]saikiran.oruganti@gmail.com

[1,2]0000-0001-9178-6448

Abstract:

Smart cities need to depend on large-scale Internet of Things (IoT) networks for critical urban functions like traffic management, air quality monitoring, smart grid, waste management and surveillance by the public. The need to keep these networks energy-efficient, secure, and reliable and to manage them on a large scale, however, poses significant challenges due to the heterogeneous and dynamic nature of these networks. In this paper, the concept of lightweight deep learning framework for intelligent clustering and routing in smart city IoT environment is proposed, which is termed as OptiMobileGAT-Net. They combine MobileNet-V2 for efficient spatial feature extraction, Graph Attention Networks (GATs) for topology-aware representation learning, and Gated Recurrent Units (GRUs) for modeling temporal dynamics of the network conditions to make adaptive routing decisions. The review of recent IoT network management approaches reveals four research gaps: high computational overhead optimization algorithms, lack of attention to energy efficiency issues, lack of focus on trust and security, and instability of RL in dynamic environments, limited adaptability of existing deep learning models for resource-limited devices.

To overcome these limitations, we introduced multi-factor node feature representation, lightweight spatial feature, topology learning based on graph attention, temporal modelling with GRUs, and a feed-forward attention optimization module, which collectively considers energy, trust, and security without iterative optimization. The framework is intended to be evaluated in a simulated 100-node smart city IoT network with malicious-node scenario. The identified research gaps and proposed methodology are presented in this paper and experimental validation and performance evaluation are presented in the subsequent stage.

Keywords: Smart Cities; IoT Networks; Graph Attention Networks; MobileNet; Secure Routing;

INTRODUCTION

Large scale Internet of Things (IoT) infrastructures have become more popular in smart cities for delivery of intelligent urban services, such as traffic control, environmental sensing, smart grid management, waste management and public safety. The applications rely on a dense network of sensors with strong interconnections that are constantly collecting and reporting data to enable real-time monitoring and decision making. However, smart city IoT systems are heterogeneous and have dynamic operating environments, which bring many challenges to the realization of energy-efficient communication, secure data transmission, reliable routing, and scalable network management [1-2].

Although significant research has been conducted on IoT clustering and routing, there are still several limitations with the current solutions. Most popular metaheuristic and fuzzy logic algorithms are based on optimizing energy, and reinforcement learning techniques tend to have a poor performance when the network conditions change rapidly. Likewise, graph-based learning models typically consume lots of computational resources and thus are not suitable for resource constrained IoT devices [3]. Additionally, the majority of the existing works focus on optimizing either one of the energy efficiency, trust evaluation

and security separately from each other and do not consider them in a unified decision making model, which is not effective in practical deployment of smart cities [4].
To overcome these challenges, The proposed architecture combines MobileNet-V2 for computationally efficient spatial feature extraction, Graph Attention Networks (GATs) for topology-aware representation learning, and Gated Recurrent Units (GRUs) for modelling temporal variations in network behaviour. A feed-forward attention optimization module further enables the joint evaluation of energy, trust, and security without relying on computationally expensive iterative optimization techniques. This paper presents the identified research gaps and the proposed methodology, while experimental validation and comprehensive performance analysis are reserved for future work .

RELATED WORK

In 2025, Darabkh and Al-Akhras developed a centrality-aware clustering and routing approach for smart city IoT sensor networks. Their framework employed dynamic cluster-head selection and role-based routing to improve network lifetime and energy utilization. Although the proposed method demonstrated enhanced energy efficiency, its dependence on optimization-based decision making increased computational complexity, thereby limiting its ability to respond efficiently to rapidly changing network conditions in large-scale urban deployments [5].
similarly in 2025, Tseng et al. presented a graph attention and reinforcement learning-based framework for intelligent urban traffic management. This combination of graph attention mechanisms and reinforcement learning enabled the model to successfully model spatial interactions and make optimal routing decisions in dynamic traffic conditions. The challenge of the training heavy RL and continual reward optimisation, however, limited its feasibility for use with resource-constrained IoT devices and edge gateways [6].
In 2024, Omrany et al. did a thorough review of smart city technologies that are based on IoT and found energy efficiency, reliability of communication, security and scalability are the main challenges in large-scale smart city IoT infrastructures. The study also noted that these goals are usually tackled individually by the existing solutions, which are not particularly successful in smart city contexts characterized by heterogeneity and continuous evolution [7].
In 2024, Nguyen et al. gave an overview of AI-driven smart city IoT systems and emphasized the expanding role of deep learning in the management of smart networks. The authors pointed out the necessity of lightweight and low computation requirement models that can simultaneously solve the routing problem, energy efficiency, trust management and secure communication issues, for practical deployment on low-power IoT devices [8].

Research gaps

The literature reveals four main research gaps using a critical analysis of the existing literature. Firstly, most of the clustering and routing technologies remain heuristic/optimization based which gives a high computational overhead and restricts the adaptability in dynamic network scenarios. Secondly, energy efficiency, trust evaluation and security are often optimized independently of each other and not jointly in a single learning framework. Third, reinforcement learning is often a time-consuming process that needs long-term training and exploration, potentially resulting in slow convergence and unstable routing decisions in dynamic environments. Third, there are a few graph-based deep learning models that require significant computational and memory resources and are therefore hard to deploy on resource constrained sensor nodes and edge devices on the IoT. Such constraints inspire the design of OptiMobileGAT-Net, a lightweight feature extraction, graph attention learning, and temporal modelling-based method to provide efficient, secure, and adaptive clustering and routing in large-scale smart city IoT networks.

PROPOSED METHODOLOGY

The proposed OptiMobileGAT-Net is a unified deep learning framework developed to improve clustering and routing in large-scale smart city Internet of Things (IoT) networks as shown in Figure 1. The framework is designed to overcome the limitations of existing approaches by simultaneously considering energy efficiency, trust, and security while maintaining low computational complexity. The network consists of N heterogeneous sensor nodes deployed across multiple smart city applications, including traffic monitoring, air quality assessment, smart grid management, waste management and public surveillance. These sensor nodes communicate through multi-hop wireless links and transmit the collected information to a central gateway for further processing and monitoring.

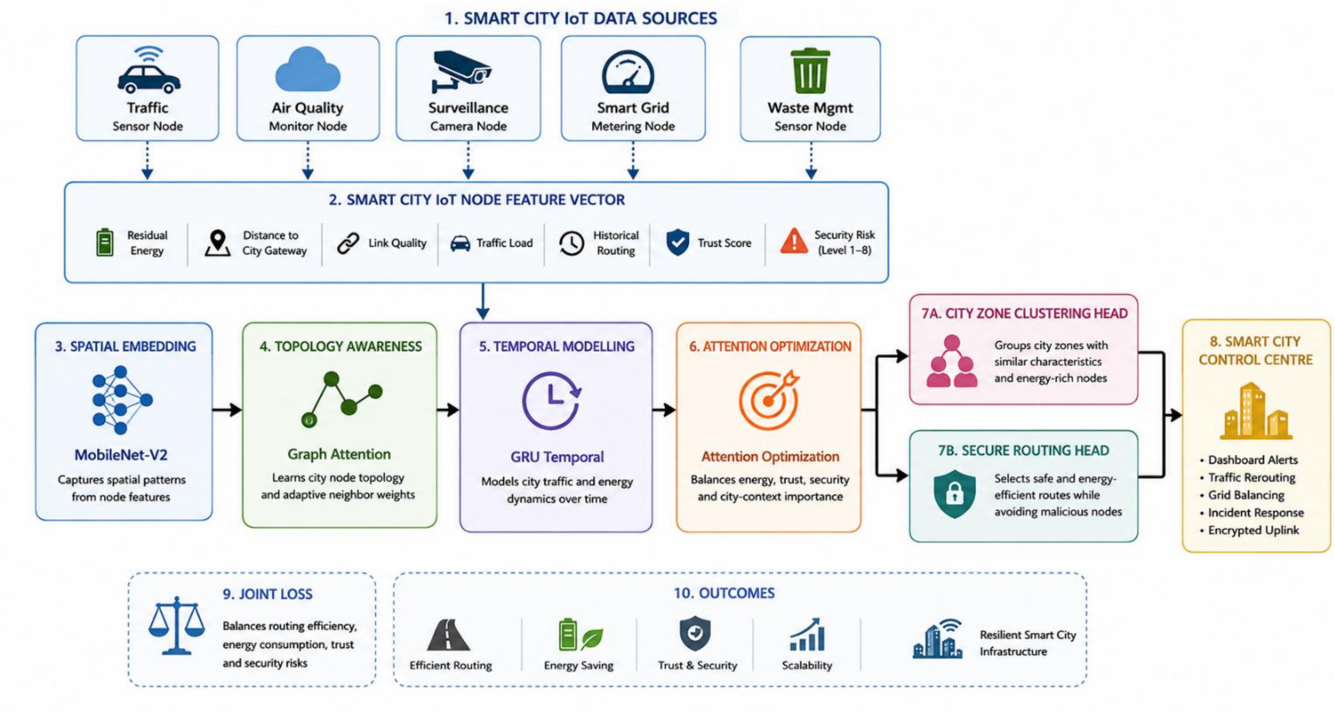

Fig. 1. OptiMobileGAT-Net pipeline for smart city IoT: city sensor node types feed a multi-factor feature vector through MobileNet-V2 spatial extraction, graph attention topology encoding, GRU temporal modeling, and attention-based optimization into city zone clustering and secure routing output heads.

The proposed framework follows an end-to-end learning strategy that eliminates the need for computationally expensive heuristic, metaheuristic, or reinforcement learning-based optimization during network operation. Initially, multiple node attributes, including residual energy, communication distance, neighbourhood connectivity, trust value and security status are combined to construct a comprehensive node feature representation. These features are processed using MobileNet-V2 to extract lightweight spatial representations with minimal computational overhead. Subsequently, Graph Attention Networks (GATs) learn the structural relationships among neighboring sensor nodes by assigning adaptive attention weights according to the network topology. The extracted spatial representations are then forwarded to Gated Recurrent Units (GRUs) to capture temporal variations arising from changes in network traffic, energy consumption, and node behaviour. Finally, a feed-forward attention optimization module jointly evaluates energy, trust, and security to determine optimal cluster-head selection and secure routing paths without relying on iterative optimization procedures.

By integrating spatial, topological, and temporal learning within a single architecture, the proposed framework addresses the major limitations identified in existing studies. The lightweight design reduces computational and memory requirements, enabling deployment on resource-constrained IoT devices and edge gateways. Furthermore, the joint optimization of energy efficiency, trust, and security provides adaptive and reliable clustering and routing decisions under dynamic smart city conditions.

A. Smart City IoT Node Feature Representation

The proposed framework begins by constructing a comprehensive feature vector for each sensor node deployed across different smart city services, including traffic monitoring, air quality sensing, surveillance, smart grid metering, and waste management. Each node is represented by a multidimensional feature vector as shown in eq. (1).

$$X_i = [E_i, d_i, LQ_i, \delta_i, H_i, T_i, S_i] \qquad (1)$$

where E_i denotes the residual energy of node i, d_i represents the distance to the city gateway, LQ_i is the communication link quality, δ_i denotes the traffic load, H_i represents the historical routing behaviour, T_i is the trust score, and S_i denotes the security risk level. These attributes collectively describe the operational status of each node and provide the input to the proposed deep learning framework [9]

B. MobileNet-V2 Spatial Feature Extraction

The generated node feature vectors are processed using MobileNet-V2 to obtain lightweight spatial representations. The network employs depth-wise separable convolutions together with pointwise convolutions and inverted residual bottleneck blocks to reduce computational complexity while preserving discriminative information. This stage transforms the original node features into compact spatial embeddings using eq. (2).

$$S_i = f_{MN}(X_i) \qquad (2)$$

which capture the relationships among node attributes while maintaining low memory and computational requirements suitable for resource-constrained IoT devices.

C. Graph Attention-Based Topology Learning

The spatial embeddings generated by MobileNet-V2 are supplied to a Graph Attention Network (GAT) to model the communication topology of the smart city IoT network. The GAT figures adaptive attention coefficients for neighbouring nodes according to their relative importance instead of assigning equal weights. Accordingly, nodes with higher trust, stronger connectivity and better communication quality exert greater influence during feature aggregation. The resulting topology-aware embedding is expressed as shown in eq. (3).

$$H_i = f_{GAT}(S_i) \qquad (3)$$

The eq. (3). enabling the framework to learn the structural characteristics of the network dynamically.

D. GRU-Based Temporal Feature Learning

Since IoT network conditions continuously change because of varying traffic loads, energy consumption and node behaviour the topology-aware features are processed using a **Gated Recurrent Unit (GRU)** network. The GRU maintains temporal memory by selectively retaining relevant historical information while discarding obsolete states.

$$h_i = f_{GRU}(H_i) \qquad (4)$$

The eq. (4) enables the framework to model temporal variations in network dynamics and generate strong temporal embeddings which improve adaptability under continuously evolving smart city environments.

E. Multi-Factor Attention Optimization

The temporal embeddings are passed to a feed-forward attention optimization module that jointly evaluates multiple routing criteria. Different conventional heuristic or metaheuristic optimization methods, this module directly learns the relative importance of energy, trust and security through trainable attention weights. The optimized node representation is obtained from eq. (5).

$$O_i = \alpha_E E_i + \alpha_T T_i + \alpha_S S_i \qquad (5)$$

whereα_E, α_T and α_S represent the learned attention weights. This strategy eliminates iterative optimization procedures while providing adaptive decision making with significantly lower computational overhead [10].

F. City Zone Clustering Head

The within each smart city zone. The clustering module selects nodes possessing high residual energy, reliable communication quality and strong trust values while maintaining balanced cluster formation throughout the network. This process reduces communication overhead, balances energy consumption among sensor nodes, and extends the operational lifetime of the overall IoT network.

G. Secure Routing Head

Following cluster establishment, the routing module determines secure multi-hop communication paths between cluster heads and the city gateway. Routing decisions are based on the optimized node representations generated by the attention module. Nodes exhibiting low trust values, high security risks or excessive congestion are excluded from route construction, thereby improving routing reliability and protecting the network against compromised or malicious nodes.

H. Joint Multi-Objective Loss Function

During model training, all components are optimized simultaneously using a joint objective function comprising routing, trust, and security losses as mentioned in eq. (6).

$$L = L_{route} + \delta_1 L_{trust} + \delta_2 L_{security} \qquad (6)$$

where L_{route} minimizes routing cost, L_{trust} improves trust prediction accuracy, and $L_{security}$ reduces the selection probability of insecure nodes. The weighting parameters δ_1, δ_2 regulate the contribution of each objective. The joint optimization strategy enables the framework to learn balanced representations that simultaneously maximize energy efficiency, routing reliability, and network security.

e optimized node representations are supplied to the clustering head to identify suitable Cluster Heads (CHs)

I. Smart City Control Centre

The final clustering and routing is broadcasted to the Smart City Control Centre where the gathered data is used for providing intelligence management in the city: traffic management, smart grid management, environmental monitoring, incident management, encrypted data communication, etc. By optimizing network operation, OptiMobileGAT-Net can help to improve energy efficiency and secure communication while delivering scalable routing performance in large-scale deployments for smart cities

CONCLUSION AND FUTURE SCOPE

This paper introduced a comprehensive deep learning based approach to the clustering and routing problem in smart city IoT network. After analysing recent studies, four major research gaps were identified: reliance on heuristic and metaheuristic optimization techniques involving high computations, separate optimization of energy efficiency, trust and security, slow convergence and instability in reinforcement learning-based techniques, and the reduced applicability of the computation-intensive architectures to the resource-constrained IoT devices. This approach is limited by its reliance on lightweight spatial feature extraction by MobileNet-V2, topology-aware representation learning using Graph Attention Networks (GATs), temporal network dynamics modelling using Gated Recurrent Units (GRUs), and the joint evaluation of energy, trust and security using a feed-forward attention optimization module. The proposed architecture is designed to be evaluated in a simulated smart city IoT scenario, having 100 heterogeneous sensor nodes deployed in a smart city application (traffic monitoring, air quality sensing, smart grid management, public surveillance, and waste management) based on which malicious nodes can be considered. Future work will include thorough experimental evaluation using performance metrics like energy consumption, routing efficiency, malicious node detection accuracy, computation overhead and packet delivery ratio, and comparing the results with the existing state of the art techniques, and implementing in a real-world smart city IoT testbed.

References

1. H. Nguyen, D. Nawara, and R. Kashef, "Connecting the indispensable roles of IoT and artificial intelligence in smart cities: A survey," Journal of Information and Intelligence, vol. 2, no. 3, pp. 261–285, May 2024.

2. C. Goumopoulos, "Smart City Middleware: A Survey and a Conceptual Framework," IEEE Access, vol. 12, pp. 4015–4047, Jan. 2024.
3. H. Omrany, K. M. Al-Obaidi, M. Hossain, N. A. M. Alduais, H. S. Al-Duais, et al., "IoT-enabled smart cities: A hybrid systematic analysis of key research areas, challenges, and recommendations for future direction," Discover Cities, vol. 1, no. 2, Art. no. 2, Mar. 2024.
4. G. Rathee, C. A. Kerrache, C. M. T. Calafate, M. Bilal, and H. Song, "SMART: A Secure Remote Sensing Solution for Smart Cities Urban Areas," IEEE Sensors Journal, vol. 24, no. 7, pp. 11553–11561, Apr. 2024.
5. K. A. Darabkh and M. Al-Akhras, "Towards Optimized IoT Sensor Networks for Smart Cities: Centrality-Aware Position-Based Occlusion-Driven and Role Dynamics Solutions for Clustering and Routing," IEEE Internet of Things Journal, early access, 2025, doi: 10.1109/JIOT.2025.3570612.
6. K.-K. Tseng, Z. Yang, H. Tang, C.-M. Chen, S. Kumari, and M. S. Hossain, "Smart IoE-Integrated Traffic Control: Dynamic Multisemantic Graph Attention and Reinforcement Learning for Optimizing Urban Mobility," IEEE Internet of Things Journal, vol. 12, no. 4, pp. 7468-7484, 2025, doi: 10.1109/JIOT.2024.3474855.
7. H. Omrany, K. M. Al-Obaidi, M. Hossain, N. A. M. Alduais, H. S. Al-Duais, et al., "IoT-Enabled Smart Cities: A Hybrid Systematic Analysis of Key Research Areas, Challenges, and Recommendations for Future Direction," Discover Cities, vol. 1, no. 2, Art. no. 2, Mar. 2024.
8. H. Nguyen, D. Nawara, and R. Kashef, "Connecting the Indispensable Roles of IoT and Artificial Intelligence in Smart Cities: A Survey," Journal of Information and Intelligence, vol. 2, no. 3, pp. 261-285, May 2024.
9. K. A. Darabkh and M. Al-Akhras, "Towards Optimized IoT Sensor Networks for Smart Cities: Centrality-Aware Position-Based Occlusion-Driven and Role Dynamics Solutions for Clustering and Routing," IEEE Internet of Things Journal, early access, 2025, doi: 10.1109/JIOT.2025.3570612.
10. H. Nguyen, D. Nawara, and R. Kashef, "Connecting the Indispensable Roles of IoT and Artificial Intelligence in Smart Cities: A Survey," Journal of Information and Intelligence, vol. 2, no. 3, pp. 261-285, May 2024.