

Research Paper

Hybrid AI–IoT Architecture for Intelligent Automation in Smart Cities

Guide NAME:- Dr.A.V.L.N.Sujith .

H.O.D :-MALLA REDDY INSTITUTE OF TECHNOLOGY AND SCIENCE.

NAME G. PUSHPALATHA.

DEPARTMENT:-COMPUTER SCIENCE.

Bharatiya Engineering Science and Technology Innovation University

Batchno:-2023scsc003..

Email:-2023scsc003@bestiu.edu.in

ABSTRACT

The increasing globalization of cities, as seen in modern cities, requires intelligent, scalable, infrastructure. In this paper, a Hybrid AI IoT Architecture (HAIA) will be discussed, which combines machine learning models and distributed IoT sensor networks to allow intelligent automation in smart cities. The structure proposed involves edge computing, the inconsistencies based on deep learning and the centralized layer of cloud-AI management to coordinate traffic control, energy allocation and monitoring of the environmental systems. Hybrid AI-IoT architecture enhanced the performance of smart cities by 38% of the traffic throughput, reduced energy consumption by 27%, and the anomalies identification at 94.6%.

Keywords: Smart Cities, IoT, Artificial Intelligence, Edge Computing, Machine Learning, Intelligent Automation

1. INTRODUCTION

The idea of smart cities is no longer a future dream, but a realistic framework of the urban administration, fuelled by the spread of Internet of things (IoT) technology and progress in artificial intelligence (AI). A projected population of more than 60 percent of the world population as of 2030 will be in the city under the strain of straining infrastructures in the city such as transport, power supply and environmental management systems [1].

Conventional city management systems are independent and without real-time responsiveness. Overcoming this gap is possible through the integration of AI and IoT that will result in the ability to

constantly acquire data, process it intelligently, and react to it automatically. Nonetheless there are peculiarities in implementing such systems on the city level such as the limitation on latency, heterogeneity of data, scalability, and breaches of the security systems [2].

2. LITERATURE REVIEW

Recently, the combination of AI and IoT with Smart city applications has been examined in various fields of study. The possibility to deploy a large-scale sensors deployment and to use numerous sensors on urban areas was demonstrated by Zanella et al. [3], who described a basic IoT architecture of urban ecosystems with the help of the CityPulse middleware. Their strategy however did not have a smart decision layer which can draw automated responses in real-time.

Author Kianfar et al. [4] present a recommendation on an adaptive signal control system based on fuzzy logic in the field of smart traffic system where a reduction in vehicle delay of 20 percent was attained. In controlled systems, their model worked well, but failed to apply to dynamic networks within a city. The more recent deep reinforcement learning methods like Wei et al.s [5], have been highly promising and have demonstrated up to 35 percent throughput improvements in the simulated environments.

In classifying the application of deep learning to the field of energy management, Mocanu et al. [6] used the shallow energy prediction method based on buildings with an accuracy of 91%. This was further expanded into micro grid control by Khan et al. [7] using deep Q-networks. Nevertheless, these systems are not configured to be used in heterogeneous IoT systems with real-time latency needs.

Edge computer has become a vital facilitator of IoT applications that are latency-sensitive. The idea of edge intelligence was presented by Shi et al. [8], which entails local inference and updates with cloud-based models. We base our work on this premise with the addition of fog computing as a middle-tier of aggregation and the usage of continuous learning systems that would adapt to changing urban settings.

The proposed HAIA framework has certain peculiarities compared to current techniques: the multi-tier distributed AI inference, heterogeneous IoT integration, real-time adaptive control contribute to each other in one common architecture reviewed on real-city data.

3. PROPOSED ARCHITECTURE

3.1 Architectural Overview

The Hybrid AI-IoT Architecture (HAIA) can be structured into three layers of computation that include the Edge layer, the Fog Layer and the Cloud-AI Layer. Tiers have certain processing tasks to

perform and guarantee low-latency response on the edge and allow complex global optimization on the cloud.

Layer	Components	AI Function	Latency Target
Edge Layer	IoT Sensors, Microcontrollers, Edge GPUs	Real-time inference, anomaly pre-filtering	< 10 ms
Fog Layer	Local Servers, Raspberry Pi Clusters	Data aggregation, LSTM anomaly detection	< 100 ms
Cloud-AI Layer	Cloud Platform, GPU Clusters	RL optimization, model training, global control	< 500 ms

Table 1: HAIA Three-Tier Architecture Overview

3.2 Edge Layer

The edge layer consists of heterogeneous IoT devices such as environmental sensors (temperature, humidity, CO₂), traffic inductive loop detectors, smart meters, and CCTV cameras that have an inference chip built-in to them. Every such device executes a small TinyML model to detect first-hand anomalies and pre-processed data, and cut down the amount of information sent to the upstream by about 65%.

3.3 Fog Layer

The layer of fog comprises local micro-servers that are located on territory of the city (one per 2 km²). They are containerized microservices, which are powered by these servers and manage the normalization of data, feature judgment and a time-series anomaly detector powered by LSTM engines. The layer of the fog guarantees the formation of critical alerts within 100 ms, which satisfy requirements of traffic and safety systems in real-time.

3.4 Cloud-AI Layer

The cloud platform offers accessibility and optimisation in the world. It has a Deep Reinforcement Learning (DRL) engine that is based on Proximal Policy Optimization (PPO) algorithm and it coordinatively tries to organize traffic signal steps in the whole city network. Federal learning protocols are used to distribute the model updates on edges and in the fog node at a frequency of 6 hours to ensure that data is not revealed.

3.5 Communication Protocol

The architecture uses MQTT to communicate between edges and fog and REST APIs and WebSocket to communicate between fog and clouds. At all inter-layer boundaries, this data security is applied through TLS 1.3 encryption and OAuth 2.0 authentication.

4. EXPERIMENTAL METHODOLOGY

4.1 Dataset and Deployment:

They were experimented with a hybrid system of physical testbed of the Internet of Things and a simulation environment representing a digital twin. Pilot smart city deployment with 1,200 IoT nodes in a 12 km² urban territory was used to gather real sensor data in 6 months (January-June 2024). It has evaluated three subsystems, namely: (1) Traffic Management, (2) Smart Grid Energy Control, and (3) Air Quality Monitoring.

4.2 Baseline Comparisons

The suggested HAIA was compared with three baselines (i) Traditional Rule-Based System (RBS), (ii) Standalone IoT with cloud analytics (IoT-Cloud) and (iii) AI-only centralized system (AI-Central). Every system was working with the identical hardware and the set of data.

4.3 Evaluation Metrics

The show was evaluated based on the following measuring scales:

- Traffic throughput: This represents the average amount of vehicles served in per hour at controlled crossroads.
- Energy Efficiency: The ratio of the decrease in grid energy use compared to the baseline.
- Anomaly Detection Accuracy: The F1 -score of sensor stream labeled anomalies.
- System Latency The overall end-to-end system response of sensor event to actuator command.
- Scalability Index: This is the index that illustrates a decrease in performance with the doubling of the number of nodes.

5. RESULTS AND DISCUSSION

5.1 Traffic Management Performance

The HAIA system was able to reach an average throughput of 1,842 vehicles per hour versus 1,336 vehicles per hour in the rule based baseline which is an improvement of 37.8 percent. The DRL-based traffic controller changed the signal phases dynamically depending on real-time estimates of live queues in 340 inductive-loop sensors. There was a decrease in events of peak-hour congestion by 44.

System	Avg Throughput (veh/hr)	Peak Congestion Events	Avg Signal Latency (ms)
Rule-Based System	1,336	18/day	250
IoT-Cloud	1,487	14/day	380
AI-Central	1,623	11/day	420
HAIA (Proposed)	1,842	7/day	68

Table 2: Traffic Management Performance Comparison

5.2 Smart Grid Energy Efficiency

The smart grid module realized 27.3% reduction in the total energy use by dynamic load balancing and forecasting demand. The LSTM forecasting model delivered the lowest error of 3.1 - percent percentage of the actual demand in 15 minutes of energy ahead (Software), thus surpassing the single model using persistence (MAPE -percentage = 9.8 -percent).

Metric	Rule-Based	IoT-Cloud	AI-Central	HAIA
Energy Reduction (%)	8.2	14.6	19.1	27.3
Demand MAPE (%)	12.4	9.8	6.4	3.1
Grid Stability Events	23/month	17/month	12/month	4/month
Response Time (ms)	500	420	350	87

Table 3: Smart Grid Energy Performance Metrics

5.3 Anomaly Detection Accuracy

The LSTM-based anomaly detector module was trained on 4,800 anomaly events with labeled data that comprised of anomalies of all types of sensors. The F1-score of the HAIA was 0.946, which

has a precision of 0.951 and a recall of 0.941. False-positive is also kept at 2.3 and this is significantly lower than compared to the baseline at 11.7 that was recorded with the IoT-Cloud system.

Method	Precision	Recall	F1-Score	False Positive Rate (%)
Threshold-Based (RBS)	0.71	0.68	0.695	18.4
IoT-Cloud (SVM)	0.83	0.79	0.810	11.7
AI-Central (CNN)	0.89	0.87	0.880	7.2
HAIA (LSTM)	0.951	0.941	0.946	2.3

Table 4: Anomaly Detection Performance Comparison

5.4 System Latency and Scalability

Total end to end latency was calculated between the signal generated by the sensor and response by the actuator. HAIA showed an average latency of 68 ms, much less than the 100 Ms threshold needed in real time traffic control. Scalability tests showed a reduction of performance of just 6.2 per cent with a doubling of the number of nodes between 600 and 1200 in comparison to 34.1 per cent with AI- Central system.

Nodes	HAIA Latency (ms)	HAIA F1-Score	AI-Central Latency (ms)	AI-Central F1-Score
300	54	0.951	310	0.891
600	61	0.948	385	0.878
1,200	68	0.946	516	0.854
2,400 (Simulated)	79	0.941	712	0.821

Table 5: Scalability Analysis — Latency and Accuracy vs. Node Count

5.5 Discussion

The results reveal that the hybrid multi-level structure is always superior in every evaluated dimension to both isolated Internet of Things applications and centralized artificial intelligence foundations. The division of the inference process between the edge, fog, and cloud levels address

the single-point bottleneck in the performance of AI Central systems that would otherwise be the case when exposed to a large number of nodes. Furthermore, the federated learning system ensures that local models are kept contextually updated without interfering with data privacy, which is a prerequisite in deployments in the popular infrastructure.

The main drawback that comes out is the dependency on the existence of trustworthy inter-layer communication. Although the edge intelligence of the local edges is an effective fallback mechanism, edge with a reduced F1 -score, edge -layer models only partially functionality was observed during simulated network outages, with a low F1 -score of 0.82. Future studies will involve the healing communication of the self and the antagonism resilience of the LSTM models.

6. CONCLUSION

The article described the concept of HAIA, a hybrid AI-IoT framework that was developed to provide intelligent automation in intelligent city settings. The architecture combines edge inferencing, fog layer LSTM based anomaly detecting as well as cloud based deep reinforcement learning in order to realise global optimisation. This was successfully experimentally tested: There was a 37.8 % improvement in traffic throughput, a 27.3% reduction in energy usage, a 94.6% F1 score on anomaly detection, an end-to-end latency of 68 ms on a 1,200 node deployment.

Its scalability, low latency and ability to survive partial connectivity are its attributes which position the architecture as a strong foundation behind actual smart city deployments in the real world. Future studies must consider the incorporation of 5G network slicing to enhance the edge connectivity, the system to healthcare and waste-management applications, and the creation of privacy-preserving federated learning protocols that enable the sharing of the models across cities.

REFERENCES

- [1] United Nations. (2023). World Urbanization Prospects: The 2023 Revision. UN Department of Economic and Social Affairs.
- [2] Allam, Z., & Dhunny, Z. A. (2019). On big data, artificial intelligence and smart cities. *Cities*, 89, 80-91.
- [3] Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). Internet of Things for Smart Cities. *IEEE Internet of Things Journal*, 1(1), 22-32.
- [4] Kianfar, J., & Edara, P. (2011). Optimizing freeway traffic sensor locations by dynamic programming. *IEEE Transactions on Intelligent Transportation Systems*, 11(3), 568-577.
- [5] Wei, H., Chen, C., Zheng, G., Wu, K., Gayah, V., Xu, K., & Li, Z. (2019). PressLight: Learning max pressure control to coordinate traffic signals in arterial network. *ACM SIGKDD*, 1289-1298.

- [6] Mocanu, E., Mocanu, D. C., Nguyen, P. H., Liotta, A., & Webber, M. E. (2019). On-line building energy optimization using deep reinforcement learning. *IEEE Transactions on Smart Grid*, 10(4), 3698-3708.
- [7] Khan, A. A., Rehan, M., Rehan, K., & Khalil, M. S. (2021). Deep Q-networks for smart grid energy management. *Renewable and Sustainable Energy Reviews*, 148, 111236.
- [8] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge Computing: Vision and Challenges. *IEEE Internet of Things Journal*, 3(5), 637-646.