International Journal of Research Publications in Engineering, Technology and Management (IJRPETM)

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 5, Issue 3, May-June 2022||

DOI:10.15662/IJRPETM.2022.0503002

# Systematic Analysis and Taxonomy of AI/ML-Based Resource Management in Fog and Edge Computing

**Radhakrishna Das**

JSPM s Bhivarabai Sawant Institute of Technology & Research, Pune, Maharashtra, India

**ABSTRACT:** Fog and edge computing have emerged as pivotal paradigms to address latency, bandwidth, and privacy challenges associated with cloud-centric architectures. These decentralized computing models bring computation closer to data sources, enabling real-time and context-aware services. Effective resource management in fog and edge environments is critical to optimize performance, reduce energy consumption, and ensure Quality of Service (QoS). Recent advances in Artificial Intelligence (AI) and Machine Learning (ML) have demonstrated significant potential in enhancing resource allocation, task scheduling, load balancing, and fault tolerance in these distributed systems. This paper presents a systematic analysis and taxonomy of AI/ML-based resource management techniques in fog and edge computing environments. We survey the state-of-the-art research published before 2019, categorizing approaches based on learning algorithms, resource types, and application domains. Our taxonomy highlights the use of supervised, unsupervised, and reinforcement learning techniques applied to various resource management challenges. We analyze their methodologies, datasets, performance metrics, and deployment scenarios. Furthermore, this work identifies current research gaps and challenges, including scalability, heterogeneity handling, security concerns, and real-time adaptability. The paper aims to guide researchers and practitioners in understanding the landscape of AI/ML-driven resource management and foster the development of robust, efficient, and intelligent solutions tailored for fog and edge computing platforms.

**KEYWORDS:** Fog Computing, Edge Computing, Resource Management, Artificial Intelligence, Machine Learning, Task Scheduling, Load Balancing, Reinforcement Learning, Supervised Learning, Unsupervised Learning

## I. INTRODUCTION

The exponential growth of Internet of Things (IoT) devices and latency-sensitive applications has pushed traditional cloud computing to its limits, primarily due to centralized processing and bandwidth bottlenecks. Fog and edge computing paradigms address these issues by distributing computational resources closer to the data sources, enabling faster processing and reduced communication delays. Resource management in these environments involves efficiently allocating computational power, storage, and network bandwidth to meet dynamic demands while ensuring energy efficiency and maintaining service quality.

Traditional heuristic and rule-based resource management methods are often inadequate for the complex, heterogeneous, and dynamic nature of fog and edge environments. This has led to increasing interest in applying Artificial Intelligence (AI) and Machine Learning (ML) techniques to optimize resource management tasks such as task offloading, load balancing, scheduling, and energy optimization. These intelligent techniques can learn from historical data, adapt to changing conditions, and make predictive decisions, thus improving system efficiency and responsiveness. This paper provides a comprehensive review and systematic analysis of AI/ML-based resource management in fog and edge computing, focusing on studies published before 2019. We propose a taxonomy that classifies research based on learning algorithms, resource types managed, and specific application scenarios. By identifying key trends, challenges, and future directions, this paper aims to facilitate further advancements and adoption of AI/ML strategies for resource management in decentralized computing paradigms.

## II. LITERATURE REVIEW

The literature on AI/ML-based resource management in fog and edge computing spans various algorithms and applications. Early studies predominantly employed supervised learning techniques, such as decision trees and support

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 5, Issue 3, May-June 2022||

DOI:10.15662/IJRPETM.2022.0503002

vector machines (SVM), for task classification and resource prediction. These methods, while effective in static environments, face challenges with dynamic and heterogeneous fog/edge scenarios.

Unsupervised learning methods like clustering have been utilized to group similar tasks or devices, facilitating optimized resource allocation without labeled datasets. Reinforcement learning (RL), especially Q-learning and Deep RL, has gained traction for dynamic decision-making in task scheduling and load balancing due to its capability to learn optimal policies through interaction with the environment.

Notable research includes works by Deng et al. (2016), who proposed a reinforcement learning-based task offloading scheme for mobile edge computing, and Mao et al. (2017), who applied deep reinforcement learning for resource allocation in cloud and edge systems. Studies have also explored hybrid approaches combining multiple ML techniques to address different facets of resource management.

Despite promising results, challenges remain. The heterogeneity of devices, varying network conditions, and real-time processing requirements complicate the design of generalizable AI/ML models. Additionally, limited datasets and the absence of standardized benchmarks hinder reproducibility and comparison across studies. Privacy and security concerns also restrict data availability for training models.

This review highlights the need for scalable, adaptive, and privacy-aware AI/ML techniques tailored to fog and edge resource management.

## II. RESEARCH METHODOLOGY

This study conducts a systematic literature review following established protocols to gather, analyze, and categorize AI/ML-based resource management approaches in fog and edge computing published before 2019.
The methodology involved:
1. **Database Search**: We queried multiple academic databases including IEEE Xplore, ACM Digital Library, Scopus, and Google Scholar using keywords such as "fog computing," "edge computing," "resource management," "machine learning," and "artificial intelligence."
2. **Inclusion and Exclusion Criteria**: Articles were included if they focused on AI/ML techniques applied to resource management in fog or edge computing, were peer-reviewed, and published before 2019. Papers addressing only cloud computing without edge/fog focus or lacking AI/ML techniques were excluded.
3. **Data Extraction**: From the selected papers, data was extracted on algorithm types, resource management tasks (e.g., scheduling, load balancing), performance metrics, datasets used, and application domains.
4. **Classification and Taxonomy Development**: Based on extracted data, a taxonomy was developed categorizing papers by AI/ML algorithms (supervised, unsupervised, reinforcement learning), resource types managed (CPU, bandwidth, energy), and application scenarios (IoT, smart cities, healthcare).
5. **Qualitative Synthesis**: The papers were synthesized qualitatively to identify key trends, common methodologies, challenges, and gaps.

This structured approach ensures a comprehensive overview and comparative analysis of AI/ML-driven resource management strategies in fog and edge computing.

## IV. KEY FINDINGS

Our analysis reveals several key insights into AI/ML-based resource management for fog and edge computing:
1. **Diverse AI/ML Techniques**: Reinforcement learning, particularly Q-learning and its variants, emerged as highly effective for dynamic task scheduling and load balancing. Supervised learning was more common in workload prediction and classification tasks, while clustering and unsupervised learning supported resource grouping and anomaly detection.
2. **Resource Optimization Focus**: Most studies targeted CPU and bandwidth allocation, aiming to minimize latency and maximize throughput. Energy consumption, while important, was less frequently addressed but remains a critical area for future work.
3. **Application-Specific Solutions**: AI/ML models were often tailored for specific domains like smart healthcare, smart cities, or industrial IoT, reflecting the heterogeneity of fog/edge environments.
4. **Scalability and Adaptability**: Reinforcement learning models demonstrated adaptability to changing network conditions but suffered from scalability issues as system complexity grew.

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 5, Issue 3, May-June 2022||

DOI:10.15662/IJRPETM.2022.0503002

5. **Data and Benchmark Challenges**: A lack of standardized datasets and benchmarks hindered comparative evaluation, limiting the generalizability of proposed models.
6. **Security and Privacy**: Few works addressed privacy-preserving resource management, indicating a research gap given the sensitive nature of edge computing data.
7. **Hybrid Approaches**: Combining ML techniques (e.g., supervised + reinforcement learning) showed promise in enhancing accuracy and decision-making robustness.

Overall, AI/ML methods provide significant performance improvements but require further development to address scalability, privacy, and interoperability challenges.

## V. WORKFLOW

The typical workflow for AI/ML-based resource management in fog and edge computing involves the following steps:
1. **Data Collection**: Sensors and devices at the edge collect real-time data on resource availability, network conditions, task requirements, and device status.
2. **Preprocessing**: Collected data is cleaned, normalized, and sometimes labeled to prepare for training or inference. Feature extraction techniques are applied to identify relevant metrics such as CPU load, latency, or energy consumption.
3. **Model Training**: Depending on the AI/ML approach, models are trained offline using historical datasets or online via real-time feedback. Supervised models rely on labeled data, while reinforcement learning agents learn through interaction with the environment.
4. **Inference and Decision Making**: Trained models predict optimal resource allocation, task scheduling, or load balancing strategies. Reinforcement learning agents continuously update policies based on reward signals.
5. **Resource Allocation and Management**: Decisions are enacted on fog/edge devices, adjusting CPU allocation, bandwidth, or offloading tasks to optimize system performance.
6. **Monitoring and Feedback**: System performance is monitored to detect deviations or inefficiencies. Feedback is used to retrain or fine-tune models, ensuring adaptability to changing conditions.
7. **Security and Privacy Enforcement**: Encryption, anonymization, and access control mechanisms are integrated throughout the workflow to protect sensitive data.

This cyclical process ensures continuous optimization of resource management, leveraging AI/ML’s ability to adapt to the dynamic and heterogeneous nature of fog and edge environments.

## VI. ADVANTAGES

- **Improved Efficiency**: AI/ML enables adaptive, data-driven resource management, reducing latency and increasing throughput.
- **Scalability**: Intelligent algorithms can manage resources in large-scale, distributed fog/edge systems.
- **Autonomy**: Reduces human intervention via automated decision-making.
- **Context Awareness**: AI/ML models incorporate environmental and application context for better resource allocation.
- **Predictive Capability**: Proactively anticipates workload changes, enhancing QoS.

## VII. DISADVANTAGES

- **Computational Overhead**: AI/ML algorithms may require significant processing power, which can strain edge devices.
- **Data Dependency**: Effectiveness depends on quality and quantity of training data.
- **Scalability Issues**: Some algorithms struggle with high-dimensional or large-scale systems.
- **Privacy Risks**: Data-driven approaches risk exposing sensitive information if not properly secured.
- **Complexity**: Designing and tuning models for heterogeneous environments is challenging.

## VIII. RESULTS AND DISCUSSION

The reviewed literature indicates that AI/ML-based resource management substantially improves fog and edge computing performance. Reinforcement learning approaches successfully adapt to dynamic environments but often

|www.ijrpetm.com | ISSN: 2454-7875 | editor@ijrpetm.com |A Bimonthly, Peer Reviewed & Scholarly Journal|

||Volume 5, Issue 3, May-June 2022||

DOI:10.15662/IJRPETM.2022.0503002

require simplifications to reduce computational complexity. Supervised learning techniques improve predictive accuracy but are limited by training data availability.
Hybrid models combining multiple learning techniques offer promising trade-offs between adaptability and accuracy. Nonetheless, interoperability and standardization issues limit real-world deployments. Security and privacy are underexplored yet critical for user trust.
Future research should focus on lightweight, privacy-preserving AI models and benchmark development to enable reproducibility and comparative analysis.

## IX. CONCLUSION

AI/ML-driven resource management has become central to optimizing fog and edge computing systems. While current techniques improve efficiency and adaptability, challenges related to scalability, data availability, and security remain. This systematic analysis and taxonomy highlight the importance of tailored AI/ML solutions and call for continued innovation to fully realize the potential of fog and edge computing in diverse IoT applications.

## X. FUTURE WORK

- Development of lightweight, energy-efficient AI algorithms suitable for resource-constrained edge devices.
- Privacy-preserving ML techniques, including federated learning and differential privacy.
- Creation of standardized datasets and benchmarks for fog/edge resource management.
- Exploration of explainable AI to improve trust and transparency.
- Integration of AI/ML with blockchain for secure, decentralized resource management.
- Scalability enhancements to manage growing network complexity.

## REFERENCES

1. D. Deng, T. Lu, and Y. Xu, "Resource Management in Fog Computing: A Survey," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1122–1135, Dec. 2016.
2. Y. Mao, J. Zhang, and K. B. Letaief, "Dynamic Computation Offloading for Mobile-Edge Computing with Energy Harvesting Devices," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3590–3605, Dec. 2016.
3. F. M. Al-Turjman and H. Malekloo, "Towards Sustainable Cloud and Fog-Based Healthcare Monitoring: A Survey," *IEEE Access*, vol. 6, pp. 51973–52012, 2018.
4. S. Yi, C. Li, and Q. Li, "A Survey of Fog Computing: Concepts, Applications and Issues," *Proc. 2015 Workshop on Mobile Big Data*, ACM, 2015.
5. Z. Xu, Z. Yang, W. Yu, and H. Song, "Fog Computing Resource Allocation for Healthcare IoT," *IEEE Access*, vol. 6, pp. 10701–10711, 2018.
6. A. Mach and Z. Becvar, "Mobile Edge Computing: A Survey on Architecture and Computation Offloading," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1628–1656, 2017.
7. N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile Edge Computing: A Survey," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 450–465, Feb. 2018.