Edge and Fog Computing for Autonomous Network Management: A Comprehensive Survey

Pasumpon Pandian A.

Professor, Department of Computer Science and Engineering, CARE College of Engineering, Trichy, India.

E-mail: pasumponp27@gmail.com

Abstract

Fog computing is becoming popular as a way to reduce the bottlenecks that occur in computing and networking when deploying large numbers of IoT devices. Edge computing utilizes autonomous network management by performing local data processing (at the edge or source) which reduces both latency and network congestion. This also enables fog computing to provide a significant addition to cloud computing (where compute, network, store, and accelerate resources are located at the edge and network tiers) through the use of distributed, cooperative, and multi-tiered deployment of these elements. The purpose of this survey is to provide one of the first comprehensive reviews of edge and fog computing technologies for supporting autonomous network management. It discusses the role that Artificial Intelligence and Machine Learning play within the context of closed-loop control systems, intent-based networking, and zero-touch operations across the fog/edge layers of the network. The application areas considered in this report include 5G and future generations of mobile networks, the IoT, software-defined networking, and network function virtualization. For each of these application areas, specific challenges related to distributed resource management, scalability, interoperability, security, privacy, and reliability are also identified and discussed in detail.

Keywords: Fog Computing, Edge computing, Network management, Internet of Things, Artificial Intelligence and Machine Learning, Mobile networks, Cloud Computing.

DOI: https://doi.org/10.36548/jismac.2026.1.005
Received: 21.01.2026, received in revised form: 18.02.2026, accepted: 05.03.2026, published: 16.03.2026

1. Introduction

The number of IoT devices, mobile users and latency-sensitive applications has created a level of complexity in how communications networks function. The requirement for low latency, real-time responsiveness, scalability, and reliability has made it increasingly difficult to use traditional cloud-centric computing models and managed networks to support these requirements. The creation of massive amounts of data at the edge of the network in significantly contributes to the issue of network congestion and increased latency because most of this data is processed centrally, adding to the number of people and devices attempting to access the data. The need for time-critical services within these limitations has prompted a movement toward more decentralized computing model.

Edge computing and fog computing are fundamental in overcoming these issues by bringing computation, storage, and control capabilities closer to the data sources and/or the end user. Edge computing allows data generated locally to be processed at or near the source of the data, thus facilitating faster decision-making and reducing the amount of backhaul traffic. Fog computing supports this paradigm by introducing an additional intermediate layer (or multi-layer), between the edge and the cloud to support distributed and cooperative management of resources across the globally distributed network. Together, these two paradigms provide flexibility and scalability for future generation services.

The primary function of cloud computing is to deliver on-demand computing resources independent of where users are located (typically far away). With the expansion of cloud computing, numerous new facets have developed regarding its use with different types of computing models. In particular, multi-tier cloud computing, edge cloud computing; mobile edge cloud computing and fog computing are all emerging complementary technologies that improve how computations utilize available resources and satisfy the needs of applications [1].

Networks are growing rapidly in size, and the way we use them is changing almost daily. This means we need new ways to manage and control these networks that require less human input—autonomous management. Autonomous network management makes use of Artificial Intelligence (AI) and machine learning technologies to perform tasks like self-configuration, self-optimization, self-healing, and self-protecting using closed-loop processes. New concepts such as Intent-Based Networking (IBN) and Zero Touch Operations (ZTO) have become necessary to effectively manage heterogeneous multi-distributed Edge/Fog Networks.

Understanding how edge/fog computing can support autonomous network management across various application areas is still under consideration. This survey explains a systematic review of architectural styles and frameworks that combine the use of edge/fog computing technologies with AI-based network management. Key areas of application such as 5G / future mobile networks, IoT devices and services, software defined networking concepts, and network function virtualization will be evaluated, highlighting the challenges that exist relative to the areas of resource management, scalability, security, privacy, interoperability, and reliability.

2. Literature Review

Multiple aspects need to be considered while practically implementing fog computing, including the design of the system, the design of applications, the implementation of software, security, the management of computing resources, and networking. The available literature addresses all these aspects in depth. The interconnectedness between these systems means that an integrated system approach must provide the identification of the functional elements and their associated interfaces across the various layers within the system The increased maturity and widespread use of virtualization technology is opening the door for multi-tenancy to occur not just in high-end computing servers but also through network equipment and ultimately user equipment. The shared network and user functions will increasingly manifest as virtual functions being executed externally to traditional application stores in utility-driven cloud-computing environments. The expansion of cloud computing globally, as a result of the emergence of universal composability (i.e. the ability to convert a traditional application into micro-service-based middleware) will only exacerbate this evolution. For micro-services that are aptly coupled with their tasks, each may require differing types and amounts of computing power, the ability to scale under load, how interactive they are, and their maximum allowable latency [1].

Fog computing is a new way to solve these problems. It allows for resource and service allocation outside of the cloud, at the end of the network (the edge of the network) where the devices are located or at a location determined by SLAs (service level agreements). While fog computing provides a method to perform processing at the edge, it also allows for interaction with the cloud. This paper provides a complete review of fog computing. We analyze the current state of fog computing against a limited number of evaluation criteria. In addition, we include information about the architectures and algorithms used in fog computing systems, as well as examples of current challenges and future research directions to address them. Finally, we also provide a review of lessons learned and discuss the potential for fog computing to play an important role in new technologies such as the tactile internet [2].

Resource Management (RM) is characterized by the heterogeneous nature of resources, the transactional workload being implemented by the user in real-time, the need to discover edge nodes, and there are multiple Quality of Service (QoS) parameters (e.g., throughput, latency, storage, etc.) authorised by service providers simultaneously, making it even more difficult than it was before. In response, researchers have begun using AI-based techniques to address these resource-related challenges. This paper reviews the aforementioned issues and challenges in terms of resource management within the fog and edge computing paradigm by providing an overview of RM-related issues and challenges and categorizing them based upon various dimensions such as computing resource provisioning, task off-loading, resource scheduling, service placement, and load balancing. In addition, the paper discusses examples of existing AI- and non-AI (traditional methods) based RM-related solutions and includes a description of various QoS metrics used, the datasets evaluated, and the advantages and disadvantages of each solution [3].

We found that the technical challenges associated with the management of scarce resources in fog and edge computing have largely been resolved through existing research; however, additional challenges still remain and will require further exploration. There is considerable use of resource-constrained devices such as wireless access points and television set-top boxes as part of fog and edge computing, which makes them inappropriate candidates for heavyweight data processing solutions like Apache Spark and deep learning libraries. The use of lighter-weight alternatives to these heavyweight methodologies (like Apache Quarks), which could potentially be used for data processing on constrained edge devices, is available, although they currently don't have enough advanced capabilities to perform advanced analytics on their own [4].

The review was based on publications between 2019 and 2024 that investigated AI-based studies of resource management, task scheduling, and load balancing techniques in the context of FC. The diversity of techniques used to improve performance in FC is explored through ML and DL as they are used for resource allocation; heuristic algorithms for task scheduling; and nature-inspired meta-heuristics for load balancing. Strengths and weaknesses of the various techniques are examined in terms of their influence on latency, energy usage, and Quality of Service (QoS). Overall, there has been considerable advancement in FC optimization through new techniques. ML and nature-inspired meta-heuristics are very promising techniques for providing resource management, task scheduling, and load balancing solutions, respectively [5].

A discussion of the characteristics and limitations of individual technologies, including but not limited to latency, bandwidth usage, security, and privacy of data, is presented. The synergy among IoT technology and cloud computing technologies is also discussed. Cloud computing is considered a backend technology that can be used to process many data streams generated by IoT devices. The review shows that there are problems with handling unreliable data and privacy protection. There is a clear need for solid security measures and regulatory guidelines for addressing these issues. In addition, the role of edge computing and IoT technologies is described, showing how edge nodes use the excess processing capabilities of IoT devices to deliver additional services. Finally, the heterogeneity of edge computing systems presents challenges for researchers to overcome and identifies computational offloading as an approach to mitigating latency in mobile edge computing [6].

The difficulties of managing networked resources due to a rapidly growing number of IoT devices create a need for innovative methods of managing networked resources, such as those based on the concept of edge computing. Edge computing removes the limitations that exist when processing the data generated by resource-constrained devices such as IoT devices and allows for offloading the computationally intensive workloads at the edge, which, in turn, reduces the amount of traffic on the core network and addresses the primary issues of privacy. The availability of edge computing offers a variety of benefits to battery-powered devices and low-latency applications [7]. As the number of IoT devices increases dramatically and disproportionately, the demand for innovative network management solutions is also increasing. Many IoT devices are not able to handle complex routing rules or can only forward a limited amount of customization due to memory limitations. Traditional networking technology is unable to effectively meet the needs of IoT devices because the constraints imposed by the way the devices operate make them challenging to use and require a reliable way to program routing and custom forwarding requirements. Traditional network management paradigms have difficulty with scalability and modularity, whereas SDN provides centralized IoT management, the ability to virtualize resources, improved innovation, and programmability.

The purpose of this survey is to provide a more detailed examination of resource management in Edge-Fog-Cloud Systems. The objective is to look at how combinatorial optimization (CO) and machine learning (ML) methods work together, where they perform best independently, and how they can be combined to address both static and dynamic issues in resource management. Additionally, this survey will take into consideration some of the evolving characteristics of contemporary systems such as changing workloads, evolving user activity, and real-time adjustments to those activities that influence the effectiveness of resource assignment and scheduling. Through a comprehensive and integrated approach, this survey will lead to a better understanding of the need for more flexible and effective resource management solutions in modern computing systems [8].

Networked computing relies heavily on security and privacy to thrive. When integrating end, edge, and cloud computing, many existing challenges also apply to each new paradigm; however, device heterogeneity, mobile support, location-awareness, and low latency services introduce new issues that must be solved. There are two primary avenues of ongoing research into privacy and safety solutions for emerging computer modelling. The first is securing and assuring the safety and privacy of the underlying computer system itself. The second is providing security and guaranteeing the confidentiality of the services delivered across this computer framework (service delivery, data processing, data transmission, data storage).[9]

There has been a dramatic increase in the number of compute-intensive applications being developed in smart cities over recent years. These applications continuously create vast quantities of data that must be processed with very low latency or have time constraints attached to their completion [10]. Edge computing offers a potential solution to address many of the latency issues that challenge developers of smart city applications but comes with its own set of complications when used. This survey focuses on how edge computing enables the achievement of the smart city vision by first detailing the evolution of edge computing paradigms and then conducting a thorough review of available literature to evaluate the different solutions available through edge computing for use in developing smart city applications.

Fog computing is a model that allows computing resources and services to be located at the edges of networks, close to users' end devices, thereby reducing latency and enabling connection to cloud computing resources. The distinction between cloud and fog computing lies in fog resources are comprised of constrained, heterogeneous nodes, which may not always maintain connectivity with one another [11]. In this environment of complexity, there is a need for defining and implementing orchestration procedures to ensure that applications and services can be delivered according to established agreements. While there have been publications regarding the orchestration of fog computing resources, there continues to be a lack of consensus in the definitions of orchestration and its intersection with other fields, including resource management and monitoring.

In recent years, the Internet of Things (IoT) has changed how we interact with physical objects in many different scenarios, but there are still challenges related to computing resources and energy use and both IoT and AIs based applications experience similar issues. The cloud offers solutions for these issues; however, it introduces new challenges, such as increased costs to communicate and security concerns related to the physical proximity of the endpoint devices and the cloud. Fog computing, which enables more computational resources that are closer to the source of the data at the edge of the network, presents a new solution to these problems; however, it has its own challenges, including many different types of nodes and limitations on available resources [12]. This survey provides an extensive review of recent literature regarding scheduling in fog computing and compares various scheduling methods based on critical characteristics of fog computing (including real-world, performance metrics and simulation tools used to evaluate scheduling methods), and identifies the advantages and disadvantages of each tool.

Although Internet of Things (IoT) technology has become a necessary means of delivering data for electronic services, it typically does not provide sufficient resources for hosting application services directly. Fog Computing (FC) is a viable means of expanding the capabilities of IoT devices via their interaction with cloud systems because it is capable of integrating and working with both centralized cloud computing and distributed cloud computing at the network edge. As a result of these two relationships, computing, storage, and networking resources and services can be distributed along the continuum from Cloud systems to Things. Therefore, the benefits of cloud computing (CC) are brought even closer to the end-user devices through the fog computing paradigm. A review paper on FC's use in supporting

IoT devices and services is presented in this article. Principles and literature defining FC are discussed, and six distinct application domains of IoT are explored to provide examples of how to use FC. Additionally, FC's expansion into the network edge raises additional challenges while influencing existing methodologies traditionally utilized in cloud-based systems [13].

Table 1. Comparative Analysis of Previous Studies on Edge–Fog Computing for Autonomous Network Management

Ref.	Authors / Year	Focus Area	Method / Approach	Key Contributions	Limitations
[1]	Habibi et al., 2020	Fog Computing Architecture	Architectural survey of fog computing frameworks	Provides comprehensive architecture models and discusses scalability and resource distribution in fog environments	Limited discussion on AI-driven autonomous network management
[2]	Mouradian et al., 2017	Fog Computing Overview	Evaluation of fog computing architectures and algorithms	Identifies research challenges and future directions in fog computing systems	Lacks detailed implementation strategies for real-time applications
[3]	Walia et al., 2023	Resource Management in Fog/Edge	AI-based techniques for resource allocation and task scheduling	Reviews QoS parameters, datasets, and AI-based solutions for resource management	Complexity of AI models and computational overhead in resource-constrained devices
[4]	Hong & Varghese, 2019	Fog/Edge Resource Management	Survey on architectures, infrastructure, and algorithms	Discus the resource discovery, service placement, and load balancing strategies	Lightweight analytics solutions for edge devices still limited
[5]	Alsadie, 2024	AI Techniques in Fog Computing	Machine Learning and Deep Learning	Demonstrates effectiveness of ML, DL, and meta-heuristics for task	High computational requirements

			methods for optimization	scheduling and load balancing	and training complexity
[6]	Kuchuk & Malokhvii, 2024	IoT–Cloud–Fog Integration	Review of integration frameworks	Explains how edge nodes utilize IoT processing capabilities and reduce latency	Data privacy and reliability concerns remain unresolved
[7]	Rafique et al., 2020	IoT with SDN and Edge Computing	SDN-based programmable networking framework	Enables scalable IoT management and improves network programmability	Traditional IoT devices still limited by hardware constraints
[8]	Boubaker et al., 2025	Resource Management in Edge–Fog–Cloud	Combination of Machine Learning and optimization methods	Proposes adaptive resource allocation strategies for dynamic workloads	Real-time deployment challenges and computational complexity
[9]	Ren et al., 2019	End–Edge–Cloud Computing Paradigm	Comparative analysis of computing paradigms	Discusses orchestration and coordination across cloud, edge, and fog layers	Lack of standardized orchestration frameworks
[10]	Khan et al., 2020	Edge Computing for Smart Cities	Survey of edge-enabled smart city applications	Shows how edge computing supports low-latency services in smart city infrastructure	Deployment complexity and data management issues
[11]	Costa et al., 2022	Fog Orchestration	Survey on orchestration mechanisms	Highlights the importance of resource orchestration in distributed fog environments	No consensus on standardized orchestration definitions
[12]	Chuan et al., 2025	Scheduling in Fog Computing	Comparative study of scheduling algorithms	Evaluates performance metrics and simulation tools for fog scheduling	Scalability and heterogeneous device management challenges

[13]	Puliafito et al., 2019	Fog Computing for IoT	Survey of fog-based IoT application domains	Explains how fog computing extends cloud capabilities closer to IoT devices	Limited focus on autonomous network management

The summary table 1 compares several major research studies about the architecture of edge and fog computing and how they will help to manage autonomous networks.

Although several surveys have examined fog computing, edge computing, and resource management individually, important gaps remain in the literature. First, most existing surveys focus primarily on infrastructure architecture or resource management techniques, without explicitly analyzing how these technologies enable autonomous network management frameworks. Second, many studies review edge computing or fog computing independently, while limited work investigates the integrated edge–fog–cloud continuum required for distributed intelligence in next-generation networks. Third, the role of AI-driven network automation mechanisms, such as intent-based networking and zero-touch operations, has not been sufficiently analyzed in the context of distributed edge–fog environments. Finally, previous studies often evaluate optimization algorithms or scheduling methods but do not provide a comprehensive synthesis linking architectural models, AI-based management mechanisms, and application domains such as 5G, IoT, and software-defined networking.

3. Architecture and System Models

This section describes the architectural framework and system models used to analyze edge–fog–cloud computing environments for autonomous network management. The architecture consists of distributed computing layers that collaboratively process data generated by IoT devices while enabling low-latency network control and resource management.

The system model considers three primary layers:

- Edge layer
- Fog layer
- Cloud layer

Each layer provides different computational capabilities and management functions, enabling efficient distribution of workloads and network intelligence across the infrastructure.

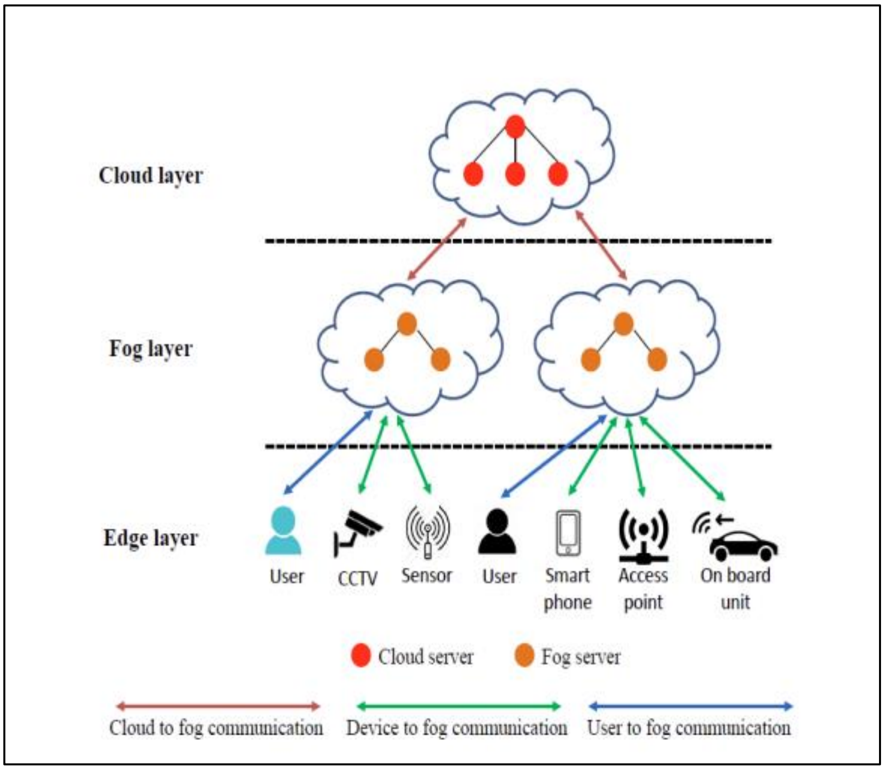

Figure 1. Three-tier Cloud–Fog–Edge Computing Architecture [14]

The bottom level of the diagram shows the edge level of the architecture, which consists of both end-use devices (e.g., users, CCTV cameras, sensors, smartphones, etc..) and data generating devices (e.g., access points, onboard units in vehicles, etc..). These devices generate high volumes of data and have a need for immediate processing or a fast reaction time. If these devices send all their data directly to the cloud, latency and network congestion will increase dramatically.

The second level in the architecture, the fog level, serves as an intermediary between the edge and cloud. This level is comprised of various fog servers processing and forwarding data closer in proximity to the devices than the cloud. The fog layer aggregates data from multiple nearby edge devices; executes computed, filtered or contextually / real-time aware services, and processes the data for each of these edge devices prior to sending the processed data to the cloud. The fog level reduces the total communication load on the cloud, thereby enabling faster reaction times for low-latency applications.

At the top of the diagram, the cloud layer consists of distributed servers that offer high capacity for processing and storage. In this layer, globally, it is responsible for long-term data storage, analyzing global data, training large data sets, and coordinating the activities of large-scale networks. The cloud communicates with the fog layer to communicate with the largest number of edge devices directly.

Arrows are used to show how the different types of communication occur:

- Cloud to fog communication for global control and coordination,
- Fog to edge communication for local data processing, and
- User to fog communication for low-latency services.

Overall, the diagram demonstrates how fog computing serves as a bridge from the edge to the cloud and provides improved scalability, latency, and efficiency.

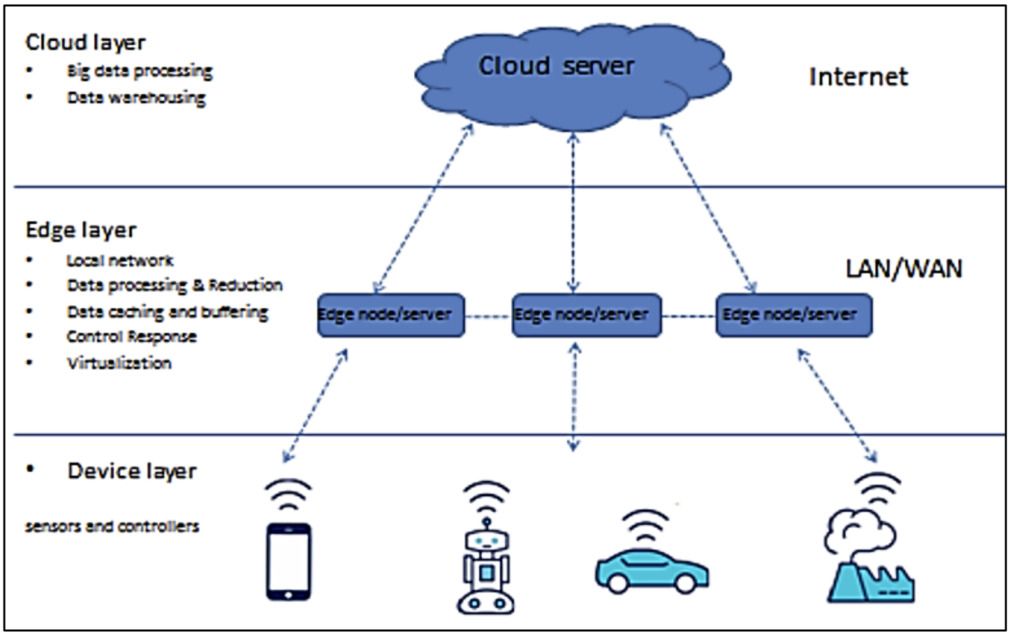

Figure 2. Edge Computing Architecture Diagram [15]

Edge computing architecture refers to a decentralized IT framework that puts processing, storage, or application services close to the source of the data (e.g., IoT, local servers) instead of using a traditional centralised cloud model for everything. Due to its decreased latency, reduced bandwidth, and real-time decision-making, it offers significant advantages. The first layer, device layer (bottom) shows various end devices including phones, robots, vehicles, industrial machines, sensors, and controllers, that continuously generate data and require instant responses at the "edge" to operate. These devices generally have limited processing and storage capabilities.

The next layer, edge layer, comprises numerous edge nodes or edge servers connected by a local (LAN) or wide area (WAN) network to provide local processing of data. This reduces the amount of incoming and outgoing data sent to and from the cloud by providing real-time control responses at the device layer level. This enables lower latencies for devices and reduces the volume of data that needs to be sent to the cloud. In addition, the ability to deploy applications and services in a virtualized manner at the edge of the environment allows for greater flexibility in deploying them near the end user.

The highest layer in the cloud system is the cloud layer, which contains the cloud server (central location, very high processing power, very high storage capacity). The cloud provides the processing capabilities to carry out the large amounts of computing power required for processing large data sets (data warehouses) and global data analysis by providing the computing resources that process the large amounts of data generated. The cloud acts as a repository for the data that has been aggregated and/or chosen to send to the Edge node(s), as well as a place to store control policies and updated model(s) to be used on the edge node.

In addition, the arrows indicating bidirectional communications show that devices will use the Edge node(s) close to them to receive low latency services and Edge nodes will communicate with the Cloud for large volume processing and long-term functional capabilities of the Edge node(s). This shows how cloud computing can be enhanced with edge computing through the implementation of temporary intelligence sources to allow more effective, scalable, and faster cloud computing when data is generated at the cloud generating edge of the cloud.

Figures 1 and 2 illustrate the layered deployment model commonly used in edge–fog–cloud systems. Rather than representing isolated architectures, these figures highlight the hierarchical distribution of computing resources and their functional roles in enabling scalable and low-latency autonomous network management. Artificial intelligence and machine learning enabling autonomy within the edge and fog layers of a network. At the edge and fog layers, artificial intelligence and machine learning models are used to provide self-configuration, self-optimisation, self-healing, and self-protection capabilities through continuous learning and adaptation based on network telemetry data as well as changes in conditions [5]. Additionally, technologies like intent-based networking and zero-touch operations have the potential to minimize human intervention by converting high-level service intentions into automated actions on the network.

4. Mathematical Modelling and Performance Metrics

To evaluate the effectiveness of edge–fog–cloud architectures in supporting autonomous network management, mathematical models and performance metrics are used to analyze system performance. These models help quantify network latency, resource utilization, and workload distribution across different layers of the architecture.

4.1 Mathematical Model of Task Processing

Let the system consist of three computing layers: edge layer (E), fog layer (F), and cloud layer (C).

A task generated by an IoT device can be represented as:

$$T_i = (D_i, C_i)$$

where:

- D_i= size of input data for task i
- C_i= computational requirement of task i

The total service delay L_i for a task processed in the system can be expressed as:

$$L_i = T_{trans} + T_{proc} + T_{queue}$$

where:

- T_{trans}= transmission delay
- T_{proc}= processing delay
- T_{queue}= waiting time in the processing queue

4.2 Transmission Delay

Transmission delay between nodes can be calculated as:

$$T_{trans} = \frac{D_i}{B}$$

where:

- D_i= data size
- B= available network bandwidth

4.3 Processing Delay

Processing delay depends on the computational capacity of the node:

$$T_{proc} = \frac{C_i}{f}$$

where:

- C_i= required CPU cycles
- f= processing capability of the computing node

In edge–fog–cloud systems, the objective of task offloading algorithms is to minimize the overall delay:

$$\min \sum_{i=1}^{N} L_i$$

where Nrepresents the total number of tasks generated by edge devices.

4.4 Resource Utilization Model

The resource utilization of a computing node can be defined as:

$$U = \frac{R_{used}}{R_{total}}$$

where:

- R_{used}= resources currently in use
- R_{total}= total available resources

Efficient autonomous network management aims to maximize resource utilization while avoiding overload conditions.

5. AI/ML-Based Autonomous Network Management Architecture

Figure 3 illustration depicts the principle of enabling networks to be self-sufficient using AI and ML technology through minimal human input. In the depiction, there is a communications infrastructure, represented by both an antenna and a satellite dish, as the central theme of the various types of currently used telecommunications infrastructures, including 5G, IoT, etc., and future versions of 6G. Surrounding this center of the diagram are many capabilities that utilize the power of artificial intelligence (AI) and demonstrate how ML algorithms assist with networking management functions.

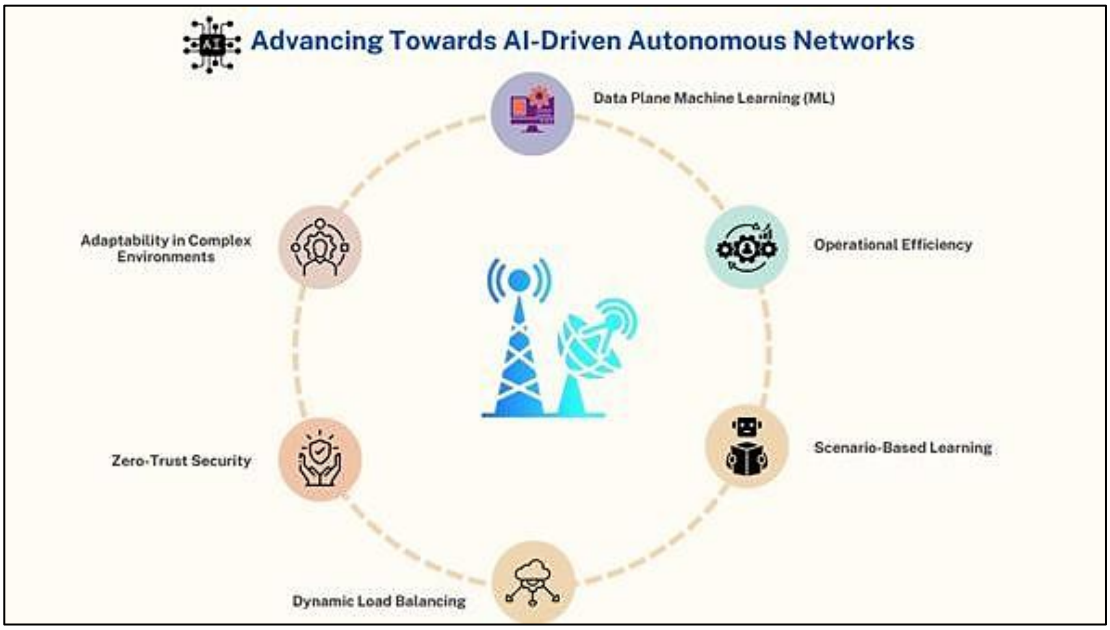

Figure 3. AI-Based Autonomous Network Management Architecture Diagram [19]

5.1 Network Data Plane Machine Learning (ML)

In this section, there are many examples of machine learning being applied directly to the network data plane where packets of data will flow. These examples of machine learning techniques and models will evaluate real-time network traffic patterns to determine where network anomalies exist, provide recommendations on how to geolocate packets of information, and validate that the resources were allocated dynamically and optimally to reduce the volume of congested data movement through the network.

5.2 Operational Efficiency

AI will be leveraged through algorithms to automate management activities that fall under standard operating procedures for the ongoing management of networks, including monitoring, configuration, and optimization of systems and processes. This will enable the network operator to perform these tasks with no human input, significantly reducing the cost of operating large distributed networks.

5. 3 Learning Through Scenarios

Scenario-based learning is an AI model's ability to use data collected from past network activities to improve performance. By using information gathered from incidents such as failures, traffic spikes, and security breaches in their training, AI can predict if an event is likely to happen again and will adjust the network configuration automatically.

5.4 Load Balancing Dynamically

AI will help distribute load within the network. It will move traffic around various computers (or nodes) on the network to reduce congestion and slowdowns. AI will continuously assess network conditions and will reroute traffic dynamically to maintain optimal throughput, reduce latency, and ensure balanced use of computing resources.

5. 5 Zero Trust Security

As shown in the diagram, all devices or connections to the network are verified at all times before being granted access using AI-enabled security mechanisms. The AI uses traffic patterns to identify suspicious behavior, security threats, or any type of anomaly as they occur.

5.6 Adaptability in Complex Environments

Modern networks consist of heterogeneous devices with different protocols, services, etc. AI will help the network continually adapt to changing conditions such as fluctuating workloads, mobile devices, and varying demands of applications.

6. Comparative Analysis

By leveraging AI (Artificial Intelligence), ML (Machine Learning) and SDN (Software Defined Networking) to autonomously manage networks via IoT (Internet of Things), we can create self-configuring, self-healing, self-optimizing networks that require very little or no human input and can be very difficult or impossible to manage manually. The ability of these systems to monitor traffic in real-time, perform predictive maintenance and dynamically reallocate resources to handle the complexity associated with IoT infrastructures enables IT departments to deliver highly performance networks that operate with low latency and high energy efficiency [8].

Autonomous network management [16] will play an increasingly important role for 6G networks because their complexity will make it impossible for a single person or a team of people to effectively manage them. Therefore, organizations must become more proactive in their approach to IT management by moving from reactive firefighting to focusing on strategic initiatives that provide the greatest value to their organization.

Table 2. Feature Comparison of Fog and Edge Computing for Autonomous Network Management

Features	Fog Computing	Edge Computing
Computing Path [9]	Intermediate layer between cloud and edge	At or very close to data sources (devices)
Latency [8]	Medium to low	Very low
Network Traffic [1]	Reduced backhaul through aggregation	Minimal backhaul traffic
Scalability [11]	High through distributed coordination	Limited by device and node capacity
Role in Autonomous Network Management [16]	Distributed coordination, policy enforcement, resource orchestration	Real-time monitoring, fast control loops, local self-healing
Artificial Intelligence and Machine Learning Deployment [15]	Inference and partial learning	Lightweight inference and fast adaptation
Intent-Based Networking (IBN) [17]	Policy enforcement and coordination	Local intent execution
Zero-Touch Operations (ZTO) [18]	Distributed automation	Device-level automation

Resource Management [4]	Distributed and cooperative resource management	Local resource optimization
Security and Privacy [3]	Improved privacy through data filtering	Enhanced privacy by local data processing
Key Challenges [12]	Heterogeneity, orchestration complexity	Resource constraints, limited processing power

Table. 2 shows a comparison between fog computing and edge computing in terms of key characteristics. Fog provides an intermediary relationship between the cloud and edge devices and reducing overall network traffic while also supporting extensive scalability for high volume traffic and providing resources for coordination, coordination policies, and the overall resource management process. There is an overlap in some of the categories of fog and edge computing; both are valuable technologies for 5G network management, IoT device aggregation, and Software Defined Networking (SDN) and Network Functions Virtualization (NFV) coordination [7]. The biggest challenge to Fog technology is the heterogeneity of devices in use and the need to develop a common framework for orchestration.

According to various studies, measurable improvements in latency have been observed when implementing edge and fog computing architectures compared to traditional centralized cloud computing models. Specifically, automated or AI-based resource management in fog and edge environments can significantly reduce service delay by processing data closer to the source and minimizing backhaul communication [3], [4]. A range of studies on AI-driven resource management, task scheduling, and load balancing techniques in fog and edge computing environments report average reductions in network latency of approximately 30–45% compared with cloud-only architectures [5], [8]. Additionally, research on SDN-based edge computing frameworks for IoT systems demonstrates that local data processing and reduced backhaul traffic can lead to latency improvements of approximately 40–55% in distributed network environments [7], [10]. Similarly, machine learning-based scheduling and resource allocation mechanisms in fog computing environments have been shown to reduce latency by 35–50% in latency-sensitive applications such as IoT systems and smart city infrastructures [8], [10]. Collectively, these studies indicate that integrating edge and fog layers into distributed cloud architectures can significantly reduce response time compared with centralized cloud models. However, the level of improvement depends on factors such as workload characteristics, resource management strategies, network topology, and application requirements.

Edge Computing differs from Fog because it is designed to operate directly that is very close where the data is generated, such as from an endpoint device (e.g., sensor/device). The result is a greater reduction in latency, reduced backhaul traffic, and increased data privacy by processing the data locally. Edge computing is suited for applications that require real-time monitoring, rapid control, local self-healing capabilities, and latency-critical applications. However, Edge Computing suffers from resource-constrained systems and thus has a smaller capacity for processing relative to Fog Computing technology.

Table 3. Applications of Edge and Fog Computing for Autonomous Network Management [17]

Applications	Description
5G and Future Mobile Networks	It supports ultra-low latency services, dynamic resource allocation, network slicing, and self-healing required for 5G and emerging 6G networks through autonomous control.
Internet of Things (IoT)	It enables local data processing, real-time decision-making, reduced backhaul traffic, and scalable management of massive IoT deployments.
Software-Defined Networking (SDN)	Facilitates centralized and distributed control, programmable networks, and automated traffic management using AI-driven policies.
Network Function Virtualization (NFV)	It supports flexible deployment of virtual network functions at the edge and fog layers, improving agility and reducing operational costs.
Smart Cities	It enables real-time monitoring, traffic control, public safety systems, and energy management with low latency and high reliability.
Autonomous Network Management	Provides self-configuration, self-optimization, self-healing, and self-protection using AI/MLart, intent-based networking, and zero-touch operations.
Latency-Critical Applications	It supports applications such as autonomous vehicles, industrial automation, and real-time video analytics by processing data close to the source.

Table 3 shows that the applications of edge and fog computing can also support autonomous connectivity networks across different industry types, particularly within 5G and future mobile networks, by providing ultra-low latency capabilities, network slicing for the delivery of services, and self-healing capabilities for rigorous performance requirements. In the Internet of Things (IoT), edge computing allows for local processing to minimize backhaul traffic, enabling real-time decisions for extensive deployments of numerous devices. The implementation and availability of software-defined networks (SDNs) have been enhanced by AI-enabled policies governing programmable and automated control; therefore, SDNs offer flexible and efficient solutions. Network function virtualization (NFV) can enhance the deployment of network functions and applications closer to the end users, reducing costs and increasing overall system agility for service delivery [10].

7. Discussion

The paper elaborates on edge and fog computing's ability to provide autonomous network operations management by addressing the challenges of existing cloud-based network management systems. [3] Both edge and fog computing increase the amount of computation, storage, and control located near the data source or edge device, respectively, providing a reduction in latency, reducing network congestion, and fostering real-time decision-making. These advantages make edge and fog computing particularly applicable to the complex and dynamic environments of 5G networks, IoT ecosystems, smart cities, and time-sensitive applications. Additionally, by integrating AI and machine learning, these technologies are empowered with additional autonomy via closed-loop control systems, intent-based networking and zero-touch solutions providing secure self-configuration, self-optimization, self-healing, and self-protecting capabilities.

By using fog computing, there are many different types of devices and platforms. Because fog nodes are typically composed of various types of hardware, use different types of operating systems, and have different types of connectivity options, orchestration and interoperability are necessary. [4] Creating standards for integration, resource management, and service placement in distributed fog environments is a significant research challenge. Additionally, fog systems will require additional management overhead due to their multi-layered and distributed nature.

Within edge computing resource constraints are the major barrier to entry. Edge devices are typically limited in computation power per unit of memory, storage, and energy resources compared to traditional computers and servers; therefore, they cannot efficiently support small-scale compute-intensive Artificial Intelligence (AI) models (and/or perform large scale analysis) [8]. This necessitates an appropriate balance of task offloading between the edge, fog, and cloud layers for effective operation. Additionally, scalability presents a major challenge as it continues to be difficult to manage and maintain a large number of Edge devices in an autonomous manner.

There are many different security issues and problems regarding security and privacy when utilizing Edge and Fog computing [12]. The process of edge computing can help a user by protecting both the security and privacy of information, while at simultaneously creating a greater number of points for a malicious entity to attack and gain access to the data by widening the attack surface area for many distributed nodes. In order to achieve secure communication, authentication, and data protection among edge, fog, and cloud systems; we need to implement effective and lightweight security measures.

The article states that edge and fog computing are critical enablers for the autonomous management of networks; however, there are still many unresolved issues regarding the management of resources, [17] orchestration of resources, interoperability between different platforms, systems, devices, scalability, and security, which need additional research for large scale autonomous deployment of wireless networks.

8. Future Scope

A broad and exciting scope exists for edge and fog computing's future perspectives on automated network management, with networks evolving toward 6G technologies. The complexity, density, and variety of devices and services within a network will continue to increase; therefore, edge and fog computing are enabling network autonomy and self-management. In the future, it can focus on developing advanced lightweight, energy-efficient artificial intelligence (AI) and machine learning (ML) models that provide continuous learning at both the edge and fog levels. These models will enhance real-time decisions based on making use of very limited resources. It is currently focused on developing advanced AI and ML models that are lightweight, energy-efficient, and capable of continuous learning at edge and fog levels. These models will enable improved decision-making in real time while operating within the constraints of resources. In the future, the availability of combined federated learning and edge-fog-cloud systems in order to perform a local model training on edge nodes and aggregation of the model parameters from fog/cloud to produce a global model as part of distributed learning. The reduction in communication overhead associated with the distributed learning will provide organizations with improved privacy capabilities regarding data and more scalable intelligent network management capabilities when deploying large-scale IoT devices.

9. Conclusion

The comprehensive survey illustrates that by combining AI and machine learning with edge and fog computing, additional self-configuring, self-optimizing, self-healing, and self-protecting functions have been facilitated. The concepts of intent based networking (IBN) and zero touch (ZT) limit the need for human involvement in enhancing operational efficiencies. Comparisons of fog computing are shown to be extremely important for distributed control and scalability, whereas edge computing has proven to be superior for ultra-low latency processing and real time control. Collectively, edge and fog, when combined with cloud computing, provide a modular, multi-tiered architecture to facilitate the requirements of more complex environments. The results of the application analysis illustrate that edge and fog computing tend to perform well across multiple sectors, including 5G/future mobile networks, IoT, SDN, NFV, smart cities and any other applications requiring low latency. These include limitations regarding available resources at the edge; heterogeneous and complex orchestration of resources in fog environments, scalability, and increasing vulnerabilities to security and privacy due to additional points of attack. In conclusion, edge and fog computing represent critical components within an overall solution for automating the management of networks. Although substantial progress has been achieved, there are still numerous areas that require additional research toward the resolution of the longstanding problems associated with standards, resource management, interoperability, scalability, and security.

References

[1]. Habibi, Pooyan, Mohammad Farhoudi, Sepehr Kazemian, Siavash Khorsandi, and Alberto Leon-Garcia. "Fog Computing: a Comprehensive Architectural Survey." IEEE access 8 (2020): 69105-69133.

[2]. Mouradian, Carla, Diala Naboulsi, Sami Yangui, Roch H. Glitho, Monique J. Morrow, and Paul A. Polakos. "A Comprehensive Survey on Fog Computing: State-of-the-art and Research Challenges." IEEE communications surveys & tutorials 20, no. 1 (2017): 416-464.

[3]. Walia, Guneet Kaur, Mohit Kumar, and Sukhpal Singh Gill. "AI-empowered fog/Edge Resource Management for IoT Applications: A Comprehensive Review, Research Challenges, and Future Perspectives." IEEE Communications Surveys & Tutorials 26, no. 1 (2023): 619-669.

[4]. Hong, Cheol-Ho, and Blesson Varghese. "Resource Management in Fog/Edge Computing: a Survey on Architectures, Infrastructure, and Algorithms." ACM computing surveys (csur) 52, no. 5 (2019): 1-37.

[5]. Alsadie, Deafallah. "A Comprehensive Review of AI Techniques for Resource Management in Fog Computing: Trends, Challenges, and Future Directions." IEEE Access 12 (2024): 118007-118059.

[6]. Kuchuk, Heorhii, and Eduard Malokhvii. "Integration of IoT with Cloud, Fog, and Edge Computing: a Review." Advanced Information Systems 8, no. 2 (2024): 65-78.

[7]. Rafique, Wajid, Lianyong Qi, Ibrar Yaqoob, Muhammad Imran, Raihan Ur Rasool, and Wanchun Dou. "Complementing IoT Services through Software Defined Networking and Edge Computing: A Comprehensive Survey." IEEE Communications Surveys & Tutorials 22, no. 3 (2020): 1761-1804.

[8]. Boubaker, Nour El Houda, Karim Zarour, Nawal Guermouche, and Djamel Benmerzoug. "A Comprehensive Survey on Resource Management for IoT Applications in edge-fog-cloud Environments." IEEE Access (2025).

[9]. Ren, Ju, Deyu Zhang, Shiwen He, Yaoxue Zhang, and Tao Li. "A Survey on end-edge-Cloud Orchestrated Network Computing Paradigms: Transparent Computing, Mobile Edge Computing, Fog Computing, And Cloudlet." ACM Computing Surveys (CSUR) 52, no. 6 (2019): 1-36.

[10]. Khan, Latif U., Ibrar Yaqoob, Nguyen H. Tran, SM Ahsan Kazmi, Tri Nguyen Dang, and Choong Seon Hong. "Edge-computing-enabled Smart Cities: A Comprehensive Survey." IEEE Internet of Things journal 7, no. 10 (2020): 10200-10232.

[11]. Costa, Breno, Joao Bachiega Jr, Leonardo Rebouças De Carvalho, and Aleteia PF Araujo. "Orchestration in Fog Computing: A Comprehensive Survey." ACM computing surveys (CSUR) 55, no. 2 (2022): 1-34.

[12]. Chuan, Wu Chia, Shams Ul Arfeen Laghari, Selvakumar Manickam, Erum Ashraf, and Shankar Karuppayah. "Challenges and Opportunities in Fog Computing Scheduling: A Literature Review." IEEE Access 13 (2025): 14702-14726.

[13]. Puliafito, Carlo, Enzo Mingozzi, Francesco Longo, Antonio Puliafito, and Omer Rana. "Fog Computing for the Internet of Things: A Survey." ACM Transactions on Internet Technology (TOIT) 19, no. 2 (2019): 1-41.

[14]. Weng, Chi-Yao, Chun-Ta Li, Chin-Ling Chen, Cheng-Chi Lee, and Yong-Yuan Deng. "A Lightweight Anonymous Authentication and Secure Communication Scheme for Fog Computing Services." IEEE Access 9 (2021): 145522-145537.

[15]. Kadhum, Methaq, Saher Manaseer, and Abdel Latif Abu Dalhoum. "Cloud-edge Network Data Processing based on User Requirements using Modify Mapreduce Algorithm and Machine Learning Techniques." International Journal of Advanced Computer Science and Applications 10, no. 12 (2019).

[16]. Ullah, Inam, Ali Arishi, Sushil Kumar Singh, Faisal Alharbi, Anwar Hassan Ibrahim, Muhammad Islam, Yousef Ibrahim Daradkeh, and Chang Choi. "Autonomous Network Management for 6G Communication: a comprehensive survey." Digital Communications and Networks (2025).

[17]. Dlamini, Sabelo & Mwangama, Joyce & Ventura, Neco & Magedanz, Thomas. (2018). Design of an Autonomous Management and Orchestration for Fog Computing. 10.1109/ICONIC.2018.8601272.

[18]. Basheer, Hanaa, and May Itani. "Zero touch in fog, IoT, and MANET for Enhanced Smart City Applications: A survey." Future Cities and Environment 9, no. 1 (2023).

[19]. https://www.linkedin.com/pulse/next-gen-autonomous-network-ai-saraswathi-singaravelu-2iyjc