SAGE Science Publishing

AI-Based Intrusion Detection and DDoS Mitigation in Fog Computing: Addressing Security Threats in Decentralized Systems

Kaushik Sathupadi[1]

[1]Staff Engineer, Google LLC, Sunnyvale, CA

Abstract

Fog computing is a decentralized paradigm designed to bring computation and services closer to the edge of the network. It has emerged as a promising solution for latency-sensitive applications. However, this architectural shift introduces a set of security challenges that are distinct from traditional centralized cloud environments. Traditional centralized security models are often inadequate for fog environments due to their reliance on centralized data processing, which contrasts with the distributed, heterogeneous, and latency-sensitive nature of fog computing. Distributed Denial of Service (DDoS) attacks, unauthorized intrusions, data breaches, malware propagation, and privacy threats are problematic in fog computing due to its decentralized structure, resource constraints, and the heterogeneous nature of fog nodes. This paper focuses on identifying the critical security threats that fog computing environments face, with a special emphasis on DDoS attacks and other forms of intrusion. In light of these challenges, the role of Artificial Intelligence (AI)-driven solutions in mitigating these security risks is also examined. This study discussed how AI-based techniques, including machine learning (ML), deep learning (DL), and reinforcement learning (RL), offer innovative approaches for real-time threat detection, anomaly recognition, and adaptive mitigation strategies. Deploying AI-based models in fog environments presents challenges such as limited computational resources, latency concerns, and energy constraints. This paper also discusses how AI can be used to enhance security in fog computing while addressing the inherent vulnerabilities of decentralized systems.

Keywords: AI-driven security, DDoS attacks, fog computing, intrusion detection, machine learning, security challenges, threat mitigation

Introduction

Fog computing, often referred to as edge computing, represents a significant shift from the traditional centralized cloud model by enabling data processing closer to where the data is generated. This decentralized approach introduces a layered architecture, with the fog layer acting as an intermediary between the cloud and the edge devices. The key advantage of fog computing lies in its ability to reduce latency and optimize bandwidth usage, making it especially useful for applications that require real-time or near-real-time responses. By processing data locally or near the source, fog computing mitigates the delays associated with transmitting data to distant cloud servers. This makes it ideal for applications such as autonomous vehicles, smart cities, healthcare systems, and industrial automation, where the speed of data processing directly impacts performance and safety (Zhang *et al.* 2018).

The architecture of fog computing can be viewed as comprising three primary layers: the edge layer, the fog layer, and the cloud layer. The edge layer consists of devices such as sensors, actuators, cameras, and IoT devices that generate vast amounts of data. These edge devices typically have limited computational power and are primarily responsible for capturing raw data in real time. While some minimal processing may occur at the edge, such as data filtering or aggregation, most of the heavy lifting in terms of computation and analytics is handled by the fog or cloud layers (Huang *et al.* 2017). The edge devices, therefore, serve as the initial point of data collection but depend on higher layers for more complex processing tasks.

Above the edge layer is the fog layer, which is central to fog computing. This layer includes intermediary nodes such as routers, gateways, switches, and local servers that are strategically located closer to the edge devices. These fog nodes possess more significant computational power and storage capacity than edge devices and can perform advanced data processing, analytics, and decision-making tasks. For example, in an autonomous vehicle network, fog nodes might process real-time data from vehicle sensors to detect obstacles, optimize routes, or initiate safety measures without waiting for communication with a remote cloud server. This ability to process data locally, within milliseconds, is critical in applications where even slight delays can have catastrophic consequences, such as in vehicle-to-vehicle communication systems or healthcare monitoring devices. The fog layer ensures that critical decisions are made closer to the data source, reducing round-trip times to the cloud and enabling more efficient use of bandwidth by transmitting only relevant data to the cloud for further storage or analysis.

Table 1 Applications of Fog Computing in Various Domains

Domain	Key Technologies	Data Sources	Fog Computing Role	Benefits
Smart Cities	Intelligent Transportation, Environmental Monitoring, Energy Management	Road sensors, traffic cameras, environmental sensors, energy meters	Local processing for real-time decision-making	Traffic optimization, air quality monitoring, energy distribution (Yi *et al.* 2015a)
Autonomous Vehicles	V2V and V2I communication, LiDAR, radar, GPS	Vehicle sensors, roadside infrastructure	Local processing for split-second decision-making	Obstacle avoidance, improved traffic flow, reduced bandwidth usage (Dastjerdi *et al.* 2016)
Healthcare	Wearable devices, telemedicine systems	Heart monitors, glucose sensors, blood pressure monitors	Local data processing for real-time health monitoring	Early detection of health issues (Quy *et al.* 2022), reliable telemedicine services (Mutlag *et al.* 2019)
Industrial Automation	Industrial IoT devices, sensors, cameras, robotic systems	Manufacturing equipment, production line sensors	Real-time processing for predictive maintenance and quality control	Reduced downtime, enhanced productivity, improved product quality

The cloud layer functions as a more centralized resource, providing large-scale storage, extensive computational capabilities, and long-term data analysis. In fog computing architectures, the cloud typically handles tasks that are less time-sensitive, such as historical data analysis, machine learning model training, and global network management. Data from fog nodes that is not time-critical or that has already undergone initial processing can be transmitted to the cloud for further refinement. For instance, data from a factory's sensors could be aggregated and analyzed over time in the cloud to predict equipment failure trends or optimize production processes. The cloud's vast computational resources make it well-suited for handling these kinds of large-scale, non-real-time tasks, though the fog layer reduces the cloud's burden by offloading time-critical computations to the edge.

In traditional cloud computing architectures, data generated at the edge must be sent to a centralized data center for processing, which introduces delays due to transmission distances and network congestion. In contrast, fog computing processes data closer to the source, drastically reducing the round-trip time for data to travel from the edge to the cloud and back. This is important in applications like autonomous driving, where even a few milliseconds of delay in data processing can mean the difference between avoiding or causing an accident (Dsouza *et al.* 2014).

Additionally, fog computing offers bandwidth optimization. The ever-increasing number of IoT devices generating massive amounts of data can overwhelm networks if all this data is transmitted to the cloud for processing. By performing data processing locally at the fog layer, only the most relevant and refined data is transmitted to the cloud, reducing the load on the network. For example, in a smart city, data from traffic sensors can be processed locally to manage traffic lights in real time, while only long-term traffic trends are sent to the cloud for further analysis.

Another technical advantage is improved security and privacy. Fog computing enables localized processing, which can limit the amount of sensitive data sent over potentially insecure wide-area networks (WANs). In healthcare applications, for example, patient data collected from wearable devices can be processed locally at the fog layer to detect abnormalities without transmitting sensitive health information to distant cloud servers. This reduces the risk of data breaches and enhances compliance with privacy regulations such as HIPAA (Health Insurance Portability and Accountability Act).

Smart cities represent one of the most prominent applications of fog computing, leveraging this architecture to manage large-scale, real-time data from a wide array of urban systems. A smart city integrates technologies such as intelligent transportation systems, environmental monitoring, and energy management, all of which generate enormous quantities of data that must be processed quickly to ensure efficient operation. Fog computing enables local processing of this data, allowing for real-time decision-making without overloading centralized cloud servers. For instance, in intelligent traffic management systems, data from road sensors, traffic cameras, and connected vehicles can be analyzed at fog nodes to optimize traffic light patterns, reduce congestion, and enhance public safety. Similarly, fog nodes can analyze data from environmental sensors to monitor air quality, noise levels, or water pollution in real time, allowing city officials to respond quickly to potential health hazards.

In smart grids, fog computing enables the real-time monitoring and control of electricity distribution, helping to balance supply and demand dynamically. Fog nodes installed at substations can process data from energy meters and sensors to detect grid anomalies, optimize energy distribution, and integrate renewable energy sources more effectively. By processing data locally, fog computing reduces the communication latency between different components of the grid, improving reliability and reducing the risk of power outages.

Autonomous vehicles rely heavily on fog computing to handle the enormous volumes of data generated by sensors such as cameras, LiDAR, radar, and GPS systems. These sensors produce real-time data about the vehicle's surroundings, including information about nearby objects, road conditions, and traffic patterns. The processing of this data must occur almost instantaneously to ensure the safe operation of the vehicle. Fog computing enables vehicles to process sensor data locally at fog nodes, often located within the vehicle itself or at nearby infrastructure, such as traffic lights or road signs. This local processing allows autonomous vehicles to make split-second decisions, such as

avoiding obstacles or adjusting speed based on traffic conditions, without the delays that would occur if data were sent to a cloud server for processing.

Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication systems also benefit from fog computing. In these systems, vehicles and roadside infrastructure exchange data to improve safety and traffic flow. Fog nodes situated along roadways or within traffic management centers can process and relay this data in real time, enabling faster decision-making and reducing the likelihood of accidents. Moreover, fog computing reduces the amount of data that needs to be transmitted to the cloud, decreasing the bandwidth required for V2V and V2I communications.

In healthcare, fog computing supports critical applications that require real-time monitoring and analysis of patient data. Wearable devices such as heart monitors, glucose sensors, and blood pressure monitors continuously generate patient data, which must be processed in a timely manner to detect abnormalities and alert healthcare providers. By processing this data locally at the fog layer, healthcare systems can identify potential health issues more quickly and intervene before conditions worsen. For instance, a fog node could analyze data from a heart monitor to detect irregular heartbeats and immediately notify a doctor or emergency services, potentially saving a patient's life.

Fog computing also plays a role in telemedicine, where patients in remote areas may lack reliable access to cloud-based healthcare services. By deploying fog nodes in remote healthcare facilities, patient data can be processed locally, reducing the reliance on distant cloud servers and ensuring that critical medical information is analyzed quickly. This local processing enhances the reliability of telemedicine systems and ensures that patients in underserved areas receive timely care.

In industrial settings, fog computing is driving the transformation of factories into smart, highly automated environments. Industrial IoT (IIoT) devices, such as sensors, cameras, and robotic systems, generate large amounts of data that must be processed in real time to monitor equipment, optimize production processes, and detect potential faults. Fog computing enables real-time data processing at the edge of the network, allowing factories to respond quickly to changes in operating conditions without relying on remote cloud servers.

For example, fog nodes can monitor data from manufacturing equipment to detect signs of wear or malfunction, enabling predictive maintenance. By identifying potential failures before they occur, factories can reduce downtime, increase productivity, and lower maintenance costs. Fog computing also supports quality control systems by analyzing data from sensors and cameras on the production line, ensuring that products meet quality standards and reducing the risk of defects.

Problem Statement

As fog computing expands, the increase in connected devices and fog nodes presents significant security challenges. Unlike traditional cloud computing, where data is processed and managed in centralized data centers with well-established security mechanisms, fog computing operates in a decentralized manner. This distributed architecture exposes the system to a range of security threats, primarily due to the broader attack surface and the complexity of securing nodes that are physically and geographically dispersed (Abbasi and Shah 2017).

One of the primary concerns in fog computing is the heterogeneity of devices within the network. Fog environments involve a wide variety of devices, from powerful gateways and servers to resource-constrained IoT devices. This diversity creates inconsistencies in the ability to apply uniform security protocols. Many IoT devices, which often form the edge layer of fog computing, are limited in terms of computational power, memory, and energy resources. As a result, these devices are typically incapable of executing advanced security mechanisms such as robust encryption or intrusion detection systems. Attackers can exploit these less secure devices, gaining access to the network through its weakest points. Moreover, the sheer volume of devices complicates the management of security updates, patches, and vulnerability assessments, making it difficult to maintain consistent security across the entire system. The diagram in figure 1 is illustrating the different layers in a fog computing environment (Cloud Layer, Fog Layer, Edge Layer), emphasizing the security challenges at each layer, focusing on the vulnerabilities of the edge devices. The figure also illustrates how attackers exploit the weakest devices in the system (Yi *et al.* 2015b).

Unlike centralized cloud data centers, which are typically housed in secure facilities with restricted access, fog nodes are often deployed in locations that are accessible to the public or semi-public environments. Examples include smart traffic lights, energy meters in smart grids, or industrial equipment located in open areas. This proximity to end users and the physical accessibility of the devices make them more vulnerable to tampering, theft, or physical attacks. An attacker with physical access to a fog node could manipulate hardware components or install malicious software, compromising the node's security and potentially affecting the larger network. Additionally, side-channel attacks, which exploit information leaked through physical characteristics such as power consumption or electromagnetic emissions, can be a concern for fog nodes deployed in unsecured environments. Figure 2 illustrates the physical security risks associated with fog nodes deployed in public or semi-public environments, such as smart traffic lights, energy meters, and industrial equipment. The figure 2 also highlights potential physical and side-channel attack vulnerabilities that arise from their exposure and proximity to end users.

In a centralized cloud environment, data is processed in a single location where stringent access control, encryption, and monitoring protocols can be applied. In fog computing, data is processed at multiple locations, including edge devices and fog nodes, before being forwarded to the cloud. This creates numerous potential points of failure where data can be intercepted, tampered with, or exposed to unauthorized access. The risk of data breaches increases as data moves between different fog nodes and the cloud, especially if secure communication channels are not consistently enforced across the system. Ensuring data privacy and integrity across a dispersed fog network is much more challenging compared to a centralized cloud system, where all traffic and storage can be monitored from a single control point (Wang *et al.* 2015).

In traditional cloud environments, authentication is typically managed by a centralized authority, ensuring that all devices and users are verified before accessing the system. However, in a fog computing architecture, where thousands of devices and fog nodes are distributed across a wide area, managing authentication becomes more complex. Each fog node must be capable of verifying the legitimacy of the devices and data it interacts with, which can lead to inconsistent trust models if not properly coordinated. The lack of a centralized authentication authority in fog

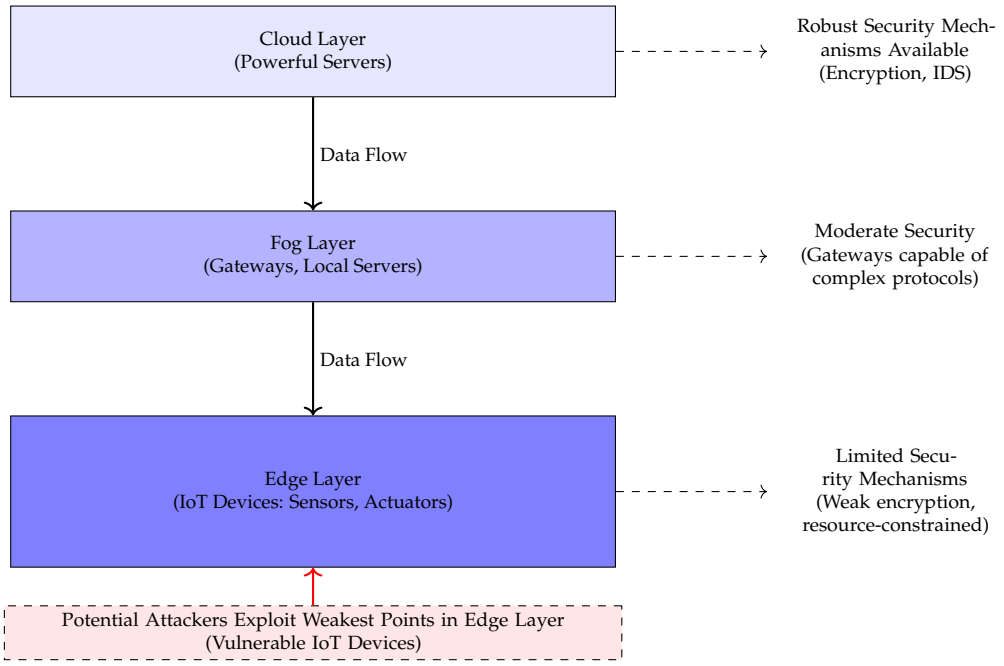

Figure 1 Security Concerns in Fog Computing: Heterogeneous Device Capabilities and Vulnerabilities

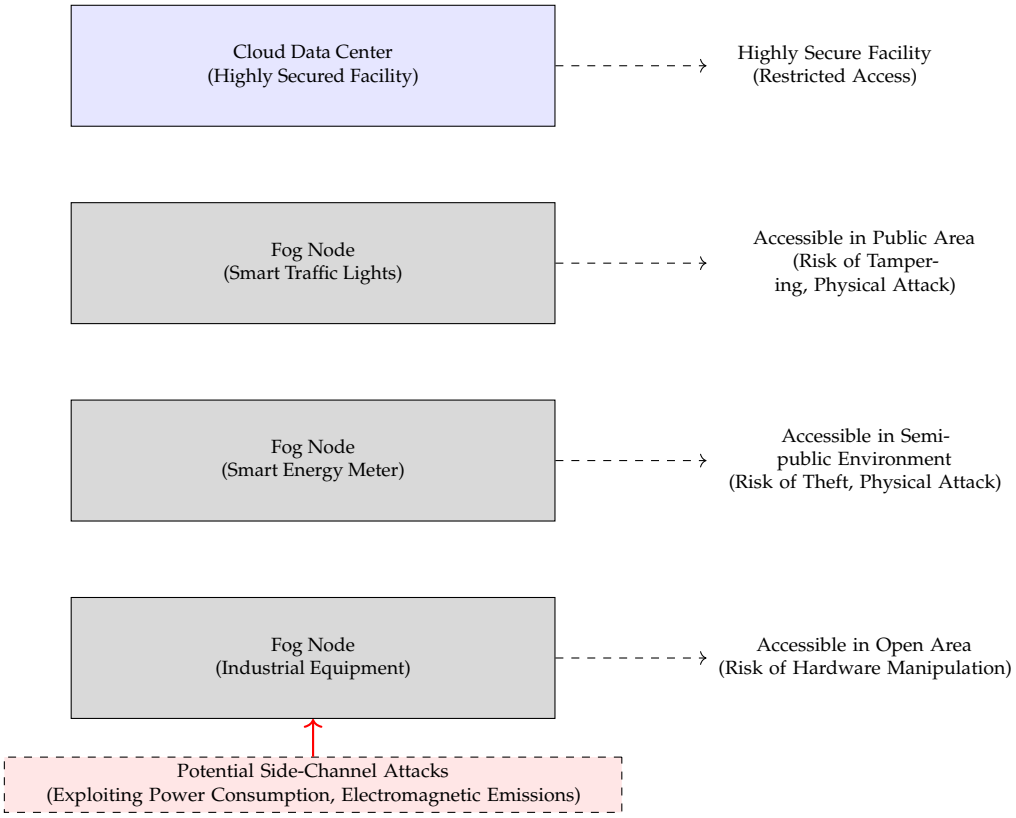

Figure 2 Physical Security Challenges in Fog Computing: Public Accessibility and Risk of Tampering

computing raises the risk of compromised devices joining the network, leading to issues such as data tampering, unauthorized access, or the spread of malware.

In a DDoS attack, multiple compromised devices flood a target system with traffic, overwhelming its resources and causing it to fail. The large number of interconnected fog nodes and edge devices in a fog computing environment provides a wide range of targets for attackers. Because fog nodes often operate in less secure, distributed environments and may have limited computational resources, they are more vulnerable to such attacks than the heavily fortified data centers used in traditional cloud computing. Once compromised, these nodes can be used

as part of a botnet to launch attacks on other parts of the fog or cloud infrastructure, leading to widespread disruption. Furthermore, since fog computing relies on real-time data processing for critical applications like healthcare, industrial automation, or autonomous vehicles, the consequences of a successful DDoS attack could be catastrophic, potentially leading to system outages or failures in safety-critical applications.

In a cloud environment, security patches can be deployed centrally, ensuring all servers and applications are updated simultaneously. In contrast, in fog computing, updates must be applied to a large number of geographically dispersed fog nodes and edge devices. Ensuring that every device is kept up to date with the latest security patches is logistically challenging, especially in environments where devices are not consistently online or are difficult to physically access. The delay or failure to apply necessary patches in a timely manner leaves the entire system exposed to vulnerabilities, as attackers can exploit outdated software or known security flaws in devices that have not been updated.

In addition to technical vulnerabilities, privacy concerns are also heightened in fog computing. With data being processed and stored across multiple nodes close to the data sources, ensuring that sensitive information is properly protected is more difficult. Personal or sensitive data may be exposed to more entities along the processing chain, increasing the risk of unauthorized access or data leakage. For instance, in healthcare applications, data from patient monitoring devices might be processed at fog nodes near hospitals or clinics. If these nodes are not adequately secured, sensitive patient information could be intercepted by attackers or unintentionally exposed to unauthorized personnel.

Data moving between fog nodes, edge devices, and the cloud must be transmitted securely, as these communications are often a prime target for attackers. Man-in-the-middle (MITM) attacks, where an attacker intercepts and potentially alters data being exchanged between devices, are a significant threat in such distributed systems. In the absence of end-to-end encryption or secure communication protocols, data flowing between fog nodes can be intercepted or compromised, leading to loss of data integrity and confidentiality. Ensuring consistent security across all communication channels, while maintaining low latency for time-sensitive applications, is a major challenge in fog computing environments.

Fog Computing Architecture and Unique Security Challenges

Fog computing operates as a decentralized computational layer between IoT (Internet of Things) devices and centralized cloud data centers, providing enhanced computational, storage, and networking capabilities closer to the data sources. By processing data locally on fog nodes—devices that vary in complexity from edge servers to IoT sensors—fog computing reduces the need for constant data transfer to the cloud. This proximity to the network edge lowers latency, which is important for time-sensitive applications such as autonomous vehicles, industrial automation, and healthcare monitoring. Furthermore, fog computing reduces bandwidth usage by filtering or processing data closer to its origin, sending only necessary information to the cloud. Despite these advantages, the architectural shift from cloud to fog introduces a set of new, complex security challenges that stem primarily from its decentralized, heterogeneous, and dynamic nature (Lee *et al.* 2015) (Mukherjee *et al.* 2017).

The foremost issue in fog computing is decentralization, which distinguishes it from traditional cloud computing, where centralized control and security protocols are uniformly applied. In a cloud environment, data and computational tasks are typically handled by a centralized set of servers within secure, physically protected data centers. This allows for the consistent implementation of security measures, policies, and updates across the entire infrastructure. In contrast, fog computing operates in a distributed manner, with potentially hundreds or thousands of fog nodes deployed across geographically dispersed locations. These nodes operate independently, and their autonomy makes it difficult to enforce uniform security policies. As a result, ensuring that all nodes adhere to the same security standards is a considerable challenge. Moreover, fog nodes may operate in environments that are not physically secure, such as public spaces, industrial sites, or remote locations. This exposes them to a higher risk of physical tampering and unauthorized access. Attackers can exploit these decentralized nodes to breach the system, often targeting less secure or poorly monitored components of the network. Without a centralized entity to monitor and enforce security protocols, fog computing environments are inherently more vulnerable to cyberattacks.

Another critical security issue in fog computing is the heterogeneity of nodes within the architecture. Fog nodes range from highly capable edge servers with significant computational power to resource-constrained devices, such as IoT sensors with limited processing and storage capacity. This disparity creates a fragmented environment where different devices require tailored security mechanisms, further complicating the implementation of a unified defense strategy. For instance, a fog node in the form of an edge server may be able to support robust encryption, firewalls, and intrusion detection systems. On the other hand, an IoT sensor or small embedded device may lack the necessary resources to implement the same security measures. This makes the weaker nodes attractive targets for attackers, who can use them as entry points to the larger network. Moreover, heterogeneity in hardware and software platforms creates compatibility issues for security solutions. Fog nodes often run on diverse operating systems, communication protocols, and network topologies, making it difficult to deploy a single security strategy that works across the entire infrastructure. This uneven security posture across fog nodes significantly increases the likelihood of breaches, as attackers are more likely to find vulnerable entry points (Lee *et al.* 2015).

Fog computing environments often include mobile nodes, such as those found in vehicles, drones, or wearable devices, which continuously change their network location and connectivity. The high mobility of these nodes creates dynamic network topologies, introducing additional complexities in maintaining a secure connection. Mobile fog nodes must establish secure communication with other nodes and devices as they move through different network regions, each of which may have varying security policies and levels of trust. Ensuring the integrity and confidentiality of data exchanged between mobile nodes and the rest of the fog network is a significant challenge, especially when these nodes encounter fluctuating network conditions, such as weak signal strength or intermittent connectivity (Ni *et al.* 2017). Attackers can exploit these conditions, targeting mobile nodes during handoffs or when the connection is weakest, to intercept or manipulate data. The frequent movement of fog nodes also complicates the application of traditional security techniques that rely on static configurations, such as fixed encryption keys or predefined access control policies. As a result, mobile nodes

are often more vulnerable to man-in-the-middle attacks, eavesdropping, and unauthorized access than stationary devices.

Another major concern in fog computing is the increased attack surface that results from the vast number of distributed fog nodes. In a traditional cloud environment, the attack surface is primarily limited to the data centers and the communication channels between the client and the cloud. However, in fog computing, the attack surface expands significantly as each fog node represents a potential point of vulnerability. With fog nodes deployed across diverse and often unsecured locations, attackers have numerous entry points through which they can launch attacks. These attacks can range from Distributed Denial of Service (DDoS) attacks, where compromised nodes flood a target node with traffic, to malware infections that spread across interconnected nodes. Additionally, the physical deployment of fog nodes in public or less-controlled environments exposes them to tampering, where malicious actors can physically compromise the devices to gain unauthorized access or manipulate data. As the number of fog nodes increases, so does the challenge of monitoring and securing all potential entry points, making it harder to detect and mitigate threats in real-time (Kunal *et al.* 2019).

The nature of applications in fog computing introduces low-latency and real-time constraints, which add another layer of complexity to security. Many fog computing applications, such as real-time healthcare monitoring, smart grids, and autonomous driving systems, require immediate data processing and response. Any delay caused by security mechanisms could degrade the performance of these latency-sensitive applications, leading to undesirable consequences, such as reduced decision-making accuracy or system malfunctions. Traditional security protocols, such as computationally intensive encryption algorithms or extensive logging and monitoring systems, often introduce latency, making them unsuitable for real-time environments. As a result, there is a constant trade-off between ensuring robust security and maintaining the performance characteristics that fog computing was designed to enhance. Security solutions that introduce significant overhead in terms of processing time or network traffic can disrupt the real-time functioning of fog applications, creating an additional challenge for designing and implementing effective security strategies.

Fog computing systems must address the resource-constrained nature of many nodes. Unlike cloud data centers, which have vast computational resources, many fog nodes, IoT devices, operate with limited processing power, memory, and energy. This resource limitation makes it difficult to implement complex security mechanisms on all fog nodes. For example, while data encryption is a fundamental security measure, the computational cost of encryption can be prohibitive for small IoT devices that must prioritize power efficiency. Similarly, resource-constrained devices may not be able to support continuous monitoring or advanced anomaly detection algorithms, leaving them more vulnerable to attacks. Attackers can exploit these resource limitations by overwhelming fog nodes with computationally intensive tasks or launching energy-draining attacks that cause the devices to shut down. This challenge is compounded by the fact that fog computing environments often operate in harsh or remote conditions where power and connectivity are limited, further restricting the ability of nodes to perform continuous security checks or updates (Khan *et al.* 2017).

Table 2 Comparison of Security Challenges in Fog and Cloud Computing

Security Challenge	Fog Computing	Cloud Computing
Decentralization	Highly decentralized	Centralized control
Node Heterogeneity	High (IoT to edge servers)	Low (homogeneous data centers)
Attack Surface	Large and distributed	Centralized
Latency Sensitivity	High	Moderate to low
Mobility	High (mobile nodes)	Low (stationary data centers)
Resource Constraints	Significant	Minimal

Table 3 Security Issues and Factors in Fog Computing

Security Issue	Exacerbating Factors
Decentralization	Lack of centralized control, geographic dispersion
Heterogeneity of Nodes	Varying computational power, incompatible platforms
High Mobility	Constantly changing network topologies, weak handoffs
Increased Attack Surface	Numerous vulnerable nodes, public deployment
Low Latency Requirements	Strict performance constraints, real-time processing
Resource Constraints	Limited power and processing capabilities

Main Security Threats in Fog Computing

1. Distributed Denial of Service (DDoS) Attacks

Distributed Denial of Service (DDoS) attacks present a significant challenge to fog computing environments due to the decentralized and resource-constrained nature of fog nodes. Fog computing extends cloud services closer to end-users by distributing resources across a wide array of devices and nodes located near the network's edge. While this architecture provides benefits such as reduced latency and enhanced scalability, it also exposes vulnerabilities that make fog systems susceptible to DDoS attacks. By overwhelming these nodes with excessive traffic or computational requests, attackers can severely degrade the performance of fog services, rendering them unavailable to legitimate users. Unlike traditional cloud architectures, fog systems typically lack the centralized security controls and expansive resources available to counteract such attacks, heightening the impact of DDoS vectors on overall service availability (Thota *et al.* 2018) (Zhang *et al.* 2018).

A DDoS attack involves distributing malicious traffic from multiple sources to a targeted network, system, or application with the goal of overwhelming its processing or bandwidth capacity. In a fog computing context, this can manifest in several ways, primarily through network flooding attacks, resource

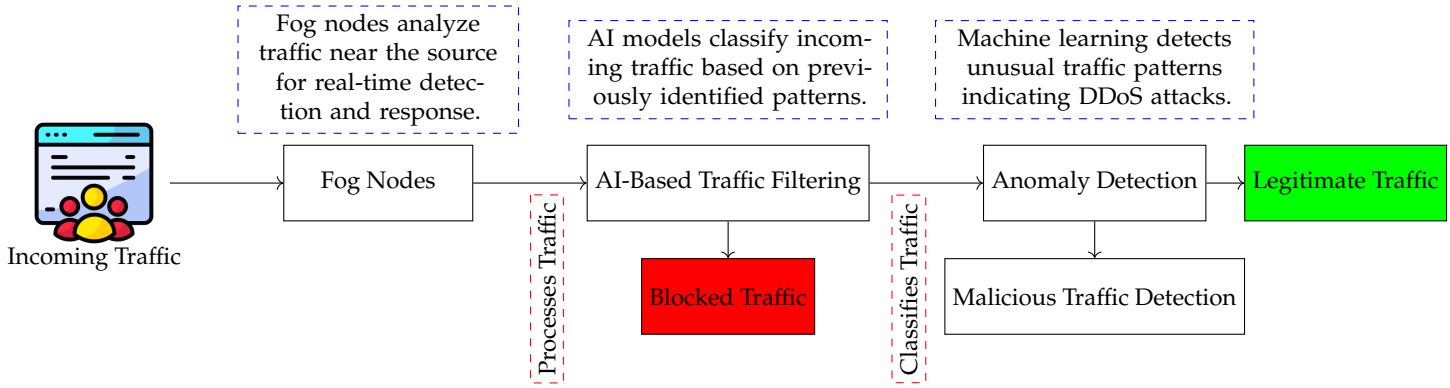

Figure 3 AI-based DDoS Mitigation in Fog Environments: Traffic Filtering and Anomaly Detection

exhaustion, and application-level disruptions. Each of these vectors exploits inherent weaknesses in fog nodes, their limited computational resources, storage capacity, and network bandwidth. The distributed and heterogeneous nature of fog nodes further complicates the defense against DDoS attacks, as each node may be vulnerable to different types of attacks depending on its specific configuration, hardware, and location within the network topology (Kunal *et al.* 2019).

Network flooding attacks are one of the most common forms of DDoS attacks in fog computing. Attackers attempt to inundate the fog network with a massive volume of data packets, overwhelming the bandwidth available to the fog nodes. This form of attack, often executed through protocols such as UDP flooding, SYN flooding, or DNS amplification, targets the limited network capacity of fog nodes, leading to congestion and subsequent service degradation. Since fog nodes typically operate with constrained bandwidth to prioritize localized data processing, a network flooding attack can quickly incapacitate their ability to handle incoming and outgoing traffic, effectively disrupting communication between the fog nodes, end-users, and the cloud infrastructure.

Unlike centralized cloud data centers, fog nodes are generally low-power devices with limited computational capabilities, such as embedded systems, IoT gateways, and edge routers. Attackers can exploit this constraint by generating resource-intensive requests that quickly exhaust the fog nodes' CPU, memory, or storage resources. When a fog node becomes overloaded, it can experience performance degradation or complete system failure, resulting in the disruption of services that depend on the node for local processing, storage, or data forwarding. This form of attack is devastating in real-time applications, such as smart cities or autonomous vehicles, where service delays or outages can lead to critical failures or safety risks (Huang *et al.* 2017).

Application-level DDoS attacks target specific applications running on fog nodes rather than the underlying network infrastructure. These attacks exploit vulnerabilities in the applications themselves, such as unoptimized code, inefficient algorithms, or unprotected endpoints, to consume an excessive amount of computational resources. For instance, attackers may target smart city sensors that rely on fog nodes for data aggregation and processing, overwhelming the application with spurious requests or malformed data that cause it to malfunction or become unresponsive. This form of attack is difficult to mitigate because it operates at a higher layer in the network stack, making traditional network-based defenses less effective.

Given the scale, diversity, and distributed nature of fog environments, the use of AI-based techniques has become essential in combating DDoS attacks. Traditional DDoS mitigation approaches, which rely on fixed traffic rules or manual interventions, are often inadequate in the dynamic and resource-constrained landscape of fog computing. Machine learning (ML) and artificial intelligence (AI) provide the ability to detect and respond to DDoS attacks in real-time, significantly reducing the potential damage that can be inflicted on fog nodes and their associated services. AI techniques can analyze patterns of network traffic, user behavior, and resource consumption to differentiate between legitimate users and malicious actors.

Traffic filtering is one AI-driven strategy that is effective in fog environments. By training machine learning models to recognize common patterns associated with DDoS attacks, such as unusually high traffic volumes or abnormal packet headers, fog nodes can automatically classify incoming traffic and block malicious packets before they reach their intended target. These models are often trained on large datasets of historical network traffic, allowing them to identify both known attack vectors and emerging threats. Traffic filtering is especially important in fog computing, where nodes must make rapid decisions with minimal computational overhead to prevent service degradation.

Anomaly detection is another AI-based technique that is widely used for DDoS mitigation in fog computing. Unlike traditional signature-based detection systems, which rely on predefined rules to identify threats, anomaly detection models are trained to recognize deviations from normal traffic patterns. When a DDoS attack begins, the surge in malicious traffic often causes noticeable shifts in the statistical properties of network traffic, such as increased packet loss, abnormal latency, or sudden spikes in bandwidth usage. Machine learning algorithms can detect these anomalies in real-time and trigger automated defense mechanisms, such as rate limiting or traffic rerouting, to mitigate the attack's impact on the fog network.

Collaborative defense mechanisms represent a more advanced AI-based strategy for mitigating DDoS attacks in fog computing. In these systems, multiple fog nodes share threat intelligence and defensive strategies in real-time, creating a decentralized and adaptive security framework. By pooling their resources and knowledge, fog nodes can develop a collective defense strategy that is more robust and resilient than the sum of its parts. AI algorithms play a crucial role in this process by facilitating the rapid exchange of information between nodes and optimizing the overall defense strategy based on the latest data. For example, if one node detects an emerging attack pattern, it can immediately notify other nodes in the network,

allowing them to preemptively adjust their defenses. This form of collaborative defense is valuable in fog computing, where the distributed nature of the architecture makes it difficult to deploy centralized security controls.

To illustrate the role of AI in DDoS mitigation in fog computing, consider the following table, which summarizes various AI-based techniques and their specific applications in defending against different types of DDoS attacks:

The ability to rapidly detect and mitigate DDoS attacks is critical in fog computing due to the time-sensitive nature of many applications that rely on these systems. For instance, in smart city environments, fog nodes process data from a variety of sensors and IoT devices, enabling real-time decision-making for tasks such as traffic management, energy distribution, and public safety. A DDoS attack targeting these nodes could lead to widespread service disruptions, undermining the reliability of the entire smart city infrastructure. Therefore, AI-driven defense mechanisms must operate with minimal latency and computational overhead to ensure the continuous availability of fog services (Guan *et al.* 2018).

Reinforcement learning represents another promising AI-driven approach to DDoS mitigation in fog computing. In this method, fog nodes learn to adapt their defense strategies over time by interacting with the network environment. Using reinforcement learning algorithms, fog nodes can identify which defensive actions are most effective against specific types of attacks, allowing them to refine their responses based on evolving attack patterns. For example, a fog node might initially implement a simple traffic filtering rule but gradually improve its defense strategy by learning from the outcomes of previous attacks. This adaptive approach enables fog nodes to stay ahead of attackers, who are continually developing new techniques to bypass traditional security measures.

2. Unauthorized Access and Intrusions

Unauthorized access and intrusions represent significant threats within fog computing, where attackers seek to manipulate, disrupt, or exfiltrate sensitive data by exploiting vulnerabilities in fog nodes. These intrusions can also facilitate more extensive infiltration of the network, enabling adversaries to compromise additional nodes, spread malware, or disable critical services. Fog computing, with its decentralized nature and reliance on heterogeneous devices, presents numerous attack surfaces (Alharbi *et al.* 2018) (Stojmenovic *et al.* 2016). The security challenges are magnified by the limited computational resources available on many participating devices, making the prevention and detection of unauthorized access a complex and urgent issue.

Several prevalent techniques are employed by attackers to gain unauthorized access to fog computing systems, each exploiting specific vulnerabilities inherent in the architecture. These techniques are effective due to the real-time processing demands, resource constraints, and distributed nature of fog computing, which create ample opportunities for exploitation.

A major form of attack is the man-in-the-middle (MitM) attack, which capitalizes on insecure communication channels between fog nodes and connected IoT devices. Since fog nodes frequently communicate over open or weakly encrypted networks, attackers positioned between the devices can intercept, modify, or steal data being transmitted. In fog environments, where data may include sensitive information such as financial transactions, healthcare records, or industrial control signals, MitM attacks can have damaging consequences. Such attacks can persist undetected for long periods, allowing adversaries to subtly alter communications, gradually exfiltrate data, or inject malicious payloads into the network, potentially undermining the integrity of the entire system.

Many IoT devices that form the endpoints of fog computing systems are designed with minimal security capabilities due to their resource limitations. These devices, which may include sensors, actuators, and other peripheral devices, are often deployed in large numbers and operate in diverse and sometimes hostile environments. Once an attacker successfully compromises one of these devices, they can use it as a gateway to access the broader fog network. The compromised device can serve as a launch point for further attacks, allowing the adversary to infiltrate additional fog nodes, propagate malware, or disrupt services. The exploitation of edge devices highlights the fragility of fog computing networks, where the compromise of even a single, seemingly low-priority device can have cascading effects throughout the entire infrastructure.

Many fog nodes, those operating in resource-constrained environments, utilize inadequate or outdated authentication methods. These mechanisms may rely on static passwords, weak encryption, or simplified identity verification protocols, all of which make it easier for attackers to bypass authentication barriers. Once authenticated, the attacker can assume the identity of legitimate users or devices, gaining full access to the system's capabilities and data. In the context of fog computing, where sensitive data and critical services are handled in real-time, the risks associated with weak authentication are magnified. Unauthorized access to fog nodes can lead to severe breaches of data confidentiality, integrity, and availability, undermining the trustworthiness of the entire network.

To address the vulnerabilities associated with unauthorized access, the deployment of AI-based intrusion detection systems (IDS) came as a promising solution for fog computing environments. These systems leverage advanced machine learning and deep learning models to identify suspicious activity and potential intrusions in real-time. Given the diverse range of devices, data types, and communication protocols involved in fog computing, traditional security measures often fall short. AI-based IDS, however, are capable of analyzing vast amounts of data and detecting complex patterns that may indicate an ongoing intrusion or anomaly.

Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have demonstrated considerable effectiveness in identifying abnormal behavior in fog networks. These models can process large streams of data from various fog nodes and detect subtle deviations from established behavioral norms, which may signal the presence of an intrusion. For instance, an IDS may analyze network traffic patterns and identify unusual communication flows between nodes or detect anomalies in data payloads that suggest tampering. Given the real-time nature of fog computing, these systems must operate with minimal latency while maintaining high accuracy in detecting threats.

AI-driven solutions extend beyond detection to also include active prevention measures against unauthorized access. One of the key approaches involves behavioral analysis. Machine learning algorithms are trained on historical data to understand the normal behavior of both users and devices within the fog network. Once this baseline is established, the system continuously monitors ongoing activity, raising alerts or initiating defensive actions when significant deviations from normal behavior are detected. This type of behavioral analysis is effective in fog

Table 4 AI-based Techniques for DDoS Mitigation in Fog Computing

AI-based Technique	Application in Fog Computing
Traffic Filtering	Classification of incoming traffic based on known attack patterns, allowing fog nodes to block malicious packets and reduce congestion.
Anomaly Detection	Real-time identification of unusual traffic patterns that signal the start of a DDoS attack, enabling automated defense responses such as traffic rate limiting or rerouting.
Collaborative Defense	Coordination of multiple fog nodes to share threat information and optimize collective defense strategies using AI algorithms.
Reinforcement Learning	Adaptive defense strategies learned over time by fog nodes to respond to evolving DDoS attack patterns.

Table 5 Comparison of Traditional vs. AI-driven DDoS Mitigation Techniques

Technique	Traditional Approach	AI-driven Approach
Traffic Classification	Rule-based filtering based on static signatures	Dynamic classification using machine learning models trained on historical attack data
Anomaly Detection	Manual identification of abnormal traffic patterns	Automated detection using real-time machine learning algorithms
Response Adaptation	Fixed defensive measures, often predefined	Adaptive defense strategies based on reinforcement learning and real-time feedback
Collaboration	Limited or no information sharing between nodes	Real-time collaborative defense mechanisms facilitated by AI algorithms

computing, where the diversity of connected devices makes it difficult to rely on static security rules or signatures to identify threats.

In addition to behavioral analysis, real-time authentication systems enhanced by AI offer another layer of protection against unauthorized access. These systems employ continuous authentication methods that go beyond traditional login credentials. For example, AI-based systems can authenticate users based on their interaction patterns, such as typing speed, touch behavior, or mouse movements. By continuously monitoring and authenticating users and devices throughout a session, the system can detect unauthorized access attempts that may occur after an initial login or when a legitimate user's credentials have been compromised.

Another crucial aspect of AI-driven security solutions is the facilitation of threat intelligence sharing across fog networks. Given the distributed nature of fog computing, individual nodes may be more vulnerable to isolated attacks. However, when fog nodes collaborate and share information about intrusion attempts, they can collectively enhance the overall security of the network. AI systems can automate the process of threat intelligence sharing, enabling fog nodes to quickly exchange data about suspicious activity, attack patterns, and mitigation strategies. This collaborative approach not only improves the speed and efficiency of threat detection but also enables the network to adapt more rapidly to security threats.

The integration of AI into fog computing security is not without its challenges. Training machine learning models to accurately identify threats requires access to large amounts of data, which may not always be readily available in a fog environment. Moreover, adversaries are continually developing more sophisticated techniques to evade detection by AI-based systems, necessitating ongoing updates to algorithms and models. Despite these challenges, AI-driven intrusion detection and prevention systems represent a significant advancement in protecting fog computing infrastructures from unauthorized access.

To further illustrate the efficacy of AI in securing fog networks, recent research has demonstrated that hybrid models, combining both supervised and unsupervised learning techniques, can offer improved accuracy in detecting intrusions. Supervised learning models are effective when labeled datasets are available, allowing the system to learn from known attack patterns. Unsupervised learning models, on the other hand, are adept at identifying novel threats by detecting anomalies in data without prior knowledge of attack signatures. By integrating these approaches, fog networks can be fortified against both known and emerging threats, enhancing their resilience to unauthorized access.

3. Data Breaches and Privacy Threats

Fog computing, as a decentralized extension of cloud computing, often operates at the edge of networks, where sensitive data is processed and transmitted closer to the point of generation. This architecture offers significant benefits in terms of low-latency processing and efficient bandwidth utilization, for time-sensitive applications such as healthcare, smart cities, and

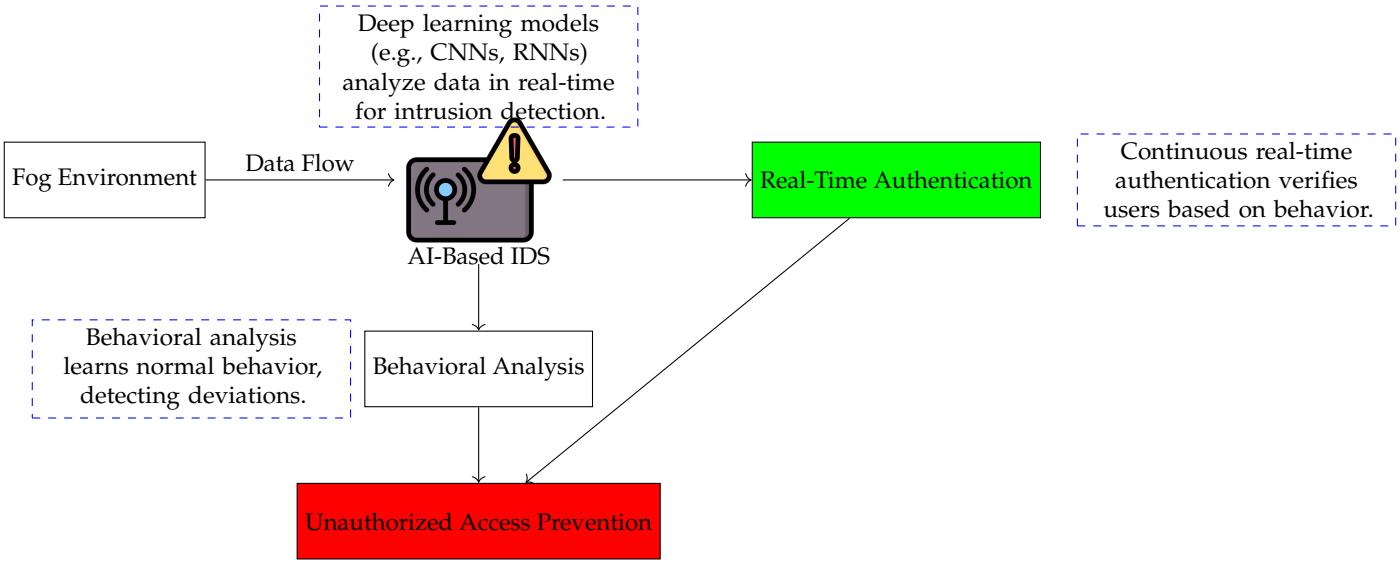

Figure 4 AI-based Intrusion Detection in Fog Networks: Behavioral Analysis and Real-Time Authentication

Intrusion Technique	**Description**
Man-in-the-Middle (MitM) Attack	Exploits insecure communication channels to intercept, alter, or steal data transmitted between fog nodes and IoT devices.
Compromised Edge Devices	Attackers gain access to the fog network by compromising vulnerable IoT devices, using them as gateways for broader network infiltration.
Weak Authentication Mechanisms	Exploits inadequate or outdated authentication methods to gain unauthorized access to fog nodes.

Table 6 Common Intrusion Techniques in Fog Computing

AI-Based Security Solution	**Functionality in Fog Networks**
Behavioral Analysis	Monitors and analyzes the normal behavior of users and devices, raising alerts when deviations occur.
Real-Time Authentication	Continuously authenticates users and devices based on behavioral biometrics, reducing the risk of unauthorized access.
Threat Intelligence Sharing	Facilitates collaboration between fog nodes, allowing the exchange of threat data to enhance overall network security.

Table 7 AI-Driven Security Solutions for Fog Computing

autonomous vehicles. However, the proximity of fog nodes to IoT devices and other edge components presents critical vulnerabilities. A breach in a fog network could lead to the exposure of personal or organizational data, potentially violating privacy regulations like the General Data Protection Regulation (GDPR) or the California Consumer Privacy Act (CCPA), and compromising mission-critical systems. The decentralized nature of fog computing, combined with the diversity of devices and nodes involved, complicates efforts to secure data and prevent breaches, underscoring the need for advanced techniques to mitigate these risks.

Data breaches in fog computing environments can occur under various scenarios, each exposing weaknesses in the architecture or security protocols. These breaches can arise from insufficient encryption, poor access controls, or direct manipulation of data integrity. Each scenario illustrates the vulnerabilities inherent in fog computing, when deployed in resource-constrained environments where security measures are often deprioritized in favor of performance or efficiency (Bhat and Kavasseri 2023).

One of the primary concerns is insufficient encryption. Encryption is a fundamental security measure that ensures data remains secure during transmission and storage. However, many fog nodes, especially those with limited computational resources, may rely on weak or outdated encryption algorithms, or worse, no encryption at all. In these cases, attackers can easily intercept data while it is being transmitted between fog nodes and IoT devices. This lack of robust encryption becomes especially problematic when the data being transmitted contains sensitive information, such as medical records, financial data, or proprietary industrial information (Stojmenovic and Wen 2014). A

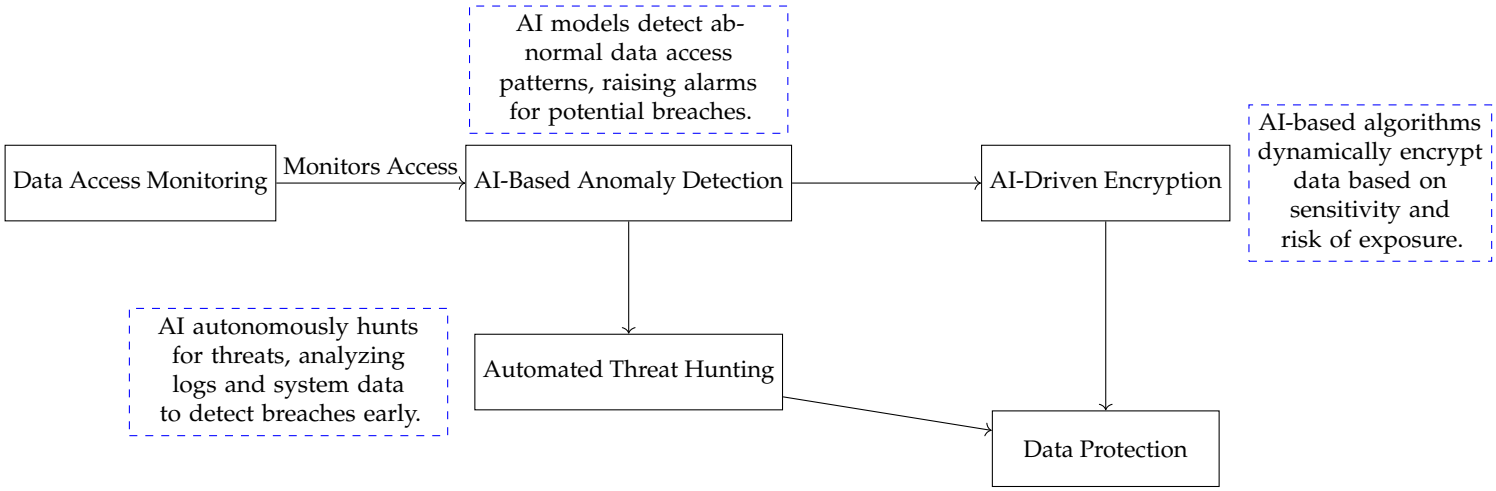

Figure 5 AI Techniques for Preventing Data Breaches: Anomaly Detection, Automated Threat Hunting, and AI-Driven Encryption

breach that exploits weak encryption could allow adversaries to access, copy, or even alter the data, leading to severe privacy violations and operational disruptions.

Another common breach scenario stems from weak access controls. Access control mechanisms are crucial for ensuring that only authorized users and devices can access data stored or processed at fog nodes. However, many fog networks employ inadequate access control protocols, in edge environments where simplicity and low cost often take precedence over security. Without strong authentication and authorization procedures in place, unauthorized users can gain access to sensitive data, including personal information, intellectual property, or critical operational data. The consequences of weak access controls are far-reaching, as they allow attackers to not only steal data but also potentially take control of fog nodes, further compromising the integrity of the entire network (Rahman and Wen 2018).

Data integrity attacks represent another significant threat in fog computing environments. In this scenario, attackers target the integrity of the data being processed at fog nodes, altering it to introduce false information into the system. Such attacks can have catastrophic consequences, in applications where accurate data is critical to decision-making. For example, in healthcare monitoring systems, compromised data could lead to incorrect diagnoses or treatments, endangering patient lives. Similarly, in traffic control systems, data integrity attacks could disrupt the flow of traffic, leading to accidents or gridlock. These attacks are insidious because they may not always be immediately apparent, allowing incorrect or manipulated data to propagate through the network and influence decisions or actions based on false information.

As traditional security mechanisms often struggle to cope with the dynamic and distributed nature of fog computing, artificial intelligence (AI) has emerged as a powerful tool for identifying and preventing data breaches. AI's ability to process vast amounts of data in real-time, detect patterns, and adapt to new threats makes it well-suited to the unique challenges of securing fog networks.

One of the key applications of AI in preventing data breaches is anomaly detection in data access. AI models can be trained to learn normal patterns of data access across the fog network, identifying legitimate behaviors of users and devices interacting with the system. By continuously monitoring these patterns, AI-based systems can detect and flag anomalies that may indicate potential breaches. For instance, if an unauthorized user attempts to access restricted data or if a legitimate user exhibits behavior that deviates from their established pattern (such as accessing a large amount of sensitive data at an unusual time), the AI system can raise an alarm or take automated corrective actions. This capability is valuable in fog computing environments, where the diversity of devices and users makes it challenging to define static security rules.

In addition to anomaly detection, automated threat hunting is another area where AI significantly enhances the security of fog networks. AI systems can autonomously analyze system logs, network traffic, and other relevant data to identify early indicators of potential breaches. By continuously scanning for suspicious activity, AI can identify threats before they escalate into full-scale data breaches. Automated threat hunting is especially useful in fog environments where manual monitoring of all nodes and devices is impractical due to the sheer scale and distributed nature of the network. AI can sift through large volumes of data far more efficiently than human analysts, allowing for faster detection and response to potential breaches.

Common encryption methods, while effective, can be computationally expensive and may not be feasible for all fog nodes, those with limited processing power. AI-based algorithms can enhance these methods by dynamically adjusting the level of encryption based on the sensitivity of the data and the context in which it is being processed. For example, an AI system might apply stronger encryption to highly sensitive data, such as medical records, while using lighter encryption for less critical information. Similarly, AI can enable real-time data masking, where sensitive data is obscured or anonymized depending on the access level of the requesting device or user. These dynamic security measures help ensure that data remains protected without overburdening the computational resources of fog nodes.

AI-driven solutions offer significant advantages in securing fog networks, but they also come with their own set of challenges. One issue is the need for large datasets to effectively train AI models. In fog computing environments, especially in highly distributed networks, obtaining and labeling sufficient data for training can be difficult. Moreover, AI systems are not immune to adversarial attacks, where malicious actors deliber-

Breach Scenario	Description
Insufficient Encryption	Fog nodes with weak or no encryption are vulnerable to data interception during transmission, allowing attackers to access sensitive information.
Weak Access Controls	Without robust access control mechanisms, unauthorized users can gain access to sensitive data or control fog nodes, leading to data theft or system compromise.
Data Integrity Attacks	Attackers manipulate the data being processed, introducing false information into critical applications, potentially causing serious operational disruptions.

Table 8 Common Data Breach Scenarios in Fog Computing

ately introduce misleading data or exploit weaknesses in the AI's decision-making processes. As a result, ongoing refinement and adaptation of AI models are essential to maintaining their effectiveness against evolving threats.

Hybrid AI models, which combine supervised and unsupervised learning techniques, are emerging as a effective approach in preventing data breaches in fog networks. Supervised learning requires labeled data and is useful for detecting known attack patterns, while unsupervised learning can detect new, unknown threats by identifying anomalies in the data. By integrating these two approaches, hybrid models offer a more comprehensive solution for securing fog environments. Supervised learning helps the system recognize and respond to established threats, while unsupervised learning ensures that novel threats are also detected, even in the absence of explicit attack signatures.

4. Malware and Ransomware

Malware and ransomware pose significant threats to fog computing environments, because fog nodes often have limited computational resources, restricting their ability to run complex security software. These malicious programs can infiltrate fog nodes, enabling attackers to control or spy on them, and in the case of ransomware, lock down critical services by encrypting data until a ransom is paid. The decentralized nature of fog computing, where a large number of nodes are distributed across the network, increases the vulnerability of the system. An attack on one node can spread rapidly, affecting other nodes and disrupting vital services, such as healthcare, smart grids, and industrial control systems (Puthal *et al.* 2019) (Thota *et al.* 2018).

Fog computing networks, with their interconnected nodes, are vulnerable to malware propagation. Many fog nodes lack robust security protocols, making them susceptible to malicious software. Once a node is compromised, malware can spread laterally through the network, infecting other nodes and devices that rely on the fog infrastructure. This widespread infection can degrade the performance of the entire system, resulting in disruptions to services that depend on the fog network for real-time processing and communication.

One of the major types of malware that poses a threat to fog environments is worm-like malware. Worms are self-replicating programs that exploit vulnerabilities in fog nodes and other connected devices, allowing them to spread without any user interaction. Once a worm infects one node, it can quickly propagate to other nodes in the network, exploiting unpatched systems and outdated security measures. This kind of malware can cause extensive damage, as it often moves undetected, taking advantage of the fog network's decentralized structure. The rapid spread of worm-like malware can lead to widespread service outages, in critical infrastructures like healthcare, transportation, and energy management.

Ransomware attacks are another significant threat in fog computing. In a ransomware attack, the attacker encrypts data stored on fog nodes, rendering critical services inoperable until a ransom is paid to restore access. This type of attack can be devastating in fog environments where real-time data processing is crucial. For example, if a fog node supporting a smart healthcare system is compromised, it could lead to the inability to monitor patient health data in real-time, potentially putting lives at risk. Similarly, an attack on a fog node managing traffic control systems could cause gridlock or accidents. In such cases, organizations may feel compelled to pay the ransom to restore services quickly, further incentivizing these attacks (Stojmenovic and Wen 2014) (Puthal *et al.* 2019).

Given the dynamic and distributed nature of fog computing, traditional security solutions often struggle to keep pace with the evolving threat landscape. Artificial intelligence (AI) provides a more adaptive and scalable solution for detecting and preventing malware and ransomware in fog environments. AI-driven security systems can analyze vast amounts of data in real-time, detect suspicious behavior, and take preemptive action to block threats before they spread throughout the network.

One of the methods AI employs for malware detection is signature-based detection. This technique involves comparing files or network activity against known malware signatures. AI enhances this process by automating the detection and response, allowing it to scan large volumes of data quickly and accurately. However, this method is limited to detecting previously identified malware, making it less effective against new or evolving threats.

To address these limitations, AI can be used to identify zero-day threats by employing behavioral analysis. Unlike signature-based methods, which rely on predefined malware signatures, behavioral analysis involves learning the typical patterns of operation within the network and identifying anomalies that may indicate the presence of malware. For instance, if a fog node begins exhibiting abnormal communication patterns or attempts to access restricted data, AI can flag this behavior as potentially malicious. This approach is useful in detecting novel malware strains or ransomware that have not yet been formally identified.

Fog computing networks consist of multiple, decentralized nodes that must work together to ensure overall network security. AI can enable these nodes to share information about detected malware, allowing the network to respond more quickly to new threats. For example, if one node identifies a suspicious activity, it can alert other nodes to implement similar defenses, thereby preventing the malware from spreading. This collaborative approach helps to create a more resilient fog network, where the system can defend itself more effectively against large-scale malware and ransomware attacks.

AI Technique	Functionality in Data Breach Prevention
Anomaly Detection	Monitors data access patterns and flags suspicious activity that deviates from normal behavior, preventing unauthorized data access.
Automated Threat Hunting	Continuously scans system logs and network traffic to detect early signs of potential data breaches, allowing for preemptive action.
Encryption and Data Masking	Dynamically applies encryption and data masking techniques based on data sensitivity, protecting information from unauthorized access.

Table 9 AI Techniques for Preventing Data Breaches in Fog Computing

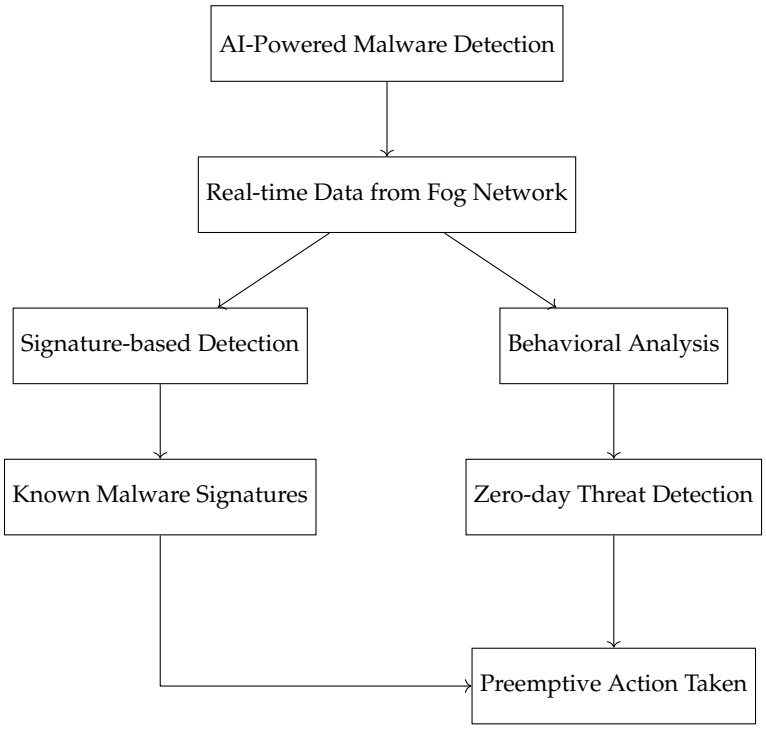

Figure 6 AI-Powered Malware Detection and Prevention in Fog Computing

Malware type	Description
Worm-like malware	Malware that rapidly spreads through interconnected fog nodes by exploiting vulnerabilities in unpatched or outdated systems, causing widespread disruption.
Ransomware	Malware that encrypts data on fog nodes, rendering services inoperable until a ransom is paid to regain access to the data.

Table 10 Types of malware in fog computing environments

AI-powered security solutions are highly effective in combating these types of attacks, but they are not without challenges. One of the key limitations of AI is the need for large datasets to train machine learning models. In a fog computing environment, collecting sufficient data across the distributed network can be difficult, especially when considering the variety of devices and applications involved. Moreover, adversaries are continually developing new methods to bypass AI defenses, which means that AI models must be regularly updated and retrained to maintain their effectiveness.

Despite these challenges, AI remains one of the most promising tools for securing fog computing environments against malware and ransomware. By leveraging techniques such as signature-based detection, behavioral analysis, and collaborative defense, AI systems can provide real-time protection and adapt to emerging threats more effectively than traditional security solutions.

Conclusion

This paper focuses on understanding the security threats inherent to fog computing environments, how their decentralized structure makes them vulnerable to attacks that are less common in centralized cloud systems. Moreover, it explores the challenges of mitigating these threats in a setting where traditional security measures are insufficient, given the distributed and resource-limited nature of fog nodes. Decentralization makes managing security more difficult. Unlike cloud systems with centralized control, fog computing involves geographically dispersed nodes. This distribution complicates the enforcement of consistent security measures.

Heterogeneous nodes add to the challenge. Fog nodes range from powerful servers to resource-constrained IoT devices. Implementing uniform security mechanisms is difficult since many nodes lack the capability to support advanced protection.

High mobility in some environments, like vehicles or drones, further complicates security. The constantly changing network topologies create new attack vectors.

Fog computing also increases the system's attack surface. A large number of nodes, each with different vulnerabilities, broadens the range of potential threats.

Low latency and real-time constraints in fog applications, like healthcare monitoring and autonomous driving, limit the extent of security mechanisms. Solutions must operate without introducing delays.

AI-based techniques offer dynamic, scalable security solutions. Real-time threat detection and mitigation, even in resource-constrained environments, address these security challenges effectively. DDoS attacks are a major concern in fog computing. The decentralized architecture and resource limitations of fog nodes make them vulnerable. Attackers overwhelm networks or services with excessive traffic, causing service disruptions.

In fog environments, network flooding attacks are common. Attackers flood the network with traffic, overwhelming the bandwidth of fog nodes. This leads to service outages.

Resource exhaustion is another DDoS method. Fog nodes have limited computational power, and attackers send resource-intensive requests to exhaust these resources, causing performance degradation or system failure.

Application-level DDoS attacks target specific applications on fog nodes. Attackers exhaust application resources, causing malfunctions and service unavailability.

AI plays a crucial role in defending against DDoS attacks in fog networks. Machine learning algorithms analyze network traffic to distinguish between legitimate users and attackers. AI also helps fog nodes adapt and improve their defense strategies by learning from evolving attack patterns.

AI-driven DDoS mitigation strategies include traffic filtering, anomaly detection, and collaborative defense. Traffic filtering allows AI models to classify and block malicious traffic based on identified patterns. Anomaly detection models identify unusual traffic patterns signaling a potential DDoS attack. Fog nodes share threat information in real-time, creating a collaborative defense strategy across the network. Intrusion attacks involve unauthorized access to fog nodes. Attackers manipulate or steal data, disrupt services, and further infiltrate the network.

Common techniques include man-in-the-middle (MitM) attacks, where attackers exploit insecure communication channels between fog nodes and IoT devices. Attackers can intercept or alter transmitted data. Another method involves compromising edge devices. Attackers exploit vulnerabilities in less secure IoT devices to gain access to the entire fog network. Weak authentication mechanisms also pose a risk. Many fog nodes use inadequate authentication, making unauthorized access easier.

AI-based intrusion detection systems (IDS) enhance fog network security by identifying suspicious behavior in real-time. Deep learning models, such as convolutional neural networks and recurrent neural networks, analyze vast amounts of data to detect complex patterns of anomalous behavior.

AI solutions for preventing unauthorized access include behavioral analysis, real-time authentication, and threat intelligence sharing. Machine learning algorithms monitor the behavior of users and devices, raising alerts when abnormal activity occurs. Real-time authentication continuously verifies users based on behavioral biometrics, reducing the risk of unauthorized access. Threat intelligence sharing allows fog nodes to collaborate and respond more efficiently to threats by sharing. Fog computing often processes sensitive data at the edge, making data breaches a significant concern. A breach exposes personal data, violates privacy regulations, and compromises critical systems.

Data breaches in fog computing occur due to insufficient encryption. Many fog nodes do not use strong encryption, leaving data vulnerable to interception. Weak access controls also contribute to breaches. Without proper access control, unauthorized users can access sensitive information stored or processed at fog nodes. Data integrity attacks are another concern. Attackers tamper with the data being processed, feeding false information into applications like healthcare monitoring or traffic control systems.

AI techniques are crucial in identifying and preventing data breaches. Anomaly detection monitors data access patterns and raises alarms when abnormal behavior, like unauthorized access attempts, is detected. Automated threat hunting uses AI to analyze logs and system data to identify potential breaches before they escalate. AI-based algorithms also enhance traditional encryption methods, dynamically encrypting data depending on its sensitivity and risk of exposure. Malware and ransomware pose significant threats to fog computing. Fog nodes often lack the resources to run complex security software. Malware can control or spy on fog nodes, while ransomware locks down critical services until a ransom is paid.

Malware spreads rapidly in fog environments. Worm-like malware exploits vulnerabilities in interconnected fog nodes, spreading through unpatched systems. Ransomware encrypts data on fog nodes, allowing attackers to demand a ransom to restore access to critical services.

AI helps detect malware by identifying known malware signatures through signature-based detection. AI-based systems also detect and block zero-day threats by analyzing behavior, making it essential in defending against previously unknown malware. AI-driven collaborative defense mechanisms enable fog nodes to share malware-related information and coordinate defense strategies across the network. Insider threats arise when authorized users intentionally or unintentionally compromise system security. Managing insider threats in decentralized fog environments is complex. Authorized users often have legitimate access, making it difficult to detect when they become a threat.

Insiders may unintentionally expose systems to risks through negligence or misconfiguration of security settings. Malicious insiders can cause significant damage by leaking data or facilitating attacks, such as distributed denial-of-service (DDoS) attacks.

AI-based techniques like user and entity behavior analytics (UEBA) detect insider threats by monitoring user behavior and comparing it to established norms. Machine learning models identify suspicious actions, such as unusual login times or access to files outside of routine behavior, helping administrators mitigate risks quickly.

References

Abbasi BZ, Shah MA. 2017. Fog computing: Security issues, solutions and robust practices. In: . pp. 1–6. IEEE.

Alharbi S, Rodriguez P, Maharaja R, Iyer P, Bose N, Ye Z. 2018. Focus: A fog computing-based security system for the internet of things. In: . pp. 1–5. IEEE.

Bhat S, Kavasseri A. 2023. Enhancing security for robot-assisted surgery through advanced authentication mechanisms over 5g networks. European Journal of Engineering and Technology Research. 8:1–4.

Dastjerdi AV, Gupta H, Calheiros RN, Ghosh SK, Buyya R. 2016. Fog computing: Principles, architectures, and applications, In: , Elsevier. pp. 61–75.

Dsouza C, Ahn GJ, Taguinod M. 2014. Policy-driven security management for fog computing: Preliminary framework and a case study. In: . pp. 16–23. IEEE.

Guan Y, Shao J, Wei G, Xie M. 2018. Data security and privacy in fog computing. IEEE Network. 32:106–111.

Huang C, Lu R, Choo KKR. 2017. Vehicular fog computing: architecture, use case, and security and forensic challenges. IEEE Communications Magazine. 55:105–111.

Khan S, Parkinson S, Qin Y. 2017. Fog computing security: a review of current applications and security solutions. Journal of Cloud Computing. 6:1–22.

Kunal S, Saha A, Amin R. 2019. An overview of cloud-fog computing: Architectures, applications with security challenges. Security and Privacy. 2:e72.

Lee K, Kim D, Ha D, Rajput U, Oh H. 2015. On security and privacy issues of fog computing supported internet of things environment. In: . pp. 1–3. IEEE.

Mukherjee M, Matam R, Shu L, Maglaras L, Ferrag MA, Choudhury N, Kumar V. 2017. Security and privacy in fog computing: Challenges. IEEE Access. 5:19293–19304.

Mutlag AA, Abd Ghani MK, Arunkumar Na, Mohammed MA, Mohd O. 2019. Enabling technologies for fog computing in healthcare iot systems. Future generation computer systems. 90:62–78.

Ni J, Zhang K, Lin X, Shen X. 2017. Securing fog computing for internet of things applications: Challenges and solutions. IEEE Communications Surveys & Tutorials. 20:601–628.

Puthal D, Mohanty SP, Bhavake SA, Morgan G, Ranjan R. 2019. Fog computing security challenges and future directions [energy and security]. IEEE Consumer Electronics Magazine. 8:92–96.

Quy VK, Hau NV, Anh DV, Ngoc LA. 2022. Smart healthcare iot applications based on fog computing: architecture, applications and challenges. Complex & Intelligent Systems. 8:3805–3815.

Rahman G, Wen CC. 2018. Fog computing, applications, security and challenges, review. International Journal of Engineering & Technology. 7:1615–1621.

Stojmenovic I, Wen S. 2014. The fog computing paradigm: Scenarios and security issues. In: . pp. 1–8. IEEE.

Stojmenovic I, Wen S, Huang X, Luan H. 2016. An overview of fog computing and its security issues. Concurrency and Computation: Practice and Experience. 28:2991–3005.

Thota C, Sundarasekar R, Manogaran G, Varatharajan R, Priyan M. 2018. Centralized fog computing security platform for iot and cloud in healthcare system, In: , IGI global. pp. 365–378.

Wang Y, Uehara T, Sasaki R. 2015. Fog computing: Issues and challenges in security and forensics. In: . volume 3. pp. 53–59. IEEE.

Yi S, Hao Z, Qin Z, Li Q. 2015a. Fog computing: Platform and applications. In: . pp. 73–78. IEEE.

Yi S, Qin Z, Li Q. 2015b. Security and privacy issues of fog computing: A survey. In: . pp. 685–695. Springer.

Zhang P, Zhou M, Fortino G. 2018. Security and trust issues in fog computing: A survey. Future Generation Computer Systems. 88:16–27.