Adaptive Fog Computing Architecture Based on IoT Device Mobility and Location Awareness

Narek Naltakyan
Institute of Information and Telecommunication Technologies and Electronics, National Polytechnic University of Armenia, Armenia

Keywords: Fog Computing, IoT Mobility, Location Awareness, Adaptive Architecture, Dynamic Orchestration, Edge Computing.

Abstract: In the modern world, we are increasingly seeing the involvement of IoT devices in our daily lives, whether in the form of sensors, smart homes, smart cars, or several other devices and equipment. Fog computing is a concept that is intended to stimulate the development of such systems. However, a problem arises when the operation of a given device or equipment depends on the location. This paper considers a new adaptive fog computing architecture that dynamically manages IoT device mobility through intelligent location awareness and geographic area orchestration. In the proposed system, devices are divided into two parts according to location: static and dynamic, which allows optimizing resource allocation and service approach. The architecture features include dynamic geographic area division, adaptive scaling of fog nodes, and the use of a smart location service that can work both in the cloud and in the fog, depending on the number and density of user requests. The system automatically scales infrastructure based on real-time demand and the geographic distribution of devices, while maintaining quality of service.

1 INTRODUCTION

With the rapid development of technology, IoT devices have become widespread in areas such as smart homes and cities, autonomous vehicles, factories, and health monitoring systems. Current forecasts indicate that by the end of 2025, 75 billion IoT devices will be connected, creating an unprecedented volume of data that requires real-time processing and context-aware responses. Traditional cloud-centric architectures face serious challenges in supporting this scale of distributed, mobile, and latency-sensitive applications. Fog computing as a paradigm was created to address these challenges by bringing computing resources closer to data sources and end-users. However, existing implementation schemes assume network topologies and static device locations, failing to adapt to the dynamic nature of modern IoT device deployments, where device locations are variable and cannot be considered constant, changes in network conditions, and, in addition, differences in service requirements over time and geography.

1.1 Problem Statement

When using fog computing architectures, we face several important issues when it comes to mobile IoT devices:

- **Static Resource Allocation:** Existing systems pre-allocate resources based on patterns, rather than dynamically adapting to the geographic location of devices and changes in their location.
- **Location Agnostic Service Management:** In existing architectural approaches, all devices are considered static in terms of location, i.e., they do not change location over time.
- **Limited Scalability Mechanisms:** Another issue is the lack of mechanisms for dynamically scaling the infrastructure based on real-time demands and geographic service distribution.
- **Inefficient Mobile Device Handling:** In existing systems, there is no separate distinction between static and dynamic IoT devices in terms of location, which leads to additional costs and insufficient support for high mobility applications.

Naltakyan, N.
Adaptive Fog Computing Architecture Based on IoT Device Mobility and Location Awareness.
DOI: 10.5220/0013783900004000

In *Proceedings of the 17th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2025) - Volume 2: KEOD and KMIS*, pages 500-505
ISBN: 978-989-758-769-6; ISSN: 2184-3228

1.2 Research Objectives

This research aims to develop an adaptive fog computing architecture with the following primary objectives:

- Design an intelligent location-aware fog orchestration system that dynamically manages fog nodes based on real-time device distribution and demand patterns.
- Implement different strategies depending on the type of IoT device deployment or location, which will allow you to optimize resource utilization while maintaining service quality across diverse mobility scenarios.
- Implement infrastructure scaling mechanisms, which should be adaptive and adjust the geographic coverage of fog nodes based on the number of requests and the required resources.

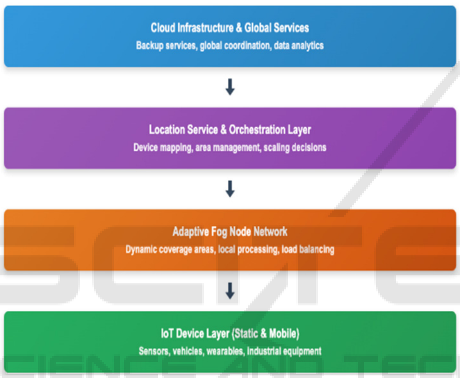

Figure 1: Hierarchical Adaptive Fog Computing Architecture.

- **Cloud Layer:** Provides global coordination and backup services.
- **Location Service:** Intelligent orchestration and device mapping.
- **Fog Layer:** Distributed processing nodes with adaptive coverage.
- **IoT Layer:** Heterogeneous devices with varying mobility patterns.

1.3 Key Contributions

This scientific paper includes several significant contributions to fog computing research and practice:

Novel Architectural Framework: The architectural model discussed in this article is a hierarchical location-aware fog computing architecture that intelligently orchestrates resources based on IoT device mobility patterns and geographic services distribution.

Dynamic Area Management Algorithm: We have developed an algorithm that adaptively changes the geographical area under the responsibility of a fog node based on real-time requirements. In this architecture, 3 algorithms for distributing responsibilities are proposed:

1. Using a load balancer and dividing requests into two or more nodes.
2. Increasing the scope of responsibility if the number of requests is less than X.
3. Adding a new fog node and dividing the geographical area under the responsibility of the nodes if the count of requests is greater than X.

Differentiated Mobility Services: A service differentiation strategy has been developed, where the key argument is the type of IoT device location - static or dynamic - that optimizes the system performance, taking into account the specific requirements of the devices.

Comprehensive Performance Validation: Results of extensive tests are provided, which show a significant performance improvement over existing static fog computing approaches across multiple evaluation metrics and deployment scenarios.

2 RELATED WORKS

2.1 Fog Computing Fundamentals

Fog computing builds on cloud computing by moving it to the network's edge. This enables distributed computing, meaning that data processing, storage, and network services are performed closer to IoT devices and end users, rather than just on centralized cloud servers. Bonomi and his colleagues were the first to introduce the idea of fog computing as an approach aimed at solving common problems associated with cloud systems: latency, network bandwidth, and reliability. The OpenFog Consortium, now part of the Industrial Internet Consortium, has created standardized models that describe fog computing as a system architecture that is distributed across multiple layers of cloud, fog computing, and IoT devices.

Recent progress in fog technologies has emphasized optimizing how resources are used, service deployment, and combining with new technologies such as 5G networks and artificial intelligence. However, much of the current research is based on stable and immobile networks and doesn't address the changing challenges posed by mobile IoT devices.

2.2 Mobility-Aware Fog Computing

Limited research has addressed mobility considerations in fog computing architectures. Qayyum et al. proposed a mobility-aware hierarchical fog framework for Industrial IoT that considers device mobility in fog node placement decisions. Their work focuses primarily on industrial environments with predictable mobility patterns rather than general-purpose mobile IoT scenarios.

Mahmud et al. developed a mobility-aware fog computing framework that uses prediction algorithms to anticipate device movement and proactively allocate resources. While this approach shows promise, it lacks comprehensive geographic area management and does not address dynamic infrastructure scaling based on mobility patterns.

Singh et al. conducted a comprehensive survey of mobility-aware fog computing in dynamic networks, identifying key challenges including mobile node utilization, service continuity, and resource optimization. Their analysis reveals significant gaps in current approaches, particularly in adaptive infrastructure management and location-aware service orchestration.

2.3 Location-Aware Computing Systems

In the world of information technology, location awareness has been widely studied in the context of mobile computing, from purely context-aware applications to location-based services. But while traditional location-aware systems rely on the location of individual devices, they do not focus on the systematic geographic optimization of distributed resources.

Kalaria et al. have developed adaptive context-aware access control systems that use location information to make security decisions in IoT environments. However, even in this case, there is no consideration of resource management or dynamic infrastructure scaling, which can be problematic when there are a large number of users and/or data flows.

The concept of Location-Aware Phase Frames, studied by Ahmad et al., proposed a lightweight Location-Aware Phase Frame (LAFF) for QoS optimization. However, the work focuses on the selection of phase nodes based on the user location, but lacks dynamic area management and adaptive scaling mechanisms.

2.4 Dynamic Resource Discovery vs. Adaptive Geographic Orchestration

There are significant differences between our proposed architecture and existing dynamic resource discovery (DRD) approaches in the field of fog computing. Although both address resource management challenges, they operate at fundamentally different architectural levels and address different problem domains.

Dynamic Resource Discovery Scope: Existing DRD research focuses on runtime discovery and allocation of computational resources within fixed infrastructure topologies. Xu et al. developed DRAM for load balancing through "static resource allocation and dynamic service migration" to existing fog nodes. Similarly, FogBus2 implements dynamic resource discovery mechanisms to assist new entities in joining established fog systems. These approaches work reactively within predetermined infrastructure boundaries.

Architectural Variation: Our research addresses fundamental, broader problem areas. Our proposed architecture addresses “How should I optimally organize the geographical coverage” and “When should I expand or contract the service areas”, while DRD answers “Which fog nodes are available” and “How should I distribute the load among the available nodes”.

Complementary Technologies: Our proposed architecture and DRD extensions can be used together to enable DRD to identify existing fog nodes, while our architecture can use dynamic resource discovery as a component technology within a broader geographic orchestration framework. Our location service can use DRD mechanisms to identify existing fog nodes, while our area management algorithms determine optimal geographic boundaries and scaling decisions.

Novel Geographic Intelligence: Our research introduces geographic area management as a primary optimization dimension, which existing DRD approaches do not address. Current resource discovery focuses on computational resource allocation without considering the spatial organization of fog infrastructure or the geographic optimization of service coverage areas.

3 SYSTEM ARCHITECTURE

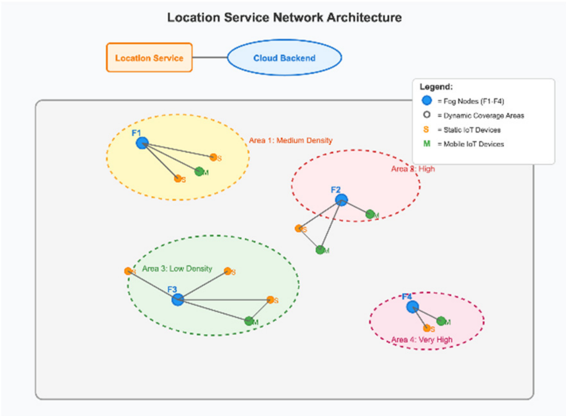

Figure 2: Dynamic Geographic Area Coverage with Fog Nodes and IoT Devices.

Key Features:

- **F1:** Medium density area with balanced static/mobile devices.
- **F2:** High-density area with contracted coverage for efficiency.
- **F3:** Low-density area with extended coverage for broad service.
- **F4:** Very high-density area with minimal coverage, ready for scaling.

3.1 Architectural Overview

Our adaptive fog computing architecture implements a three-tier system: IoT devices<=>fog infrastructure<=>cloud resources. An orchestration is performed through an intelligent location service that provides load balancing, area assignment among fog nodes, and dynamic resource management.

The main innovation of the architecture is the dynamic allocation of responsibility areas across geographic areas to fog nodes based on real-time device distribution, demand patterns, and infrastructure capacity. This enables a fog environment that adapts to real-time requirements while maintaining quality of service guarantees across different IoT deployment scenarios.

Key Architectural Components:

1. **Location Service Layer:** Provides centralized intelligence for device-to-fog mapping, geographic area management, and infrastructure scaling decisions
2. **Fog Orchestration Layer:** Manages distributed fog nodes, handles service placement, and coordinates resource allocation across geographic areas
3. **Device Management Layer:** Implements differentiated services for static and dynamic IoT devices with optimized communication protocols
4. **Adaptive Scaling Controller:** Monitors system performance and triggers infrastructure scaling events based on demand patterns and service quality metrics

3.2 Location Service Architecture

The location service is the cornerstone of the above architecture, enabling intelligent zoning and rezoning over time. This service provides communication between IoT devices and the fog infrastructure, informing IoT devices over time whether their responsible fog node has changed or not, while continuously optimizing the zoning.

Adaptive Service Deployment: The location service itself applies an adaptive deployment strategy based on the number and density of requests. It is initially deployed in the cloud and runs from the cloud for low-density deployments, but migrates to distributed fog nodes as requests grow.

Geolocation Management: The service maintains real-time state information, including device locations, fog node capabilities, current area boundaries, and historical demand patterns. This enables decisions to be made for area adjustments and scaling events.

Query Optimization: For dynamic devices that require frequent location updates, the service includes a query optimization algorithm that attempts to perform predictive caching based on historical data, direction, and speed of movement.

3.3 Location Service Architecture

The dynamic area management system represents a novel contribution that enables fog nodes to automatically adjust their geographic coverage areas based on real-time demand characteristics and service quality requirements.

Area Sizing Algorithm: The coverage area of each fog node is dynamically scaled. When there are few requests, it is expanded to a specified maximum size that is allowed for that fog node, and is reduced when the number of requests exceeds the specified size. This algorithm balances the load on the fog nodes and optimizes resource utilization.

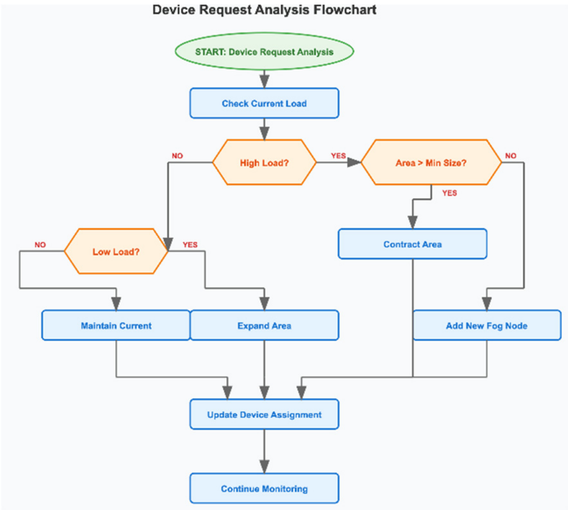

Figure 3: Area Adaptation Algorithm Flow.

Minimum Area Constraints: The system has a minimum area size constraint, when the area of responsibility of a given fog node shrinks to a defined minimum, but the number of requests does not decrease, then another fog node and load balancer are added, added, and both fog nodes are responsible for the given area.

Border Coordination: Neighboring fog nodes coordinate area boundary adjustments to avoid gaps in coverage and service continuity

3.4 Differentiated Device Management

The architecture is based on device classification and management strategies that allow optimizing system performance by segmenting devices according to location characteristics.

Static Device Protocol: Devices that do not change their location during operation will be called static in our architecture, such as sensors, fixed infrastructure, and permanent installations.

Dynamic Device Protocol: Devices that change their location during operation will be called dynamic in this architecture. These are: vehicles, wearable devices, portable sensors. I also find it noteworthy to note that the frequency of requests increases depending on the speed of the device's movement, since the device needs to "know" which node is responsible for a given area.

Predictive Assignment: For devices with predictable paths, such as public transit, delivery vehicles, commuters, the system uses predictive fog node assignment to minimize handover delay and improve service continuity.

4 DYNAMIC ORCHESTRATION ALGORITHMS

4.1 Geographic Area Optimization Algorithm

The geographic area optimization algorithm continuously modifies the coverage areas of fog nodes to balance quality of service, efficient resource use, and infrastructure efficiency. It operates in real time based on metrics such as request rate, response latency, and fog node load levels.

Mathematical Formulation:

For each fog node i, the optimal area size A_i is determined by:

A_i = f(R_i, C_i, L_i, Q_target)
Where:

- R_i = request frequency in area i
- C_i = computational capacity of fog node i
- L_i = average latency from area boundary to fog node i
- Q_target = target service quality threshold

Area Adjustment Decision Logic:

IF (R_i > threshold_high AND A_i > A_min):
  REDUCE area size by scaling factor α
ELIF (R_i < threshold_low AND utilization_i < utilization_target):
  EXPAND area size by scaling factor β
ELIF (A_i ≤ A_min AND R_i > threshold_critical):
  TRIGGER fog node addition

4.2 Location Service Placement Algorithm

The location service placement algorithm selects where to deploy location service components based on client distribution, request patterns, and network topology characteristics.

Cloud vs. Fog Placement Decision:

IF (client_density < threshold_fog AND query_frequency < threshold_freq):
  DEPLOY location service in cloud
ELSE:

DEPLOY location service in fog infrastructure IMPLEMENT regional service distribution

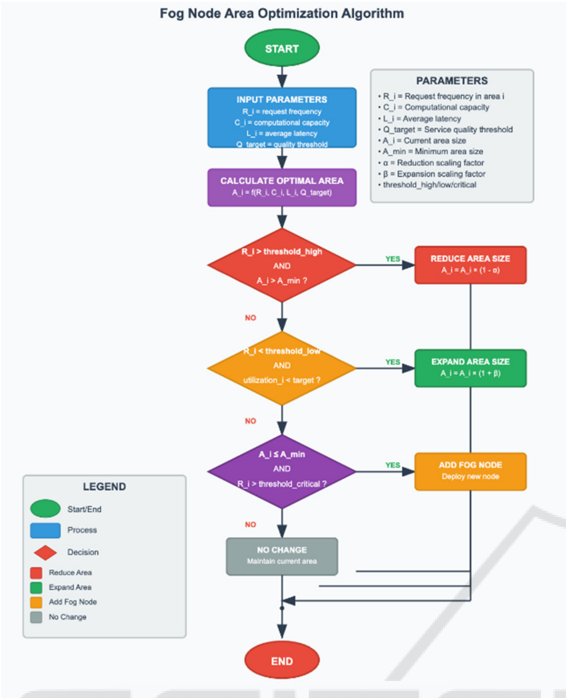

Figure 4: Area optimization algorithm.

4.3 Device Assignment Optimization

Static Device Assignment:

fog_node = SELECT fog_node_i WHERE: device_location ∈ coverage_area_i AND capacity_available_i > service_requirements AND latency(device, fog_node_i) < latency_threshold

Dynamic Device Assignment:

fog_node = SELECT fog_node_i WHERE: predicted_location ∈ coverage_area_i AND handoff_probability < handoff_threshold AND service_continuity_score > continuity_threshold

REFERENCES

Ahmad, N., Hussain, M., Riaz, N., Subhani, F., Haider, S., Alamri, K. S., & Alasmary, W. (2021). Enhancing SLA compliance in Cloud computing through efficient resource provisioning. Future Generation Computer Systems, 118, 365-380.

Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012). Fog computing and its role in the internet of things. In Proceedings of the first edition of the MCC workshop on Mobile cloud computing (pp. 13-16). ACM.

Kalaria, R., Gupta, A., Singh, R., & Sharma, P. (2020). Adaptive context-aware access control framework for IoT environments. IEEE Access, 8, 42462-42475.

Mahmud, R., Kotagiri, R., & Buyya, R. (2018). Fog computing: A taxonomy, survey and future directions. In Internet of everything (pp. 103-130). Springer.

Qayyum, T., Malik, A. W., Khattak, M. A. K., Khalid, O., & Khan, S. U. (2018). FogNetSim++: A toolkit for modeling and simulation of distributed fog environment. IEEE Access, 6, 63570-63583.

Singh, S., Sharma, P. K., Yoon, B., Shojafar, M., Cho, G. H., & Ra, I. H. (2020). Convergence of blockchain and artificial intelligence in IoT network for the sustainable smart city. Sustainable Cities and Society, 63, 102364.

Xu, R., Wang, Y., Cheng, Y., Zhu, Y., Xie, Y., Sani, A. S., & Yuan, D. (2021). Improved particle swarm optimization based workflow scheduling in cloud-fog environment. In International Conference on Business Process Management (pp. 337-347). Springer.