Title	自律分散モバイルシステム環境におけるAAAの実現
Author(s)	髙田, 敦生
Citation	
Issue Date	2026-03
Type	Thesis or Dissertation
Text version	author
URL	https://hdl.handle.net/10119/20537
Rights	
Description	Supervisor:宇多　仁, 先端科学技術研究科, 修士(情報科学)

Implementation of AAA in Autonomous Distributed Mobile System Environments

2310085 Atsuki Takata

Modern mobile communication systems have evolved into large-scale and highly complex infrastructures that are indispensable to both daily life and industrial activities. As mobile networks increasingly serve as social and economic foundations, stable and continuous service operation has become a critical requirement. With the ongoing expansion and diversification of mobile systems, it is now evident that improving overall availability and flexibility requires strong guarantees of service continuity, particularly under failure or partition scenarios.

Service continuity in mobile systems is fundamentally supported by two key functional domains Authentication, Authorization, and Accounting(AAA) and mobility management. In this context, service continuity refers to a state in which (1)information regarding who used which services and to what extent is guaranteed by AAA mechanisms, (2)terminal mobility, including transitions between base stations, is correctly tracked by mobility management functions, and (3)communication paths are successfully established and maintained based on these pieces of information. Thus, ensuring service continuity in mobile systems requires both AAA and mobility management to operate correctly and consistently. From this perspective, a mobile system itself can be regarded as a large-scale authentication and state management architecture.

In current standard 5G Core Networks (5GC), UE (User Equipment) context information, such as authentication state and session state, is managed by Network Functions (NFs) including the Access and Mobility Management Function (AMF) and the Session Management Function (SMF). These NFs are implemented following a microservice-based architecture. During registration and data session establishment, the Core Network (CN) processes UE requests received via the Radio Access Network (RAN), retrieves subscriber information, performs authentication and authorization, and generates various context states required for mobility and session management. Each NF stores and manages its own context locally and maintains consistency through inter-NF message exchanges.

As a result, normal communication services can be provided only when context consistency across NFs is preserved. If inconsistencies occur, additional control procedures are triggered to resolve them, increasing inter-NF communication. Such behavior may further amplify congestion and potentially lead to service degradation or disconnection. The RAN and CN are connected through a carrier-managed transport network, and all UE requests traverse the RAN before being processed by the CN. Within the CN, subscriber information is retrieved and processed, and context states are incrementally generated and updated. Authentication, mobility management, and session control are all executed based on these context states. Because each NF manages context locally, maintaining inter-NF consistency is a prerequisite for stable service operation.

Subscriber information obtained from centralized databases such as the Unified Data Repository (UDR), as well as UE context information maintained by NFs, are both essential to service continuity. Subscriber information mainly consists of static data, including authentication keys and service profiles, and is used to verify the legitimacy of a UE. In contrast, UE context information contains dynamic state, such as authentication progress and session status, and is frequently updated as the UE moves. Due to its dynamic nature, context information requires rapid access and timely updates while maintaining consistency in order to ensure service continuity. Therefore, the method used to manage UE context information has a significant impact on system performance and availability.

We further analyze existing context management mechanisms in standard 5G Core Networks in comparison with conventional approaches. Through this comparative analysis, we discuss the limitations of current context management designs and derive the requirements for an appropriate UE context management system in modern mobile networks.

first work has addressed the availability of subscriber information through mechanisms such as the Replication Layer approach. The Replication Layer improves availability by placing cache servers between the Unified Data Repository (UDR) and the Core Network (CN), enabling replication and distributed management of subscriber information. When the UDR becomes unavailable, subscriber data can be retrieved from the Replication Layer, thereby maintaining service continuity.

In this study, we focus on improving the availability of UE context information. Based on an analysis of communication failures observed in mobile networks and existing approaches to context availability, we propose a distributed UE context management scheme aimed at enhancing service continuity under failures and network partitions.

The proposed Context Share method realizes distributed UE context management using a Distributed Hash Table (DHT). Context Share adopts a locality-aware, two-layer hierarchical DHT architecture consisting of Regional Clusters and a Core Cluster. This design is inspired by hierarchical DHT approaches that balance locality and scalability. Each Regional Cluster is organized on a per AD-RAN basis and manages UE contexts within the corresponding AD-RAN domain. The Core Cluster aggregates multiple Regional Clusters and enables UE context sharing across AD-RAN domains.

With this hierarchical structure, handovers occurring within an AD-RAN can be completed entirely inside a Regional Cluster. Inter-AD-RAN handovers are processed through the Core Cluster, which locates and forwards UE context information to the appropriate Regional Cluster. This design localizes most context access and update operations, reducing unnecessary long-distance communication and centralized processing. The feasibility of this approach has been confirmed through both system design and implementation.

We evaluated the Context Share from three complementary perspectives: operational overhead under normal conditions, communication continuity under failure or partition scenarios, and scalability characteristics derived from a theoretical model. To assess operational overhead, we compared the processing time of major procedures, including registration and session establishment, among the proposed method, a conventional ADM-based system, and a standard 5GC configuration. The results indicate that the proposed method introduces no significant additional delay and does not degrade system availability under normal operation.

To evaluate robustness, we conducted intra-region and inter-region handover experiments under partitioned network conditions. Procedure success rates and U-Plane communication continuity were measured using throughput traces. The proposed method maintained high procedure success rates and continuous U-Plane communication even when failures or network partitions occurred. These results demonstrate that Context Share effectively preserves service continuity in adverse conditions.

To understand scalability behavior, we analyzed how the number of DHT nodes affects load distribution and lookup path length using a mathematical model. The results indicate that Context Share maintains efficient operation even at large network scale and supports scalable UE context management through hierarchical placement.

Overall, the evaluation results demonstrate that the Context Share provides an effective and practical approach to locality-aware distributed UE context management with strong robustness and scalability properties.

As future work, it is necessary to evaluate tighter integration between the Replication Layer and the Context Share in order to achieve efficiency closer to real operational environments. In addition, further enhancements for fault tolerance should be investigated, including mechanisms for replicating UE context across multiple DHT nodes within the Context Share architecture. Such extensions are expected to improve resilience against node failures and network disruptions.

For theoretical model evaluation, parameter settings should be aligned more closely with real deployment conditions. This requires constructing a model that reflects realistic inter-node latency and per-node processing capacity. Developing such a parameterized and environment-aware model will enable more accurate estimation of optimal configurations and performance at practical scale.