AN APPROACH BASED ON THE SDN/NFV ECOSYSTEM FOR MOBILITY MANAGEMENT IN IOT

FULVIO YESID VIVAS CANTERO

Doctoral Thesis in Telematics Engineering

Advisor:
Oscar Mauricio Caicedo Rendón
Ph.D. in Computer Science

University of Cauca
Faculty of Electronic and Telecommunications Engineering
Department of Telematics
Line of Research in Advanced Services of Telecommunications
Popayán, December 2022

FULVIO YESID VIVAS CANTERO

AN APPROACH BASED ON THE SDN/NFV ECOSYSTEM FOR MOBILITY MANAGEMENT IN IOT

Thesis presented to the Faculty of Electronic
and Telecommunications Engineering of the
University of Cauca to obtain the degree of

Doctor in:
Telematics Engineering

Advisor:
Prof. Dr. Oscar Mauricio Caicedo Rendón

Popayán
2022

The final version replaces this page with a copy of the act of public defense signed by the advisors and the evaluation panel members.

I dedicate this thesis to everyone that I love, that supported and loved me through distance and time to come this far.

Especially, I dedicate this thesis to my wife, Claudia Huila, And to my songs Gilber Santiago and Matthew Esteban. whose unconditional love and support have given me the strength to work every day.

To my parents, Leonor Elvira (QEPD) and Jesus Hildebrando, whose dedication and sacrifice have been the base for my progress.

“I would give everything I know for half of what I ignore.”

(René Descartes)

Acknowledgements

I want to thank my family, whose endless love and support have been with me every day, even in the distance and in adversity.

I am very grateful to the distinguished faculty members (professors and staff), and Professor Oscar Mauricio Caicedo Rendón. As my advisor, he shared a detailed guidance and encouragement throughout the course of preparing for and conducting the research. Their belief that it was, indeed, possible to finish kept me going. Certainly, today I am a better researcher because of their influence.

My thanks to the Department of Telematics of the University of Cauca, to all the professors and staff who helped and allowed me to develop my doctoral studies.

Thanks to all my friends in the 116 IPET room at the University of Cauca, for all those off-topic discussions, jokes, and laughs during the stressful days.

Finally, I thank to the InnovAccion Cauca project of the Colombian Science, Technology, and Innovation Fund (SGR- CTI - ID-3848) under Grant 01-2016.

Structured abstract

Background. The Internet of Things (IoT) integrates things, processes, and people that interact with communication networks offering their functionalities, services, and personalized experiences to the users. The IoT environment considers factors such as device constraints, type of information exchanged, mode of operation of devices, type of technology used, high mobility, and high density of end devices. In this way, when the user is continually entering and leaving a network coverage area, the Handover Management (HM) must provide seamless connectivity, enable timely and reliable delivery of services, and offer Quality of Service (QoS). However, current HM approaches present some limitations. First, they consider insufficient criteria, which may lead to unnecessary and frequent handovers due to a partial network view to make appropriate decisions. Second, the high complexity of mechanisms used in network selection can lead to low network performance (decreasing the throughput and increasing the packet loss in the network) and even service disruption. Thus, handover is a significant issue in IoT environments and 5G networks.

Aims. This thesis introduces an approach based on an Software-Defined Networking (SDN) and Network Function Virtualization (NFV) ecosystem for mobility management in IoT. This objective is three-fold: *(i)* proposes an information model for supporting mobility management in IoT by considering an SDN/NFV ecosystem, *(ii)* proposes a communication model for supporting mobility management in IoT by considering an SDN/NFV ecosystem, and *(iii)* build up a prototype per proposed model and evaluate its efficiency, regarding packet loss, handover delay, and the false handover indication to meet the QoS requirements of the end-user applications.

Methods. This thesis proposes an approach based on the SDN/NFV ecosystem for mobility management in IoT. Three main components integrate the proposed mobility management approach. First, a model for network selection based on multi-criteria and supervised learning, named NetSel-RF. This model uses criteria from different sources (network, user preferences, end devices, and applications) and applies Random Forest (RF) as a supervised learning technique to select the best network. Second, a semantic knowledge-based approach for HM, named SIM-Know, improves decision-making during handover. This approach proposes a Semantic Information Model (SIM) and its distributed instantiation in various network entities as Knowledge Base Profile (KBP). SIM and KBP offer local and global knowledge to make contextual and proactive decisions during the handover. Third, an autonomous and cognitive HM approach (ZTHM-5G) optimizes the handover procedure. ZTHM-5G introduces an

Autonomic Knowledge Base Profile (AKBP) based on a Cognitive Closed Loop (CCL) to generate local intelligence, a semantic and goal-oriented communication model that exchanges local and global intelligence among network entities, and Multi-Agent System (MAS) that provides a new distributed, scalable, and personalized HM. These components reduce interactions and signaling message size during handover.

Results. The evaluation of the thesis uses handover evaluation metrics (numbers of handover, ping-pong, and instantaneous throughput) and network performance metrics (throughput, delay, jitter, and packet loss) and compared to HM approaches Strongest Signal First (SSF) and Analytic Hierarchy Process (AHP)-Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS). The results show that the NetSel-RF model reduces the number of handovers, ping-pong, and instantaneous throughput. Furthermore, our model is proactive; because the mobile selects a new AP while connected to another one. On the other hand, SIM-Know improves the throughput, delay, jitter, and packet loss due to its context-aware, cognition, and proactivity capabilities; further, it decreases the number of handovers and instantaneous throughput. In turn, ZTHM-5G result reveals that the handover-related signaling cost is lower than traditional HM approaches.

Conclusions. Multiple criteria from different information sources (network, user preferences, end device, and application) improve decision-making during handover. The representation of these multi-criteria uses SIM to provide context awareness, cognition, and proactivity to the handover procedure. Additionally, learning techniques improve HM to make cognitive and appropriate decisions. Finally, a semantic and goal-oriented communication model (SIM-based, network performance goal, and QoS goals) reduces the exchange of signaling messages in number and size. Therefore, the solutions proposed in this thesis using an Information and Communication model are attractive for HM.

Keywords: Mobility Management, Handover Management, Information Model, Communication Model, Multi-Agent Systems.

Resumen estructurado

Antecedentes. El Internet de las Cosas (IoT) integra cosas, procesos y personas que interactúan con redes de comunicación ofreciendo sus funcionalidades, servicios y experiencias personalizadas a los usuarios. El entorno de IoT considera factores como las limitaciones de los dispositivos, el tipo de información intercambiada, el modo de operación de los dispositivos, el tipo de tecnología utilizada, la alta movilidad y la alta densidad de dispositivos finales. De este modo. Cuando el usuario entra y sale continuamente de un área de cobertura de la red, la gestión de traspaso (Handover Management - HM) proporciona una conectividad perfecta, permite una entrega de servicios oportuna y fiable y ofrece calidad de servicio (QoS). Sin embargo, los enfoques actuales de HM presentan algunas limitaciones. En primer lugar, consideran criterios insuficientes que pueden dar lugar a traspasos innecesarios y frecuentes debido a una visión parcial de la red para tomar decisiones adecuadas. En segundo lugar, la alta complejidad de los mecanismos utilizados en la selección de red puede conducir a un bajo rendimiento de la red (disminución del rendimiento y aumento de la pérdida de paquetes en la red) e incluso interrupción del servicio. Por lo tanto, el traspaso es un problema importante en entornos IoT y redes 5G.

Objetivos. Esta tesis introduce un enfoque basado en un ecosistema de redes definidas por software (SDN) y virtualizacion de funciones de red (NFV) para la gestión de la movilidad en IoT. Este objetivo se divide en tres partes: *(i)* proponer un modelo de información para soportar la gestión de la movilidad en IoT considerando un ecosistema SDN/NFV, *(ii)* proponer un modelo de comunicación para soportar la gestión de la movilidad en IoT considerando un ecosistema SDN/NFV, y *(iii)* construir un prototipo según el modelo propuesto y evaluar su eficiencia, con respecto a la pérdida de paquetes, el retraso del traspaso y la falsa indicación del traspaso para cumplir con los requisitos de calidad de servicio (QoS) de las aplicaciones del usuario final.

Métodos. Esta tesis propone un enfoque basado en el ecosistema SDN/NFV para la gestión de la movilidad en IoT. Tres componentes principales integran el enfoque de gestión de movilidad propuesto. Primero, un modelo de selección de redes basado en múlticriterios y aprendizaje supervisado, llamado NetSel-RF. Este modelo utiliza criterios de diferentes fuentes (red, preferencias del usuario, dispositivos finales y aplicaciones) y aplica Random Forest (RF) como técnica de aprendizaje supervisado para seleccionar la mejor red. Segundo, un enfoque de gestión de traspaso basado en semántica y conocimiento, llamado SIM-Know, para mejorar la gestión de traspaso. Este enfoque propone un modelo de información semántico (Semantic Information Model -

SIM) y su instanciación distribuida en varias entidades de red como perfiles de base de conocimiento (Knowledge Base Profile - KBP). SIM y KBP ofrecen conocimiento local y global para tomar decisiones contextuales y proactivas durante el traspaso. Tercero, un enfoque de gestión de traspaso autónomo y cognitivo (ZTHM-5G) optimiza el procedimiento de traspaso. ZTHM-5G introduce un perfil de base de conocimiento autónomo (Autonomous Knowledge Base Profile - AKBP) basado en un ciclo cerrado cognitivo (Cognitive Closed Loop - CCL) para generar inteligencia local, un modelo de comunicación semántico y orientado a objetivos que intercambia inteligencia local y global entre entidades de la red y un sistema multi-agente (Multi-Agent System - MAS) que provee una nueva gestion de traspaso distribuida, escalable y personalizable. Estos componentes reducen las interacciones y el tamaño de los mensajes de señalización durante el traspaso.

Resultados. La evaluación de la tesis utiliza métricas de evaluación de traspaso (número de traspasos, ping-pong y throughput instantáneos) y métricas de rendimiento de la red (throughput, delay, jitter, and packet loss) y se compara con enfoques de gestión de traspaso (SSF y AHP-TOPSIS). Los resultados muestran que el modelo NetSel-RF reduce la cantidad de traspasos, ping-pong y throughput instantáneos. Además, nuestro modelo es proactivo; porque el móvil selecciona un nuevo AP mientras está conectado a otro. Por otro lado, SIM-Know mejora en throughput, delay, jitter, and packet loss debido a sus capacidades de conocimiento del contexto, cognición y proactividad, además disminuye el número de traspasos y el numero de throughput instantáneos. A su vez, el resultado de ZTHM-5G revela que el costo de señalización relacionado con el traspaso es menor que los enfoques de gestión de traspaso tradicionales.

Conclusiones. Los múltiples criterios provenientes desde diferentes fuentes de información (red, preferencias del usuario, dispositivo finales y aplicación) mejoran la toma de decisiones durante el traspaso. La representación de estos multi criterios usan SIM para proveer conciencia de contexto, cognición y proactividad al procedimiento de traspaso. Adicionalmente, las técnicas de aprendizaje mejoran la gestión de traspaso para tomar decisiones cognitivas y apropiadas. Finalmente, un modelo de comunicación semántico y orientado a objetivos (basado en SIM, y objetivos de desempeño de red y calidad de servicio) reduce el intercambio de mensajes de señalización tanto en número y tamaño. Por lo tanto, las soluciones propuestas en esta tesis utilizando un modelo de Información y Comunicación resultan atractivas para la gestión de traspaso.

Palabras clave: Gestión de movilidad, Gestión de traspaso, Modelo de Información, Modelo de Comunicación, Sistema Multi-Agente.

Contents

List of Figures

List of Tables

Listings

Acronyms

AHP	Analytic Hierarchy Process
AI	Artificial Intelligence
$AKBP$	Autonomic Knowledge Base Profile
ANM	Autonomic Network Management
AP	Access Point
BS	Base Station
CCL	Cognitive Closed Loop
CIM	Common Information Model
DL	Description Logic
HM	Handover Management
IoT	Internet of Things
KBP	Knowledge Base Profile
LLF	Least Loaded First
$MADM$	Multiple Attributes Decision Algorithms
MAS	Multi-Agent System
MD	Mobile Device
ML	Machine Learning
MM	Mobility Management
$mMTC$	massive Machine-Type Communications
NFV	Network Function Virtualization
NN	Neural Network

OWL	Web Ontology Language
RF	Random Forest
RL	Reinforcement learning
$RSSI$	Received Signal Strength Indication
SDN	Software-Defined Networking
SIM	Semantic Information Model
SSF	Strongest Signal First
$TOPSIS$	Technique for Order of Preference by Similarity to Ideal Solution
$URLLC$	Ultra-Reliable and Low-Latency Communication

Chapter 1

Introduction

1.1 Problem statement

The Internet of Things (IoT) integrates things, processes, and people that interact with communication networks offering their functionalities, services, and personalized experiences to the users [1,2]. In an IoT environment, the things consist of heterogeneous devices, static or mobile. In this sense, in IoT, the Mobility Management (MM) provides seamless connectivity [3], enables timely and reliable delivery of services [4,5], and offers Quality of Service (QoS) [6]. In IoT, MM consider factors such as devices constraints (in processing, storage capacity, communication, and energy consumption), type of information exchanged, mode of operation of devices (*e.g.*, sleeping), and type of technology used (*e.g.*, WiFi) [7–10]. Some IoT scenarios in which MM is essential are, first, in a hospital when a patient moves freely, needing continuous monitoring of vital signs across the connected medical devices. As these devices must send data in (near) real-time for early diagnosis and treatment of acute issues, mobility must be guaranteed. Second, in events or conferences, each person with a wearable personalizes his/her information and knowledge services by all the areas visited. Since these wearables transmit essential information for generating customized user experiences, mobility must also be supported.

In the aforementioned scenarios, it is necessary to consider two facts: on the one hand, when a person moves along with a Mobile Device (MD), more than one network may appear; therefore, such MD could disconnect from the current Access Point (AP) and connect to another one. Indeed, in IoT, it is essential to manage connection requests for billions of devices efficiently and reliably, and manage the rapid topological changes caused by the unavailability of some MDs in, for instance, sleep mode operation [3,11,12]. On the other hand, the connection of any MD to the network must adapt dynamically (*i.e.*, real-time interactive services require consistent network capacity) to different APs in the same or different network [13,14] to meet QoS [15,16]. The process that keeps the connection active of an MD when moves from an AP to another one, is named HM. Handover consists of three phases called: handover initiation, network selection, and handover execution [17]. It is noteworthy that the optimization of handover is directly related to meet QoS requirements of end-users [6, 18]. Thus, handover is considered a significant issue in IoT.

In IoT, from a general perspective, the handover presents some related limitations. First, the insufficient criteria to make appropriate decisions about, for instance, when starting?, which AP select? and so on. Second, the high complexity of mechanisms used in AP selection leads to low network performance and even service disruption. The insufficiency is associated with needing more criteria for a complete network view [19]. The high complexity is associated with interrelating QoS and network criteria, targeting a trade-off to select the best AP [20]. The service disruption is the discontinuity in the delivery of services during handover (e.g., handover timing impacts delay-sensitive applications) [21, 22]. This thesis focuses on problems of insufficient criteria to make decisions about handover and the service disruption during handover because they affect network connections and their performance directly and, as a result, QoS as a whole.

In the literature, the insufficient criteria to make handover decisions have been addressed using single-criterion, estimated-criterion, and multi-criteria techniques. The works [23–25] have used a single-criterion (i.e., RSSI) that is directly related to QoS. The drawbacks of these works are false handover indication, wrong selection of the network, and handover decisions realized from a constrained perspective (*e.g.*, wireless links) instead of the overall network state. In the works [16, 26, 27], the authors have applied the technique estimated-criterion using a window-based mechanism that continually evaluated the criterion by improving its estimation. However, these works share shortcomings related to a significant number of messages to estimate the criterion and induce potential delays or losses in the network. In turn, the works [20, 28–32] have used multi-criteria to obtain several criteria that represent an entire network view. Nonetheless, these works based on multi-criteria increase network signaling traffic, do not provide details about how the criteria are collected, and criteria may be in conflict between themselves.

In the literature, the service disruption during handover has been addressed using techniques such as fast-handover, seamless handover, and virtualization of network and device. The works in [33,34] have used fast-handover in networks with multiple APs to change the channel between multiple RF channels and continue with their connection without service disruption. These works increase collisions when many MDs are active simultaneously, and each AP in the network needs information about its neighbors. The works in [6, 35–37] have used a seamless handover based on the SDN and NFV implementing a Virtual AP (VAP) abstraction to control association and re-association processes of an MD with an AP. The drawbacks of these works are the multiple applications running on a single controller make prioritization difficult and the absence of mechanisms implemented in MDs that can gather information from users, end-devices, and applications. Furthermore, these works do not consider QoS negotiation during and after the handover. The work in [38] has used network-device-virtualization to build up the SDN control plane in APs, the SDN data plane in MDs, and the corresponding communication by the OpenFlow protocol. This SDN-based work overloads the communication channel with signaling (i.e., the OpenFlow protocol transports packets of topology discovery among sensor nodes) and lacks cooperation between controllers (e.g., load balancing to neighbor controllers, which have spare resources), leading to a simplified view of the network, and affecting handover performance regarding, for instance, latency and delay.

This doctoral thesis argues that the shortcomings of the insufficient criteria to make decisions about the handover and the service disruption during the handover are related to the lack of an MM approach based on an Information Model and a Communication Model supported by an SDN/NFV ecosystem. An Information Model would assist in establishing a shared characterization of the network, simplify the development of functionalities for MM (e.g., criteria transformation methods) and facilitate the deployment of policies on the network to meet QoS. In turn, a Communication Model would aid in delineating the exchange of information in IoT and reduce network signaling traffic. An SDN/NFV ecosystem would assist in increasing the efficiency and network agility to address the dynamic changes depending on traffic flow, application-specific requirements, and mobility of the devices in the IoT environment from a logically centralized point of view. To sum up, considering the shortcomings above, this doctoral thesis raises the following research question:

How to carry out efficiently mobility management in IoT to meet QoS?

1.2 Hypothesis

To address the research question stated in Section 1.1, this doctoral thesis raises the following hypothesis:

An SDN/NFV ecosystem allows performing mobility management efficiently in IoT to meet QoS. This doctoral thesis argues that an SDN/NFV ecosystem would define a programmable SDN plane by virtualizing the network entities to support MM in IoT. The SDN controller would gather all the necessary information from other components in the system to realize global optimization of MD associations with network monitoring techniques. Network virtualization would enable a dynamic attachment of MDs to multiple networks allowing them to move from one location to another. Also, SDN/NFV would enhance the QoS provisioning capability in MDs connected with APs [39–41].

Global network state awareness comprises the monitoring and gathering of network state information. The global information determines the selection criteria more efficiently to carry out the MD association in a centralized manner. In this sense, an Information Model would facilitate the discovery and regular access to criteria of multiple sources by using a common model at a syntactic and semantic level. Also, this model would provide a well-designed structure to describe the criteria in the resources domain and requirements from MDs. Furthermore, user requirements should be involved in such a model, and a similarity degree between user requests, MDs, and APs should also be provided. The Information Model should also be lightweight to reduce traffic and processing time [42–45]. In the same way, provisioning an efficient and optimized Communication Model would reduce bandwidth needs without affecting related resources such as battery life, energy consumed for processing, and memory size. Besides, this model would improve the QoS provision by reducing the chance of service interruption while keeping network signaling traffic with a short delay, and small cost [2, 46–49].

1.3 Objectives

1.3.1 General Objective

To introduce an approach based on an SDN/NFV ecosystem for mobility management in IoT.

1.3.2 Specific Objectives

- To propose an information model for supporting mobility management in IoT by considering an SDN/NFV ecosystem.
- To propose a communication model for supporting mobility management in IoT by considering an SDN/NFV ecosystem.
- To build a prototype per the proposed models and evaluate its efficiency regarding packet loss, handover delay, and false handover indication to meet the QoS requirements of the end-user applications.

1.4 Contributions

The scientific research process conducted during this thesis led to introduce an approach based on an SDN/NFV ecosystem for MM in IoT. The expected research contributions of this doctoral thesis are to provide a MM approach formed by:

- An information model for supporting HM. This model enables context-aware and multicriteria handover decisions and distributed knowledge base to incorporate cognition in HM.
- A communication model for supporting HM. This model delineates the exchange of local and global intelligence while reducing the interactions and size of the signaling messages.
- A prototype per model and its corresponding evaluation include SIM-Know and ZTHM-5G. SIM-Know improves HM by including SIM that provides local and global knowledge to make contextual and proactive handover decisions. ZTHM-5G optimizes the handover procedure by reducing interactions and the size of the signaling messages using an autonomous and cognitive HM approach from the Autonomic Network Management (ANM) point of view.

1.5 Scientific production

Two published journal papers and one journal paper under revision report the major contributions achieved during this thesis to the scientific community.

- F. Y. Vivas, O. M. Caicedo, and J. C. Nieves, "A Semantic and Knowledge-Based Approach for Handover Management," published in Sensor MDPI, 2021 [50]. Ranking: JCR Q1, SJR Q1, Publindex A1. Contribution: Information model based on semantic for MM.
- D. A. Embus, A. J. Castillo, F. Y. Vivas, O. M. Caicedo, and A. Ordóñez, "NetSel-RF: A Model for Network Selection Based on Multi-Criteria and Supervised Learning," published in Applied Sciences, 2020 [51]. Ranking: JCR[1] Q2, SJR[2] Q2, Publindex[3] A1. Contribution: Network Selection using Machine Learning (ML).
- F. Y. Vivas and O. M. Caicedo, "ZTHM-5G: Zero-Touch Handover Management in 5G," Submitted. Contribution: Communication model for autonomous and cognitive MM from an ANM point of view.

Furthermore, one paper at national conference reports to the scientific community the contributions achieved in collaboration with other researchers. This paper is presented below.

[1]Quartile from Journal Citation Reports (JCR)
[2]Quartile from SCImago Journal Rank (SJR)
[3]Bibliographic index from COLCIENCIAS, Colombia

- F. Samboni, S. Bedoya, F. Y. Vivas, and O. M. Caicedo, "MEC IoT: Monitorización de estructuras civiles en el contexto IoT," published in the proceedings of the 2017 IEEE Colombian Conference on Communications and Computing (COLCOM) [52]. Ranking: H5-index 12. Contribution: the structural health monitoring system in IoT and cloud computing.

Appendix A lists the four published and under revision papers in inverse chronological order.

1.6 Methodology and organization

The research process that guided the development of this thesis is based on a typical scheme of the scientific method [53]. Figure 1.1 depicts the phases of the scientific research process: Problem Statement, Hypothesis Construction, Experimentation, Conclusion, and Publication. Problem Statement, for identifying and establishing the research question. Hypothesis Construction, for formulating the hypothesis and the associated fundamental questions. In addition, this phase aims to define and carry out the conceptual and technological approaches. Experimentation, for testing the hypothesis and analyzing the evaluation results. Conclusion, for outlining conclusions and future works. Note that Hypothesis Construction has feedback from Experimentation and Conclusion. Publication, for submitting and publishing papers for renowned conferences and journals. The writing of the dissertation document also belongs to this last phase.

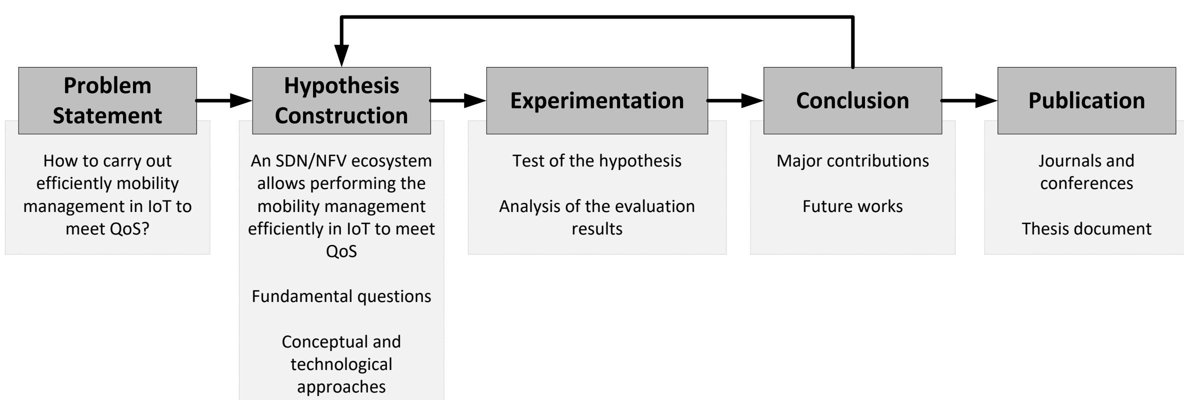


Figure 1.1: Thesis phases

The organization of this document reflects the phases of the methodology.

- This introductory chapter presents the problem statement, raises the hypothesis, exposes the objectives of this thesis, summarizes the contributions, lists the scientific production, and describes the overall structure of this dissertation.

- **Chapter 2** reviews the main concepts and research related to IoT, MM, network management models, SDN, and NFV.
- **Chapter 3** introduces a model for network selection based on multi-criteria and supervised learning.
- **Chapter 4** presents a semantic and knowledge-based approach for HM that introduces an information model and their instances distributed in the network entities.
- **Chapter 5** introduces an autonomous and cognitive HM approach that introduces autonomic agents based on CCL and delineates a semantic and goal-oriented communication model.
- **Chapter 6** presents conclusions about the research question and hypothesis and proposes some future works.

Chapter 2

State-of-the-art

This chapter presents the central concepts of this thesis as follows. The first section introduces a description of MM and especially HM (2.1). The second section provides an IoT overview, discussing the challenges for MM (2.2). The third section presents the models needed to describe network management approaches (2.3). The fourth (2.4) and fifth (2.5) sections describe the SDN and NFV ecosystem. The section 2.6 provides a literature review of HM in wireless and mobile networks, focusing on seminal works based on information and communication models.

2.1 Mobility Management

In MM, an MD moves within a single AP or across many APs. Therefore, the network must provide to each MD a global identifier, location management, and HM [4, 54, 55]. A global identifier is used to recognize an MD in the network. Location management allows the network to track the locations of MDs between consecutive communications. HM keeps the connection active of any MD when it moves from an AP to another one located at the same (horizontal handover) or different network (vertical handover) [56, 57]. The handover process usually involves the transmission of packets, resulting in increased signaling costs in the network. The signaling cost is the mobility signaling overhead incurred during a handover. According to [58–60], the signaling message delivery cost result of computing the product of the number of network hops, the size of the mobility signaling message, and the weighting factors in the network [61]. The more the number of handover occurrences, the higher the signaling overhead.

HM consists of three phases: initiation, selection, and execution [17]. The handover initiation periodically collects different parameters. These parameters are transferred among network entities to discover the network environment and trigger handover. Handover Trigger is when the connectivity between an MD and its current AP drops below a particular level (threshold). The network selection phase determines an appropriate target AP, radio link transfer, and channel assignment to continue connectivity and meet QoS [62]. In the literature, network selection is performed using either a network-centric or a user-centric approach. In a network-centric approach, a centralized entity assigns APs to MDs in the service area. However, in network-centric, all wireless networks are involved and, so, the communication overhead increases significantly. On the other hand, in the user-centric approach, the MD is responsible for running network selection algorithms. However, in this approach, the connection of the MD to an AP affects the network performance because MDs do not know the network load. Furthermore, the high resource consumption of MDs can lead to a decrease in their lifetime.

The handover execution phase establishes sessions after the AP change, allocates new addresses, delivers stored packets, and routes packets [63]. This phase interrupts the data flow to the user due to network change and signaling overhead. This interruption results in the reduction of the user throughput and increases latency [64]. In the literature, there are two kinds of MM protocols [14]. On the one hand, host-based protocols must modify the protocols stack and change the addressing in MDs. On the other hand, network-based protocols need a central entity to know the entire network and manage mobility. Up to date, IoT solutions prefer the network-based MM protocols since improving the handover, signaling, and QoS for real-time applications.

HM includes a control mechanism [5] that can be classified as: Network-Controlled HandOver (NCHO), Mobile-Controlled HandOver (MCHO), Mobile-Assisted HandOver (MAHO) and Network-Assisted HandOver (NAHO). In NCHO, the network starts and controls the handover; operators usually adopt it for load balancing and traffic management. In MCHO—traditionally used by IEEE 802.11 technologies—the MD initiates and controls the handover. In MAHO, the MD helps in the handover process controlled by the network; MAHO is typically used in cellular networks. In NAHO, the network helps with the handover process controlled by the MD; NAHO is intended for heterogeneous wireless networks.

2.2 Internet of Things

An IoT environment (Figure 2.1) represents a smart environment where different kinds of MDs and processes are continuously working in a (or several) communication networks to make inhabitants' lives more comfortable [56, 65]. In such an environment, MDs generate traffic that may exceed the traffic generated by humans. Furthermore, this environment may be organized by different application domains, also called verticals, such as health [9], industrial [8], and agricultural [66]. Each vertical has its data model and authentication method. Indeed, using different data representation models represents a barrier to information exchange and retrieval; note that the traffic produced by MDs is heterogeneous syntactic and semantically.

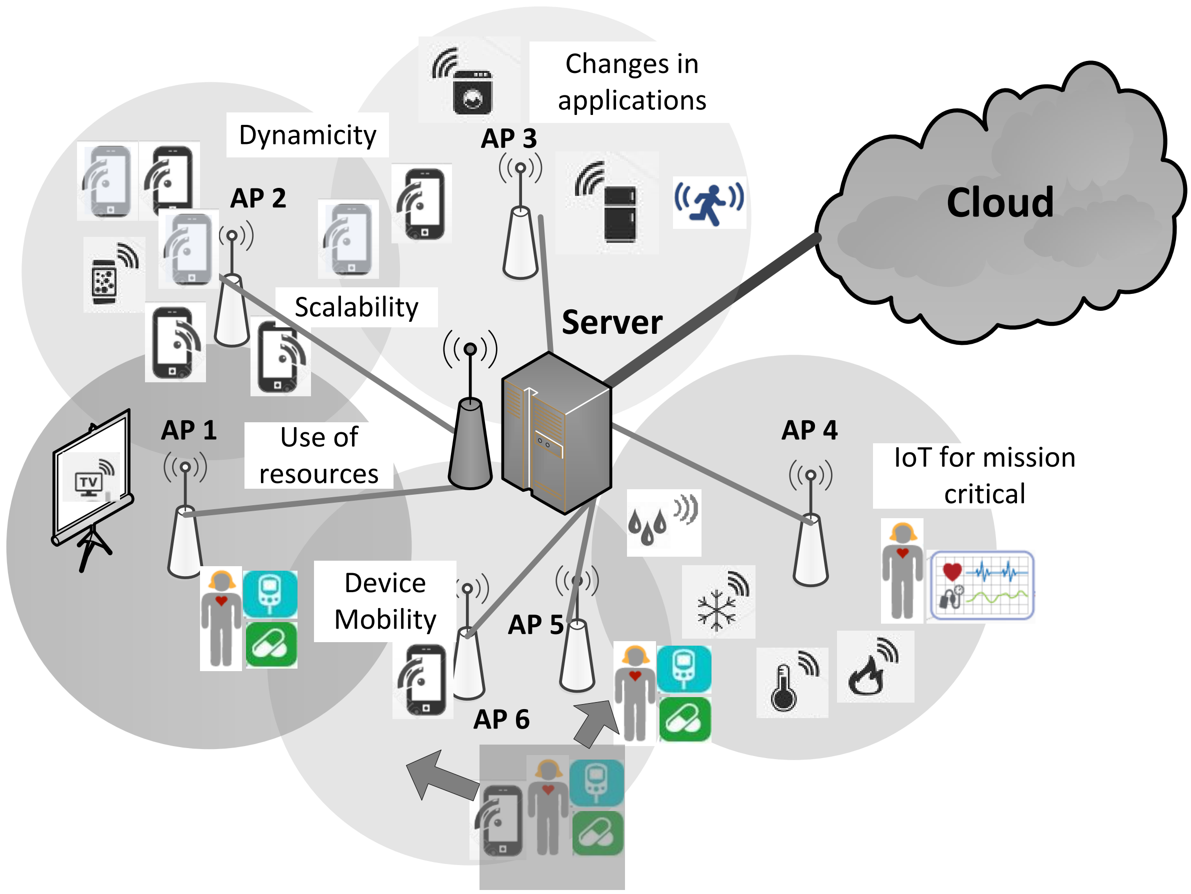


Figure 2.1: IoT environment

Connectivity plays an essential role in IoT since IoT consists of diverse MDs with different network connection times [67]. IoT applications have different requirements, especially regarding response time, identifying two classes of applications. On the one hand, near real-time applications (delay-sensitive like electroencephalography, tractor beam game, and vehicular monitoring). On the other hand, delay-tolerant applications such as video surveillance and object tracking [68]. It is important to highlight that each application needs to be treated separately to offer differentiated QoS levels [69].

An IoT infrastructure consists of many APs, where MDs can connect and disconnect constantly. Thus, the efficient handling of mobility is crucial for the overall performance of any IoT application. Hence, the carrying out of a seamless handover is needed between APs [7] [10]. In IoT, MM aims at achieving service discovery, improving connectivity, and optimizing QoS support for differentiated services in the network. However, because the absence of an information model, the description of MDs and their capabilities are heterogeneous syntactically and semantically so that software agents cannot perform tasks such as automatic discovery and orchestration of devices and services.

2.3 Network Management Models

Network management employs a variety of protocols, tools, applications, and devices for monitoring and controlling network resources. A network management system is described using the Open System Interconnection (OSI) network management model, which comprises, in turn, four significant models: Information, Organizational, Communication, and Functional. The Information Model specifies the information base useful to describe the managed objects and their relationships. This model represents a general abstraction that can be specialized in different domains, regardless of technologies and implementations. The Organizational Model defines the managers, agents, and managed objects. The Communication Model delineates the information exchanged between the managed systems with a complete network view. The Functional Model organizes functional areas of network management [70] [71].

The current approaches use ANM based on a set of CCL (Fig. 2.2) that meet specific intents or policies [72]. CCL controls the status of a managed entity according to an operator-specified desired goal. Therefore, ANM allows the network entities to control their context, adapt to changes in the environment, and self-govern their behavior to achieve specific goals [73].

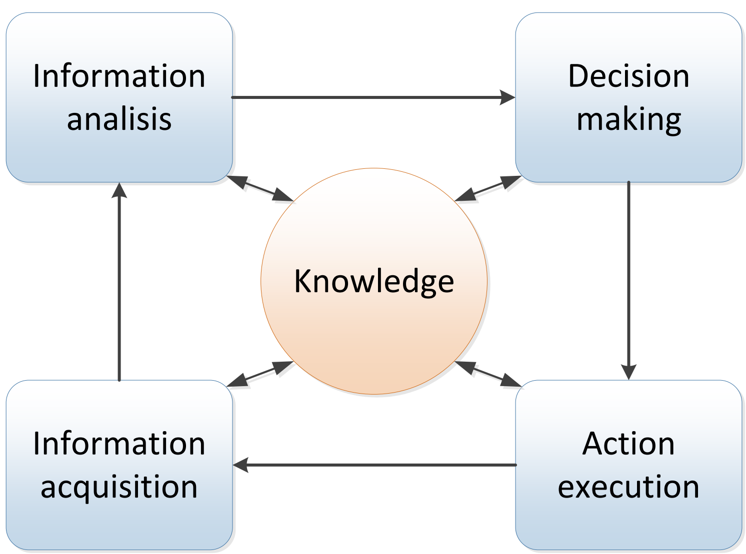


Figure 2.2: Cognitive Control Loops

MAPE (Monitor, Analyze, Plan, and Execute) is the architecture for autonomic computing proposed by IBM [74]. MAPE is a CCL to *Monitor* continuously network entities gathering changes in the network and its environment; *Analyze* the collected data to achieve the goals; *Plan* actions to meet the desired goal reconfiguration if goals cannot be achieved; and *Execute* those actions and observe the results without human intervention.

In MM, an information model characterizes information collected from the network and the interaction between the users and their MDs and services [75]. Furthermore, the design of a common format (*e.g.*, ontologies) would simplify the development of functionalities (e.g., topology discovery), which should allow selecting the best AP and satisfy QoS [43–45]. A communication model delineates the exchange of information in an IoT environment [47, 63]. This model can support MM efficiently by a mechanism to share information in multiple domain networks while keeping signaling traffic with a short delay and small cost.

2.4 Software-Defined Networking

SDN is an emerging network paradigm (Figure 2.3) that allows dynamic and flexible network operations by decoupling the network control plane from the data plane [76]. The control plane implements a logically centralized controller (or network operating system) to simplify policy enforcement and network configuration. SDN achieves network programmability, where the network operator programs the controller to manage devices automatically located at the data plane and optimize the use of network resources [40]. Briefly, SDN focuses on improving network performance regarding network management, control, and data handling. The SDN architecture allows for classifying the IoT applications to define flow-forwarding policies and improve QoS. Furthermore, a dynamic load balancing optimizes the usage of the wireless link to face topology changes caused by MDs. Notably, the centralized control plane of SDN allows for generating optimized strategies to handle mobility events such as the localization of MDs and the user mobility patterns [77].

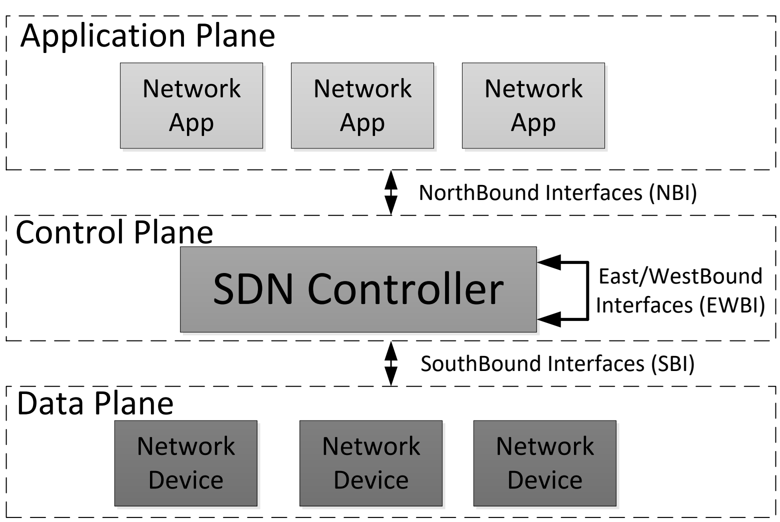


Figure 2.3: SDN architecture

Recently, SDN has been used for supporting handover. For instance, the abstraction of Physical AP (PAP) provides a platform to facilitate a programmable control plane and MM functions. In this sense, an SDN Controller handles each domain of a cluster of MDs, providing visibility of all network traffic. The Logical AP (LAP) is a logical entity that resides in the PAP with an extended SDN abstraction. Furthermore, LAP acts as an SDN agent that enables a dynamic mapping configuration between the controllers and MDs. In particular, LAP assigns a unique BSSID per client (MD) targeting to generate beacon and acknowledgment frames, establishes the initial configuration parameters, such as IP address, MAC address, and SSID, of MDs, and holds the information into OpenFlow tables. The LAP prioritizes critical/control traffic to meet QoS [36].

2.5 Network Functions Virtualization

NFV is an initiative of the European Telecommunications Standards Institute (ETSI), aiming to enhance the delivery of network services by separating network functions from the hardware they run [78] (Figure 2.4). NFV is a network architecture that aims to deploy Network Services (NSs) flexibly and dynamically [39,79]. NSs are complete end-to-end functionalities offered by network operators. These functionalities are delivered by composing Network Functions (NFs) by a process called Network Service Chaining (NSC) or Service Function Chaining (SFC) [80]. Virtualized Network Functions (VNFs) are essential elements based on computing resources.

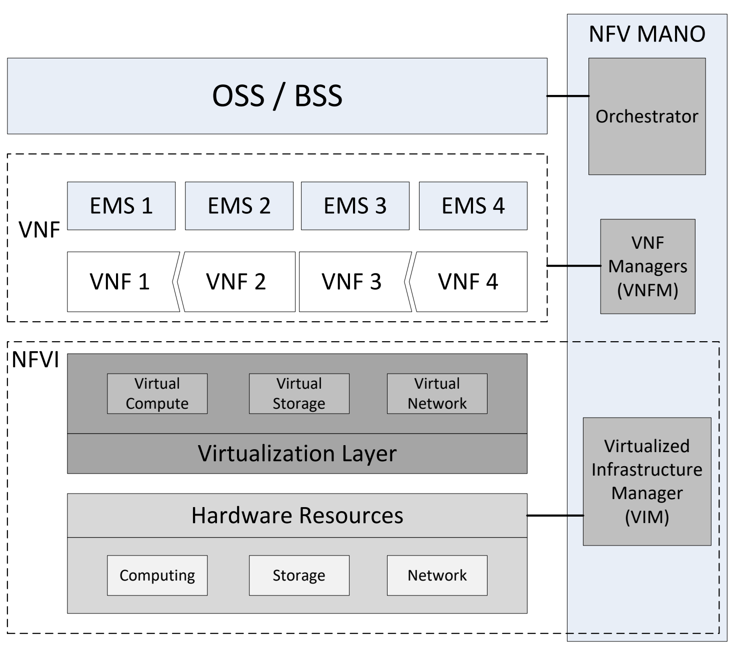


Figure 2.4: NFV architecture

In NFV, the Management and Orchestration Framework (MANO) [37] is responsible for managing NSs, NFs, NSC, and physical and virtual infrastructures. NFV MANO is formed by an Orchestrator for composing software resources and virtualized hardware; a Virtualized Network Functions Manager (VNFM) for managing the lifecycle of VNFs; and a Virtualized Infrastructure Manager (VIM) for virtualizing and managing network resources [78, 79].

NFV is related to the building blocks for virtual networks characterized by highly dynamic network environments like IoT. NFV allows optimizing the allocation of the available network resources to improve the NSs delivery. Furthermore, NFV provides differentiated services to users, enabling diversified QoS for several usage scenarios, such as MM and MANO for SDN. MM involves NF control and orchestration to achieve optimized functionality. NFV in IoT would allow the transition from a network of entities to a network of functions. MANO for SDN is an essential element for monitoring, configuring, and controlling SDN because its use can bring from NFV benefits like scalability, elasticity, and workability [78].

An SDN/NFV ecosystem (Figure 2.5) combines the abstraction of functions by NFV and the abstraction of the network by SDN to increase the efficiency and network agility of IoT applications. This ecosystem presents scalable distribution capabilities of NFV through VNF and virtualized resources depending on traffic flows and application-specific requirements. SDN in this ecosystem allows the configuration flexibility of physical and virtual resources [78, 81].

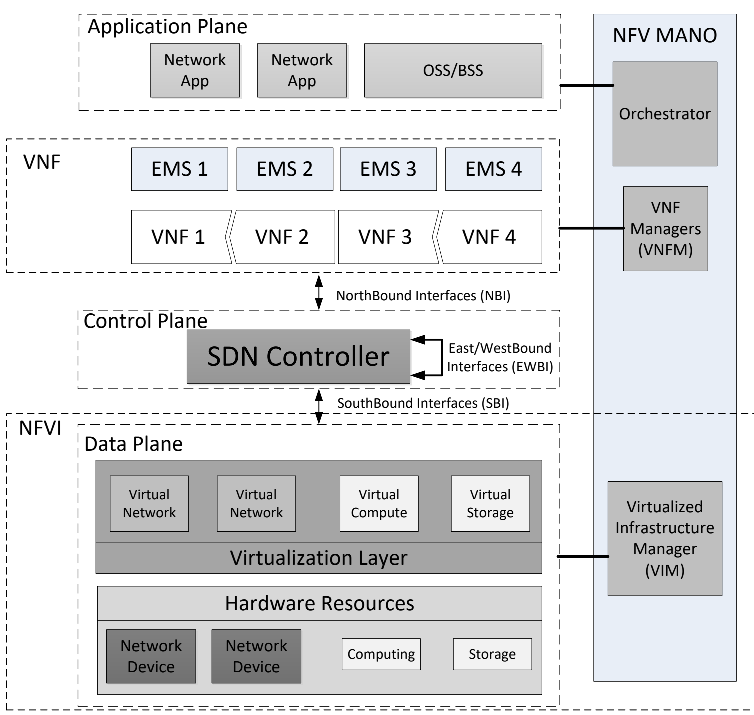


Figure 2.5: SDN/NFV ecosystem

An SDN/NFV ecosystem provides differentiated services to users, enabling diversified QoS for the scenario. In this ecosystem, NFV allows the virtualization of the AP and VNFs, such as optimizing handover, controlling policies for connecting AP and signaling, and making decisions in the handover procedure. Furthermore, SDN implies a logically centralized network control plane, which allows the implementation of sophisticated mechanisms for traffic control and dynamic resource management [36, 37, 40].

2.6 Mobility Management in IoT

This section presents the related work of this thesis divided into Criteria for MM and seamless mobility. The criteria provide a network view to make appropriate decisions. Investigations about seamless mobility are pivotal to providing access to the user to services everywhere, service continuity, and experience contextualized and personalized.

2.6.1 Criteria for Mobility Management

In the works [24, 25, 29], the authors have proposed Simple Additive Weighting (SAW), which processes the data rate of different applications to obtain a predefined handover threshold. In the same direction, the works [24, 25] used the TOPSIS to select an AP based on data rate. In the work [29], the authors proposed the Artificial Bee Colony optimization algorithm that uses as a criterion the Bit Error Rate (BER) to select the optimal AP with minimum handover delay. However, these works share shortcomings related to the significant number of messages needed to estimate the data rate, leading to adding network signaling traffic. Furthermore, the data rate is a criterion from the user perspective that lacks information from the network perspective. Note that an IoT network should adjust to allow some users to associate and meet QoS without degrading network performance.

In the work [16], the authors have proposed as a criterion the bandwidth reservation for maintaining a proper handover policy and allowing the users to migrate from one AP to another one based on their QoS requirements. The drawback of this work is that using a bandwidth reservation criterion by service reduces the number of simultaneous traffic for other service classes resulting in QoS degradation. This thesis argues that a single criterion is insufficient to take a handover decision and meet QoS.

In the work [26], the authors have proposed a linear regression model to select the best criterion to trigger handover. They consider a set of criteria: signal strength, data rates, delay, and associations. The authors conclude that the Signal-to-Noise Ratio (SNR) is the best criterion for a handover trigger. The drawback of this work is related to the linear regression model used, which depends on the quality of collected information. Using a method to transform criteria to normalized values would improve the predictive performance of the handover trigger algorithms.

In [20], the authors have developed a genetic algorithm (GA) that applies the best combination of weights to QoS parameters (network coverage area, battery power requirement, network latency, and monetary cost) to select the best network connection according to user preferences. The proposed GA avoids the slow and massive computations associated with direct search techniques, thus, reducing the computation time. Note that this GA was not evaluated in a real environment and only used estimated values.

In the work [28], the authors have used Markov Decision Process (MDP) to formulate the network selection problem. They consider criteria such as type of service (i.e., a combination of reliability, latency, and data rate), monetary cost from the user perspective, and network conditions parameters such as available bandwidth and network latency. The authors applied a GA to find a set of optimal decisions that ensure the best trade-off between the QoS of the connection to the network (reward function) and overloading the network with signaling traffic (cost function). The drawback of this work is that it does not specify details about measuring parameters on the user side. Furthermore, many iterations of the proposed GA must run in the central controller to select the best AP, and the network-based criteria may conflict with the quality criteria.

In the work [30], the authors have developed a Quantum-inspired Immune Clonal Algorithm to perform horizontal and vertical handover. They consider criteria network-related or user-related, economic or non-economic, objective or subjective, accurate or fuzzy. Also, in this work, the criteria are descriptions based on fuzzy mathematics methods which can adjust to different network conditions. However, these criteria are hard to be accurate, requiring a data model. Note that data models are narrowly related to Information Models.

In the work [31], the authors have used an AP selection algorithm based on the Fittingness Factor (FF) concept. This algorithm considers the provided quality assessment in AP and the required quality assessment in MD as criteria. In [32], the authors have proposed the Network Fittingness Factor metric, which considers QoS requirements by joining total flows, active flows, and bandwidth efficiency (i.e., the current network capacity and the quality of the connectivity). The drawback of this work is that it does not specify details to collect parameters from the end-users. Furthermore, transforming QoS requirements into bit-rate metrics can represent a loss of crucial characteristics to meet QoS.

Table 2.1 presents the investigations related to insufficient criteria to make decisions about the handover, revealing that notwithstanding their contributions, they share some shortcomings: a) incomplete representation of the IoT handover process, b) inappropriate communication management to share information in multiple domain networks; and c) scarcity of a framework that characterizes MM in IoT to meet QoS. This thesis argues to overcome these shortcomings and optimize the IoT handover procedure. First, an Information Model is needed to establish a shared characterization of IoT and simplify the functionalities development for MM (e.g., criteria transformation methods). Second, a Communication Model is required to facilitate MM by sharing information in multiple networks and reducing network signaling traffic. Third, an SDN/NFV ecosystem is demanded to deliver high levels of automation and flexibility to improve real-time monitoring and connectivity of MDs with AP.

Table 2.1: Related work - criteria for mobility management

Paper	Information Model	Communication Model	Single-criterion	Estimate-criteria	Multi-criteria	SDN/NFV
[24, 25]		✓	✓			
[29]		✓			✓	
[16]		✓		✓		
[26]		✓		✓		
[20]	✓				✓	
[28]		✓			✓	
[30]	✓				✓	
[31, 32]					✓	✓
This Thesis	✓	✓			✓	✓

During the thesis, various research papers related to HM were explored based on single-criterion and multicriteria. About those papers, we describe the control mechanism, the method used to make a decision, and the wireless technology used [72, 82]. Table 2.2 presents the reviewed work.

Table 2.2: Related Work - methods and technology for HM

Work	Description	Making-Decision		Wireless
		Control	Method	Technology
[83]	A Software-Defined Networking (SDN) controller uses a fuzzy system to score candidate networks for staying in the current network or connecting to a better one	NCHO	Fuzzy Logic	Fifth Generation (5G)
[84]	An algorithm is proposed to reduce the handovers by multicriteria decision-making algorithms improved with a context-aware and threshold-based scheme	MCHO	TOPSIS, SAW	5G, Long-Term Evolution (LTE), WLAN
[85]	A fuzzy logic and reinforcement learning-based mechanism is introduced to address unnecessary and frequent handovers by adjusting HandOver Margin (HOM) and Time-To-Trigger (TTT)	NCHO	Fuzzy Q-Learning	LTE
[22]	A solution, based on SDN, Binary Integer Linear Programming (BILP), user criteria and network packet error rate data, is proposed to rank candidate Base Stations (BSs) and to enhance the handover selection phase	NCHO	BILP	LTE
[86]	A framework, based on data analytics, context extraction, user profiling and pre-processing contextual information, is presented to score the available BSs and to improve network access selection	NAHO	Fuzzy Logic	5G
[87]	The AP or BS selection is improved by using AHP for weighting selection criteria coming from the user and networks' context and TOPSIS for ranking the available networks	NCHO	AHP-TOPSIS	LTE, WLAN
[88]	A mechanism is proposed for selecting the radio access network that best meets the end-user needs by considering the on/off state and battery level of the mobile device and the available bandwidth in the target and serving network	MCHO	Policy	LTE, WLAN
[89]	A versatile modeling methodology is introduced for evaluating proactive and reactive vertical handover approaches	NCHO	Policy	5G, LTE, WLAN
[90]	Two co-operating algorithms with adaptive thresholds are introduced for performing network selection while avoiding network congestion and meeting user preferences regarding monetary cost, QoS, security and energy consumption	NAHO	Policy	5G
[91]	A multiattribute decision handover making scheme, centered in the triggering phase and based on SDN and Fuzzy Logic, is proposed for increasing the network throughput and reducing unnecessary handovers and total handover delay in femto-access points and device-to-device communications	NAHO	TOPSIS, Fuzzy Logic, AHP	LTE, WLAN

2.6.2 Seamless Mobility in IoT

In the work [33], the authors have proposed a Multichannel VAP (mVAP) to support seamless handover in networks with APs operating on multiple channels. MDs choose new APs based on messages exchanged between APs using Inter-AP Protocol. APs provide MDs with information on possible new APs to change channels and continue with their connections without service disruption. This work increases the collisions when many MDs are active simultaneously because of the overhead generated by handling a VAP per AP, making this solution few scalable. Furthermore, each AP requires a list of its neighbors to send the scan request messages, which may lead to needing more storage capacity per AP.

In [34], the authors have proposed the framework BIGAP with a single global BSSID and different RF channels for all co-located APs. BIGAP consists of two components. The first one resides at APs. It collects statistics and executes BIGAP controller commands (*e.g.,* trigger the RF channel switching in an MD). The second component is the BIGAP controller, which aims to provide a global view of the network state. This controller also facilitates the handover operations between the serving and target AP. Unfortunately, BIGAP needs a sufficiently large number of available channels to make a collision-free channel assignment. Note that since this work uses only network-based criteria, it may cause degradation in the network performance.

In [35], the authors have introduced Odin, a Light Weight AP (LWAP) that runs over a controller to associate or disassociate MDs with APs. LWAP offers a dedicated logical connection per MD with a unique BSSID. Odin ensures seamless handover to reduce the delay. The drawbacks of this work are: *i)* the network attachment processes increase workload in APs, *ii)* the complex task of prioritizing applications in the controller; and *iii)* the handover process depends on a single criterion, which could lead to a load imbalance situation, inhibiting to achieve an optimal network performance.

In [36], the authors constructed a logical AP (LAP) in IEEE 802.11 WLAN using an SDN/NFV abstraction. LAP acts as a VAP, an abstraction of a PAP. This LAP also provides auxiliary network functions like disabling LAP and collecting the neighboring PAP information to provide a gateway between the SDN controller and MDs. The evaluation results of this work evidence a decrease in false handover indications, latency, and ping-pong ratio. However, the handover process incorporates additional workload values in PAP that cause low network performance and avoids satisfying the QoS requirements of users.

In the work [6], the authors have proposed UbiFlow to manage mobility by deploying a network of distributed SDN controllers and OpenFlow switches. The authors implement an optimal assignment algorithm for AP selection on these controllers based on the network status analysis and flow requests. Furthermore, each controller maintains a finger table to achieve critical scalable look-up in this overlay structure. The handover process localizes the destination controller using a quick look-up in its finger table, which further saves the communication cost and improves the efficiency of handover. It is noteworthy that. First, this work has not been evaluated in a real environment. Second, the assignment process triggered at the end of each time window would lead to low network performance.

In [37], the authors have proposed an IoT architecture formed by three layers named, perception layer (i.e., sense and collect data from MDs), network layer (i.e., provide connectivity to MDs using different technologies) and application layer (i.e., IoT applications). The network layer supported by controllers is programmable by SDN and NFV. In particular, an SDN Controller modifies the flow tables of OpenFlow switches to assist the MDs mobility. The data plane virtualizes an IoT gateway to maintain MDs current session continuity. The Virtualized Gateway manages the dynamic attachment of MDs to multiple APs. This architecture does not consider mechanisms in the perception layer that may lead to enrichment MM. Thus, the proposed MM of this work cannot gather valuable information from users and MD, change parameters of association and re-association, and re-configure network access.

In SDN-WISE [38], the authors have proposed a solution based on SDN for Wireless Sensor Networks (WSN). SDN-WISE defines two kinds of nodes. The first one is the sensor node running in the data plane. Second, the sink node represents the gateways between the sensor nodes and the controller running at the control plane. In this solution, a WISE-Visor establishes the communication between the layers using OpenFlow and reduces the exchanged information between sensor nodes and SDN controllers. The drawbacks of this work are. First, implementing the solution is impractical since it requires a high level of programmability, unknowing that MDs have constrained resources. Second, it overloads the communication channel with network signaling traffic.

Table 2.3 presents the investigations related to the service interruption during handover, revealing that notwithstanding their contributions, they share some shortcomings: a) network attachment processes increase workload in APs, b) lack of QoS negotiation during handover; and c) do not specify an exact timing procedure for a handover to minimize service disruptions. Considering these shortcomings, this thesis argues that to optimize the IoT handover process. First, an Information Model is needed to characterize user QoS requirements and to provide a similarity degree between QoS requirements and the available devices/services during and after handover. Second, a Communication Model is required to minimize the inter-packet delay of individual flows, improving the QoS provisioning by reducing the chance of service disruption. Third, an SDN/NFV ecosystem is demanded to provide efficient, proactive, fine-grained, QoS-aware, and seamless MM, manner according to the current connectivity options (i.e., wireless networks (WiFi) and network devices (AP/BS)).

Table 2.3: Related work - seamless mobility in IoT

Paper	Information Model	Communication Model	QoS	Fast-Handover	Seamless Handover	SDN	NFV
[33]		✓	✓	✓			
[34]	✓		✓	✓			
[35]		✓			✓	✓	
[36]		✓			✓	✓	✓
[6]	✓	✓			✓	✓	
[37]		✓	✓		✓	✓	✓
[38]		✓	✓			✓	
This Thesis	✓	✓	✓		✓	✓	✓

HM in 5G networks becomes more complicated with many issues and challenges. 5G combines ultra-dense network scenarios and radio access technologies with short coverage areas increasing network signaling traffic due to frequent handovers. At the same time, the users contain a large MDs number and present high mobility raising the workload within network entities due to many handovers (e.g., massive Machine-Type Communications (mMTC)). Additionally, the growing applications number with increasingly strict restrictions in terms of QoS (e.g., Ultra-Reliable and Low-Latency Communication (URLLC)) impose conditions on HM to reduce the delay from minimizing the signaling generated [92]. Therefore, HM from an ANM point of view offers seamless mobility where users achieve a seamless, contextualized, and personalized experience to access services everywhere [93] [94].

Autonomous and cognitive HM approaches achieve automatic and adaptive management using ML mechanisms to analyze different data from multicriteria and use new network management structures. [95] outlines the handover procedure from an ANM point of view to assess robustness and self-optimization. [96] proposes an autonomous framework with five functionalities: context-aware interface, 5G access point, macro cloud unit, control interface, and enabling platform. The macro cloud drive behaves as an autonomous, self-healing module. This framework uses the Fuzzy AHP (FAHP) to consolidate the criteria weights and the Efficacy Coefficient Method-based TOPSIS (ECM-TOPSIS) technique to select the network.

The work in [97] introduces a CCL and cross-layer access network selection framework. The CCL comprises three modules: perception, decision, and execution to select the access network adaptively. The decision-making uses AHP to calculate the criteria weights and TOPSIS to rank the candidate networks. [98] enables a Software-Defined Wireless Network (SDWN) to manage and control the network autonomously. Decision-making uses multiple criteria and Fuzzy Logic (FL) techniques with adaptive hysteresis values taking into account the Quality Of Experience (QoE). [99] presents an algorithm based on Reinforcement learning (RL) to trigger the handover and select the network. This algorithm uses multicriteria and applies State-Action-Reward-State-Action (SARSA) to learn an optimal handover policy. However, the lack of personalizing the user actions during handover increases the network signaling traffic generated.

2.7 Final remarks

This chapter presented the central concepts of this thesis as well as the related work. From the literature review, we concluded that the investigations based on multicriteria provide a global network view to make appropriate decisions in handover. In turn, the works based on seamless mobility delineate the information exchange during handover to reduce service interruption time. However, the insufficient criteria and the incomplete representation of the handover process drive to wrong selection of the network and handover decisions realized from a constrained perspective. Moreover, network attachment processes increase workload in APs and service disruption time.

Chapter 3

A Model for Network Selection Based on Multi-Criteria and Supervised Learning

HM plays a fundamental role in wireless communications. HM is the process by which a MD maintains an active connection when the user roams from the coverage area of one network to another [100]. HM comprises three phases [101] [102]: initiation, selection, and execution. Handover Initiation collects all information required to identify and determine the neighbor networks, their parameters, and their available services. Network Selection selects the best available network by taking into account diverse parameters and metrics. Handover Execution establishes the connection and releases resources. This chapter focuses on the network selection phase, which is crucial to ensure service continuity, provide QoS, and satisfy (QoE) [103] [104].

Traditional approaches for Network Selection use a single-criterion to decide the next network (*e.g.,* AP or BS) to which MDs must connect as they move [105]. SSF is the most widely used traditional approach, which performs the selection by comparing Received Signal Strength Indication (RSSI) of the current network with available adjacent networks [100]. Other approaches based on a single-criterion use the available bandwidth [106], network load [107], or Signal to Interference Noise Ratio (SINR) [108] to select the network. As these approaches only use one criterion, they ignore important information, leading to instantaneous throughput, unnecessary handovers, and even service disruption [109] [110] [105].

Other approaches use multi-criteria to solve the shortcomings mentioned above. The selection criteria can be classified in Network, User, Device, and Application [111]. Network criteria include parameters that describe the wireless network characteristics (e.g., coverage area and RSSI). User criteria comprise parameters related to user preferences. Device criteria encompass parameters that provide information about the MD, such as battery consumption and location. Application criteria cover parameters related to QoS requirements of real or non-real-time applications, like the bandwidth consumed by such applications. Some multicriteria approaches (*a.k.a.,* Multi-Attribute Decision-Making (MADM) are: SAW, Gray Relational Model, and TOPSIS [103] [112]. These approaches have some shortcomings related to skipping diverse selection criteria, resulting in wrong network selection [105]. Other relevant multicriteria approaches use techniques like mathematical models [113], Neural Network (NN) [111] [114], Fuzzy Logic (FL) [115] [116], and Game Theory [105] to perform network selection. However, these techniques can decrease the throughput during the handover [117]. Hybrid solutions combine multicriteria approaches [118]. However, their major problem is the high computational complexity.

In this chapter, we introduce NetSel-RF, a multicriteria model based on supervised learning (i.e., RL), to overcome the abovementioned shortcomings and efficiently select WiFi networks. To create this model, we constructed a dataset, performed data preparation, carried out feature selection, and applied different supervised ML techniques. The remainder of this chapter is as follows. Section 3.1 introduces the network selection approach based on multicriteria and ML. Section 3.2 presents the efficiency of NetSel-RF in comparison with AHP-TOPSIS and SSF. Finally, Section 3.3 contains some concluding remarks.

3.1 NetSel-RF Model

This section presents our classification model for optimizing the network selection in a WiFi-based environment. In particular, this section presents the motivation, methodology, interpretation of data, data preparation, and model construction.

3.1.1 Motivation

Figure 3.1 shows a user with a MD ($dev1$) moving through the hall. This $dev1$ needs to be always connected to the best network. As mentioned above, traditional handover approaches consider only RSSI for making decisions [119]; however, his approach produces frequent disconnection and inefficient seamless handovers, leading to handover operation failures [105].

Several approaches have addressed these problems using multicriteria considering the handover decision criteria related to QoS and QoE [117]. Nevertheless, current approaches based on multicriteria can perform wrong network selection by skipping one or more criteria related to the network, devices, users, or application parameters. In this chapter, we argue that a supervised learning model based on multicriteria allows the network selection to be optimized by considering the criteria of networks, devices, users, and applications. In this way, such a model can learn the best network to satisfy

QoS and QoE. Additionally, supervised learning allows the new network to be selected proactively. The MDs performs a handover before the network link quality degrades.

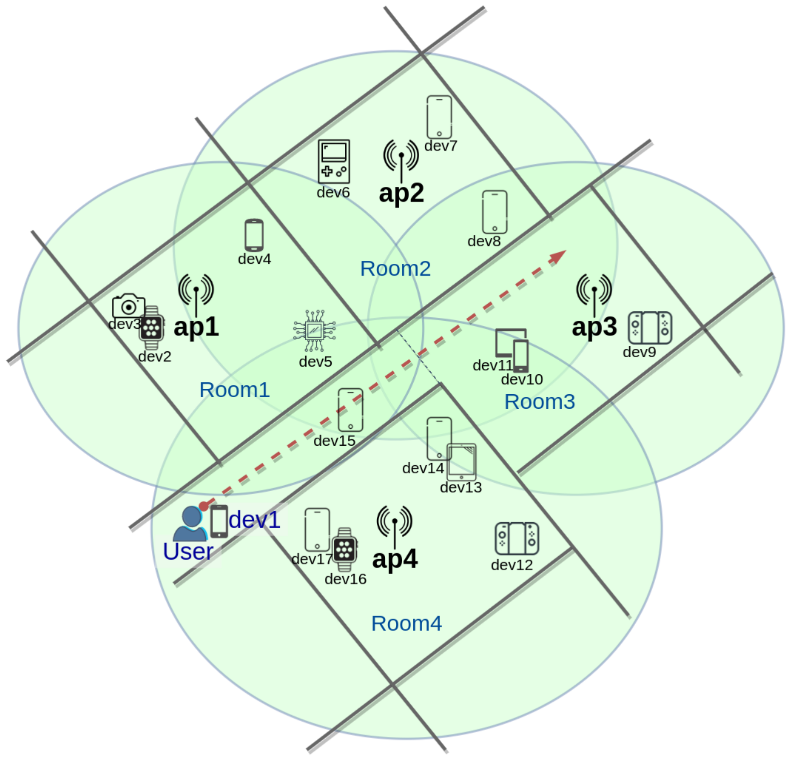


Figure 3.1: Motivation Scenario.

3.1.2 Methodology

We followed the Cross Industry Standard Process for Data Mining methodology (CRISP-DM) [120] to build our classification model. The aim was to optimize the network selection by choosing the best AP available in Wi-Fi networks. We carried out three CRISP-DM steps: data interpretation, data preparation, and modeling. Data interpretation generates, gathers, and defines the initial dataset that includes the set of data and features related to selecting the best AP to perform a handover. Data preparation covers all activities necessary to construct the final dataset from the initial dataset. Data preparation tasks include attribute selection as well as the transformation and cleaning of data. Modeling refers to the assessment of several supervised ML algorithms to choose the one that best foresees the AP to make the handover decision. The selection metrics are the True Positive Rate (TPR) [121], False Positive Rate (FPR) [122], Matthews Correlation Coefficient (MCC) [123], and Time of Classification (T_C) [121].

3.1.3 Data Interpretation

This phase builds an initial dataset with the appropriate features (i.e., parameters) to select the best available AP in a WiFi network. This phase includes the following tasks: first, an experimental scenario is designed and implemented. Second, the relevant parameters/attributes to optimize the AP selection are chosen. Third, the target variable is defined. Fourth, the dataset is filled by capturing the attributes when the MDs are moving and have at least two APs in range.

Figure 3.2 shows the scenario built to obtain the initial dataset. This scenario includes four APs (described in Table 3.1), together with 24 MDs ($devs$). In particular, $dev1$, $dev2$, $dev3$, and $dev4$ perform a circular movement, and $dev5$, $dev6$, $dev7$, and $dev8$, carry out a linear movement. In turn, $dev9$ to $dev24$ perform a random movement. All devices perform the movement in 500 steps at a speed of 1 m/s and communicate at 2462 GHz. The scenario described was deployed by using the Mininet-WiFi [124] emulator in a virtual machine with Ubuntu Server 16.04.

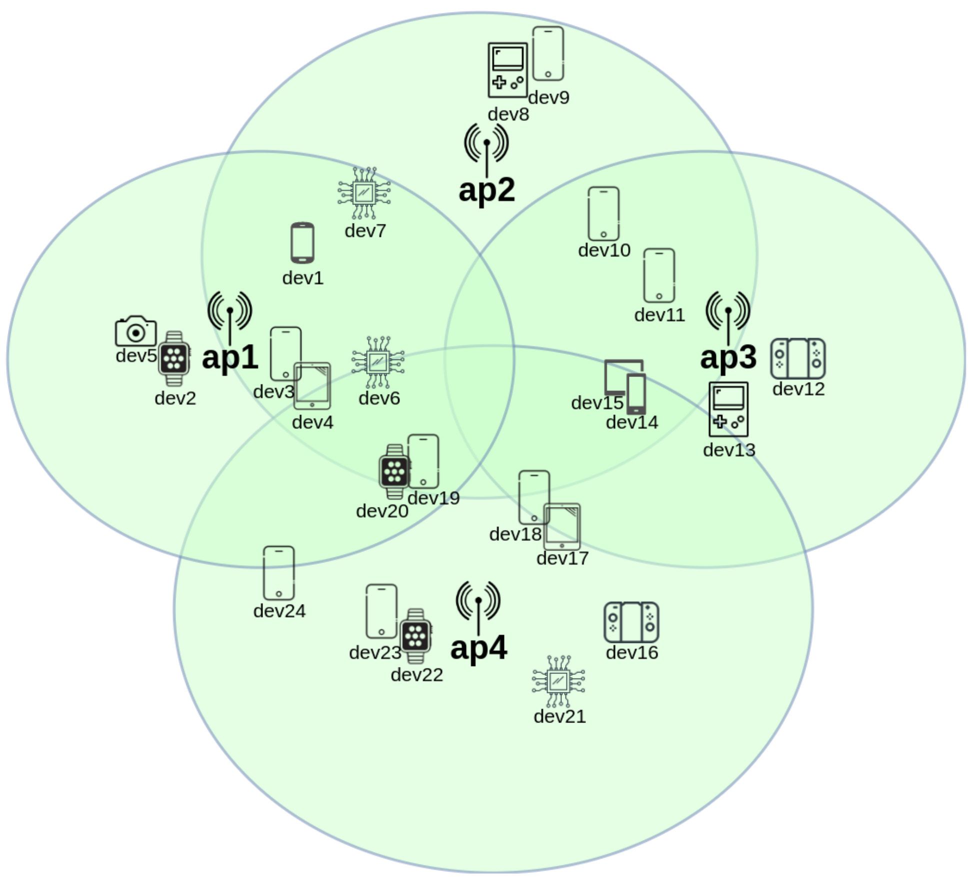


Figure 3.2: Dataset scenario set-up.

In the second step, we investigated the most used parameters related to the network selection phase in the literature. Table 3.2 briefly describes the parameters chosen for

Table 3.1: Access Point (AP) setup.

AP	SSID	Position [x,y,z] (m)	Range (m)	Channel	Total_Users_Support
ap1	ap1-ssid	[50.0,150.0,0]	45	1	15
ap2	ap2-ssid	[90.0,180.0,0]	50	7	18
ap3	ap3-ssid	[130.0,150.0,0]	45	11	15
ap4	ap4-ssid	[90.0,90.0,0]	57	11	20

our initial dataset and classifies them according to their type: network, device, application, and user. In the third task, we defined AP_target as the dependent parameter/attribute (also known as output variable) to predict; this variable is categorical (i.e., ap1, ap2, ap3 and ap4) because it allows the prediction of the AP to which the MD should connect taking into account the user preference. In the fourth task, we filled the dataset and used TOPSIS to complete the AP_target parameter. We gathered the data when MDs ($devs$) moved and had at least two candidate APs in range.

Table 3.2: Parameter description.

Proposal	Criterion	Parameter	Description
[113, 115, 116]	Network	AP	Indicates which are the candidate APs that are in range of the MD.
		RSSI	Reference scale for measuring the power level of the signals received by a MD and determining if the signal is sufficient to get a good wireless connection.
		AP Ocupation	Percentage of users connected to the AP concerning the entire capacity of users support the AP.
[113]	Devices	Distance	Distance between MD and candidate AP.
[125]	Device and Application	Battery consumption	Estimated discharge percentage of the MD when it is connected to the candidate AP taking into account the applications used by the user and the distance to the AP.
[126]	Application	Power consumption	Battery consumption in the MD due to the type of application that users are using in the handover moment.
[112]	User	User Preference	User preference between good signal quality, lower battery consumption or good QoS.

The result of this phase is the initial dataset with five input parameters/attributes and one output parameter. The input parameters are as follows: AP available, RSSI, AP occupation, battery consumption of devices, distance from the MD to the AP, and power consumption of applications. The output parameter (dependent variable) is the selected AP. To create this dataset, we leveraged the global view of the network offered by Mininet-WiFi. This emulator provides a global view, following the Software-Defined Network concept [127]. With Mininet-WiFi, it is possible to collect parameters from MD and APs [128]. Listing 3.1 shows code excerpts to illustrate how the aforementioned parameters were collected.

Listing 3.1: Parameter collection.

```
//AP available
dev.params["apsInRange"]
dev.params["associatedTo"]

//RSSI
ap_dis = dev.get_distance_to(ap_name)
dev_rssi = dev.get_rssi(ap_name,0,ap_dis)

//AP occupation
ap_num_dev = ap_name.params["associatedStations"]
ap_max_dis.params['maxDis']
ap_ocu = ((ap_num_dev*100.)/ap_max_dis)

//battery consumption of devices
dev_app = dev.params.get("app")
ap_dis = dev.get_distance_to(ap_name)
bat_temp = constant_battery_discharge * ap_dis * dev_app

//distance from the device to the AP
dev.get_distance_to(ap_temp)

//power consumption of applications
dev.params["pow"]
```

3.1.4 Data Preparation

This phase analyzes and processes the initial dataset in order to build the final dataset. The data preparation involves the assessment of the quality of the dataset and the reduction of the dataset's dimension. Firstly, we captured the data from the Mininet Wi-Fi emulator. We noted that the data for APs in the range of MDs were correct in the dataset. For APs which were out of range, there were missing values in the dataset. Therefore, we decided to substitute the missing values with the maximum value of each parameter. In particular, we assigned -100 dBm for RSSI, 100% for AP occupation, and 1% for battery consumption. These maximum values in the dataset represent the non-selection of an AP outside the range of the MD.

Secondly, we selected the most relevant features with the aim of reducing the dataset dimension and computational complexity. To determine the final dataset features, we applied three ML classification algorithms to all datasets resulting from the combination of all features of the initial dataset. In particular, we used Hoedding Tree (HT) [129], RF [129], and Support Vector Machine (SVM) [130] because the output variable of these methods is categorical (i.e., ap1, ap2, ap3 and ap4). HT is an incremental decision tree learner for large data streams, which assumes that the data distribution is not changing over time [129]. Additionally, HT uses the simple idea that a small sample can often be sufficient to choose an optimal splitting attribute, which is supported mathematically by the Hoeffding bound that quantifies the number of observations needed to estimate some statistics within a prescribed precision [131]. RF consists of a large number of individual decision trees that operate as an ensemble. Each tree predicts a class, and the class with the most votes becomes the prediction [132]. SVM outputs an optimal N-dimensional hyperplane (N—the number of labels) that categorizes the output variable given labeled training data [130].

We used accuracy [133] and MCC as metrics in this feature selection task because they are typical in classification problems. Accuracy is the ratio between the number of correct predictions and all predictions made [122]. MCC is a measure of the quality of classifications; it takes into account true and false positives and negatives. MCC is, in essence, a correlation coefficient value between -1 and 1. A coefficient of 1 represents a perfect prediction, 0 is an average random prediction, and -1 is an inverse prediction [123].

For the sake of readability, Figure 3.3 depicts only the evaluation results of four tested datasets regarding accuracy and MCC. Dataset 1, in purple, includes RSSI, AP occupation, and battery consumption. Dataset 2, in green, considers RSSI, AP occupation, battery consumption, and distance. Dataset 3, in blue, comprises RSSI, AP occupation, battery consumption, and power consumption. Dataset 4, in orange, includes RSSI, AP occupation, battery consumption, distance, and power consumption. Dataset 1 obtained the best performance regarding the accuracy and MCC for all evaluated ML algorithms. Specifically, this dataset obtained the highest accuracy, near 99.7%, and the highest MCC score, of around 0.996; thus, Dataset 1 was selected as the final dataset.

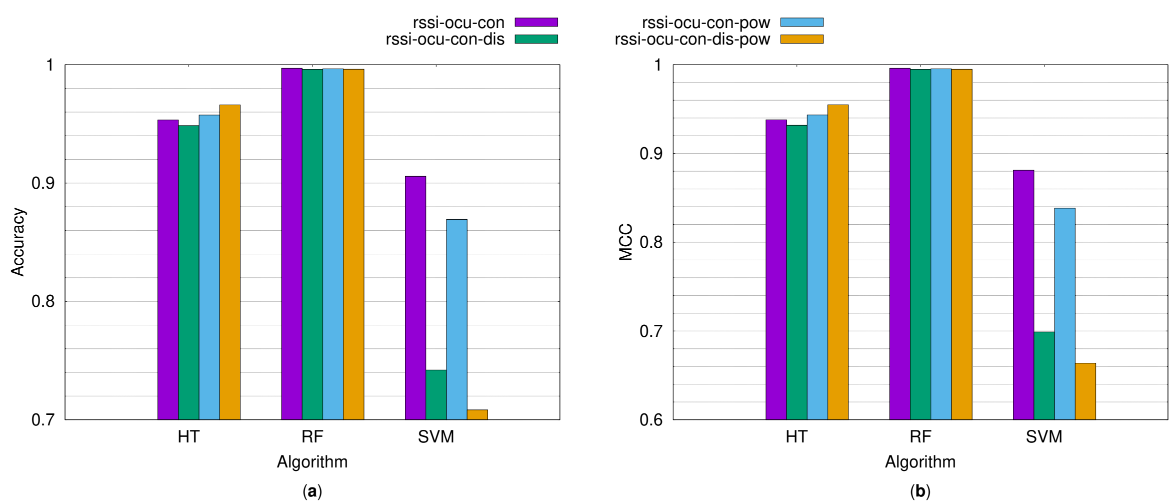


Figure 3.3: Feature Selection. (**a**) Accuracy, (**b**) MCC.

Considering the accuracy and MCC results mentioned above, we reduced the dataset dimension (i.e., resulting in lower computational complexity) from five to four input features; the final dataset (Dataset 1) includes the output variable AP_target and the input features: AP available, RSSI, AP occupation, and battery consumption. Note that the final dataset does not include the distance feature (no relevant); this may be due to its relationship with battery consumption. We filled the final dataset by capturing the features/parameters when the MD is in movement and has at least two APs in range; note that we assigned the maximum values to the missing data. Our final dataset comprises $10,500$ samples (Table 3.3).

Table 3.3: Excerpt of final dataset.

	station	ap1	rssi1	ocu1	con1	ap2	rssi2	ocu2	con2	ap3	rssi3	ocu3	con3	ap4	rssi4	ocu4	con4	AP_Target
1	sta1	0	−100	100	1.00	1	−69.0	66.67	0.031	1	−72.0	46.67	0.031	0	−100	100	1.00	ap3
2	sta2	0	−100	100	1.00	1	−55.0	55.56	0.010	1	−76.0	60.00	0.041	0	−100	100	1.00	ap2
3	sta3	0	−100	100	1.00	0	−100	100	1.00	1	−69.0	53.33	0.024	1	−76.00	45.00	0.052	ap3
4	sta4	0	−100	100	1.00	1	−74.0	66.67	0.046	1	−66.0	60.00	0.018	1	−77.00	40.00	0.056	ap3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
10500	sta24	0	−100.0	100.0	1.0	1	−72.0	66.67	0.0376	1	−64.0	46.67	0.0169	0	−100.0	100.0	1.0	ap3

3.1.5 Modeling

We evaluated five classification algorithms: the three used in the features selection plus Adaptive Random Forest (ARF) [134] and Hoeffding Adaptive Tree (HAT) [134]. ARF is an adaptation of RF, which includes mechanisms to adapt to different kinds of concept drifts given the same hyper-parameters. HAT adaptively learns from data streams that change over time without needing a fixed-size of the sliding window. The optimal size of the sliding window is a complicated parameter to guess for users since it depends on the rate of change of the distribution of the dataset [134]. We used these five algorithms because the target variable is categorical.

As evaluation metrics, we used the two used in the feature selection plus TPR, FPR, precision, and time of classification (T_C). TPR is the number of class members classified correctly over the total number of class members (i.e., the label AP1 is used when AP1 is the correct selection for handover). In contrast to TPR, FPR is the number of class members classified incorrectly over the total number of class members (i.e., label the AP1 is used when AP3 is the correct selection for handover) [121]. The precision is the number of class members classified correctly over the total number of instances classified as class members [122]. T_C is the time spent by an algorithm to classify the selected AP [121]. Furthermore, we used cross-validation, which consisted of randomly dividing the final dataset into two parts: 80% for training and 20% for validation.

Figure 3.4 shows the evaluation results regarding accuracy. These results reveal that RF and ARF obtain higher accuracy than the other evaluated algorithms. In particular, RF and ARF reach an accuracy of approximately 99.7% and 99.57%, respectively. These results are due to RF and ARF being able to handle large amounts of data with higher dimensionality; additionally, RF and ARF require less cleaning and pre-processing of data compared to other learning methods. Figure 3.4 depicts the evaluation results regarding MCC. These results reveal that RF and ARF also obtain a higher MCC than SVM, HT, and HAT. In particular, RF and ARF reach an MCC of approximately 0.996 and 0.994, respectively. These results are because tree-based algorithms are efficient at finding complex correlations. Therefore, RF and ARF provide a high correlation between real and predicted values for HM.

We also evaluated HT, HAT, RF, ARF, and SVM by using the confusion matrix, which is a fundamental tool to evaluate the performance of classification algorithms; this matrix allows us to determine quickly if the model is confusing different classes. In this matrix, each column represents the number of predictions per class, while each row represents the instances in the real class. TPR and FRP represent the proportion of APs selected correctly and incorrectly, respectively.

Table 3.4 presents the performance evaluation results of the five algorithms mentioned above according to metrics derived from the confusion matrix and T_C. These results show that all algorithms have a higher TPR than 90%. Again, RF and ARF behave better than the other evaluated algorithms, reaching percentages of around 99.7% and 99.57%, respectively. Furthermore, RF and ARF have a low FPR of about 0.29% and 0.43%, respectively. In addition, RF and ARF reach a high precision of around 99.68% and 99.69%, respectively. T_C is low for all tested algorithms, ranging between 0.66 and 5.146 ms. HT and SVM obtain the shortest T_C.

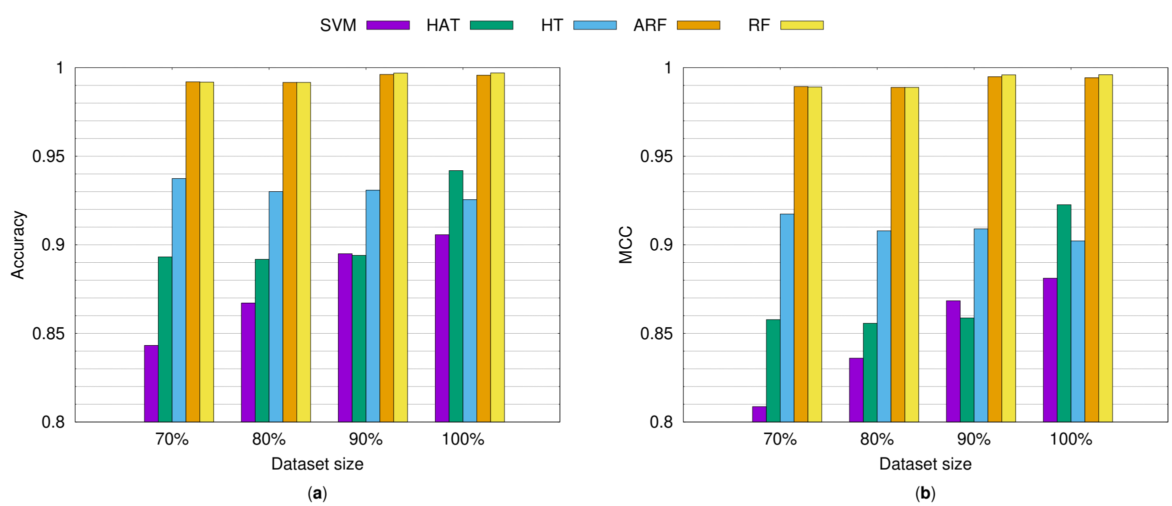


Figure 3.4: Modeling evaluation. (**a**) Accuracy, (**b**) MCC.

From the cross-validation and confusion matrix analysis, we concluded the following: (a) none of the algorithms achieve the best results for all the metrics presented, (b) RF and ARF algorithms outperform the other algorithms regarding accuracy and MCC; and (c) the algorithms that obtained better performance in the TPR metric result in a high T_C and vice-versa. Considering these conclusions, we decided to use RF for network selection due to it being the model with the best balance between accuracy, MCC, and T_C. Thus, as a result of the modeling phase, we obtained the classification model called NetSel-RF. In the next section, we compare NetSel-RF with SSF and AHP-TOPSIS.

Table 3.4: T_C and metrics derived from the confusion matrix. TPR: true positive rate; FPR: false positive rate; RF: Random Forest; ARF: Adaptive Random Forest; SVM: Support Vector Machine; HAT: Hoeffding Adaptive Tree; HT: Hoedding Tree.

Algorithms	**TPR%**	**FPR%**	**Precision%**	T_C **(ms)**
RF	99.70	0.29	99.68	3.865
ARF	99.57	0.42	99.65	5.146
SVM	90.57	9.42	92.99	1.445
HAT	94.21	5.78	75.65	1.471
HT	92.55	7.44	74.44	0.666

3.2 Network Selection

This section aims to evaluate the behavior of NetSel-RF regarding two metrics: the number of handovers and instantaneous throughput. Furthermore, this section compares NetSel-RF to SSF and AHP-TOPSIS qualitatively and quantitatively. The following subsections present the metrics, evaluation scenario, and results obtained.

3.2.1 Metrics and Evaluation Scenario

We compared NetSel-RF to SSF and AHP-TOPSIS regarding the number of handovers and instantaneous throughput. The quantity of handovers is the number of transfers an MD makes when it moves from one place to another one [112]. This number is affected by the ping-pong effect that occurs when the MD does not stay within the coverage of the selected AP and returns to the associated AP. The instantaneous throughput (throughput drops) represents the time that the number of bytes transmitted falls to zero due to handover.

With the aim of evaluating NetSel-RF, we used Mininet-WiFi to emulate a network formed by four APs with 802.11 g and 17 MD distributed in four rooms (see Figure 3.1). We analyzed the performance of NetSel-RF, SSF, and AHP-TOPSIS regarding the metrics mentioned above when $dev1$ carried out a linear movement at a speed of 1 m/s.

3.2.2 Results and Analysis

Figure 3.5 depicts how NetSel-RF, SSF, and AHP-TOPSIS behave regarding the number of handovers, revealing NetSel-RF has a higher number of unnecessary handovers than SSF and AHP-TOPSIS because it suffers from the ping-pong effect in the movements 100, 215, and 415. This ping-pong is caused by $AP3$ and $AP4$ offering identical conditions (i.e., the same RSSI level and percentage of users connected to AP) to $dev1$.

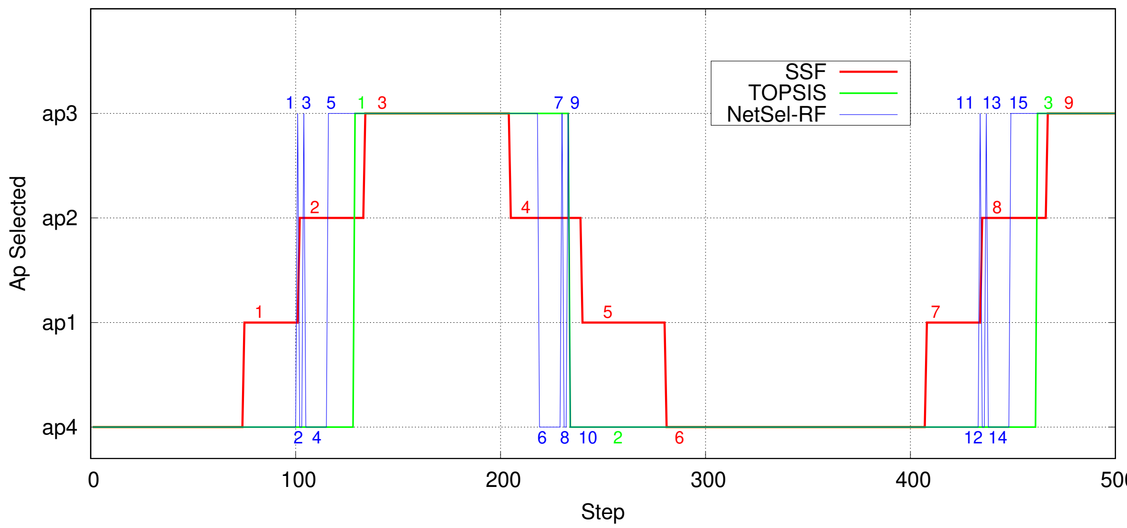


Figure 3.5: Number of handovers: NetSel-RF without movements module.

To address the ping-pong effect, we enhanced NetSel-RF with a movements module. This module uses a row vector to store the variable AP_target as well as the last seven locations (latitude and longitude) of MDs. A movement is given by a location change. Thus, the handover is performed only when AP_target has the same value for all seven movements; otherwise, the MDs remain connected to their current AP. It is important to highlight that we performed tests considering the last 3, 5, 7, and 9 moves to choose the proper number to eliminate the ping-pong effect.

Table 3.5 exemplifies two vectors containing AP_target. NetSel-RF does not perform any handover for the AP_target presented in the first row since this row jumps between $AP2$ and $AP3$. Conversely, since the AP_target is the same for the seven location changes (movements) in the second row, NetSel-RF carries out the handover. In practice, we filled this vector for all MD in movement using Mininet features: $dev.params[''associatedT$ and $dev.params[''position'']$.

Table 3.5: AP_target example.

VariableMovement	**Mov1**	**Mov2**	**Mov3**	**Mov4**	**Mov5**	**Mov6**	**Mov7**	
AP_target	AP2,loc1	AP2,loc2	AP2,loc3	AP3,loc4	AP2,loc5	AP3,loc6	AP2,loc7	X
AP_target	AP2,loc1	AP2,loc2	AP2,loc3	AP2,loc4	AP2,loc5	AP2,loc6	AP2,loc7	$\surd$

Figure 3.6 depicts the evaluation results regarding the number of handovers carried out by the enhanced NetSel-RF, AHP-TOPSIS, and SSF. These results reveal that NetSet-RF and AHP-TOPSIS outperform SSF regarding this evaluation metric. In particular, the number of handovers for RSSI, AHP-TOPSIS, and our model are 9, 3, and 3, respectively.

Figure 3.7 shows the evaluation results regarding the instantaneous throughput of

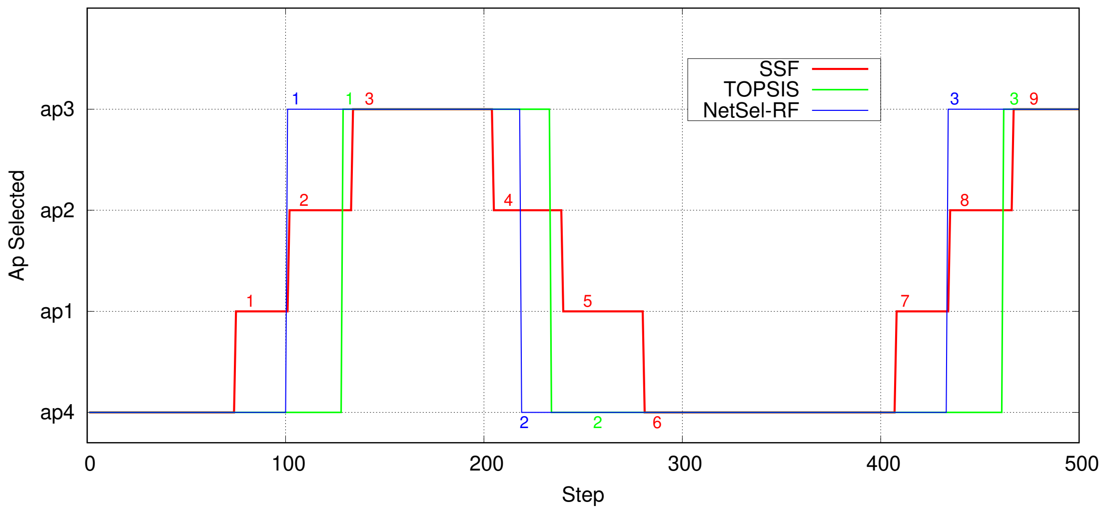


Figure 3.6: Number of handovers: NetSel-RF with movements module.

the enhanced NetSel-RF, AHP-TOPSIS, and SSF. These results reveal our model and AHP-TOPSIS outperform SSF regarding this evaluation metric. In particular, NetSel-RF and AHP-TOPSIS suffer 3 instantaneous throughput ($0 Mbps$). In turn, SSF has 7 drops. In summary, according to Figures 3.6 and 3.7, the multicriteria-based approaches behave better than SSF regarding the number of handovers and instantaneous throughput.

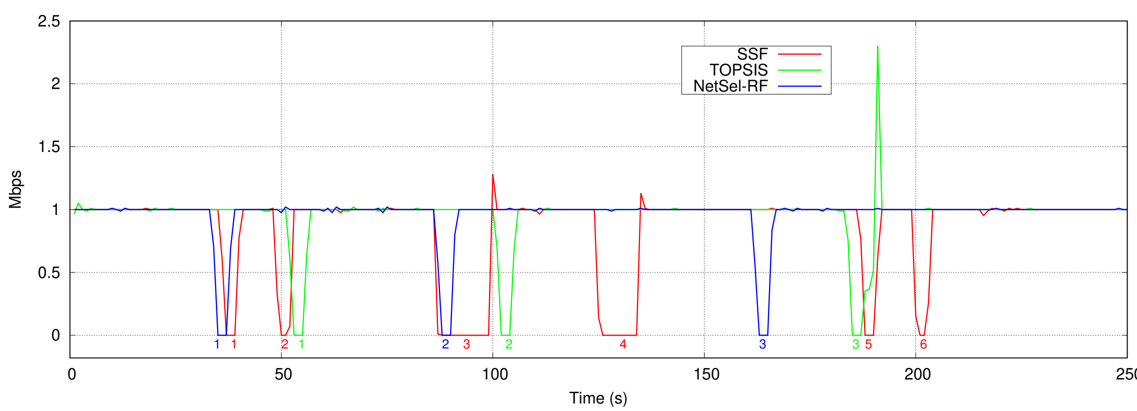


Figure 3.7: Instantaneous throughput in NetSel-RF.

Figures 3.6 and 3.7 reveal that NetSel-RF behaves similarly to AHP-TOPSIS regarding the number of handovers and instantaneous throughput. However, there is an essential difference between our model and AHP-TOPSIS. NetSel-RF performs transfers earlier than AHP-TOPSIS (Figure 3.8). This is due to the fact that our model is proactive, and AHP-TOPSIS is reactive. Thus, NetSel-RF selects a new AP before suffering a disconnection. The proactive approaches are more effective than reactive ones regarding QoS and QoE since the MDs perform the handover before the network link quality degrades, meaning that MDs are always connected [135].

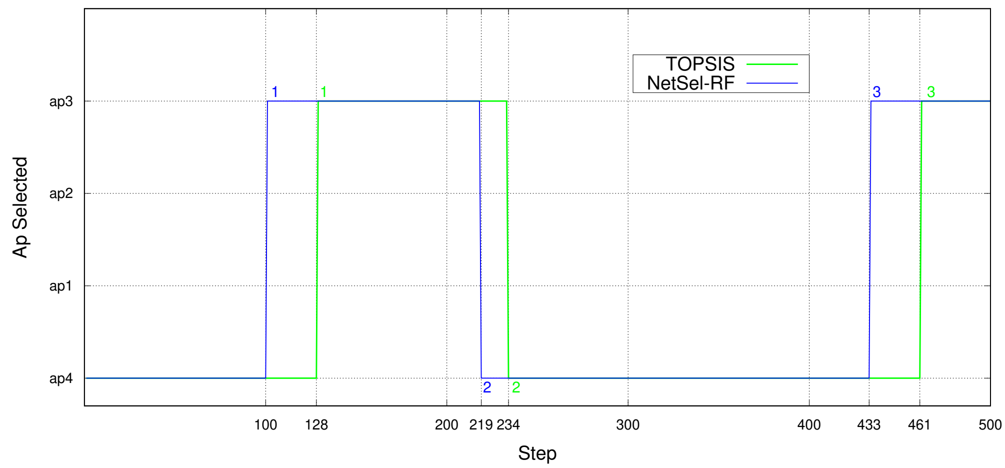


Figure 3.8: AP Selection.

3.3 Final Remarks

In this chapter, we introduced a multicriteria model based on supervised learning, called NetSel-RF, to cope with wrong network selection decisions that lead to unnecessary handovers, instantaneous throughput, and complexity in the current HM solutions. We constructed a dataset, performed data preparation, carried out feature selection, and developed our model by using RF. Our model, based on multicriteria (coming from the network, user preferences, devices, and applications) and supervised learning, outperformed the SSF approach regarding several metrics, such as the number of handovers (67%) and instantaneous throughput (50%). Moreover, our model behaved similarly to AHP-TOPSIS regarding the metrics mentioned, but selecting a new AP without waiting for the current AP is beyond the capability of the model. Our approach is proactive, and AHP-TOPSIS is reactive; therefore, NetSelf-RF makes a handover to an AP with better conditions before the MD loses connection, in contrast to AHP-TOPSIS.

Considering these results, we concluded that NetSelf-RF is an attractive and feasible solution for cognitive HM. In this sense, NetSel-RF envisions being deployed on WiFi networks following the SDN paradigm. Since the NetSel-RF model needs information from MD (e.g., battery consumption) and APs (e.g., occupation), it requires a network global view for a real deployment. SDN offers programmability, a global view, logically centralized control, and the decoupling of network control and packet forwarding. It is noteworthy that our training and validation datasets were collected by using the Mininet-WiFi emulator, which is SDN-oriented.

We analyzed the most common parameters used in the literature to build the final dataset used to train and validate NetSel-RF. The use of parameters other than those used in the data interpretation and data preparation phases is in other chapters of this thesis. It is also worthy of note that we use the parameters available in the MD and AP because, in practice, NetSel-RF is geared towards Software-Defined Wireless Networks [136–138], which is implicit from the data interpretation phase to the evaluation phase. For actual deployment, we suggest installing the NetSel-RF learning agent in the management plane [139] or in the cognitive plane [140].

Chapter 4

A Semantic and Knowledge-Based Approach for Handover Management

HM is responsible for making network (dis)connection decisions in a timely manner [54, 55]. In this sense, HM is pivotal for providing service continuity, ultra-high reliability, extreme-low latency, and meeting sky-high data rates in current and upcoming wireless communications [72, 82]. In order to achieve efficient HM, challenges need to be faced that are related to high handover rates and ping-pongs in dense communication environments, leading to an increase in both the data flow latency and the packet loss and, consequently, a reduction of the network throughput [141, 142]. Users moving at moderate-to-high speed require a seamless handover mechanism with few failures [60, 143].

In the networking literature, we find two approaches that address HM: single criterion-based and multicriteria-based. Approaches based on a single-criterion, such as SSF, usually consider only the link quality in the MD for carrying out handovers. SSF compares the RSSI of available networks and selects the network with the highest signal [144]. Single criterion-based approaches operate with a constrained network view that disregards contextual information, such as movement velocity and application requirements, leading to unnecessary and frequent handovers. These handover issues can decrease throughput, increase packet loss and even cause network service disruption [29, 145, 146]. The multicriteria-based approaches in [22, 30, 83–90] use RSSI and context information as criteria for ranking the available networks; the top-ranked network is selected by the MD for performing the connection process. These approaches disregard one or more relevant criteria, such as wireless network characteristics (e.g., coverage area), MD features (e.g., battery consumption), application requirements (e.g., real-time response), or user peculiarities (e.g., mobility pattern), leading to the handovers failure and wrong network selection, negatively impacting the network performance [24, 146]. Hybrid solutions combine multicriteria approaches [91]; however, their computational complexity is high.

This chapter presents SIM-Know, an approach for improving HM. The contributions of SIM-Know are two-fold. SIM-Know proposes a SIM that allows us to make context-aware handover decisions by considering and relating criteria from several context information domains: Network, Application, User, UserDevice, and Handover. SIM-Know also introduces a SIM-based distributed KBP that offers local and global knowledge for making contextual and proactive decisions during the handover process. We evaluated SIM-Know in an emulated wireless network. The remainder of this chapter is as follows. Section 4.1 introduces SIM-Know, including SIM and KBP. Section 4.2 presents the evaluation of SIM-Know. Section 4.3 compares SIM-Know to other related work. Finally, some remarks are presented in Section 4.4.

4.1 SIM-Know

HM allows an MD to keep an active connection when moving from one network coverage area (BS or AP coverage) to another [144]. HM comprises the initiation, selection, and execution phases [22, 147, 148]. Handover Initiation gathers all the information needed to identify and determine the neighboring networks and their current and future statuses (e.g., data about network performance and available services). Network Selection chooses the best available network from a ranking created by taking into account a single-criterion or multicriteria. Handover Execution connects and disconnects users to and from a network, involving resource allocation and releasing [51].

SIM-Know introduces SIM and KBP for improving HM. SIM allows SIM-Know to make context-aware handover decisions. The distributed KBP provides local and global knowledge to make rule-based cognitive decisions about network connection and disconnection. SIM and KBP envision diminishing the number of handovers and instantaneous throughputs and, as a consequence, have a positive impact on several network performance metrics (delay, jitter, packet loss, and throughput).

4.1.1 Semantic Information Model

SIM-Know makes appropriate and contextual handover decisions by considering criteria from several information domains (i.e., *Network*, *Application*, *User*, *UserDevice*, and *Handover*) modeled by SIM. We use the Common Information Model (CIM) and the Web Ontology Language (OWL) to carry out SIM (Figure 4.1). We adopted CIM [149] because it provides high expressiveness for modeling, management purposes, information systems, applications, and networks [71]. We used OWL [150] because it enables reasoning in the model and the sharing of knowledge among software agents [151]. In particular, SIM uses OWL classes and properties to characterize HM entirely by modeling the information domains and their relationships.

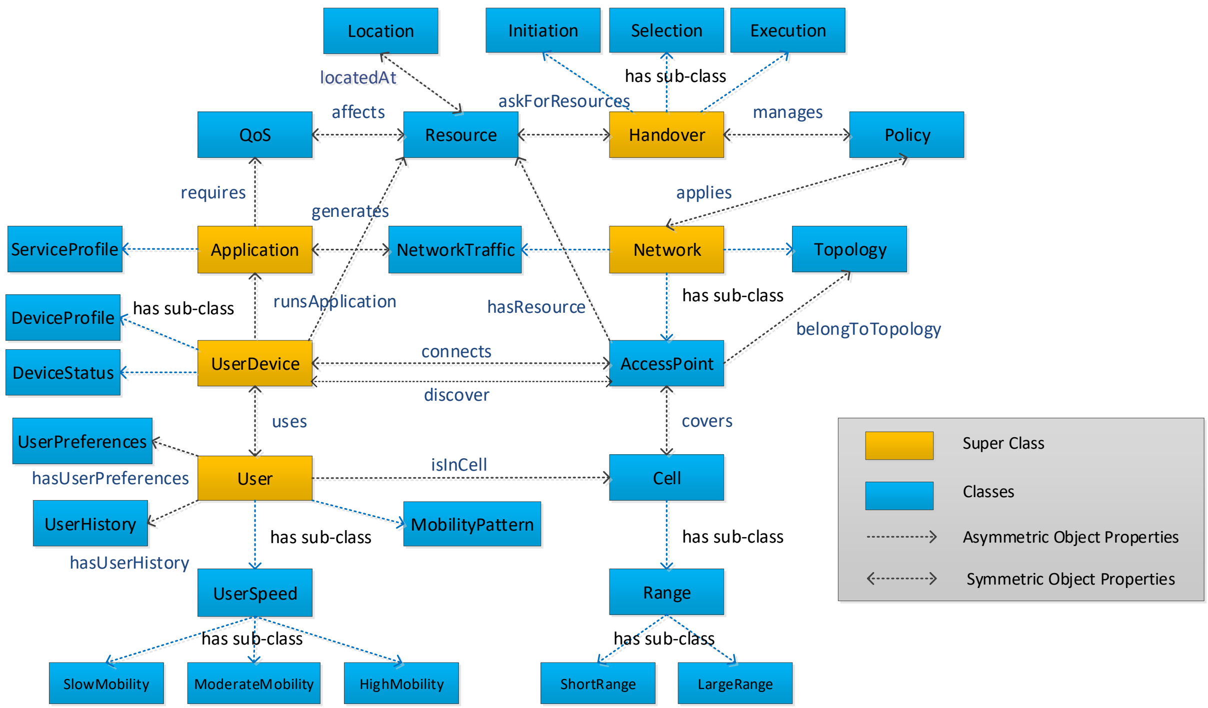


Figure 4.1: Semantic Information Model.

Figure 4.1 shows the five knowledge domains comprising SIM as superclasses, namely, Network, User, Application, UserDevice and Handover. It is worth noting that the *Handover* superclass models HM by using the *Initiation*, *Selection*, *Execution*, and *Policy* classes. The other superclasses represent the information domains containing the parameters used to improve decision-making in HM. Each superclass relates to other classes by the *has sub-class* relationship. The *Initiation* class models the initiation phase that defines when the selection phase is triggered, which, in turn, is modeled by the *Selection* class responsible for obtaining the candidate networks for performing the handover. The *Execution* class performs the handover itself since it allocates and releases AP (i.e., *AccessPoint* class) and user device (i.e., *UserDevice* class) resources, represented by the *Resource* class, which affects the QoS required (i.e., *QoS* class) by a user application (i.e., *Application* class). The *Policy* class represents the policies to apply to the *Network* class and governs the HM process. An example of a policy is to rank the candidate networks considering some criteria, such as user speeds and movement patterns.

The *Network* superclass models the characteristics and status of a network by using the *Topology*, *NetworkTraffic*, and *AccessPoint* classes. The *Topology* class represents the network's organization, including nodes and links. The *NetworkTraffic* class represents data and control traffic passing by the network. The *AccessPoint* class models a networking device using wireless technology; this class considers the area covered by the *Cell* class, which includes the *Range* class, which contains the *LargeRange* and *SmallRange* classes. The *isCoveredByCell*, *hasResource* and *belongToTopology* properties represent the *AccessPoint* class's relationship with the *Cell*, *Resource* and *Topology* classes, respectively. The *Resource* class models the ability to manage the resource consumption of APs located at (*Location* class) a particular network point.

The *User* superclass models the profile and behavior of the users by the *UserPreferences*, *UserHistory*, *MobilityPattern*, and *UserSpeed* classes. The UserSpeed class includes the *SlowMobility*, *ModerateMobility*, and *HighMobility* classes in order to model how fast a user moves. The *UserPreferences* class profiles the users with information related to, for instance, network preference by cost and service quality expectation. The *UserHistory* class models the historical (dis)connection of user. The *MobilityPattern* contains information about the user mobility pattern, which is predictable from their trajectory and velocity. The *Application* superclass represents user applications with the *ServiceProfile* and *QoS*. The *ServiceProfile* class models the application type (e.g., remote surgery, augmented reality, high definition video conferences). The *QoS* class allows the representation of a set of QoS requirements (e.g., delay, throughput, and packet loss) for each type of application.

The *UserDevice* superclass models the end-user devices and their components by the *DeviceProfile*, *DeviceStatus*, and *Resource* classes. The *DeviceProfile* class models the device's characteristics. The *DeviceStatus* class represents the device's current status (e.g., low-battery and off-air). The *Resource* class models the ability to manage the resource consumption of MDs located at (*Location* class) a particular network point. The *UserDevice* superclass relates to the *Application* superclass via the *runsApplication* property, which allows knowledge of the applications that are running in each MD. The *isUsedByUser* property defines a relationship between *UserDevice* and *User*.

4.1.2 Knowledge Base Profile

KBP is a distributed knowledge base that intends to provide local and global knowledge that supports making rule-based cognitive decisions about network connection and disconnection processes. Figure 4.2 depicts the KBP internal structure, comprised of layers and processes. The *Semantic* layer uses SIM (the entire model or a part) to obtain information from the data included in the *Context* layer. The *Reasoning* layer obtains knowledge from the information represented by SIM. The *Adaption* process acts on the layers to maintain updated data, information, and knowledge. The *Collaboration* process enables sharing the obtained knowledge between KBP instances. Next, we detail the KBP's layers and processes.

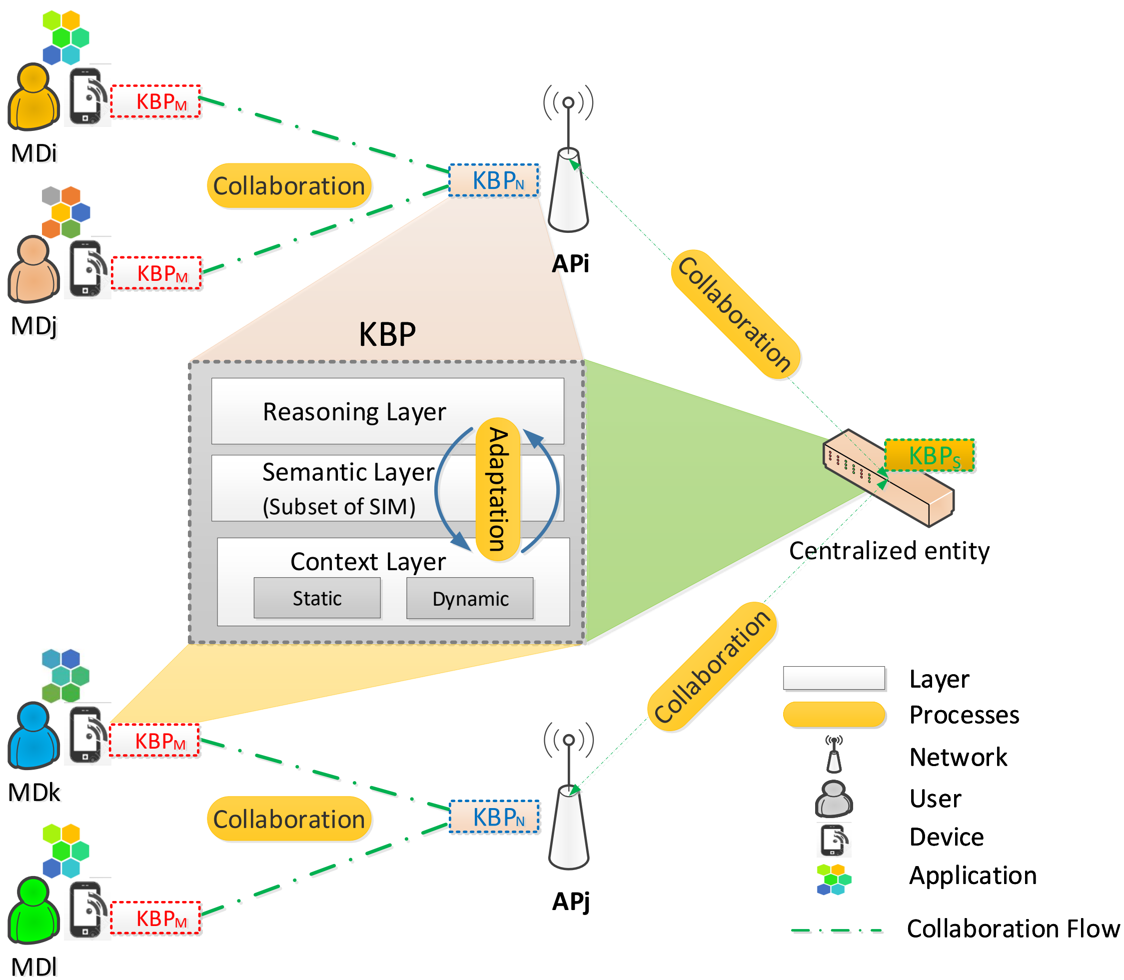


Figure 4.2: Knowledge Base Profile.

Layers

The *Context* layer includes contextual data about the user, network, device, application, and handover. Contextual data are essential for carrying out HM in environments with multiple wireless networks [88]. As in [89, 152], this layer is divided into static and dynamic sublayers. The *Static Context* sublayer involves data that do not change or rarely do; it plays a vital role in assisting with neighbor network discovery [153]. Examples of static data include the wireless technology supported by the MDs and APs, and the wireless network technology coverage area. The *Dynamic Context* sublayer serves an updated network view, including dynamic data such as the application requirements of a device needing handover and capacity available in a target AP, which enables the upper layers (*Semantic and Reasoning* layers) to realize knowledge-based handovers.

The *Semantic* layer offers a SIM instance nourished by the bottom layer's contextual data. Thus, the *Semantic* layer structures the information to achieve intelligent, timely, and context-aware HM (considering criteria from the Static and Dynamic contexts). For instance, the *DeviceProfile*, *UserPreferences*, *UserSpeed*, and *Mobility-Pattern* SIM classes can be used to build up a map of candidate networks; overall, SIM classes provide a structure to contextual data. It is worth noting that we consider three KBP flavors depending on how they instantiate SIM. KBP_N, located at any AP or BS, instantiates the superclasses *Network* and *Handover*. KBP_M, located at MDs, instantiates the superclasses *Application*, *User*, *UserDevice* and *Handover*. KBP_S is a complete KBP that can run on a logically centralized entity (e.g., a controller in a software-defined wireless network). The SIM's distribution allows any KBP (SIM-Know) to create and share local knowledge to generate global knowledge.

The *Reasoning* layer triggers the Initiation phase, selects the target network, and realizes the handover itself by inferring knowledge from SIM. We use Description Logic (DL) [154] to express in a structured and formal way the rules governing the *Reasoning* layer and, so, HM; the reasoning rules generate local and global intelligence to make autonomous handover connection decisions. Each rule has a set of conditions and settings. To illustrate how the *Reasoning* layer operates, we present some of the rules modeled to realize a policy intended to select candidate networks proactively, considering the coverage of APs and the mobility pattern of MD. For example, the Rule $APInRange$ (Listing 4.1) serves to discover neighboring networks considering RSSI.

Listing 4.1: Rule for APInRange.

$$APInRange \equiv User \sqcap \exists isInCell.(\exists covers.AccessPoint)$$

Listing 4.2 shows that the Rule $UserSpeed$ is useful for defining the speed of users. If a User is moving with $v > th_u$, he/she has a $UserHighSpeed$. If a User is moving with speed higher than th_l and lower than or equal to th_u, he/she has a $UserModSpeed$. $UserSlowSpeed$ is when the user moves with $v \leq th_l$. According to [155], th_u can be set to 50 Km/h and th_l to 10 Km/h.

Listing 4.2: Rule for UserSpeed.

$$UserHighSpeed \equiv User \sqcap (\exists hasUserSpeed.HighMobility)$$
$$UserModSpeed \equiv User \sqcap (\exists hasUserSpeed.ModerateMobility)$$
$$UserSlowSpeed \equiv User \sqcap (\exists hasUserSpeed.SlowMobility)$$

Listing 4.3 shows that Rule $APRange$ is helpful for listing the APs by coverage range. $LargeRange$ is given by $range > th_r$ and $ShortRange$ by $range \leq th_r$. According to [84], th_r can be set to 35 m for 802.11ac.

Listing 4.3: Rule for APRange.

$$APLRange \equiv AP \sqcap \exists isCoveredBy.(\exists hasRange.LargeRange)$$

$APSRange \equiv AP \sqcap \exists isCoveredBy.(\exists hasRange.ShortRange)$

Listing 4.4 presents Rule $SoJournTime$, which is useful for determining the time the user stays covered by an AP. If $UserSlowSpeed$ moves in $APLRange$, it results in $LongSojournTime$. If $UserHighSpeed$ moves in $APSRange$, it results in $SmallSojournTime$. A $MSojournTime$ happens when $UserHighSpeed$ or $UserModSpeed$ is moving in $APLRange$. If $UserSlowSpeed$ or $UserModSpeed$ moves in $APSrange$, it also results in $MSojournTime$. $SmallSojournTime$ may be more challenging than $LongSojournTime$, and $MSojournTime$ in 5G networks and beyond are characterized by small coverage areas and high-mobility.

Listing 4.4: Rule for SoJournTime.

$LongSojournTime \equiv UserSlowSpeed \sqcap (\exists APLRange)$ \\
$SmallSojournTime \equiv UserHighSpeed \sqcap (\exists APSRange)$ \\
$MSojournTime \equiv APLRange \sqcap (UserHighSpeed \sqcup UserModSpeed) \sqcup$ \\
$APSRange \sqcap (UserSlowSpeed \sqcup UserModSpeed)$

Listing 4.5 shows that Rule $CandidateAP$ is useful for creating the list of candidate APs for MDs with $MSojournTime$ or $LongSojournTime$ in the network.

Listing 4.5: Rule for CandidateAP.

$CandidateAP \equiv MSojournTime \sqcup LongSojournTime$

Listing 4.6 presents Rule $AssociateAP$, which links the MD with the first AP in the list of candidates. It is worth noting that each network administrator can define his/her own rules to manage the wireless network as he/she needs.

Listing 4.6: Rule for AssociateAP.

$AssociationToAP \equiv User \sqcap \exists Uses.UserDevice(\exists Connects.AP)$

Processes

The *Adaptation* process allows SIM-Know to dynamically adapt to the changing environments and enhance HM by modifying the content of the layers of the KBP instances. The content is modified in a bottom-up way, starting with the contextual data, followed by the SIM instances, and ending with the acquired knowledge when environmental changes happen, such as new networks appearing, dynamic traffic conditions, and variations in QoS requirements. Furthermore, this process allows the addition and updating of the *Reasoning* layer seeking to meet QoS and to preserve network performance.

The *Collaboration* process allows KBP (KBP_N, KBP_S, KBP_M, and any other profile defined to extend SIM-Know) to interchange the knowledge obtained for enhancing the decision-making in HM. For instance, the collaboration between KBP_M and KBP_N would allow for choosing the optimal and appropriate time to trigger the handover and select the most suitable access network according to the user QoS requirements and network status.

4.1.3 SIM-Know Operation

Figure 4.3 presents how SIM-Know operates in WLAN. First, KBP_S, KBP_M and KBP_N collect their *Static Context*. Second, KBP_M monitors and updates the *Dynamic Context* information related to, for instance, user speed, APs in range, and RSSI. In parallel, KBP_M requests from KBP_N the *Dynamic Context* information, which includes the associated and in-range MDs. Third, every KBP generates local knowledge based on the *Static Context* and *Dynamic Context*. For example, the local knowledge in KBP_M can be $HighMobility$, and in KBP_N can be $LargeRange$. Fourth, KBP_M launches the handover process. Fifth, the *Collaboration* process starts between the corresponding KBP_M and KBP_N and ends with sending knowledge to KBP_S. Sixth, KBP_S builds up the global network view, generates a handover policy for selecting candidate APs according to the rules defined in its *Reasoning Layer*, and sends those candidates to KBP_M. Seventh, KBP_M selects the *Target* KBP_N by using the rules defined in its *Reasoning Layer*. Eighth, KBP_M sends a Handover Request to the *Target* KBP_N, which sends back an acknowledgment to KBP_M. Ninth, KBP_M sends disconnection requests to the current *Serving* KBP_N, which, in turn, sends a disconnection acknowledgment to KBP_M. Tenth, every KBP executes the *Adaptation* process to handle the context variations dynamically.

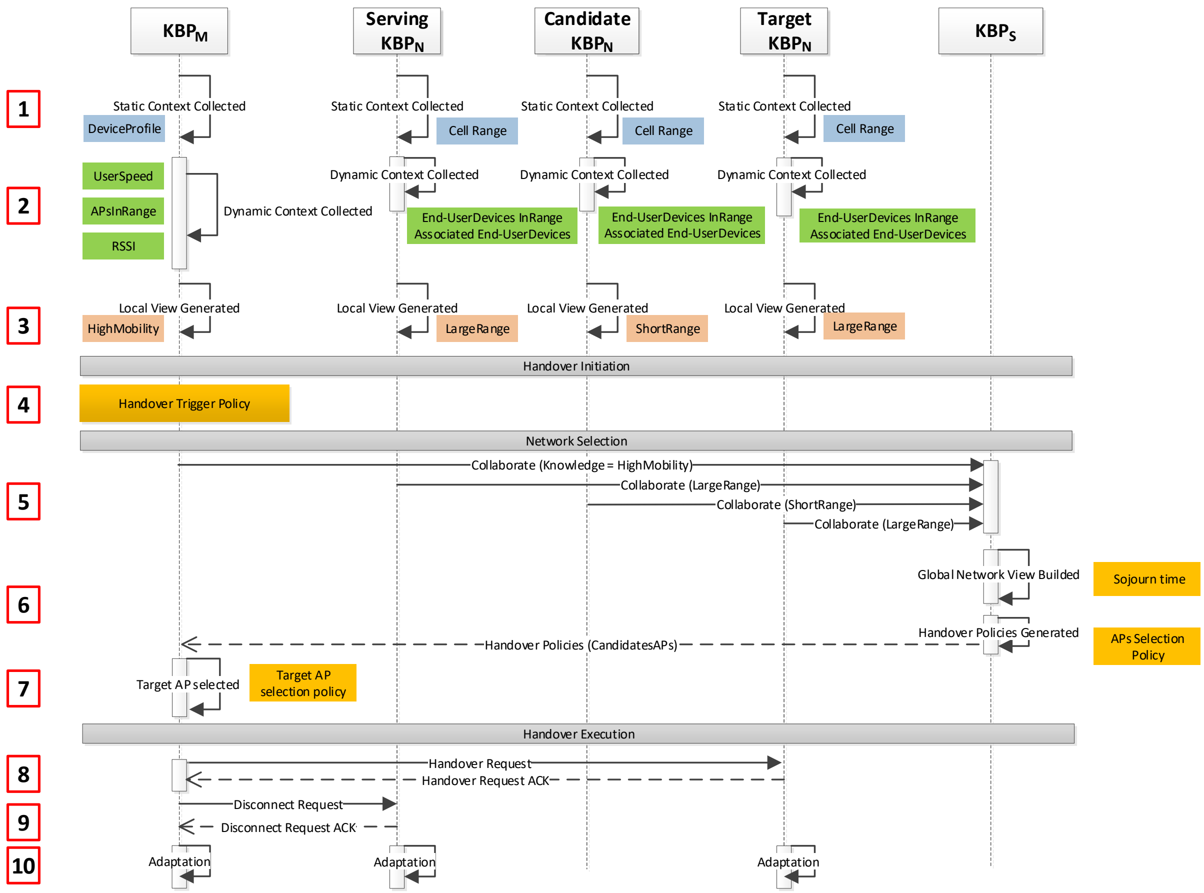


Figure 4.3: SIM-Know Operation.

Figure 4.4 shows the data format used in KBP_S to store the global knowledge built with information from KBP_N and KBP_M. The format follows the triplet (Subject, Predicate, Object) encoded in Entity Notation (EN) [156], which enables a lightweight knowledge representation for resource-constrained environments. The Subject identifies a class in SIM by the combination of ClassId (e.g., ceu101) and ClassType (e.g., *User*). Each Subject is related to various pairs, Predicate–Object. The Predicate identifies a property (e.g., UserSpeed) of the Subject while the Object provides the value of such a property (e.g., HighMobility). The Object can also be a ClassId, to represent relationships among Subjects.

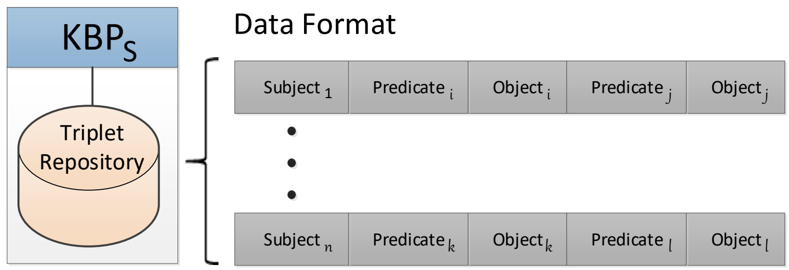


Figure 4.4: KBP_S Data Format

Figure 4.5 illustrates how SIM-Know can operate in 5G by running KBP_M in the User Equipment (UE), KBP_N in gNodeB (gNB), and KBP_S in the Core Network (CN). The handover in 5G, according to the specification 3GPP TS38.300 [157], consists of three phases: preparation (steps 0–5), execution (steps 6–8), and completion (steps 9–12).

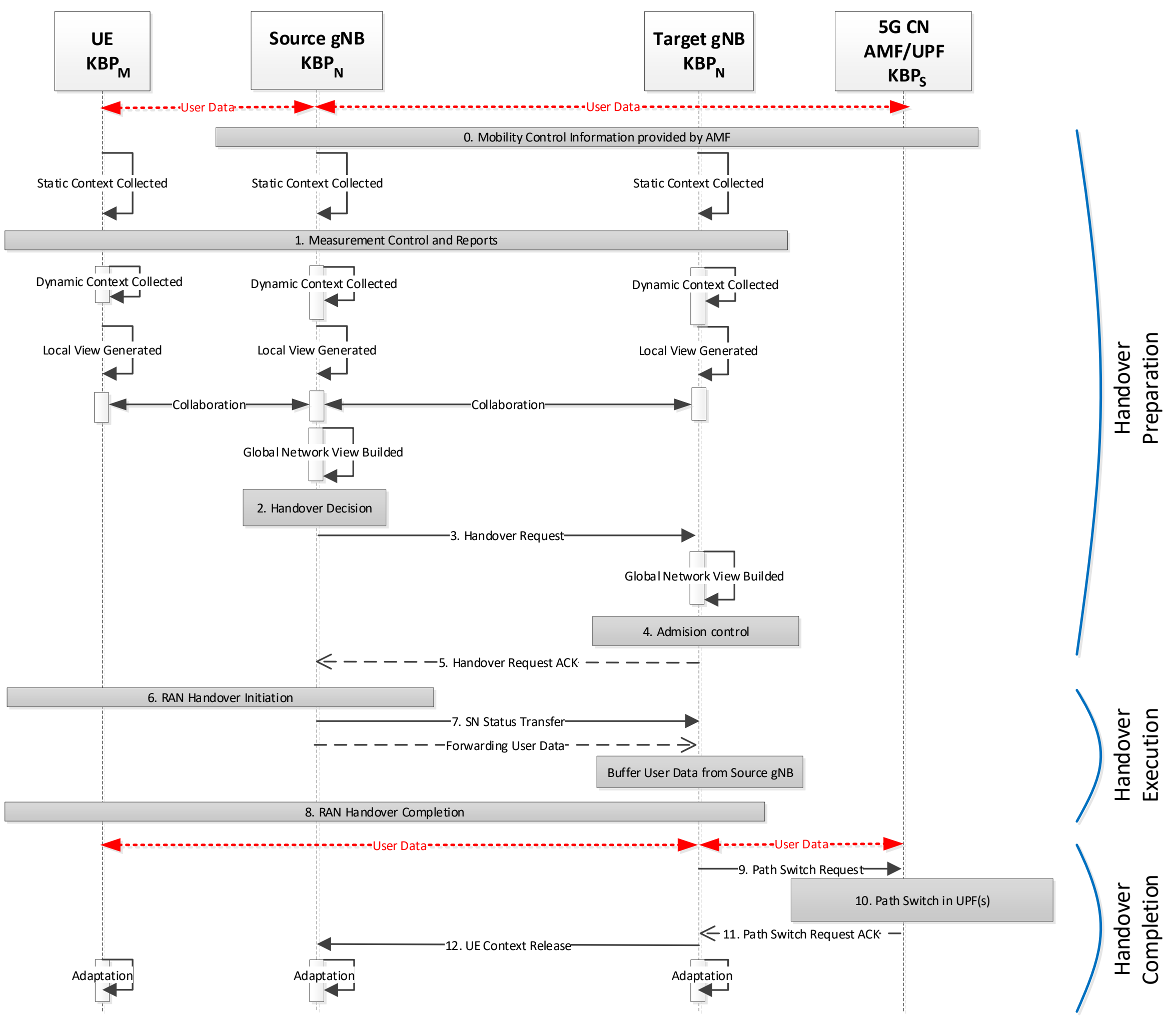


Figure 4.5: SIM-Know in 5G Intra-AMF/UPF Handover.

The steps are as follows:

- Step 0: Each KBP gathers the *Static* context.
- Step 1: KBP_M and KBP_N initialize measuring procedures to collect the *Dynamic* context, generate local knowledge, and exchange Measurement Reports by way of the *Collaboration* process. Source KBP_N builds up its local knowledge.
- Step 2: Source KBP_N makes handover decisions based on its local knowledge. gNBs are responsible for making handover decisions.
- Step 3: Source KBP_N sends a handover request message to Target KBP_N.
- Step 4: Target KBP_N executes the admission control procedure based on its local knowledge.
- Step 5: Target KBP_N sends a handover request acknowledgment to Source KBP_N.
- Step 6: Source KBP_N sends a handover command to KBP_M for handover initiation.
- Step 7: Source KBP_N sends the sequence number status transfer message to Target KBP_N. Source KBP_N may initiate data forwarding.
- Step 8: KBP_M detaches from Source KBP_N and synchronizes with Target KBP_N.
- Step 9: Target KBP_N informs KBP_S that KBP_M has changed the cell by way of the path switch request message.
- Step 10: KBP_S switches the data path towards Target KBP_N.
- Step 11: KBP_S acknowledges the path switch request message.
- Step 12: Target KBP_N informs Source KBP_N that the handover was successful and triggers the release of resources for Source KBP_N by sending a UE Context Release message. Finally, Source KBP_N releases the resources associated with KBP_M, invoking the *Adaptation* process.

4.2 Evaluation

This section presents the evaluation of SIM-Know in a WLAN, aiming to show its behavior regarding the number of handovers, instantaneous throughput, and its impact on various typical network performance metrics. Subsection 4.2.1 depicts the SIM-Know's prototype and the test environment. Subsection 4.2.2 shows the performance metrics and traffic generation.

4.2.1 Prototype and Test Environment

We implemented the SIM-Know prototype for WLAN, including KBP_M, KBP_N, and KBP_S, by using the Python programming language version 2.7. We also deployed the prototype in a Mininet-WiFi emulator [158] (Figure 4.6) running on an Ubuntu 16.04 Virtual Machine (VM) with a Core i7-3630 processor and 8 GB RAM. Mininet-WiFi adds virtual BSs and APs to classical Mininet [159] to enable the emulation of wireless network environments. The SIM-Know prototype, as well as all test scripts, are available in the project repository [160].

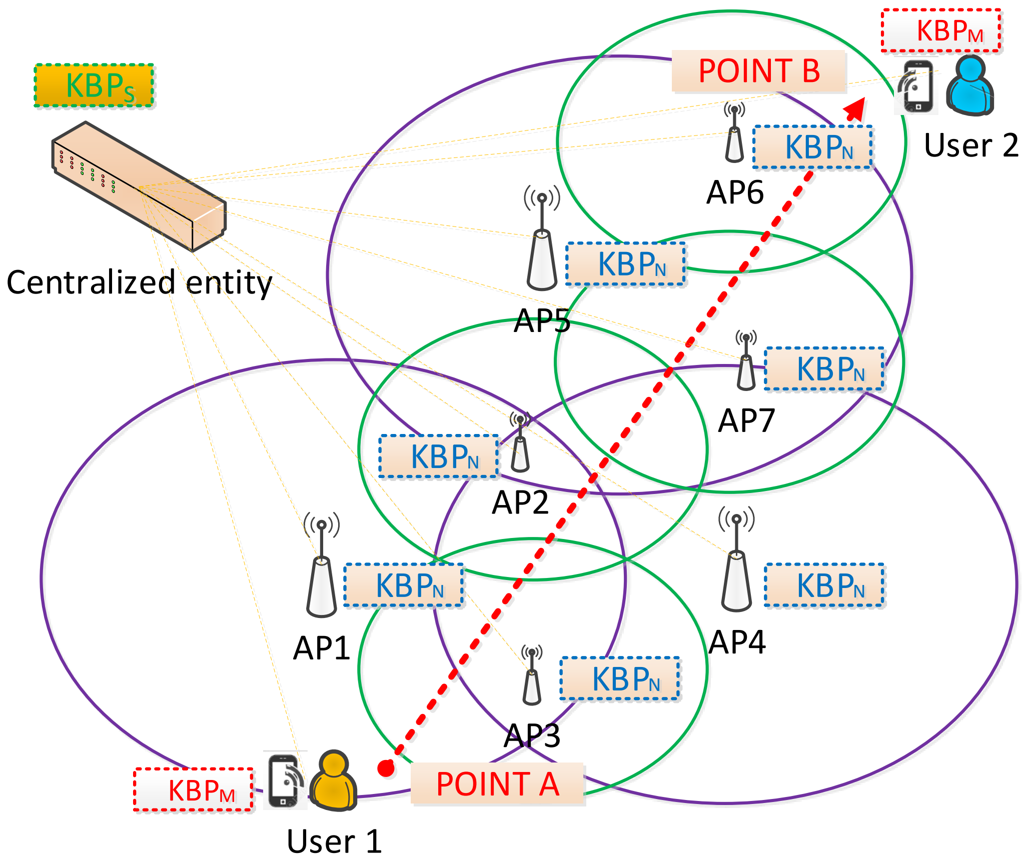


Figure 4.6: Test Environment.

Figure 4.6 shows the WLAN test scenario in which we evaluated and compared SIM-Know, SSF, and AHP-TOPSIS. The scenario, deployed in Mininet-WiFi, included seven APs, an MD associated with $User1$, and another MD linked to $User2$. In particular, we used three APs with a large coverage range (i.e., up to $75m$ for AP_1, AP_4, and AP_5 with $802.11n$) and four with a short coverage range (i.e., up to $35m$ for AP_2, AP_3, AP_6, and AP_7 with $802.11ac$).

We also analyzed the performance when the $User\ 1$ moved from Point A to Point B by following a straight line without directional change at a constant speed. We used three speeds for testing: 1.42 m/s corresponding to $SlowMobility$, 3.74 m/s to $ModerateMobility$, and 13.41 m/s to $HighMobility$. The MD associated with $User\ 1$ transmitted traffic (Voice over IP (VoIP) or Transmission Control Protocol (TCP)) to the MD linked to $User\ 2$, which was immobile. We repeated the experiments thirty-three times to obtain results with a 95% confidence level. Table 4.1 summarizes the setup of the experiments.

Table 4.1: Experiment Setup.

Parameters	**Value**
Wireless technology	802.11n, 802.11ac
Emulation area	180 × 180 m
Carrier frequency	2 GHz
Channel bandwidth	20 MHz
Transmission power of cells large-range/short-range	25/14 dBm
Path loss model from cells	Log-Distance Propagation Loss/ITU-R P1283
Emulation time for HighMobility	30 s
Emulation time for ModerateMobility	80 s
Emulation time for SlowMobility	180 s
TCP traffic	Flows with constant inter-departure time between packets (1000 pkts/s) and constant packets size (512 bytes)
VoIP traffic	Flows with audio code (G.711.2 - 84 Kbps and 50 pkts/s) transmitted using real-time protocol and voice activity detection

It is worth mentioning that the described scenario was constrained to a small number of MDs because our main objective was to show the feasibility of SIM-Know.

4.2.2 Performance Metrics and Traffic Generation

We compared SIM-Know to SSF and AHP-TOPSIS in terms of the number of handovers, number of instantaneous throughput, handover latency, and various well-known network performance metrics (throughput, delay, jitter, and packet loss) [161]. The quantity of handovers is the number of transfers an MD makes when it moves from one place to another [162]. The instantaneous throughput (throughput drops) represents the times that the number of bytes transmitted falls to zero because of a handover [51]. The handover latency is the time that elapses between the instant that the MD sends the last link-going-down message to the serving AP and the moment at which the MD establishes the connection with the target AP [82].

In the emulation experiments, scripts for generating traffic were developed using the iPerf3 [163], and D-ITG [164] tools. We used D-ITG to generate VoIP flows with audio code (G.711.2 - 84 Kbps and 50 pkts/s) transmitted using the real-time protocol and voice activity detection. We used iperf3 to generate TCP flows with constant inter-departure time between packets (1000 pkts/s) and constant packet size (512 bytes).

4.3 Results and Analysis

Table 4.2 shows that SIM-Know and AHP-TOPSIS outperformed SSF, in terms of the number of handovers and instantaneous throughput when the user moved at any speed (slow, moderate, or high). This behavior is expected because SSF is the baseline and triggers a handover when an AP with an RSSI higher than the serving AP is available.

Table 4.2: Handover Performance in SIM-Know.

Parameter	**SSF**	**AHP-TOPSIS**	**SIM-Know**
	SlowMobility		
Number of handovers	7	3	3
Number of instantaneous throughput	5	3	3
	ModerateMobility		
Number of handovers	7	3	3
Number of instantaneous throughput	5	3	3
	HighMobility		
Number of handovers	7	4	3
Number of instantaneous throughput	5	2	1

Table 4.2 also reveals that when the user moved at slow and moderate speeds, SIM-Know behaved as AHP-TOPSIS does regarding the number of handovers and instantaneous throughput. Figure 4.7 shows that SIM-Know outperformed AHP-TOPSIS in these metrics when the user moved at a high velocity. The outperformance regarding the number of handovers and instantaneous throughput was due to SIM-Know making context-aware, cognitive, and proactive handovers. Figure 4.8 corroborates that SIM-Know carried out handovers before SSF and AHP-TOPSIS did.

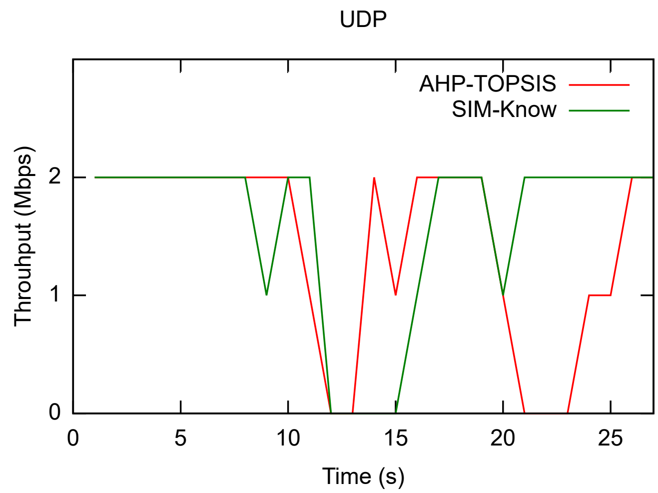


Figure 4.7: Instantaneous throughput in SIM-Know.

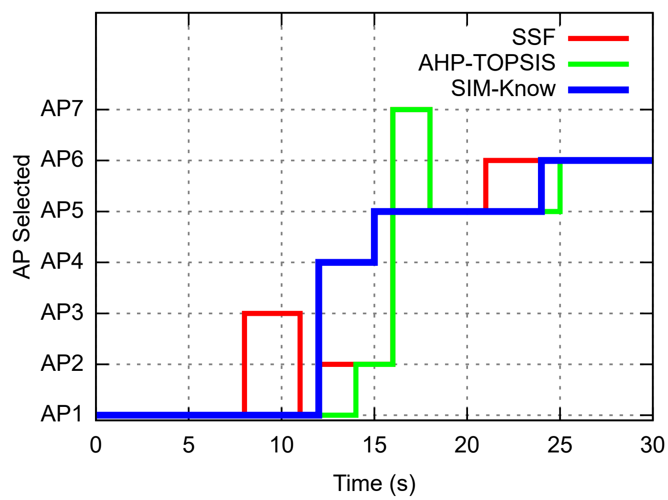


Figure 4.8: Proactivity in SIM-Know.

Figure 4.9 shows, as expected, that SIM-Know obtained a higher handover latency than SSF did, since our approach is knowledge-based and SSF makes decisions considering a single criterion. SIM-Know had 27% less handover latency than AHP-TOPSIS because, first, our approach is proactive and, according to [165], AHP-TOPSIS is reactive; the proactivity shortens the Handover Initiation phase [89]. Secondly, SIM-Know employs a rule-based reasoning method, while AHP-TOPSIS uses a complex mathematical model that requires considerable time to make handover decisions.

Next, we present how SIM-Know, SSF and AHP-TOPSIS impact various network performance metrics when the end-user device moves at $HighMobility$. Figure 4.10 depicts SIM-Know overcoming SSF and AHP-TOPSIS regarding the throughput, delay, and packet loss when the wireless network transferred VoIP/UDP traffic. In particular, the delay attained by SIM-Know was 29.28% and 23.13% lower than that achieved by

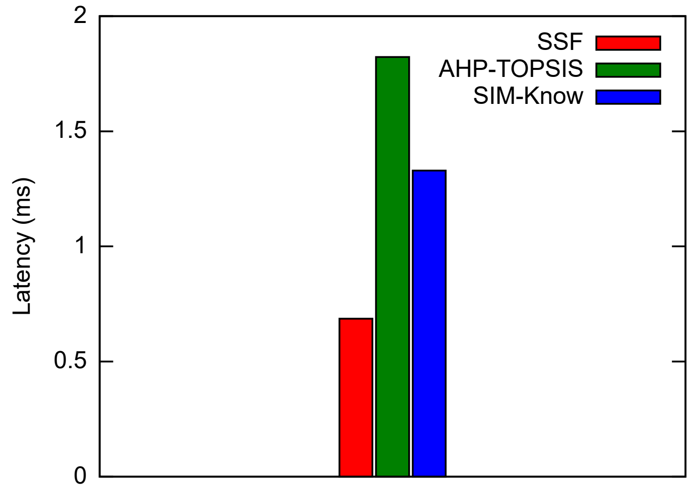
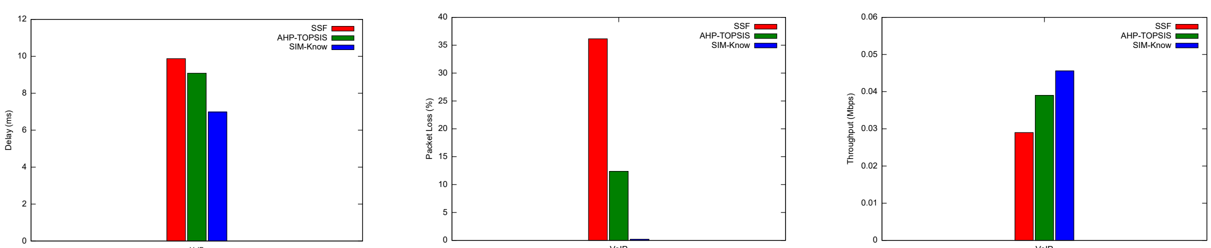


Figure 4.9: Handover Latency.

SSF and AHP-TOPSIS (Figure 4.10a). The packet loss of SIM-Know was 99.44% and 98.38% lower than that obtained by SSF and AHP-TOPSIS (Figure 4.10b). The throughput obtained by SIM-Know was 57.17% and 16.87% higher than that obtained by SSF and AHP-TOPSIS (Figure 4.10c).



(a) Delay in VoIP Traffic at $HighMobility$

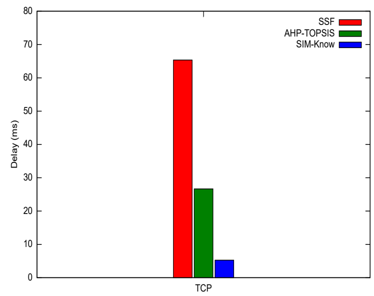
(b) Packet Loss in VoIP Traffic at $HighMobility$

(c) Throughput in VoIP Traffic at $HighMobility$

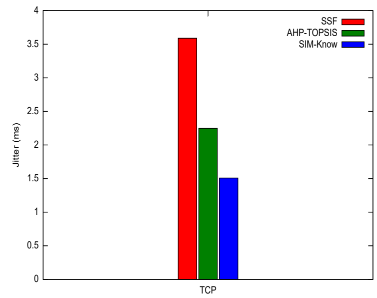
Figure 4.10: Impact on VoIP Traffic.

Figure 4.11 shows that SIM-Know outperformed SSF and AHP-TOPSIS regarding the throughput, delay, and jitter when the wireless network transferred TCP traffic. Specifically, the delay attained by SIM-Know was 91.95% and 80.27% lower than that achieved by SSF and AHP-TOPSIS (Figure 4.11a). The jitter performed by SIM-Know was 57.98% and 32.94% lower than that obtained by SSF and AHP-TOPSIS (Figure 4.11b). The throughput accomplished by SIM-Know was 80.3% and 29.22% higher than SSF and AHP-TOPSIS (Figure 4.11c).

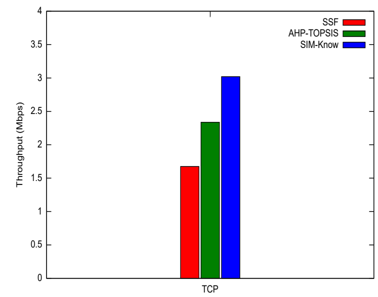
We argue that the improvement in throughput, delay, jitter, and packet loss offered by SIM-Know is due to its context-aware, cognition, and proactivity capabilities, which decreased the number of handovers and instantaneous throughput. In particular, SIM provides the context-aware capability to perform handover decisions through the com-



(a) Delay in TCP Traffic at $HighMobility$



(b) Jitter in TCP Traffic at $HighMobility$



(c) Throughput in TCP Traffic at $HighMobility$

Figure 4.11: Impact on TCP Traffic

prehensive and semantic network view given by the information domains (*Network*, *Application*, *User*, *UserDevice*, and *Handover*).

The KBP's *Reasoning* layer allows the achievement of cognitive HM. The distributed nature of KBP and its continuous updating allow the Handover Initiation phase to be proactive and operate with the local knowledge, built by KBP_M and KBP_N, and the global knowledge, available in KBP_S. It is worth noting that the above results corroborate the idea that proactive approaches are more effective than reactive ones for meeting QoS requirements. On the other hand, we consider that the handover latency of SIM-Know can be addressed by improving the data format and the communication model between KBP_S and both KBP_M and KBP_N. These improvements are out of the scope of this chapter and constitute limitations of the current SIM-Know version.

4.4 Final remarks

This chapter introduced SIM-Know, an approach that performs context-aware, cognitive, and proactive handovers. SIM-Know includes SIM to provide context-awareness to handover decisions and KBP to incorporate cognition of HM. KBP distributes knowledge (local and global) to afford the proactivity capability in HM. The evaluation results showed that thanks to the aforementioned SIM-Know capabilities, our approach overcomes SSF regarding the number of handovers and instantaneous throughputs when the user moves at any speed and, further, equals AHP-TOPSIS when it travels at low and moderate speeds. SSF outperforms SIM-Know and AHP-TOPSIS regarding the handover latency metric because SSF runs a straightforward process for making handover decisions. SIM-Know overcomes AHP-TOPSIS regarding all evaluated metrics when the user moves at high speed, positively impacting the wireless network's performance in terms of delay, throughput, packet loss, and jitter metrics. Considering these results, we concluded that SIM-Know is an attractive and feasible solution for cognitive HM.

Chapter 5

An Autonomic and Cognitive Handover Management Approach in 5G

In wireless communications, the handover aims to connect and disconnects the MD of an AP every time it leaves the network's coverage range. HM comprises three phases [147,148]: handover initiation, network selection, and handover execution. Handover initiation must discover other networks (candidate networks) to collect all information required (e.g., neighbor networks, parameters, and available services). Network selection chooses the most suitable access network (target network) according to diverse parameters and evaluation metrics. Handover execution establishes the network change and releases resources. In this sense, HM is pivotal for providing service continuity, ultra-high reliability, extreme-low latency, and meeting sky-high data rates [72,82]. However, each phase of the handover process involves signaling overhead, defined as exchanging necessary information between the MD and network to facilitate the operation. Therefore, HM is a considerably complex process since it brings heavy overheads and large workloads within network entities, but it is a crucial factor in seamless mobility [92].

HM in 5G networks turns more complicated with many issues and challenges. 5G combines ultra-dense network scenarios and different radio access technologies with short coverage areas increasing signaling overhead in the network due to frequent handovers. At the same time, the users contain a large MDs number and present high mobility raising the workload within network entities due to many handovers. Additionally, the growing applications number with increasingly strict restrictions in terms of QoS (e.g., URLLC) impose conditions on HM to reduce the delay from minimizing the network signaling traffic generated [92]. However, the future 5G network cannot reduce the handover overhead under such circumstances because it relies on a traditionally rigid and complex hierarchical sequence for a handover procedure. Thus, these issues limit the users' achieving a seamless experience, contextualized and personalized, to access services everywhere [93] [94].

Existing handover solutions based on a single-criterion (e.g., RSSI) move the MD context from one cell to another, increasing the signaling overhead [60]. Moreover, the handover decision is user-centric and uses thresholds established to select the most suitable network increasing frequent handovers as well as signaling messages [142]. The context-based approaches use multicriteria and make a decision network-centric based on MADM [97]. This approach exchanges more information among network entities and higher the signaling overhead. Intelligent and cognitive approaches use the latest ML techniques combined with context awareness to optimize the handover parameters and reduce the number of handovers. However, the lack of personalizing the user actions during handover increases the signaling generated. Thus, these approaches hardly handle the handover complexity (heavy signaling) caused by continuous changes to the network at the same time that it shares information in multiple domains of the network under a manual or semi-automatic network management approach.

This chapter presents ZTHM-5G, an autonomous and cognitive HM approach from an ANM point of view to optimize the handover procedure by reducing the interactions and size of the signaling messages. ZTHM-5G introduces an AKBP based on CCL to reduce the number of interactions (i.e., signaling messages) between network entities by controlling its context and making local decisions to be shared during handover. In turn, a semantic and goal-oriented communication model delineates the exchange of meaningful information (i.e., local and global decisions) among network entities to reduce the size of the signaling messages. Finally, a MAS provides a new distributed, scalable, and personalized HM approach. The intelligent agents effectively handle the handover complexity in 5G networks.

The contributions of ZTHM-5G are three-fold.

- AKBP – Agents with Cognitive Closed Loop based on Autonomic Knowledge Base Profile
- Multi-Agent System for distributed, scalable, and personalized HM
- Semantic and goal-oriented communication model

The remainder of this chapter is as follows: Section 5.1 introduces ZTHM-5G, including AKBP, MAS, and Semantic and goal-oriented communication model. Section 5.2 presents the usage case. Section 5.3 presents the evaluation of ZTHM-5G. Finally, some remarks are presented in Section 5.4.

5.1 ZTHM-5G: Zero-Touch Handover Management in 5G

ZTHM-5G proposes an autonomous and cognitive HM approach from an ANM point of view to optimize the handover procedure by reducing the interactions and size of the signaling messages. ZTHM-5G uses the management system decomposition that follows MAPE (Monitor-Analyse-Plan-Execute) paradigm to implement the handover procedure in a distributed way utilizing multiple CCL-driven operations. Moreover, HM distributes the operations at multiple network entities (MD, AP/BS, server centralized), providing a solid separation of management. All network entities process the HM, leading to reduced information exchange compared to the traditional and centralized handover solutions. The exchanged information among network entities has meaningful (i.e., knowledge generated with intention from raw data) since it exchanges local intelligence to build optimal global intelligence. This self-management enables the development of appropriate global policies to optimize handover, improve the whole network performance, and meet user QoS requirements.

ZTHM-5G manages the handover procedure using a cognitive and autonomous approach driven by high-level policies and rules that are flexible and adaptive. Autonomous HM ensures the user access services everywhere, achieving a seamless, personalized, and contextualized experience. ZTHM-5G introduces logical entities for monitoring, analytics, decision-making, and execution distributed and executed at various network entities. The local knowledge and decisions minimize the signaling messages between logical entities to keep management scalable and significantly reduce the reaction time of handover decisions handled locally. Therefore, ZTHM-5G is based on a model-driven approach to perform HM using information and communication models. Furthermore, our approach offers self-monitoring and self-optimization which are fundamental self-properties defined by IBM's MAPE-K.

Figure 5.1 shows an AKBP based on CCL to provide self-management in ZTHM-5G. The AKBP makes local cognitive decisions about handover and reduces the number of interactions among network entities. In turn, a semantic and goal-oriented communication model delineates the exchange of local and global decisions in ZTHM-5G, reducing the size of the signaling messages. Finally, a MAS distributes the AKBP in all network entities creating goal-oriented autonomous agents in ZTHM-5G. This approach groups autonomous agents to work together and accomplish a mission or perform tasks using their communication and coordination capabilities. Therefore, ZTHM-5G provides a new distributed, scalable, and personalized HM approach to handle the handover complexity effectively in 5G networks.

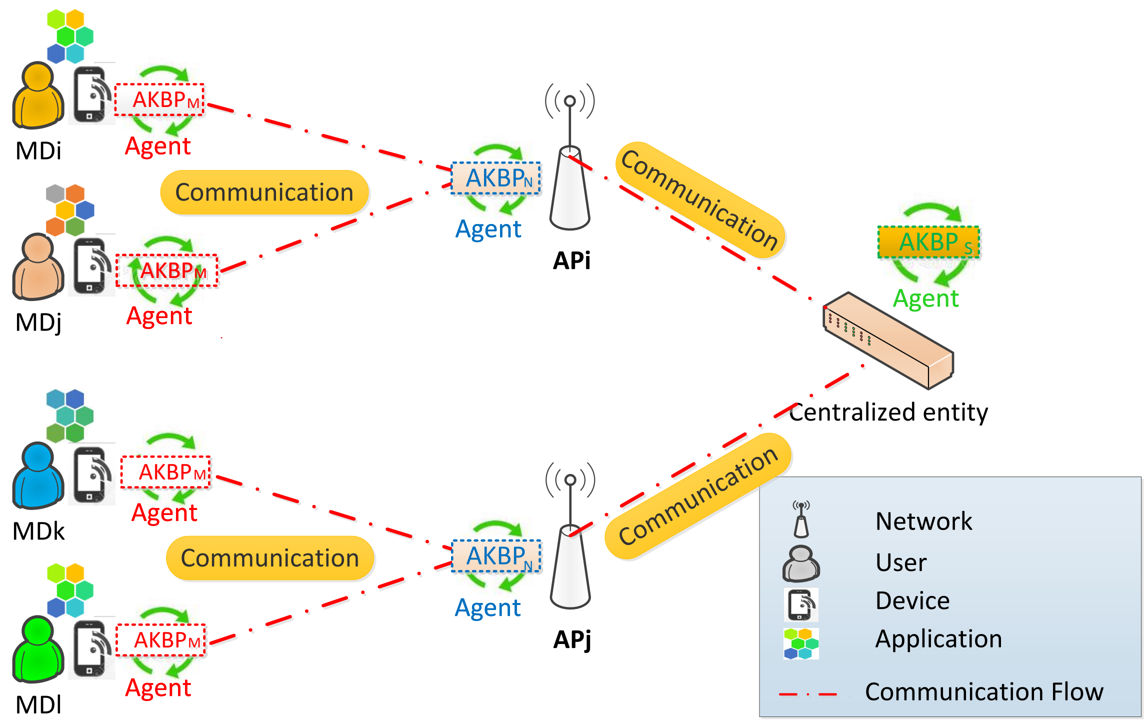


Figure 5.1: ZTHM-5G Architecture

5.1.1 Autonomic Knowledge Base Profile

AKBP is a logical entity using a MAPE-based CCL for monitoring, analytics, decision-making, and handover execution. A set of logical entities are distributed and executed at various network entities to generate local intelligence about HM. AKBP creates self-managed elements, minimizes the signaling messages between logical entities to keep management scalable, and significantly reduces the processing time of handover decisions. This self-managed enables the development of appropriate global policies to optimize the handover, improve the whole network performance, and meet user QoS requirements. Therefore, AKBP makes autonomous decisions for management actions, distributed throughout the network with policy and rules pre-set, achieving HM flexibility and adaptability.

Figure 5.2 depicts the AKBP internal structure. AKBP comprises internal components responsible for monitoring the network environment (context component), representing meaningful information (semantics component), and inferring over the data (learning component) to create a knowledge base. Additionally, the intelligence generated (reasoning component) helps in making decisions (decision component) and actuating (action and execution component) performed during the handover procedure. The context component discovers criteria/parameters by monitoring the network environment. The meaningful information obtained from the semantics component uses an information model well-structured about the handover procedure. The learning component uses an AI-driving model inferring new knowledge. The reasoning component generates local and global intelligence from the knowledge of the network entities drawing meaningful conclusions. The decision component makes decisions concerning handover initiation and network selection, and finally, the actuation component converts the decision into a set of low-level commands. All the components cooperate to achieve MAPE behavior at the network entity (MD, AP/BS, server centralized).

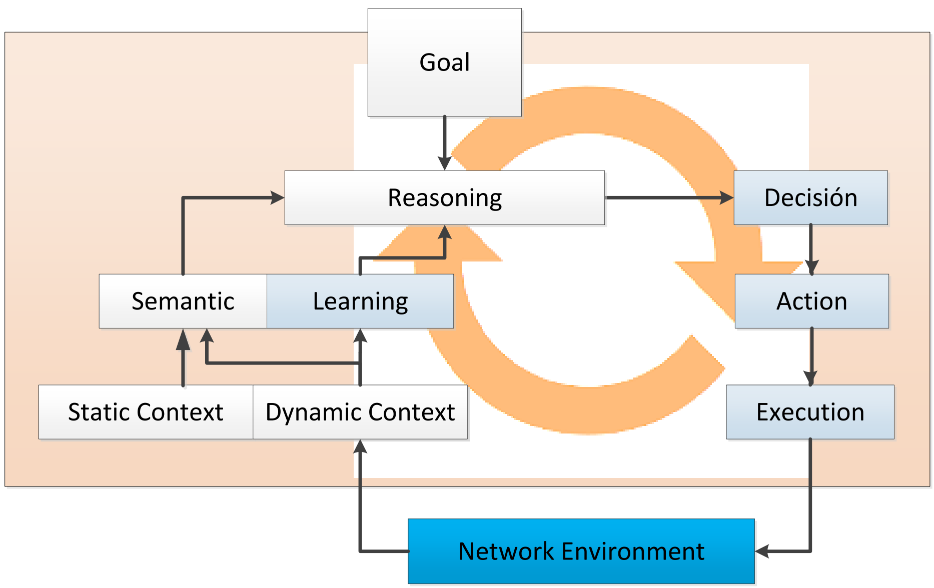


Figure 5.2: AKBP: Autonomic Knowledge Base Profile

AKBP Components

The context component discovers contextual data about the user, network, device, application, and handover to use as criteria in HM. This component monitors the network environment obtaining data that do not change (static context) and dynamic data (dynamic context). The Static Context subcomponent is vital in assisting with neighbor network discovery from its intrinsic capabilities. The Dynamic Context subcomponent updates the network view, including data such as the application requirements and capacity available in a target AP. Thus, the context component enables self-awareness to realize context-based handovers.

The semantics component represents meaningful information using an information model well-structured about the handover procedure. This information model uses a SIM (Figure 4.1 in subsection 4.1.1) [50] instance, which provides a well-designed structure to describe the criteria of multiple context sources by using a common model at a syntactic and semantic level. The Semantic component uses the context component data. Thus, the Semantic component structures the information to achieve intelligent, timely, and context-aware HM. The information model utilizes an ontology-based scheme to capture attributes and support operations of resources and services, increasing portability and reusability.

The learning component adds new information using a data-driving model to acquire actionable knowledge. This component analyses and draws inferences from patterns in a data set constructed from a dynamic context. The learning uses AI-driving algorithms, models, and methods with criteria complete and limited in the information source. An example of the learning component is, improving network selection by classification [51] or clustering methods using criteria analysis to discover candidate networks [64] [166]. Furthermore, the learning component analyzes network entities' personalized preferences to assist the semantic component in creating a knowledge base for reasoning.

The reasoning component draws meaningful conclusions about the handover procedure within the network entities using the implicit knowledge base to generate local and global intelligence. This component uses the knowledge acquired (semantic component) or learned (learning component) and takes into account the network entity preferences (goal) in the derivation of conclusions [167]. In this sense, reasoning allows consciously making sense of the handover process, triggering the events, exchanging knowledge, and determining what operations perform during the HM [168]. Therefore, the reasoning component based on new or existing information can answer questions related to handover.

The decision component makes decisions concerning HM to be implemented and executed on the appropriate network entity. This way, all network entities can know what actions to perform according to the reasoning results of the knowledge base without human intervention. This component enables the network entities to evolve and adapt to changes in either network administrator objectives, user preferences, or applications QoS requirements. For this reason, the decision introduces rules to manage the handover operations of various network entities in response to changes in the network environment. Thus, the decision component triggers the initiation phase, selects the target network, and executes the handover.

The actuation component converts the decision into a set of low-level commands guiding the behavior of network entities to HM. This component (action and execution) performs connection to the target network, disconnection to the serving network, determines the trigger time of the handover, and releases resources in the network entities. Moreover, the interaction among AKBP using the actuation component builds global intelligence to reach a specific goal in the network entity. Therefore, the actuation component executes atomic commands during the handover procedure.

The goals component is a set of high-level abstractions or intents representing user needs and business requirements. Afterward, these high-level intents are generally converted into network policies aligned with expressed goals. In this way, the network administrators define the behavior of the whole network, and the user establishes their goal according to their preferences to access services everywhere. Therefore, the goal component enables the reasoning component to draw optimal conclusions in the network entity to exhibit self-management properties in handover.

AKBP operation

AKBP operates as follows:

First, the context component monitors the network environment and collects raw data on the network entity (AP/BS and user). The data collected from the static context is done only once (configuration time), while the dynamic context continuously collects data to update the local view (run time).

Second, based on contextual information (static and dynamic), the semantic component acquires local knowledge using a SIM (Figure 4.1 in subsection 4.1.1) [50].

At the same time, the learning component helps to determine semantic states, optimize parameters (TTT and HOM), and obtain adaptive thresholds and possible decisions (preferred APs or candidates) using Artificial Intelligence (AI) techniques and methods. For example, the classification result from the dynamic context data (user speed) generates a semantic state such as high, moderate, and low mobility. Other methods (RF) generate a list of candidate networks based on various criteria to select the best network during the handover [51]. This locally learned knowledge uses a data-driven approach.

The acquired knowledge (Semantics) and the learned knowledge (Learning), together with the goals of each network entity, outline significant conclusions to make decisions reasonably. For example, user speed has three semantic states (high, moderate, and slow speed). The dynamic context obtains user speed values between 70 and 50 km/h by a classification method; we will have a semantic state of high speed. Moreover, this same knowledge uses RF to give a list of candidate networks. The goal is to restrict handovers of high-mobility users to small cells. Therefore, the reasoning component concludes that high-mobility users execute handover to large-range cells.

These conclusions are made by the decision component, which manages the handover operations through a series of rules and policies. The policy decides *i)* to filter candidate networks by cell size. *ii)* to send a message with the semantic state of high mobility. *iii)* wait a time to trigger the handover, and *iv)* update the global view.

Finally, the decisions are executed on the same network entity or shared with other entities to generate global intelligence. These actions can be reactive, proactive, or predictive and executed by the actuation component.

5.1.2 Semantic and goal-oriented communication model

The semantic and goal-oriented communication model delineates the exchange of meaningful information (i.e., local and global decisions) among network entities (i.e., AKBP) to reduce the size of the signaling messages during handover. The exchanged information among AKBP has meaningful (goal-oriented) since it exchanges semantic knowledge and enables the development of appropriate global policies to improve the whole network performance and meet user QoS requirements. Semantic communication works based on its innovative "semantic-meaning passing" concept to extract the "meanings" of sent information between a transmitter and a receiver [169]. Additionally, goal-oriented communication identifies the information strictly necessary to accomplish a goal. Therefore, combining the semantic and goal-oriented aspects in a communication model to HM helps to reduce communication resources, minimizing the number and size of the signaling messages.

Figure 5.3 shows the semantic and goal-oriented communication model comprised of four modules. The strategy goal module collects the specific goal of the network entity to optimize performance according to KPIs. In turn, the conflict management module delivers a common goal among competing goals by an agreement consensus-based. Then, the semantic communication module delineates meaningful information exchange among network entities to reduce signaling messages. Finally, the goal module represents a common goal to achieve optimal global HM.

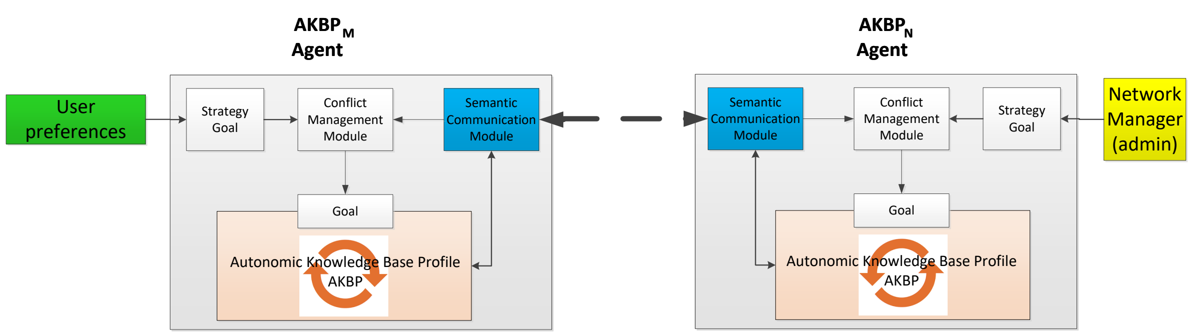


Figure 5.3: Communication Model

Modules

The modules of the communication model enable the semantics to exchange meaningful information and the goal orientation to deploy appropriate policies in HM.

The strategy goal module represents the user needs, and business requirements [170]. On the one hand, the network administrators define the behavior of the network entities to improve the whole network's performance. On the other hand, users establish their preferences to access services everywhere to meet the application's QoS requirements. This strategy goal module includes requirements, goals, and constraints that a network should meet and the outcomes that the network is supposed to deliver. Hence, the strategy goal module allows the users to achieve a seamless experience, contextualized and personalized for the network to obtain an optimal HM performance.

The conflict management module solves conflicts between different, complex, and contrary goals. Since the user demands more resources to improve application QoS, the network administrator manages resources limited and shared to optimize the performance of the whole network during handover. This module achieves an agreement consensus-based [171] to establish a common goal and guide the behavior of the network entity. Hence, the conflict management module sets a common goal and target to make actionable decisions from local optimization to global.

The semantic communication module exchanges "meanings" between a transmitter and a receiver involving generation, transfer, and interpretation of semantic information. This module reduces the size of the signaling messages in the handover procedure to enable efficient communication. The transmitter sends the relevant information in a data format pre-set based on semantics (e.g., RDF, EN) [172, 173]. In the receiver, a match between the knowledge base and received information provides the successfully interpreted information to accomplish a goal. Therefore, the semantic communication module defines a data format used to exchange meaningful information among network entities to reduce the interactions and size of the signaling messages.

The goal module contains all common expectations agreement between the user and network administrator to achieve optimal global performance in the handover. This module allows to performance of the handover procedure in the context of Autonomic Networks according to the zero-touch approach. The goal enables the reasoning component to derivate meaningful conclusions and optimize local decisions. These decisions can answer questions about the optimal time to trigger the handover and select the most suitable access network.

Operation

This section presents the operation of the communication model to enable a goal-oriented exchange of information. The operation follows the steps described below:

- First, the network entity establishes its goals evaluated during the handover procedure. For example, the user defines preferences according to applications, device status, monetary cost, and mobility pattern. These user preferences are strategic user goals for a seamless experience anywhere. AP and BS also present goals for better network performance in a shared environment.

- Second, the goals of the other network entities arrive through the semantic communication module, which reflects the goal of the network as a whole and its interactions with other users. Therefore, the network entity goals conflict with the shared goals of the entire network as they pursue very different objectives.

- Third, the conflict management module resolves the conflicts between goals fighting. This module achieves an agreement common on using the network to obtain optimal performance by serving several users with different preferences and improving user satisfaction by meeting their preferences and application QoS.
- Fourth, the result of this agreement is the goal component integrated into the CCL and AKBP so that the network entity self-manages the handover.

Data Format

The data format is based on semantics to interpret the information exchanges and use the information model introduced in the AKBP [174]. ZTHM-5G uses EN [156] as a data format to describe classes, relationships, and properties of an ontology-based scheme, which enables a lightweight knowledge representation for resource-constrained environments. EN has two representation formats: complete and short packets. Complete packets connect with ontologies to provide detailed descriptions and information. Short packets aim to send identifiers, variables, and templates.

Figure 5.4 shows the data format used to exchange global and local intelligence built with the information from AKBP. The format follows the triplet (Subject, Predicate, Object) encoded using complete packets in EN. The Subject identifies a class in SIM by the combination of ClassId (e.g., ceu101) and ClassType (e.g., *UserDevice*). Each Subject is related to various pairs, Predicate–Object. The Predicate identifies a property (e.g., BatteryStatus) of the Subject while the Object provides the value of such a property (e.g., LowBattery). The Object can also be a ClassId, to represent relationships among Subjects.

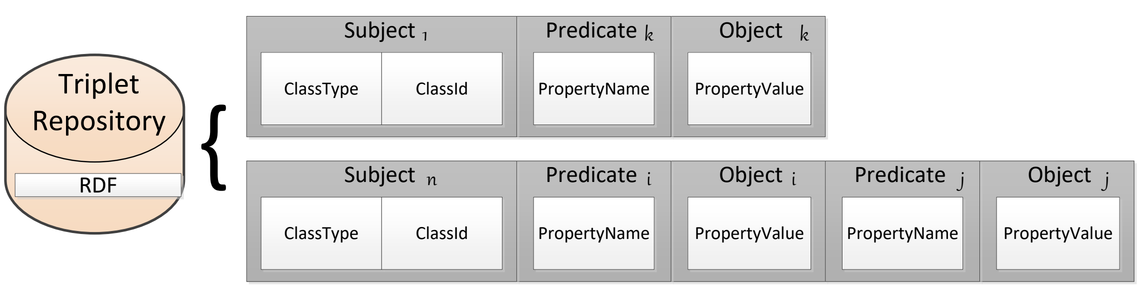


Figure 5.4: $AKBP$ Data Format

5.1.3 Multi-Agent System

MAS provides a new distributed, scalable, and personalized HM approach. The intelligent agents effectively handle the handover complexity in 5G networks and the contextualized and personalized user access to services everywhere. ZTHM-5G distributes the goal-oriented autonomous agents (i.e., AKBP) at multiple network entities ($AKBP_M$ in MD, $AKBP_N$ in AP/BS, $AKBP_S$ in server centralized) to work together and accomplish a mission or perform tasks. Therefore, MAS, with its communication and coordination capabilities provides ZTHM-5G with a solid separation of management to optimize the handover procedure.

The communication capability determines the external behavior of agents AKBP to control their interactions and communications whit other agents. ZTHM-5G delineates communication capability as a communication model to achieve a general goal. Thus, agents AKBP in HM acts proactively to collect context information and make flexible decisions. Moreover, agents AKBP react autonomously to changes in the environment, adapting their actions appropriately. For instance, by changing the capacity available in a target AP, the agent AKBP connects to the following candidate network with resources available.

The coordination capability of each agent AKBP uses collaboration to achieve its goal and maximize its utility by identifying a common goal for choosing and performing coherent actions. This cooperative behavior uses a consensus to agree on an aspect of interest, common value, or state between the agents. However, the cooperation presents issues since the agents AKBP can accept or reject the agreement. Therefore, to ensure that all agents in a network come to an agreement, the consensus has different configurations, such as leader-following, formation, synchronization in robotic arms, and state estimation in sensor networks [171,175]. Another alternative approach uses argumentation to exchange proposals, counterproposals, and reasons supporting them.

The complex process of HM includes a control mechanism traditionally centralized: NCHO, MCHO, MAHO, and NAHO. The centralized handover process delegates complete control to one network entity, bringing heavy overheads and large workloads. Unlike NCHO, MCHO, MAHO, and NAHO, ZTHM-5G adopts distributed control method by splitting and distributing the handover procedure in various simple tasks performed by the agents AKBP. Agents $AKBP_M$ and $AKBP_N$ cooperate on a common cognitive structure (Information and Communication model) using goal-oriented consensus to solve conflict situations. Therefore, ZTHM-5G obtains high reliability, fast running speed, and convenient operation to reduce signaling costs during handover.

5.1.4 ZTHM-5G Operation

Figure 5.5 presents how ZTHM-5G operates in 5G. First, every $AKBP$ specifies its own goal to be met; for example, $AKBP_M$ defines the user goal, and $AKBP_N$ establishes the network goal. Second, every $AKBP$ exchanges its goal using the communication model. Third, the conflict management module generates a common goal for each $AKBP$. Fourth, $AKBP_M$ and $AKBP_N$ collect their ***Static Context***. Fifth, $AKBP_M$ and $AKBP_N$ monitor and update the *Dynamic Context* information related to, for instance, user speed, APs in range, RSSI, and the associated and in-range MDs. Sixth, based on the *Static Context* and *Dynamic Context*, every $AKBP$ generates its local knowledge. For example, the local knowledge in $AKBP_M$ can be $HighMobility$, and in $AKBP_N$ can be $LargeRange$. Seventh, each $AKBP$ analyzes the dynamic context information to identify, classify, or predict patterns. In this way, $AKBP$ obtains learned knowledge. Eighth, every $AKBP$ generates local intelligence by drawing meaningful conclusions and making appropriate decisions. Ninth, $AKBP_M$ executes actions within its network entity or exchanges local intelligence with another network entity. Tenth, $AKBP_M$ sends local intelligence to $AKBP_N$, which builds up global intelligence. Eleventh, $AKBP_N$ performs local actions and sends them to the $AKBP_M$. Twelfth, $AKBP_M$ and $AKBP_N$ evaluate goal compliance to achieve user satisfaction and maintain network performance.

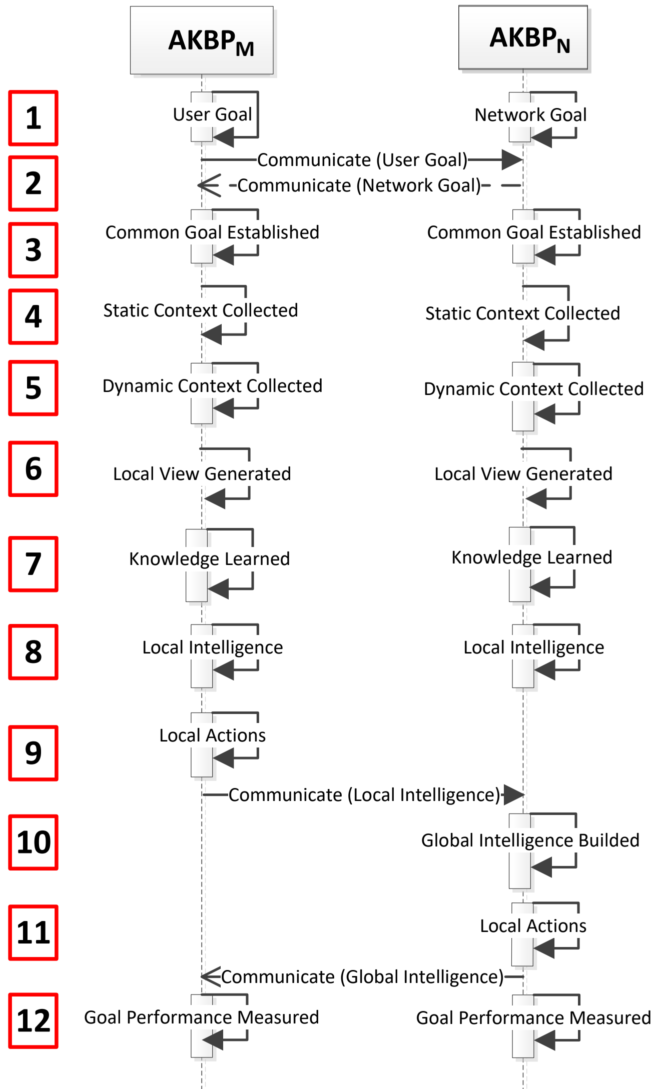


Figure 5.5: ZTHM-5G operation

5.2 Usage case: Maximization of Energy Efficiency

This section validates ZTHM-5G by a usage case. The use case examines the application of the approach proposed in the real world to improve HM. Figure 5.6 shows the usage case operation with $AKBP_M$ instantiated on the MD and $AKBP_N$ in the AP/BS. Additionally, the handover procedure covers three phases: handover initiation, network selection, and handover execution.

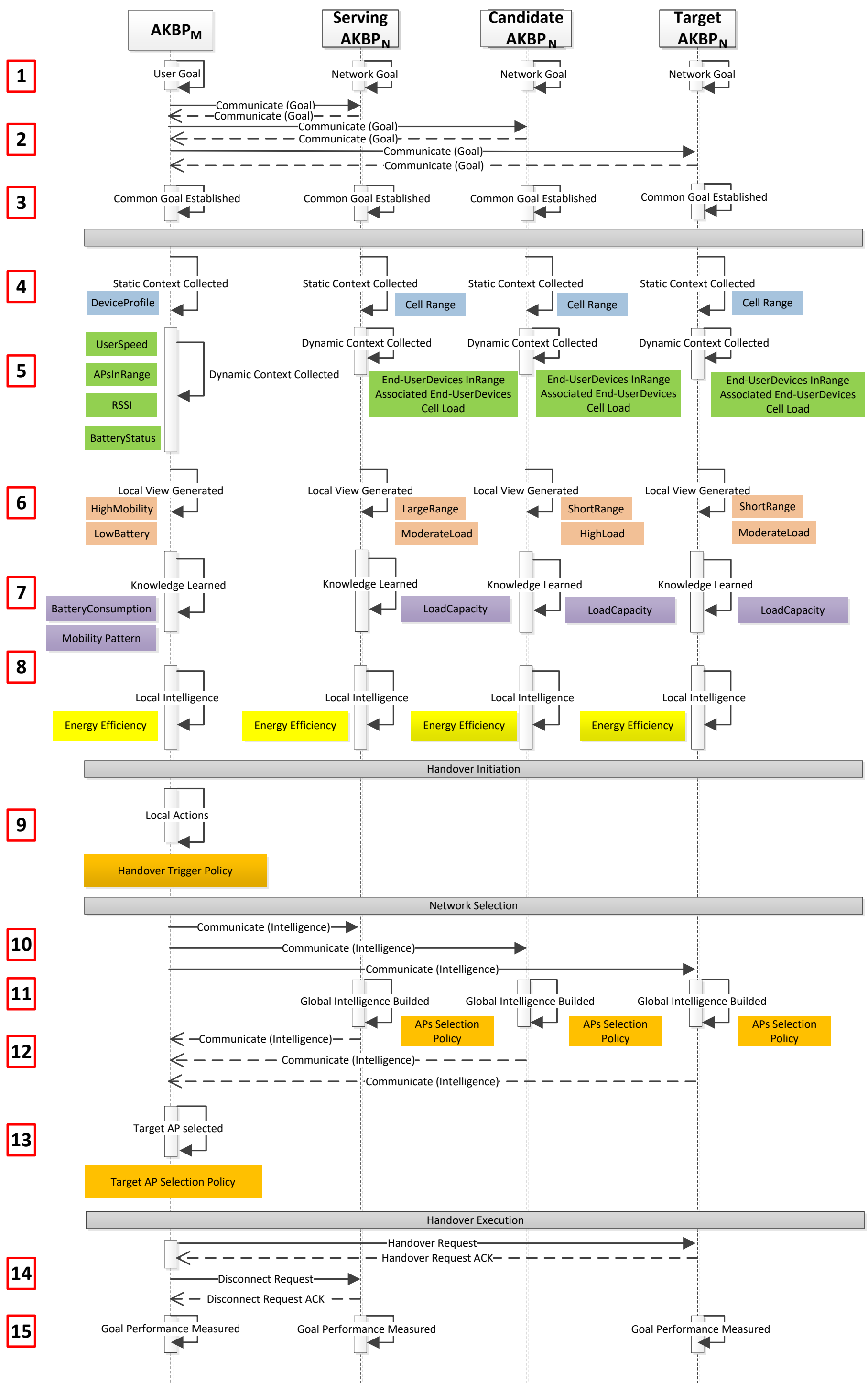


Figure 5.6: Usage case operation

5.2.1 Goal Establishing

This subsection performs the goal-Establishing process for the user and the network through an argumentation-based consensus model, which seeks a common goal.

Step 1. Setting the goals. The user sets a strategic goal for its operation: *"minimizing energy consumption"*. This goal can be accomplished during the handover by either decreasing the network discovery time, triggering the handover at the appropriate time, or selecting networks with lower ranges. The network sets its own goal, which is *"maximizing bandwidth use"*.

Step 2. Determining a common goal. The conflict management module offers a common goal resulting from a consensus between the user and the network. This common goal is *"the maximization of energy efficiency"*. The two above goals conflict since maximizing bandwidth does not regard energy consumption at the risk of shortening the MD operational time. Similarly, with power minimization, the throughput may suffer when connecting to overloaded networks (AP/BS) (i.e., many users connected to the same network or applications with higher bandwidth requirements).

Step 3. Analyzing the common goal. Each network entity analyzes the common target using the AKBP. The AKBP *i)* determines the required resources to achieve the desired goal, *ii)* determines the degree of goal attainment with the available resources, and *iii)* provides the best satisfying solution under a varying amount of resources and priorities of the goal.

5.2.2 AKBP Operation

The $AKBP$ operation describes the interaction of each component of CCL working independently $AKBP_M$ and $AKBP_N$.

Step 4. The static context component gathers static information on each network entity. The $AKBP_M$ collects the device profile that stores device capabilities and technology type. The $AKBP_N$ gathers the technology type, maximum cell load, cell range, localization, and transmitting power.

Step 5. The dynamic context component collects information on each network entity using a database to store the sample and the timestamp obtaining a time series. In the $AKBP_M$, the dynamic context comprises the battery consumption and applications data rate every time interval. The $AKBP_N$ gathers the cell load in terms of bandwidth, the number of connected MDs, and the number of discovered MDs on the range of the AP/BS.

Step 6. The knowledge component establishes semantic classes with the help of SIM for $AKBP_M$ and $AKBP_N$ (Figure 4.1 in subsection 4.1.1). The $AKBP_M$ determines the *UserDevice* and *Application* superclasses. *UserDevice* superclass has a *DeviceStatus* class which represents the state of the MD based on the battery level (e.g., LowBattery, HighBattery). *Application* superclass allows knowledge of the applications running in each MD to generate network traffic (i.e., *NetworkTraffic* class). These applications set QoS (i.e., *QoS* class) requirements (e.g., delay, throughput, and packet loss) for each type of application. For example, the Rule $APInRange$ (Listing 5.1) serves to discover neighboring networks considering RSSI.

Listing 5.1: Rule for APInRange in usage case.

$$APInRange \equiv User \sqcap \exists isInCell.(\exists\, covers.AccessPoint)$$

Listing 5.2 shows the Rule $BatteryStatus$ to define the battery level of the MD. If MD has battery level $battlev > th_{batt}$, MD has a $DevHighBattery$. If a MD has a battery level lower than th_{batt}, MD has a $DevLowBattery$.

Listing 5.2: Rule for BatteryStatus.

$$DevHighBattery \equiv Dev \sqcap (\exists\, hasBatteryStatus.HighBattery)$$
$$DevLowBattery \equiv Dev \sqcap (\exists\, hasBatteryStatus.LowBattery)$$

Listing 5.3 shows that the Rule $ScanningFrequency$ is useful to define the scanning periodicity for context information discovery. If an MD has a high scanning frequency $ScanFreq > th_u$, the MD has a $DevHighScanFreq$. If an MD has a scanning frequency lower than th_u, the MD has a $DevLowScanFreq$.

Listing 5.3: Rule for ScanningFrequency.

$$DevHighScanFreq \equiv Dev \sqcap (\exists\, hasScanFreq.HighScanFreq)$$
$$DevLowScanFreq \equiv Dev \sqcap (\exists\, hasScanFreq.LowScanFreq)$$

$AKBP_N$ defines classes like *AccessPoint*, *Cell*, and *Range*, which represent the range of the cell (i.e., Short and Large Range), and the cell load as Low, Moderate, and High. The *NetworkTraffic* class models the semantic state of the bandwidth consumed by user applications, which influences cell load.

Listing 5.4 shows that Rule $APRange$ is helpful for listing the APs by coverage range. $LargeRange$ is given by $range > th_{range}$ and $ShortRange$ by $range \leq th_{range}$. According to [84], th_{range} can be set to $35m$ for 802.11ac.

Listing 5.4: Rule for APRange in usage case.

$$APLRange \equiv AP \sqcap \exists isCoveredBy.(\exists\, hasRange.LargeRange)$$
$$APSRange \equiv AP \sqcap \exists isCoveredBy.(\exists\, hasRange.ShortRange)$$

Listing 5.5 shows that the Rule $CellLoad$ is useful for defining the cell load. If a cell has load $cellload > th_{hload}$, the cell has a $CellHighLoad$. If a cell has a load higher than

th_{lload} and lower than or equal to th_{hload}, cell has a $CellModLoad$. $CellLowLoad$ is when the cell has load $cellload \leq th_{lload}$. According to [176], th_{hload} can be set to 75% and th_{lload} to 25% load capacity.

Listing 5.5: Rule for CellLoad.

$CellHighLoad \equiv Cell \sqcap (\exists\, hasCellLoad.HighLoad)$
$CellModLoad \equiv Cell \sqcap (\exists\, hasCellLoad.ModerateLoad)$
$CellLowLoad \equiv Cell \sqcap (\exists\, hasCellLoad.LowLoad)$

Step 7. The learning component uses AI techniques and methods to learn from dynamic context information (e,g., SVM and Deep Learning).

$AKBP_M$ uses battery level data to predict battery consumption based on the applications used, the mobility pattern employed [177], and the connected APs. The battery level dataset facilitates the data analysis to create groups semantics (e.g., $LowBattery$, $HighBattery$) by classification.

SVM analyzes the user mobility (i.e., user speed), resulting in a slow-mobility mobility pattern $(UserSpeed.SlowMobility)$ *(Listing 5.6). In addition, deep learning predicts the data rate according to the user's applications leading to a high data rate* $(App.HighDataRate)$ *(Listing 5.7).*

Listing 5.6: Rule for MobilityPattern.

$UserSlowSpeed \equiv User \sqcap (\exists\, hasUserSpeed.SlowMobility)$

Listing 5.7: Rule for Application.

$AppHighDataRate \equiv UserDevice \sqcap (\exists\, hasApp.HighDataRate)$

$AKBP_N$ analyzes network bandwidth, and data performance obtained from the dynamic context and generates knowledge related to the load capacity. This load capacity predicts future changes in the environment produced during handover (i.e., MD connection and disconnection to the AP/BS).

ML or deep learning techniques analyze the load capacity (i.e., cell load), resulting in a low-cellload $(Cell.CellLowLoad)$ *(Listing 5.5).*

Step 8. The reasoning component implemented with propositional logic use DL [154] to conclude the objective of "the maximization of energy efficiency" for $AKBP_M$ and $AKBP_N$ [60, 178].

In the $AKBP_M$, the MD with low battery power prefers to connect to short-range APs (Listing 5.8). Moreover, users with bandwidth-demanding applications prefer to connect to APs with high throughput (Listing 5.9). The user speed affects the battery consumption. Therefore at low speeds ($UserSlowSpeed$, we discover more networks and consume more energy (Listing 5.10). In addition, to reduce battery consumption, the MD can reduce the frequency of monitoring ($ScanFreq$) of the battery level ($BatteryStatus$) using the results of the energy consumption prediction (Listing 5.11).

Listing 5.8: Rule for Energy Efficiency.

$HighEnergyEfficiency \equiv DevHighBattery \sqcap (\exists\, APSRange)$
$LowEnergyEfficiency \equiv DevLowBattery \sqcap (\exists\, APLRange)$
$MediumEnergyEfficiency \equiv DevLowBattery \sqcap APSRange$
$MediumEnergyEfficiency \equiv DevHighBattery \sqcap APLRange$

Listing 5.9: Rule for Energy Consumption by Applications.

$HighEnergyConsumptionApp \equiv AppHighDataRate \sqcap (\exists\, APSRange)$
$LowEnergyConsumptionApp \equiv AppLowDataRate \sqcap (\exists\, APLRange)$
$MediumEnergyConsumptionApp \equiv AppLowDataRate \sqcap APSRange$
$MediumEnergyConsumptionApp \equiv AppHighDataRate \sqcap APLRange$

Listing 5.10: Rule for Energy Consumption by User Speed.

$LowEnergyConsumptionApp \equiv UserHighSpeed \sqcap (\exists\, APLRange)$
$HighEnergyConsumptionSpeed \equiv UserSlowSpeed \sqcap (\exists\, APSRange)$

Listing 5.11: Rule for Energy Consumption by Scan Frequency.

$HighEnergyConsumptionSpeed \equiv UserDevice \sqcap (\exists\, DevHighScanFreq)$
$LowEnergyConsumptionSpeed \equiv UserDevice \sqcap (\exists\, DevLowScanFreq)$

$AKBP_N$ concludes that load capacity (i.e., number of connected devices) and cell coverage must be balanced to achieve optimal energy efficiency [177] [176] (Listing 5.12). In addition, managing the load capacity increases the number of satisfied users from a data rate perspective.

Listing 5.12: Rule for EnergyEfficiency in Access Point.

$HighEnergyEfficiency \equiv (CellLowLoad \sqcap (\exists\, APSRange)) \sqcup (CellHighLoad \sqcap (\exists\, APLRange))$
$LowEnergyEfficiency \equiv (CellHighLoad \sqcap (\exists\, APSRange)) \sqcup (CellLowLoad \sqcap (\exists\, APLRange))$
$MediumEnergyEfficiency \equiv CellModLoad \sqcap (\exists\, APLRange \sqcup \exists\, APSRange)$

5.2.3 Handover Initiation

In handover initiation, $AKBP_M$ discovers other networks (candidate networks) and triggers the handover according to criteria (e.g., RSSI, load, packet loss, throughput, etc.).

Step 9. The decision component establishes a handover initiation policy based on HOM and TTT. In addition, a dynamic contextual information monitoring policy manages battery consumption (Listing 5.13). These policies maximize energy efficiency according to the common goal.

Listing 5.13: Handover Initiation Policy.

1. To reduce scanning frequency:
   $LowEnergyConsumption \equiv UserDevice \sqcap (\exists\, DevLowScanFreq)$
2. To trigger the handover:
   Application Throughput is lower that a threshold ($th_{appdatrate}$).
3. To define candidates APs:
   $CandidateAP \equiv\ MediumEnergyEfficiency \sqcup HighEnergyEfficiency$
4. To send frame:
   $Cell, Range, ShortRange$
   $UserDevice, DeviceStatus, MediumEnergyEfficiency$

The actuation component executes the local actions to reduce the frequency of monitoring to initiate network discovery proactively using the prediction of energy consumption [178]. Additionally, MD triggers the handover sending a message to other network entities with the information about the energy efficiency and taking into account the available throughput of the applications.

5.2.4 Network Selection

Network selection chooses the most suitable access network (target network) according to diverse multicriteria and evaluation metrics.

Step 10. $AKBP_M$ initiates the information exchange to $AKBP_N$ candidates by sending EN complete packets.

Step 11. The decision component of $AKBP_N$ establishes an admission control policy and AP/BS selection using load balancing (Listing 5.14). This load balancing adapts the load capacity to changes in network traffic and the number of connected MDs. Listing 5.15 shows that Rule $CandidateAP$ is useful for creating the list of APs candidates for each MD with $MediumEnergyEfficiency$ or $HighEnergyEfficiency$ in the network.

Listing 5.14: Admission Control Policy.

1. To receive frame:
   $Cell, Range, ShortRange$
   $UserDevice, DeviceStatus, MediumEnergyEfficiency$
2. To decide on the handover:
   $CandidateAP \equiv (UserDevice \sqcap MediumEnergyEfficiency) \sqcap (Cell \sqcap APSRange)$
   $\sqcap(AccessPoint \sqcap (hasMediumEnergyEfficiency \sqcup hasHighEnergyEfficiency)$
3. To send frame:
   $Handover, Selection, Accepted$
   $AccessPoint, APStatus, HighEnergyEfficiency$

Listing 5.15: Rule for CandidateAP in usage case.

$$CandidateAP \equiv MediumEnergyEfficiency \sqcup HighEnergyEfficiency$$

Step 12. The actuation component of $AKBP_N$ executes actions that limit the cell load by denying access to MDs, managing transmission power, redirecting the antennas, and managing throughput according to the data rate. In addition, this component sends significant information related to the established policies using EN complete packets.

Step 13. $AKBP_N$ performs local actions and sends them to $AKBP_M$. Listing 5.16 presents rule $AssociateAP$, which links the MD with the first AP in the list of candidates.

Listing 5.16: Rule for AssociateAP in usage case.

$$AssociationToAP \equiv User \sqcap \exists Uses.UserDevice(\exists Connects.AP)$$

5.2.5 Handover Execution

Handover execution establishes the network change and releases resources.

Step 14. $AKBP_M$ sends a handover request to the Target $AKBP_N$ and returns a handover response message (Handover Request ACK). Additionally, $AKBP_M$ transmits a disconnect request to the Serving $AKBP_N$ and returns with a Disconnect-Request ACK message.

Step 15. $AKBP_M$ and $AKBP_N$ evaluate goal compliance to achieve user satisfaction and maintain network performance.

5.3 ZTHM-5G Evaluation

This section presents the ZTHM-5G evaluation in a WLAN for an IoT environment (i.e., mMTC), aiming to show its behavior regarding the number of handovers and instantaneous throughput, and its impact on various typical network performance metrics (e.g., delay, packet loss, and throughput). Section 5.3.1 depicts the ZTHM-5G prototype and the test environment. Section 5.3.2 shows the performance metrics and traffic generation. Section 5.3.3 presents and discusses the results of ZTHM-5G, SIM-Know, and two well-known handover solutions.

5.3.1 Prototype and Test Environment

We implemented the ZTHM-5G prototype for WLAN, including $AKBP_M$ and $AKBP_N$, by using the Python programming language version 3.0. We also deployed the prototype in a Mininet-WiFi emulator [158] (Figure 5.7) running on an Ubuntu 18.04 VM with a Core i7-3630 processor and 8 GB RAM. Mininet-WiFi adds virtual BSs and APs to classical Mininet [179] to enable the emulation of wireless network environments.

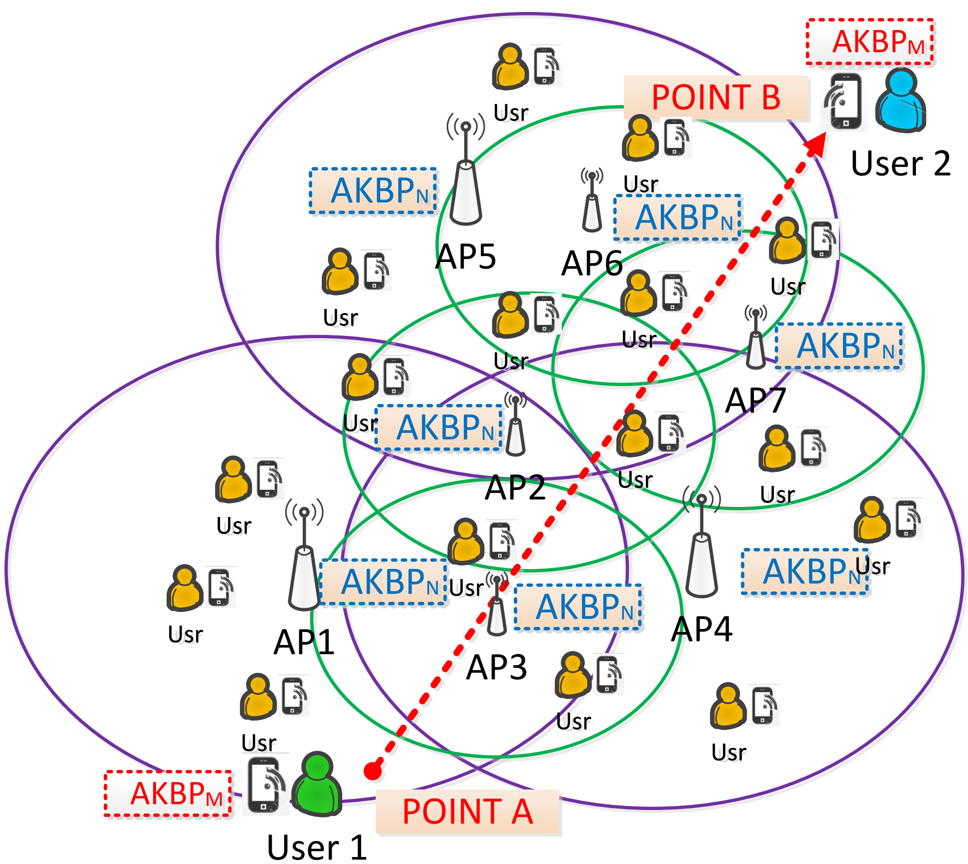


Figure 5.7: ZTHM-5G Test Environment.

Figure 5.7 shows the WLAN test scenario in which we evaluated and compared ZTHM-5G, SIM-Know, SSF, and AHP-TOPSIS. The scenario, deployed in Mininet-WiFi, included seven APs, an MD associated with $User1$, and another MD linked to $User2$.

Additionally, thirteen (13) MDs are associated with the APs and without mobility. In particular, we used three APs with a large coverage range (i.e., up to $75m$ for AP_1, AP_4, and AP_5 with $802.11n$) and four with a short coverage range (i.e., up to $35m$ for AP_2, AP_3, AP_6, and AP_7 with $802.11g$). We also analyzed the performance when the $User\ 1$ moved from Point A to Point B, following a straight line without directional change at a constant speed. We used two battery statuses for testing: $< 20\%$ corresponding to $LowBattery$, and $> 20\%$ to $HighBattery$. The MD associated with $User\ 1$ transmitted traffic (VoIP or TCP) to the MD linked to $User\ 2$, which was static. We repeated the experiments thirty-three times to obtain results with a 95% confidence level. Table 5.1 summarizes the setup of the experiments.

Table 5.1: ZTHM-5G Experiment Setup.

Parameters	**Value**
Wireless technology	802.11n, 802.11ac
Emulation area	200 $\times$ 200 m
Carrier frequency	2.4 GHz
Channel bandwidth	20 MHz
Transmission power of cells large-range/short-range	24/16 dBm
Path loss model from cells	Log-Distance Propagation Loss/ITU-R P1283
Emulation time for HighMobility	30 s
Emulation time for ModerateMobility	80 s
Emulation time for SlowMobility	180 s
TCP traffic	Flows with constant inter-departure time between packets (1000 pkts/s) and constant packets size (512 bytes)
VoIP traffic	Flows with audio code (G.711.2 - 84 Kbps and 50 pkts/s) transmitted using real time protocol and voice activity detection

5.3.2 Performance Metrics and Traffic Generation

We compared ZTHM-5G to SSF, AHP-TOPSIS, and SIM-Know regarding the number of handovers, number of instantaneous throughput, handover latency, signaling overhead, and various well-known network performance metrics (throughput, delay, jitter and packet loss) [161]. The quantity of handovers is the number of transfers an MD makes when it moves from one place to another [162]. The instantaneous throughput (throughput drops) represents the times that the number of bytes transmitted falls to zero because of a handover [51]. The handover latency is the time that elapses between the instant the MD sends the last link-going-down message to the serving AP and the moment the MD establishes the connection with the target AP [82]. Handover signaling overhead is the data generated during the handover process to facilitate the operation. However, the handover process interrupts the data flow and results in the reduction of the MD throughput [64].

In the emulation experiments, scripts for generating traffic were developed by using the iPerf3 [163], and D-ITG [164] tools. We used D-ITG to generate VoIP flows with audio code (G.711.2 - 84 Kbps and 50 pkts/s) transmitted using the real-time protocol and voice activity detection. We used iperf3 to generate TCP flows with constant inter-departure time between packets (1000 pkts/s) and constant packet size (512 bytes).

5.3.3 Results and Analysis

Figure 5.8 evidences the maximization of energy efficiency made by the ZTHM-5G approach. User goal setting and consensus with network goals drive HM toward a personalized, seamless, contextualized experience. The values of 6.37%, 9.86%, and 9.73% concerning SSF, AHP-TOPSIS, and SIM-Know confirm that ZTHM-5G meets user preferences to minimize battery consumption and access services anywhere. In addition, the cognitive and autonomous approach with high-level policies of ZTHM-5G efficiently handles the complexity of the handover process. These goal-oriented and adaptive policies achieved a balance between the user and the network.

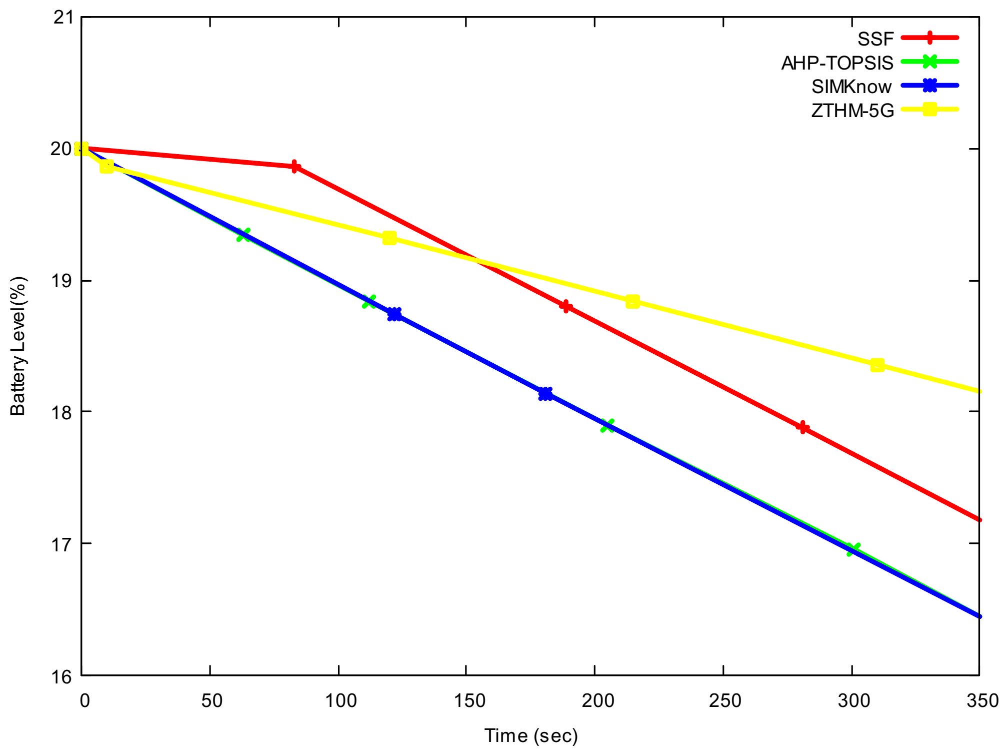


Figure 5.8: Usage Case: Maximization of Energy Efficiency.

Table 5.2 reveals that when the user moved at a slow speed, ZTHM-5G behaved as AHP-TOPSIS does concerning the number of handovers and instantaneous throughput. Furthermore, ZTHM-5G outperformed SSF on these metrics since it uses multi-criteria to make an appropriate decision. SIM-Know and ZTHM-5G behavior are similar in handovers since they use the same KBP in the semantic component. However, the goal-oriented approach of ZTHM-5G satisfies the user preferences to execute a handover when necessary. Figure 5.9 shows the first handover performed by ZTHM-5G to meet the user goal of maximizing of energy efficiency. Additionally, this Figure corroborates that ZTHM-5G carried out cognitive and proactive handovers before SSF and AHP-TOPSIS.

Table 5.2: ZTHM-5G Handover Performance.

Parameter	SSF	AHP-TOPSIS	SIM-Know	ZTHM-5G
Number of handovers	5	4	3	4
Number of instantaneous throughput	3	4	2	4

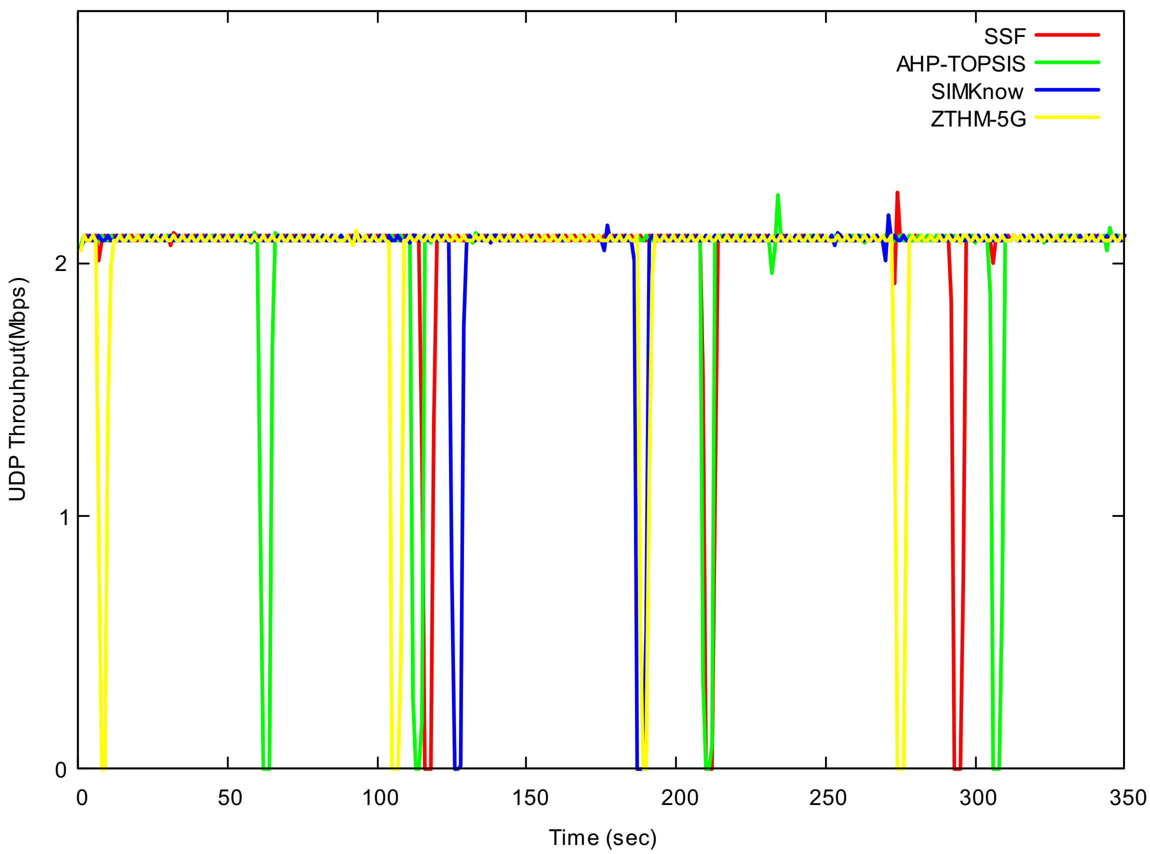


Figure 5.9: Usage Case: Throughput.

Figure 5.10 shows that ZTHM-5G selected networks with short coverage range due to its low battery level and following the user's goal. Contrary to the SIM-Know approach, which selected networks with a large coverage range, obtaining less handover. ZTMH-5G and AHP-TOPSIS selected the same APs since they use multicriteria. However, due to the reasoning and learning component, ZTHM-5G executes the handover at the appropriate time. Additionally, SSF performs a wrong selection of the cell at time T=250 seconds, decreasing user throughput (Figure 5.10). In this case, ZTHM-5G uses the context-aware to wait a time interval and select the network.

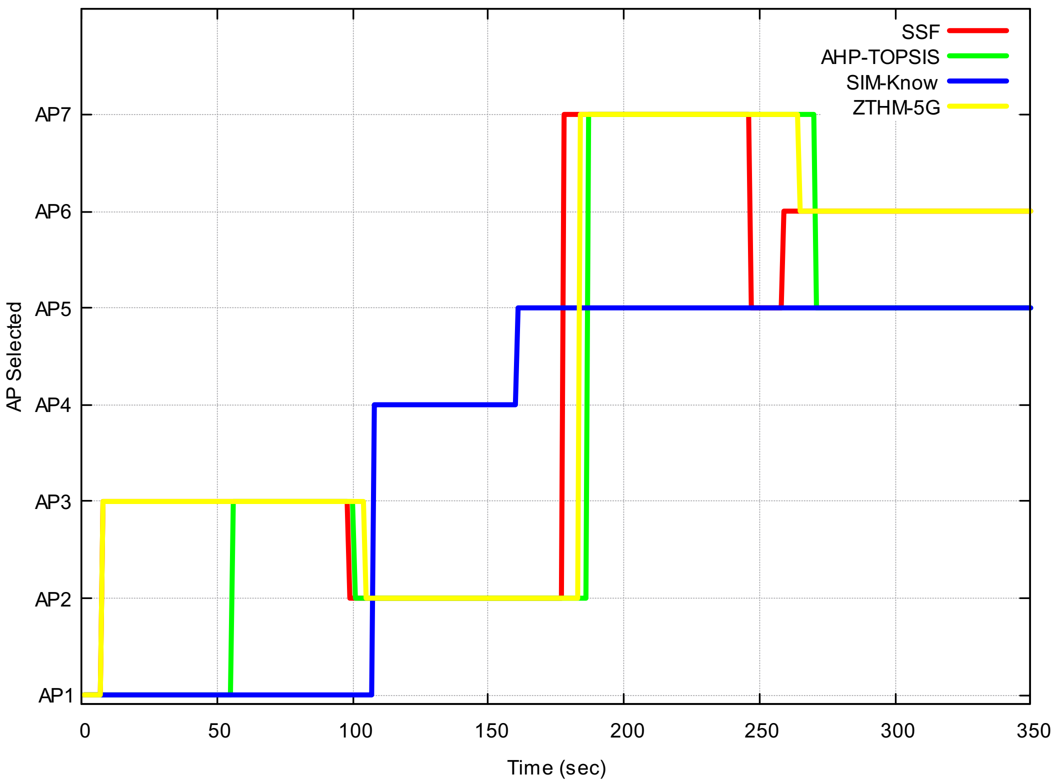
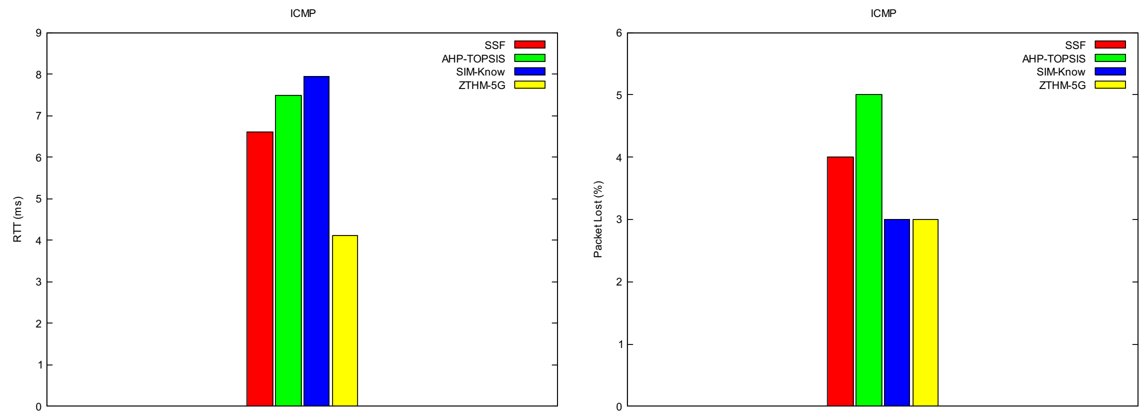


Figure 5.10: Usage Case: AP selections.

Figure 5.11 a shows the Round Trip Time (RTT) in the four evaluated approaches, obtaining that ZTHM-5G has the lowest value. This value means that ZTHM-5G uses the AKBP to perform HM internally in each network entity, reducing the signaling overhead. Furthermore, this result reveals that ZTHM-5G made efficient and appropriate network selection decisions. These selected networks make up a network topology to improve the speed and reliability of the network connection (packet loss lower than 3% see Figure 5.11b).



(a) RTT in ICMP Traffic at $Low Mobility$ (b) Packet Loss in ICMP Traffic at $Low Mobility$

Figure 5.11: Impact on ICMP Traffic.

Figure 5.12 shows, as expected that ZTHM-5G obtained higher throughput than SSF, which is corroborated at all user speeds since our approach is knowledge-based, and SSF makes decisions considering only single-criterion. ZTHM-5G had higher throughput than AHP-TOPSIS with values of 0.2% at slow speed and 1% for high speed because our approach is proactive and, according to [165], AHP-TOPSIS is reactive; proactivity shortens the handover initiation phase [89]. SIM-Know outperforms ZTHM-5G by 0.94% for high and slow user speed and 7.6% for moderate user speed. These values show that ZTHM-5G integrates more internal components in decision-making with high-level policies, increasing latency and affecting throughput. At the same time, SIM-Know uses rule-based reasoning that requires little time to execute the handover.

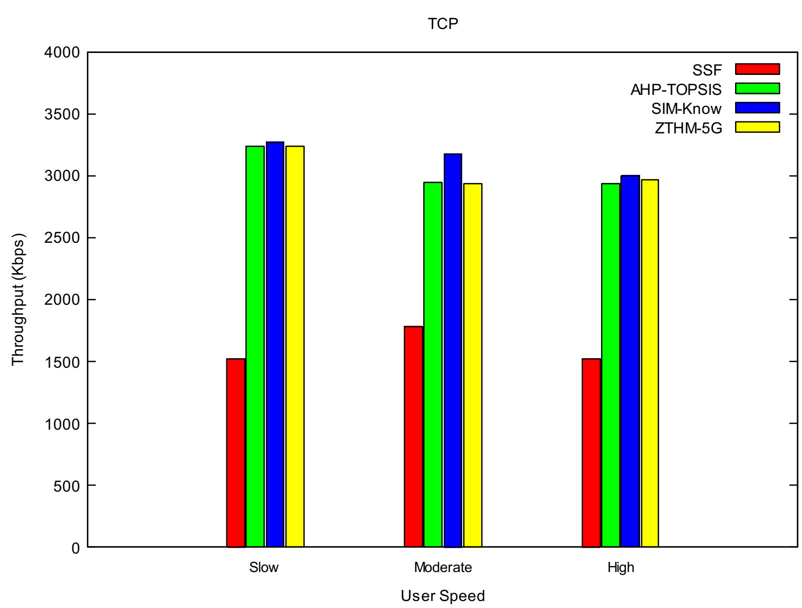
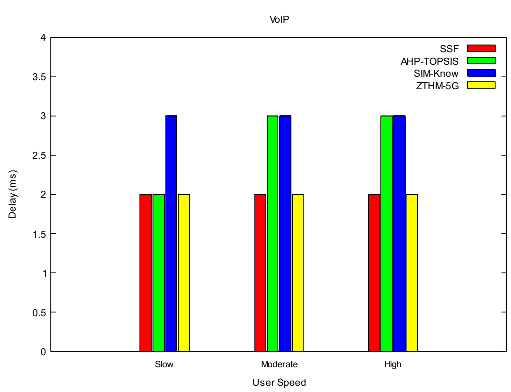
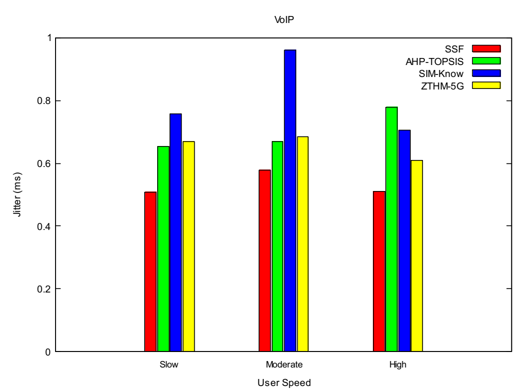


Figure 5.12: Throughput vs User Speed.

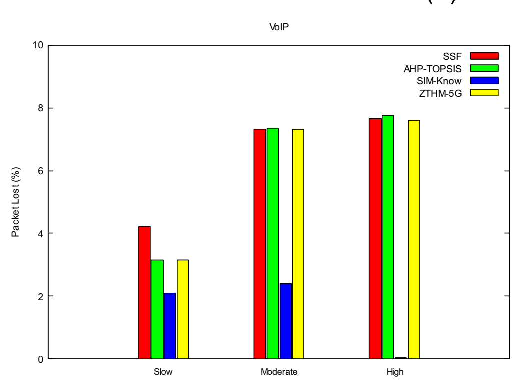
Next, we present how ZTHM-5G, SIM-Know, SSF, and AHP-TOPSIS impact various network performance metrics when the user moves with different speeds ($HighMobility$, $ModerateMobility$, and $SlowMobility$). Figure 5.13 depicts ZTHM-5G overcoming SIM-Know, SSF, and AHP-TOPSIS regarding the delay, jitter, and packet loss when the wireless network transferred VoIP/UDP traffic. In particular, the delay attained by ZTHM-5G was $6.13\%\ to\ 14.84\%$, and $21.7\%\ to\ 24.17\%$ lower than that achieved by AHP-TOPSIS and SIM-Know at all speeds. But, ZTHM-5G was a delay of $2.34\%\ to\ 4.32\%$ higher than that achieved by SSF (Figure 5.13a). The jitter obtained by ZTHM-5G was between $19.37\%\ to\ 31.43\%$ and $2.29\%\ to\ 2.09\%$ higher than that obtained by SSF, AHP-TOPSIS at all speeds. ZTHM-5G got a lower jitter between 11.62% and 28.82% than that obtained by SIM-Know (Figure 5.13b). The packet loss of ZTHM-5G was 0.52% and 0.13% for high and moderate user speed and 24.94% for slow user speed lower than that obtained by SSF. ZTHM-5G obtained a lower packet loss between 1.8% and 0.4% than that obtained by AHP-TOPSIS for high and moderate user speed, however for slow user speed, the value obtained was 0.31% higher. On the other hand, SIM-Know was much superior to ZTHM-5G in all speeds with values obtained greater than 50% (Figure 5.13c).



(a) Delay in VoIP Traffic



(b) Jitter in VoIP Traffic



(c) Packet Lost in VoIP Traffic

Figure 5.13: VoIP Traffic vs User Speed.

We argue that the improvement in throughput, delay, jitter, and packet loss offered by ZTHM-5G compared to SSF and AHP-TOPSIS is due to its context awareness, cognition, proactivity, and autonomic capabilities. In particular, ZTHM-5G uses AKBP, which provides local intelligence on each network entity and shares such intelligence through a semantic and goal-oriented communication model to generate the global intelligence needed to optimize HM. The effective management of the complexity of the handover process in ZTHM-5G is reflected in the comparison made with SSF when using a single-criterion since all the metrics were superior. Although ZTHM-5G surpassed SIM-Know in some metrics, the additional components ensure user-oriented HM regarding preferences, device status, applications, and network context.

ZTHM-5G, with the learning, decision reasoning, and actuation components integrated into a CCL, offers a robust AKBP that generates local intelligence in the network entity. This local intelligence optimizes the handover process from its context component by discovering multicriteria in the initiation phase. Later in the network selection phase, the semantics and learning component analyzes the acquired and learned knowledge to draw conclusions in the reasoning component. The decision and actuation components help in the phase of handover execution by taking local and global actions. ZTHM-5G allows HM according to an agreement between the goals proposed by the user and the network, resolving conflicts between goals and offering a balance to optimize the handover process. The semantic and goal-oriented communication model was partially evaluated with indirect metrics revealing reduced signaling overhead. In future works, more in-depth studies will be made with direct metrics that establish the number of signaling messages exchanged and their size.

5.4 Final remarks

In this chapter, we introduced ZTHM-5G, an approach that performs autonomous and cognitive handovers from an ANM point of view. ZTHM-5G includes AKBP to reduce the number of signaling messages between network entities by self-management of its context and generating local intelligence to be shared during handover. In addition, a semantic and goal-oriented communication model delineates the exchange of local intelligence reducing the size of the signaling messages. ZTHM-5G distributes the AKBP in all network entities creating goal-oriented autonomous agents using MAS. Therefore, ZTHM-5G enables the development of appropriate global policies to optimize handover, improve the whole network performance, and meet user QoS requirements.

The evaluation results showed that thanks to the aforementioned ZTHM-5G capabilities, our approach overcomes SSF regarding the number of handovers and instantaneous throughput when the user moves at any speed and, further, equals AHP-TOPSIS when it moves at low and moderate speeds. SIM-Know overcomes ZTHM-5G regarding all evaluated metrics when the user moves at a high speed. ZTHM-5G positively impacts the wireless network's performance in terms of delay, throughput, packet loss, and jitter metrics at all user speeds. ZTHM-5G and their semantic and goal-oriented communication model were partially evaluated with indirect metrics revealing reduced signaling overhead. Considering those results, we concluded that ZTHM-5G is a feasible solution for autonomous and cognitive HM since it handles the handover complexity effectively in 5G networks.

Chapter 6

Conclusions

This thesis presented the investigation carried out to verify the hypothesis: **An SDN/NFV ecosystem allows performing mobility management efficiently in IoT to meet QoS.** Considering the hypothesis, this work proposed three components to perform MM in IoT: NetSel-RF, SIM-Know, and ZTHM-5G.

This thesis concludes that an SDN/NFV ecosystem supports MM in IoT since it uses a programmable SDN plane to virtualize the network entities. The 5G system architecture supports data connectivity and services, enabling deployments to use techniques such as e.g. NFV and SDN, according to the specification 3GPP TS38.300. The SIM-Know approach proposed in this thesis can operate in 5G by running KBP_M in the MD (i.e., User Equipment (UE)), KBP_N in gNodeB (gNB), and KBP_S in the Core Network (CN). Moreover, the 5G System architecture consists of network functions (NF): Policy Control Function (PCF), Access and Mobility Management Function (AMF), Session Management Function (SMF), User Plane Function (UPF), Network Data Analytics Function (NWDAF), UE radio Capability Management Function (UCMF). Therefore, the network functions interact among them during the handover procedure (preparation, execution, and completion) to enhance the QoS provisioning capability in MDs connected to the network.

The global information allows determining the selection criteria more efficiently to carry out the MD association in a centralized manner. All approaches (NetSel-RF, SIM-Know, and ZTHM-5G) use criteria coming from the network, user preferences, devices, and applications to perform context-aware, cognitive, and proactive handovers. In Addition, SIM represented the global and local knowledge of the network state to make appropriate and contextual handover decisions. SIM provided a well-designed structure to facilitate the discovery and regular access to criteria of multiple sources using a common model at a syntactic (CIM) and semantic level (OWL). The SIM-Know approach distributed and instanced SIM in various network entities named KBP. KBP comprises layers (context, semantics, and reasoning) and processes (collaboration and adaption) to provide local and global intelligence. ZTHM-5G uses AKBP for self-management of the handover procedure, which provides local intelligence on each network entity and shares such intelligence through a semantic and goal-oriented communication model to generate the global intelligence needed to optimize HM. NetSel-RF outperforms the SSF and AHP-TOPSIS approaches regarding the number of handovers, ping-pong, and instantaneous throughput. Moreover, SIM-Know and ZTHM-5G overcome SSF and AHP-TOPSIS regarding all evaluated metrics, positively impacting the wireless network's performance in delay, throughput, packet loss, and jitter metrics.

6.1 Answers for the fundamental question

The issues related to insufficient criteria to make decisions about the handover and the service disruption during the handover are related to the need for an approach for MM. In this way, the question guided the investigation of mobility management in IoT supported by SDN/NFV ecosystem.

Fundamental question: *How to carry out efficiently mobility management in IoT to meet QoS?*

HM is pivotal for providing service continuity, ultra-high reliability, extreme-low latency, and meeting sky-high data rates in wireless communications. Current HM approaches may lead to unnecessary and frequent handovers due to a partial network view. Additionally, the wrong network selection decreases the throughput and increases packet loss. For this reason, SIM-Know improves HM by including SIM that enables context-aware and multicriteria handover decisions. In this way, SIM represented the global and local knowledge of the network state to make appropriate and contextual handover decisions. SIM is lightweight to reduce traffic and processing time. SIM-Know also introduces a SIM-based distributed KBP that provides local and global intelligence to make contextual and proactive handover decisions.

The handover procedure is traditionally rigid and with a complex hierarchical sequence most demanding in cost and time consumption. For this reason, ZTHM-5G proposed an autonomous and cognitive HM approach from an ANM point of view to optimize handover. This approach reduced the number of interactions among network entities by generating local intelligence using an AKBP. Furthermore, a semantic and goal-oriented communication model delineated the exchange of local and global intelligence while reducing the size of the signaling messages. This model improves the QoS provision by reducing the chance of service interruption while keeping network signaling traffic with a short delay and small cost. Therefore, the ZTHM-5G results reveal that the handover-related signaling cost is lower than traditional HM approaches.

6.2 Future work

During the development of this thesis, we observed interesting opportunities for further research. These opportunities are outlined as follows.

- More extensive evaluations in large emulated environments, for example, ultra-dense networks with a more significant number of users and MD, as well as applications with more stringent QoS requirements.
- Enriching the proposed solutions with SDN and NFV capabilities would address the scalability problem imposed on HM by Industrial IoT and the massive use case of 5G IoT.
- The creation of an efficient model to communicate KBP and evaluate SIM-Know when making handover decisions in scenarios with many MD, high network traffic, and high load on APs.
- Implementing and evaluating the semantic and goal-oriented communication model introduced in ZTHM-5G would favor introducing semantic services in network management.

Bibliography

[1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, “Internet of Things (IoT): A vision, architectural elements, and future directions,” *Future Generation Computer Systems*, vol. 29, pp. 1645–1660, sep 2013.

[2] E. Borgia, “The internet of things vision: Key features, applications and open issues,” *Computer Communications*, vol. 54, pp. 1–31, dec 2014.

[3] Y.-K. Hsiao and Y.-W. Lin, “A Mobility Management Scheme for Internet of Things,” in *Mobile, Ubiquitous, and Intelligent Computing*, ch. A Mobility, pp. 569–575, Springer Berlin Heidelberg, 2014.

[4] A. George, A. Kumar, D. Cavalcanti, and D. P. Agrawal, “Protocols for mobility management in heterogeneous multi-hop wireless networks,” *Pervasive and Mobile Computing*, vol. 4, no. 1, pp. 92–116, 2008.

[5] M. Zekri, B. Jouaber, and D. Zeghlache, “A review on mobility management and vertical handover solutions over heterogeneous wireless networks,” *Computer Communications*, vol. 35, pp. 2055–2068, oct 2012.

[6] D. Wu, D. I. Arkhipov, E. Asmare, Z. Qin, and J. A. McCann, “UbiFlow: Mobility management in urban-scale software defined IoT,” in *Proceedings - IEEE INFOCOM*, vol. 26, pp. 208–216, IEEE, apr 2015.

[7] A. Whitmore, A. Agarwal, and L. Da Xu, “The Internet of Things—A survey of topics and trends,” *Information Systems Frontiers*, vol. 17, pp. 261–274, apr 2015.

[8] J. Wan, S. Tang, Z. Shu, D. Li, S. Wang, M. Imran, and A. V. Vasilakos, “Software-Defined Industrial Internet of Things in the Context of Industry 4.0,” *IEEE Sensors Journal*, vol. 16, no. 20, pp. 7373–7380, 2016.

[9] F. Sallabi and K. Shuaib, “Internet of things network management system architecture for smart healthcare,” in *2016 Sixth International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, pp. 165–170, IEEE, jul 2016.

[10] J. Mineraud, O. Mazhelis, X. Su, and S. Tarkoma, "A gap analysis of Internet-of-Things platforms," *Computer Communications*, vol. 89-90, pp. 5–16, sep 2016.

[11] K. Nahrstedt, H. Li, P. Nguyen, S. Chang, and L. Vu, "Internet of mobile things: Mobility-driven challenges, designs and implementations," in *Proceedings - 2016 IEEE 1st International Conference on Internet-of-Things Design and Implementation, IoTDI 2016*, pp. 25–36, IEEE, apr 2016.

[12] T. Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014–2019 White Paper," 2011.

[13] A. Achour, L. Deru, and J. C. Deprez, "Mobility management for Wireless Sensor Networks a state-of-the-art," *Procedia Computer Science*, vol. 52, no. 1, pp. 1101–1107, 2015.

[14] S. M. Ghaleb, S. Subramaniam, Z. A. Zukarnain, and A. Muhammed, "Mobility management for IoT: a survey," *Eurasip Journal on Wireless Communications and Networking*, vol. 2016, p. 165, dec 2016.

[15] F. Ganz, R. Li, P. Barnaghi, and H. Harai, "A Resource Mobility Scheme for Service-Continuity in the Internet of Things," in *2012 IEEE International Conference on Green Computing and Communications*, pp. 261–264, IEEE, nov 2012.

[16] A. Sarma, S. Chakraborty, and S. Nandi, "Deciding Handover Points Based on Context-Aware Load Balancing in a WiFi-WiMAX Heterogeneous Network Environment," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 348–357, jan 2016.

[17] F. Siddiqui and S. Zeadally, "Mobility management across hybrid wireless networks: Trends and challenges," *Computer Communications*, vol. 29, pp. 1363–1385, may 2006.

[18] S. Tomovic, K. Yoshigoe, I. Maljevic, and I. Radusinovic, "Software-Defined Fog Network Architecture for IoT," *Wireless Personal Communications*, vol. 92, pp. 181–196, jan 2017.

[19] J. Chen, B. Liu, H. Zhou, Q. Yu, L. Gui, and X. S. Shen, "QoS-Driven Efficient Client Association in High-Density Software-Defined WLAN," *IEEE Transactions on Vehicular Technology*, vol. 66, pp. 7372–7383, aug 2017.

[20] M. D. Jaraiz-Simon, J. A. Gomez-Pulido, and M. A. Vega-Rodriguez, "Embedded intelligence for fast QoS-based vertical handoff in heterogeneous wireless access networks," *Pervasive and Mobile Computing*, vol. 19, pp. 141–155, may 2015.

[21] W. Bao, D. Yuan, Z. Yang, S. Wang, W. Li, B. B. Zhou, and A. Y. Zomaya, "Follow Me Fog: Toward Seamless Handover Timing Schemes in a Fog Computing Environment," *IEEE Communications Magazine*, vol. 55, pp. 72–78, nov 2017.

[22] L. Tartarini, M. A. Marotta, E. Cerqueira, J. Rochol, C. B. Both, M. Gerla, and P. Bellavista, "Software-defined handover decision engine for heterogeneous cloud radio access networks," *Computer Communications*, vol. 115, pp. 21–34, jan 2018.

[23] M. Khan and K. Han, "An optimized network selection and handover triggering scheme for heterogeneous self-organized wireless networks," *Mathematical Problems in Engineering*, vol. 2014, pp. 1–11, 2014.

[24] M. Khan and K. Han, "A Vertical Handover Management Scheme based on Decision Modelling in Heterogeneous Wireless Networks," *IETE Technical Review*, vol. 32, pp. 402–412, nov 2015.

[25] G. Mahardhika, M. Ismail, and K. Mat, "Multi-criteria vertical handover decision in heterogeneous network," in *2012 IEEE Symposium on Wireless Technology and Applications (ISWTA)*, (New York, New York, USA), pp. 1–4, IEEE, sep 2012.

[26] B. Ng, A. Deng, Y. Qu, and W. K. Seah, "Changeover prediction model for improving handover support in campus area WLAN," *Proceedings of the NOMS 2016 - 2016 IEEE/IFIP Network Operations and Management Symposium*, pp. 265–272, 2016.

[27] J. A. Santana, E. Macías, Á. Suárez, D. Marrero, and V. Mena, "Adaptive Estimation of WiFi RSSI and Its Impact Over Advanced Wireless Services," dec 2016.

[28] S. Goudarzi, W. H. Hassan, M. H. Anisi, and S. A. Soleymani, "MDP-Based Network Selection Scheme by Genetic Algorithm and Simulated Annealing for Vertical-Handover in Heterogeneous Wireless Networks," *Wireless Personal Communications*, vol. 92, pp. 399–436, jan 2017.

[29] M. Khan, S. Din, M. Gohar, A. Ahmad, S. Cuomo, F. Piccialli, and G. Jeon, "Enabling multimedia aware vertical handover Management in Internet of Things based heterogeneous wireless networks," *Multimedia Tools and Applications*, pp. 1–23, may 2017.

[30] X. Wang, D. Qu, K. Li, H. Cheng, S. K. Das, M. Huang, R. Wang, and S. Chen, "A flexible and generalized framework for access network selection in heterogeneous wireless networks," *Pervasive and Mobile Computing*, vol. 40, pp. 556–576, sep 2017.

[31] A. Raschella, F. Bouhafs, M. Seyedebrahimi, M. Mackay, and Q. Shi, “A centralized framework for smart access point selection based on the Fittingness Factor,” in *2016 23rd International Conference on Telecommunications, ICT 2016*, pp. 1–5, IEEE, may 2016.

[32] A. Raschella, F. Bouhafs, M. Seyedebrahimi, M. Mackay, and Q. Shi, “Quality of Service Oriented Access Point Selection Framework for Large Wi-Fi Networks,” *IEEE Transactions on Network and Service Management*, vol. 14, pp. 441–455, jun 2017.

[33] M. E. Berezin, F. Rousseau, and A. Duda, “Multichannel Virtual Access Points for Seamless Handoffs in IEEE 802.11 Wireless Networks,” in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, pp. 1–5, IEEE, may 2011.

[34] S. Zehl, A. Zubow, and A. Wolisz, “BIGAP - A seamless handover scheme for high performance enterprise IEEE 802.11 networks,” in *Proceedings of the NOMS 2016 - 2016 IEEE/IFIP Network Operations and Management Symposium*, pp. 1015–1016, IEEE, apr 2016.

[35] L. Suresh, J. Schulz-Zander, R. Merz, and A. Feldmann, “Demo: Programming Enterprise WLANs with Odin,” in *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication - SIGCOMM ’12*, (New York, New York, USA), p. 279, ACM Press, 2012.

[36] S. M. M. Gilani, T. Hong, W. Jin, G. Zhao, H. M. Heang, and C. Xu, “Mobility management in IEEE 802.11 WLAN using SDN/NFV technologies,” *Eurasip Journal on Wireless Communications and Networking*, vol. 2017, p. 67, dec 2017.

[37] M. Ojo, D. Adami, and S. Giordano, “A SDN-IoT Architecture with NFV Implementation,” in *2016 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, IEEE, dec 2016.

[38] L. Galluccio, S. Milardo, G. Morabito, and S. Palazzo, “SDN-WISE: Design, prototyping and experimentation of a stateful SDN solution for WIreless SEnsor networks,” in *Proceedings - IEEE INFOCOM*, vol. 26, pp. 513–521, IEEE, apr 2015.

[39] N. Bizanis and F. A. Kuipers, “SDN and Virtualization Solutions for the Internet of Things: A Survey,” *IEEE Access*, vol. 4, no. 99, pp. 5591–5606, 2016.

[40] T. Bakhshi, “State of the art and recent research advances in software defined networking,” *Wireless Communications and Mobile Computing*, vol. 2017, pp. 1–35, jan 2017.

[41] A. Abdelaziz, A. Tan Fong, A. Gani, S. Khan, F. Alotaibi, and M. Khurram Khan, "On Software-Defined Wireless Network (SDWN) Network Virtualization: Challenges and Open Issues," *Computer Journal*, vol. 60, no. 10, pp. 1510–1519, 2017.

[42] A. Martínez, "An ontology-based approach toward the configuration of heterogeneous network devices," *TDX (Tesis Doctorals en Xarxa)*, no. April, 2015.

[43] M. Ganzha, M. Paprzycki, W. Pawłowski, P. Szmeja, and K. Wasielewska, "Towards common vocabulary for IoT ecosystems—preliminary considerations," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10191 LNAI, pp. 35–45, 2017.

[44] G. Bajaj, R. Agarwal, P. Singh, N. Georgantas, and V. Issarny, "A study of existing Ontologies in the IoT-domain," jul 2017.

[45] N. Kolbe, A. Zaslavsky, S. Kubler, J. Robert, and Y. Le Traon, "Enriching a situation awareness framework for IoT with knowledge base and reasoning components," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10257 LNAI, pp. 41–54, 2017.

[46] A. Hakiri, P. Berthou, A. Gokhale, and S. Abdellatif, "Publish/Subscribe enabled software defined networking for efficient and scalable IoT comunications," *IEEE Communications Magazine*, no. September, pp. 48–54, 2015.

[47] S.-M. Chun, H.-S. Kim, and J.-T. Park, "CoAP-Based Mobility Management for the Internet of Things," *Sensors*, vol. 15, pp. 16060–16082, jul 2015.

[48] F. Van den Abeele, J. Hoebeke, G. K. Teklemariam, I. Moerman, and P. Demeester, "Sensor Function Virtualization to Support Distributed Intelligence in the Internet of Things," *Wireless Personal Communications*, vol. 81, no. 4, pp. 1415–1436, 2015.

[49] J. Kim, F. Filali, and Y. B. Ko, "A lightweight CoAP-based software defined networking for resource constrained AMI devices," *2015 IEEE International Conference on Smart Grid Communications, SmartGridComm 2015*, pp. 719–724, 2016.

[50] F. Y. Vivas, O. M. Caicedo, and J. C. Nieves, "A Semantic and Knowledge-Based Approach for Handover Management," *Sensors*, vol. 21, p. 4234, jun 2021.

[51] D. A. Embus, A. J. Castillo, F. Y. Vivas, O. M. Caicedo, and A. Ordóñez, "NetSel-RF: A Model for Network Selection Based on Multi-Criteria and Supervised Learning," *Applied Sciences*, vol. 10, p. 4382, jun 2020.

[52] F. Samboni, S. Bedoya, F. Y. Vivas, and O. M. Caicedo, "MEC IoT: Monitorización de estructuras civiles en el contexto IoT," in *2017 IEEE Colombian Conference on Communications and Computing (COLCOM)*, pp. 1–6, IEEE, aug 2017.

[53] S. Crawford and L. Stucki, "Peer review and the changing research record," *J. Am. Soc. Inf. Sci.*, vol. 41, pp. 223–228, Mar. 1990.

[54] S. K. Ghosh and S. C. Ghosh, "A predictive handoff mechanism for 5G ultra dense networks," in *2017 IEEE 16th International Symposium on Network Computing and Applications (NCA)*, vol. 2017-Janua, pp. 1–5, IEEE, oct 2017.

[55] D. Calabuig, S. Barmpounakis, S. Gimenez, A. Kousaridas, T. R. Lakshmana, J. Lorca, P. Lunden, Z. Ren, P. Sroka, E. Ternon, V. Venkatasubramanian, and M. Maternia, "Resource and Mobility Management in the Network Layer of 5G Cellular Ultra-Dense Networks," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 162–169, 2017.

[56] D. J. Cook and S. K. Das, "How smart are our environments? An updated look at the state of the art," *Pervasive and Mobile Computing*, vol. 3, pp. 53–73, mar 2007.

[57] S. Ferretti, V. Ghini, and F. Panzieri, "A survey on handover management in mobility architectures," *Computer Networks*, vol. 94, pp. 390–413, jan 2016.

[58] D. Guo, L. Tang, X. Zhang, and Y. C. Liang, "Joint Optimization of Handover Control and Power Allocation Based on Multi-Agent Deep Reinforcement Learning," *IEEE Transactions on Vehicular Technology*, vol. 69, pp. 13124–13138, nov 2020.

[59] C.-M. Huang, D.-T. Dao, and M.-S. Chiang, "A Bursty Multi-node Handover scheme for mobile internet using the partially Distributed Mobility Management (BMH–DMM) architecture," *Telecommunication Systems*, vol. 69, pp. 113–130, sep 2018.

[60] S. M. A. Zaidi, M. Manalastas, H. Farooq, and A. Imran, "Mobility Management in Emerging Ultra-Dense Cellular Networks: A Survey, Outlook, and Future Research Directions," *IEEE Access*, vol. 8, pp. 183505–183533, sep 2020.

[61] A. Yazdinejad, R. M. Parizi, A. Dehghantanha, and K.-K. R. Choo, "Blockchain-enabled Authentication Handover with Efficient Privacy Protection in SDN-based 5G Networks," *IEEE Transactions on Network Science and Engineering*, pp. 1–1, may 2020.

[62] D. Charilas and A. Panagopoulous, “Multiaccess Radio Network Enviroments,” *IEEE Vehicular Technology Magazine*, vol. 5, pp. 40–49, dec 2010.

[63] I. Elgendi, K. S. Munasinghe, and A. Jamalipour, “Mobility management in three-Tier SDN architecture for DenseNets,” *IEEE Wireless Communications and Networking Conference, WCNC*, vol. 2016-September, no. Wcnc, pp. 0–5, 2016.

[64] M. S. Mollel, A. I. Abubakar, M. Ozturk, S. F. Kaijage, M. Kisangiri, S. Hussain, M. A. Imran, and Q. H. Abbasi, “A Survey of Machine Learning Applications to Handover Management in 5G and Beyond,” *IEEE Access*, vol. 9, pp. 45770–45802, 2021.

[65] S. Guoqiang, C. Yanming, Z. Chao, and Z. Yanxu, “Design and implementation of a smart IoT gateway,” in *Proceedings - 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, GreenCom-iThings-CPSCom 2013*, pp. 720–723, IEEE, aug 2013.

[66] H. Lin, K. Cai, H. Chen, and Z. Zeng, “The Construction of a Precise Agricultural Information System Based on Internet of Things,” *International Journal of Online Engineering (iJOE)*, vol. 11, p. 10, nov 2015.

[67] M. Aazam, M. St-Hilaire, C.-H. Lung, I. Lambadaris, and E.-N. Huh, “IoT Resource Estimation Challenges and Modeling in Fog,” in *Fog Computing in the Internet of Things*, pp. 17–31, Cham: Springer International Publishing, 2018.

[68] L. F. Bittencourt, J. Diaz-Montes, R. Buyya, O. F. Rana, and M. Parashar, “Mobility-Aware Application Scheduling in Fog Computing,” *IEEE Cloud Computing*, vol. 4, no. 2, pp. 26–35, 2017.

[69] S. Tyagi, A. Darwish, and M. Y. Khan, “Managing Computing Infrastructure for IoT Data,” *Advances in Internet of Things*, vol. 4, no. July, pp. 29–35, 2014.

[70] M. Subramanian., *Network Management - Principles and Practice*. Addison Wesley, 2000.

[71] F. Estrada-Solano, A. Ordonez, L. Zambenedetti Granville, and O. Mauricio Caicedo Rendon, “A Framework for SDN Integrated Management based on a CIM Model and a Vertical Management Plane A Framework for SDN Integrated Management based on a CIM Model and a Vertical Management Plane,” *Computer Communications*, vol. 102, pp. 150–164, apr 2016.

[72] A. Stamou, N. Dimitriou, K. Kontovasilis, and S. Papavassiliou, “Autonomic Handover Management for Heterogeneous Networks in a Future Internet Context: A

Survey," *IEEE Communications Surveys and Tutorials*, vol. 21, pp. 3274–3297, oct 2019.

[73] N. F. S. D. Sousa and C. E. Rothenberg, "CLARA: Closed Loop-based Zero-Touch Network Management Framework," *2021 IEEE Conference on Network Function Virtualization and Software Defined Networks, NFV-SDN 2021 - Proceedings*, pp. 110–115, 2021.

[74] S. R. White, J. E. Hanson, I. Whalley, D. M. Chess, and J. O. Kephart, "An architectural approach to autonomic computing," in *International Conference on Autonomic Computing, 2004. Proceedings.*, pp. 2–9, May 2004.

[75] Q. Zhou, C. X. Wang, S. McLaughlin, and X. Zhou, "Network virtualization and resource description in software-defined wireless networks," *IEEE Communications Magazine*, vol. 53, pp. 110–117, nov 2015.

[76] H. Kim and N. Feamster, "Improving network management with software defined networking," *IEEE Communications Magazine*, vol. 51, pp. 114–119, feb 2013.

[77] Y. Wang, J. Bi, and K. Zhang, "Design and Implementation of a Software-Defined Mobility Architecture for IP Networks," *Mobile Networks and Applications*, vol. 20, pp. 40–52, feb 2015.

[78] R. Mijumbi, J. Serrat, J. L. Gorricho, N. Bouten, F. De Turck, and R. Boutaba, "Network function virtualization: State-of-the-art and research challenges," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 1, pp. 236–262, 2016.

[79] B. Han, V. Gopalakrishnan, L. Ji, and S. Lee, "Network function virtualization: Challenges and opportunities for innovations," *IEEE Communications Magazine*, vol. 53, no. 2, pp. 90–97, 2015.

[80] Q. Zhou, C. X. Wang, S. McLaughlin, and X. Zhou, "Network virtualization and resource description in software-defined wireless networks," *IEEE Communications Magazine*, vol. 53, pp. 110–117, nov 2015.

[81] N. Omnes, M. Bouillon, G. Fromentoux, and O. Le Grand, "A programmable and virtualized network & IT infrastructure for the internet of things: How can NFV & SDN help for facing the upcoming challenges," in *2015 18th International Conference on Intelligence in Next Generation Networks, ICIN 2015*, pp. 64–69, IEEE, 2015.

[82] M. Tayyab, X. Gelabert, and R. Jantti, "A Survey on Handover Management: From LTE to NR," *IEEE Access*, vol. 7, pp. 118907–118930, 2019.

[83] P. Zhao, X. Yang, W. Yu, J. Lin, and D. Meng, "Context-Aware Multi-Criteria Handover with Fuzzy Inference in Software Defined 5G HetNets," in *2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, 2018.

[84] G. Gaur, T. Velmurugan, P. Prakasam, and S. Nandakumar, "Application specific thresholding scheme for handover reduction in 5G Ultra Dense Networks," *Telecommunication Systems*, vol. 76, pp. 97–113, jan 2021.

[85] R. D. Hegazy, O. A. Nasr, and H. A. Kamal, "Optimization of user behavior based handover using fuzzy Q-learning for LTE networks," *Wireless Networks*, vol. 24, pp. 481–495, feb 2018.

[86] S. Barmpounakis, A. Kaloxylos, P. Spapis, P. Magdalinos, N. Alonistioti, and C. Zhou, "Data Analytics for 5G Networks: A Complete Framework for Network Access Selection and Traffic Steering," *International Journal on Advances in Telecommunications*, vol. 11, no. 3 & 4, pp. 101–114, 2018.

[87] A. Habbal, S. I. Goudar, and S. Hassan, "A Context-aware Radio Access Technology selection mechanism in 5G mobile network for smart city applications," *Journal of Network and Computer Applications*, vol. 135, pp. 97–107, jun 2019.

[88] A. Monteiro, E. Souto, R. Pazzi, and M. Nogueira, "Context-aware network selection in heterogeneous wireless networks," *Computer Communications*, vol. 135, pp. 1–15, feb 2019.

[89] A. Stamou, N. Dimitriou, K. Kontovasilis, and S. Papavassiliou, "Context-aware handover management for HetNets: Performance evaluation models and comparative assessment of alternative context acquisition strategies," *Computer Networks*, vol. 176, p. 107272, may 2020.

[90] I. Modeas, A. Kaloxylos, L. Merakos, and D. Tsolkas, "An adaptive and distributed network selection mechanism for 5G networks," *Computer Networks*, vol. 189, p. 107943, apr 2021.

[91] M. Sadik, N. Akkari, and G. Aldabbagh, "SDN-based handover scheme for multitier LTE/Femto and D2D networks," *Computer Networks*, vol. 142, pp. 142–153, sep 2018.

[92] A. Rafiq, M. Afaq, and W.-C. Song, "Mobility Performance Enhancement in Small Cells Cluster of 5G Network: A Handover Overhead Reduction Approach," in *2020 21st Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 373–376, IEEE, sep 2020.

[93] M. A. Adedoyin and O. E. Falowo, "Combination of ultra-dense networks and other 5G enabling technologies: A survey," 2020.

[94] I. Shayea, M. Ergen, M. Hadri Azmi, S. Aldirmaz Colak, R. Nordin, and Y. I. Daradkeh, "Key Challenges, Drivers and Solutions for Mobility Management in 5G Networks: A Survey," *IEEE Access*, vol. 8, pp. 172534–172552, 2020.

[95] A. Stamou, N. Dimitriou, K. Kontovasilis, and S. Papavassiliou, "Autonomic Handover Management for Heterogeneous Networks in a Future Internet Context: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3274–3297, 2019.

[96] B. Priya and J. Malhotra, "5GAuNetS: an autonomous 5G network selection framework for Industry 4.0," *Soft Computing*, vol. 24, pp. 9507–9523, jul 2020.

[97] Y. Zhong, H. Wang, and H. Lv, "A cognitive wireless networks access selection algorithm based on MADM," *Ad Hoc Networks*, vol. 109, p. 102286, dec 2020.

[98] O. A. Aldhaibani, M. H. AL-Jumaili, A. Raschella, H. Kolivand, and A. P. Preethi, "A centralized architecture for autonomic quality of experience oriented handover in dense networks," *Computers and Electrical Engineering*, vol. 94, no. October 2020, 2021.

[99] Q. Liu, C. F. Kwong, S. Zhou, T. Ye, L. Li, and S. P. Ardakani, "Autonomous Mobility Management for 5G Ultra-Dense HetNets via Reinforcement Learning with Tile Coding Function Approximation," *IEEE Access*, vol. 9, pp. 97942–97952, 2021.

[100] A. Jain and S. Tokekar, "Application based vertical handoff decision in heterogeneous network," *Procedia Computer Science*, vol. 57, pp. 782–788, 2015.

[101] Y. K. Salih, O. H. See, R. W. Ibrahim, S. Yussof, and A. Iqbal, "An overview of intelligent selection and prediction method in heterogeneous wireless networks," *Journal of Central South University*, vol. 21, no. 8, pp. 3138–3154, 2014.

[102] Z. Han, T. Lei, Z. Lu, X. Wen, W. Zheng, and L. Guo, "Artificial intelligence-based handoff management for dense wlans: A deep reinforcement learning approach," *IEEE Access*, vol. 7, pp. 31688–31701, 2019.

[103] M. Drissi, M. Oumsis, and D. Aboutajdine, "A multi-criteria decision framework for network selection over lte and wlan," *Engineering Applications of Artificial Intelligence*, vol. 66, pp. 113–127, 2017.

[104] M. Lahby, "Enhancing modeling for network selection using graph theory in beyond 4g networks," *International Journal of Business Data Communications and Networking (IJBDCN)*, vol. 16, no. 1, pp. 48–69, 2020.

[105] E. Malathy and V. Muthuswamy, "State of art: vertical handover decision schemes in next-generation wireless network," *Journal of Communications and Information Networks*, vol. 3, no. 1, pp. 43–52, 2018.

[106] C. Chi, X. Cai, R. Hao, and F. Liu, "Modeling and analysis of handover algorithms," in *IEEE GLOBECOM 2007-IEEE Global Telecommunications Conference*, pp. 4473–4477, IEEE, 2007.

[107] K. Nahida, C. Yin, Y. Hu, Z. A. Arain, C. Pan, I. Khan, Y. Zhang, and G. S. Rahman, "Handover based on ap load in software defined wi-fi systems," *Journal of Communications and Networks*, vol. 19, no. 6, pp. 596–604, 2017.

[108] G. A. F. M. Khalaf and H. Z. Badr, "A comprehensive approach to vertical handoff in heterogeneous wireless networks," *Journal of King Saud University-Computer and Information Sciences*, vol. 25, no. 2, pp. 197–205, 2013.

[109] M. Khan, S. Din, M. Gohar, A. Ahmad, S. Cuomo, F. Piccialli, and G. Jeon, "Enabling multimedia aware vertical handover management in internet of things based heterogeneous wireless networks," *Multimedia Tools and Applications*, vol. 76, no. 24, pp. 25919–25941, 2017.

[110] S. Din, A. Ahmad, A. Paul, and S. Rho, "Mgr: Multi-parameter green reliable communication for internet of things in 5g network," *Journal of Parallel and Distributed Computing*, vol. 118, pp. 34–45, 2018.

[111] N. M. Alotaibi and S. S. Alwakeel, "A neural network based handover management strategy for heterogeneous networks," in *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, pp. 1210–1214, IEEE, 2015.

[112] H.-W. Yu and B. Zhang, "A heterogeneous network selection algorithm based on network attribute and user preference," *Ad Hoc Networks*, vol. 72, pp. 68–80, 2018.

[113] P. Goyal, D. Lobiyal, and C. Katti, "Dynamic user preference based network selection for vertical handoff in heterogeneous wireless networks," *Wireless Personal Communications*, vol. 98, no. 1, pp. 725–742, 2018.

[114] Z. Ali, N. Baldo, J. Mangues-Bafalluy, and L. Giupponi, "Machine learning based handover management for improved qoe in lte," in *NOMS 2016-2016 IEEE/IFIP Network Operations and Management Symposium*, pp. 794–798, IEEE, 2016.

[115] M. Khan, A. Ahmad, S. Khalid, S. H. Ahmed, S. Jabbar, and J. Ahmad, "Fuzzy based multi-criteria vertical handover decision modeling in heterogeneous wireless networks," *Multimedia Tools and Applications*, vol. 76, no. 23, pp. 24649–24674, 2017.

[116] R. Goyal, T. Goyal, S. Kaushal, and H. Kumar, "Fuzzy ahp based technique for handover optimization in heterogeneous network," in *Proceedings of 2nd International Conference on Communication, Computing and Networking*, pp. 293–301, Springer, 2019.

[117] M. Khan and K. Han, "A vertical handover management scheme based on decision modelling in heterogeneous wireless networks," *IETE Technical Review*, vol. 32, no. 6, pp. 402–412, 2015.

[118] M. Sadik, N. Akkari, and G. Aldabbagh, "Sdn-based handover scheme for multi-tier lte/femto and d2d networks," *Computer Networks*, vol. 142, pp. 142–153, 2018.

[119] S. Goudarzi, W. H. Hassan, M. H. Anisi, A. Soleymani, M. Sookhak, M. K. Khan, A.-H. A. Hashim, and M. Zareei, "Abc-pso for vertical handover in heterogeneous wireless networks," *Neurocomputing*, vol. 256, pp. 63–81, 2017.

[120] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer, R. Wirth, *et al.*, "Crisp-dm 1.0: Step-by-step data mining guide," *SPSS inc*, vol. 16, 2000.

[121] N. Williams, S. Zander, and G. Armitage, "A preliminary performance comparison of five machine learning algorithms for practical ip traffic flow classification," *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 5, pp. 5–16, 2006.

[122] S. Raschka, "An overview of general performance metrics of binary classifier systems," *arXiv preprint arXiv:1410.5330*, 2014.

[123] D. Chicco and G. Jurman, "The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation," *BMC genomics*, vol. 21, no. 1, p. 6, 2020.

[124] H. T. Larasati, F. H. Ilma, B. Nuhamara, A. Mustafa, R. Hakimi, and E. Mulyana, "Performance evaluation of handover association mechanisms in sdn-based wireless network," in *2017 3rd International Conference on Wireless and Telematics (ICWT)*, pp. 103–108, IEEE, 2017.

[125] M. F. Tuysuz and R. Trestian, "Energy-efficient vertical handover parameters, classification and solutions over wireless heterogeneous networks: a comprehensive survey," *Wireless Personal Communications*, vol. 97, no. 1, pp. 1155–1184, 2017.

[126] I. Chamodrakas and D. Martakos, “A utility-based fuzzy topsis method for energy efficient network selection in heterogeneous wireless networks,” *Applied Soft Computing*, vol. 11, no. 4, pp. 3734–3743, 2011.

[127] H. Kim and N. Feamster, “Improving network management with software defined networking,” *IEEE Communications Magazine*, vol. 51, no. 2, pp. 114–119, 2013.

[128] R. R. Fontes, S. Afzal, S. H. Brito, M. A. Santos, and C. E. Rothenberg, “Mininet-wifi: Emulating software-defined wireless networks,” in *2015 11th International Conference on Network and Service Management (CNSM)*, pp. 384–389, IEEE, 2015.

[129] A. Kumar, P. Kaur, and P. Sharma, “A survey on hoeffding tree stream data classification algorithms,” *CPUH-Research Journal*, vol. 1, no. 2, pp. 28–32, 2015.

[130] U. Maulik and D. Chakraborty, “Remote sensing image classification: A survey of support-vector-machine-based advanced techniques,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 1, pp. 33–52, 2017.

[131] S. Hoeglinger and R. Pears, “Use of hoeffding trees in concept based data stream mining,” in *2007 Third International Conference on Information and Automation for Sustainability*, pp. 57–62, IEEE, 2007.

[132] P. A. A. Resende and A. C. Drummond, “A survey of random forest based methods for intrusion detection systems,” *ACM Computing Surveys (CSUR)*, vol. 51, no. 3, pp. 1–36, 2018.

[133] R. Boutaba, M. A. Salahuddin, N. Limam, S. Ayoubi, N. Shahriar, F. Estrada-Solano, and O. M. Caicedo, “A comprehensive survey on machine learning for networking: evolution, applications and research opportunities,” *Journal of Internet Services and Applications*, vol. 9, no. 1, 2018.

[134] A. Bifet and R. Gavaldà, “Adaptive learning from evolving data streams,” in *International Symposium on Intelligent Data Analysis*, pp. 249–260, Springer, 2009.

[135] E. Zeljković, J. M. Marquez-Barja, A. Kassler, R. Riggio, and S. Latré, “Proactive access point driven handovers in ieee 802.11 networks,” in *2018 14th International Conference on Network and Service Management (CNSM)*, pp. 261–267, IEEE, 2018.

[136] S. Costanzo, L. Galluccio, G. Morabito, and S. Palazzo, “Software defined wireless networks: Unbridling sdns,” in *2012 European Workshop on Software Defined Networking*, pp. 1–6, IEEE, 2012.

[137] L. Tartarini, M. A. Marotta, E. Cerqueira, J. Rochol, C. B. Both, M. Gerla, and P. Bellavista, “Software-defined handover decision engine for heterogeneous cloud radio access networks,” *Computer Communications*, vol. 115, pp. 21 – 34, 2018.

[138] H. Moura, A. R. Alves, J. R. Borges, D. F. Macedo, and M. A. Vieira, “Ethanol: A software-defined wireless networking architecture for ieee 802.11 networks,” *Computer Communications*, vol. 149, pp. 176–188, 2020.

[139] F. Estrada-Solano, A. Ordonez, L. Z. Granville, and O. M. C. Rendon, “A framework for sdn integrated management based on a cim model and a vertical management plane,” *Computer Communications*, vol. 102, pp. 150–164, 2017.

[140] S. Ayoubi, N. Limam, M. A. Salahuddin, N. Shahriar, R. Boutaba, F. Estrada-Solano, and O. M. Caicedo, “Machine learning for cognitive network management,” *IEEE Communications Magazine*, vol. 56, no. 1, pp. 158–165, 2018.

[141] S. Barmpounakis, A. Kaloxylos, P. Spapis, and N. Alonistioti, “Context-aware, user-driven, network-controlled RAT selection for 5G networks,” *Computer Networks*, vol. 113, pp. 124–147, feb 2017.

[142] N. Akkari and N. Dimitriou, “Mobility management solutions for 5G networks: Architecture and services,” *Computer Networks*, vol. 169, p. 107082, mar 2020.

[143] T. Bilen, B. Canberk, and K. R. Chowdhury, “Handover Management in Software-Defined Ultra-Dense 5G Networks,” *IEEE Network*, vol. 31, pp. 49–55, jul 2017.

[144] A. Jain and S. Tokekar, “Application Based Vertical Handoff Decision in Heterogeneous Network,” in *Procedia Computer Science*, vol. 57, pp. 782–788, Elsevier, jan 2015.

[145] S. Din, A. Ahmad, A. Paul, and S. Rho, “MGR: Multi-parameter Green Reliable communication for Internet of Things in 5G network,” *Journal of Parallel and Distributed Computing*, vol. 118, pp. 34–45, aug 2018.

[146] E. M. Malathy and V. Muthuswamy, “State of Art: Vertical Handover Decision Schemes in Next-Generation Wireless Network,” *Journal of Communications and Information Networks*, vol. 3, pp. 43–52, mar 2018.

[147] Y. K. Salih, O. H. See, R. W. Ibrahim, S. Yussof, and A. Iqbal, “An overview of intelligent selection and prediction method in heterogeneous wireless networks,” *Journal of Central South University*, vol. 21, pp. 3138–3154, aug 2014.

[148] Z. Han, T. Lei, Z. Lu, X. Wen, W. Zheng, and L. Guo, "Artificial Intelligence-Based Handoff Management for Dense WLANs: A Deep Reinforcement Learning Approach," *IEEE Access*, vol. 7, pp. 31688–31701, 2019.

[149] Distributed Management Task Force, Inc, "The cim standard: Common information model." https://www.dmtf.org/standards/cim, 2021. Accessed: 2021-04-25.

[150] W3C Recommendation, "Owl web ontology language." https://www.w3.org/TR/owl-features/, 2004. Accessed: 2021-04-25.

[151] J. C. Nieves, A. Espinoza, Y. K. Penya, M. O. de Mues, and A. Peña, "Intelligence distribution for data processing in smart grids: A semantic approach," *Engineering Applications of Artificial Intelligence*, vol. 26, pp. 1841–1853, sep 2013.

[152] C. Bettini, O. Brdiczka, K. Henricksen, J. Indulska, D. Nicklas, A. Ranganathan, and D. Riboni, "A survey of context modelling and reasoning techniques," *Pervasive and Mobile Computing*, vol. 6, pp. 161–180, apr 2010.

[153] A. Uzun, M. von Hoffen, and A. Küpper, "Enabling semantically enriched data analytics by leveraging topology-based mobile network context ontologies," in *Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS14)*, WIMS '14, (New York, NY, USA), Association for Computing Machinery, 2014.

[154] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider, *The Description Logic Handbook*. Cambridge: Cambridge University Press, 2007.

[155] K. C. Silva, Z. Becvar, E. H. Cardoso, and C. R. Francěs, "Self-tuning handover algorithm based on fuzzy logic in mobile networks with dense small cells," in *IEEE Wireless Communications and Networking Conference, WCNC*, vol. 2018-April, pp. 1–6, Institute of Electrical and Electronics Engineers Inc., jun 2018.

[156] X. Su, J. Riekki, and J. Haverinen, "Entity Notation: enabling knowledge representations for resource-constrained sensors," *Personal and Ubiquitous Computing*, vol. 16, pp. 819–834, oct 2012.

[157] 3GPP, "5G; NR; NR and NG-RAN Overall description; Stage-2," Technical Specification (TS) 38.300, 3rd Generation Partnership Project (3GPP), 01 2021. Version 16.4.0.

[158] R. Fontes, S. Afzal, S. Brito, M. Santos, and C. Esteve Rothenberg, "Mininet-WiFi: emulating Software-Defined wireless networks," in *Network and Service Management (CNSM), 2015 11th International Conference on*, (Barcelona, Spain), pp. 384–389, Nov 2015.

[159] B. Lantz, B. Heller, and N. McKeown, “A network in a laptop: Rapid prototyping for software-defined networks,” in *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, Hotnets-IX, (New York, NY, USA), Association for Computing Machinery, 2010.

[160] “Sim-know project repository.” https://github.com/fyvivas/SIMKnow, 2021. Accessed: 2021-06-11.

[161] A. Hanemann, A. Liakopoulos, M. Molina, and D. M. Swany, “A study on network performance metrics and their composition,” *Campus-Wide Information Systems*, vol. 23, pp. 268–282, jan 2006.

[162] H.-W. Yu and B. Zhang, “A heterogeneous network selection algorithm based on network attribute and user preference,” *Ad Hoc Networks*, vol. 72, pp. 68–80, apr 2018.

[163] J. Dugan, S. Elliott, B. A. Mah, J. Poskanzer, and K. Prabhu, “iperf3, tool for active measurements of the maximum achievable bandwidth on ip networks.” https://iperf.fr/, 2012. Accessed: 2021-04-25.

[164] S. Avallone, S. Guadagno, D. Emma, A. Pescape, and G. Ventre, “D-ITG distributed Internet traffic generator,” in *First International Conference on the Quantitative Evaluation of Systems, 2004. QEST 2004. Proceedings.*, pp. 316–317, IEEE, 2004.

[165] A. Habbal, S. Goudar, and S. Hassan, “Context-aware Radio Access Technology Selection Approach in 5G Ultra Dense Networks,” *IEEE Access*, vol. 5, pp. 1–1, 2017.

[166] J. Tanveer, A. Haider, R. Ali, and A. Kim, “An Overview of Reinforcement Learning Algorithms for Handover Management in 5G Ultra-Dense Small Cell Networks,” *Applied Sciences (Switzerland)*, vol. 12, no. 1, 2022.

[167] U. Furbach, S. Hölldobler, M. Ragni, C. Schon, and F. Stolzenburg, “Cognitive Reasoning: A Personal View,” *KI - Künstliche Intelligenz*, vol. 33, pp. 209–217, sep 2019.

[168] S. Kapoor, *Learning and Reasoning Strategies for User Association in Ultra-dense Small Cell Vehicular Networks*. PhD thesis, University of York, 2019.

[169] M. Kountouris and N. Pappas, “Semantics-Empowered Communication for Networked Intelligent Systems,” *IEEE Communications Magazine*, vol. 59, pp. 96–102, jun 2021.

[170] Z. Peng, B. Xu, A. M. Gates, D. Cui, and W. Lin, “A Study of a Multi-Agent Organizational Framework with Virtual Machine Clusters as the Unit of Granularity in Cloud Computing,” *The Computer Journal*, vol. 60, pp. 1032–1043, aug 2016.

[171] Y. Li and C. Tan, “A survey of the consensus for multi-agent systems,” *Systems Science & Control Engineering*, vol. 7, pp. 468–482, jan 2019.

[172] A. I. Maarala, X. Su, and J. Riekki, “Semantic Reasoning for Context-Aware Internet of Things Applications,” *IEEE Internet of Things Journal*, vol. 4, no. 2, pp. 461–473, 2017.

[173] F. Shi, Q. Li, T. Zhu, and H. Ning, “A survey of data semantization in internet of things,” *Sensors (Switzerland)*, vol. 18, p. 313, jan 2018.

[174] N. F. Saraiva de Sousa, D. L. Perez, C. E. Rothenberg, and P. H. Gomes, “End-to-end service monitoring for Zero-touch networks,” *Journal of ICT Standardization*, vol. 9, no. 2, pp. 91–112, 2021.

[175] A. Amirkhani and A. H. Barshooi, *Consensus in multi-agent systems: a review*, vol. 55. Springer Netherlands, 2022.

[176] I. Shayea, M. Ergen, A. Azizan, M. Ismail, and Y. I. Daradkeh, “Individualistic Dynamic Handover Parameter Self-Optimization Algorithm for 5G Networks Based on Automatic Weight Function,” *IEEE Access*, vol. 8, pp. 214392–214412, 2020.

[177] A. Mohajer, M. Bavaghar, and H. Farrokhi, “Mobility-aware load Balancing for Reliable Self-Organization Networks: Multi-agent Deep Reinforcement Learning,” *Reliability Engineering and System Safety*, vol. 202, no. May, 2020.

[178] S. A. Khwandah, J. P. Cosmas, P. I. Lazaridis, I. A. Glover, Z. D. Zaharis, and N. R. Prasad, “Energy Efficient Mobility Enhancement in LTE Pico–Macro HetNet Systems,” *Wireless Personal Communications*, vol. 109, pp. 1491–1502, dec 2019.

[179] R. Dos Reis Fontes, C. Campolo, C. Esteve Rothenberg, and A. Molinaro, “From Theory to Experimental Evaluation: Resource Management in Software-Defined Vehicular Networks,” *IEEE Access*, vol. 5, pp. 3069–3076, 2017.

Appendix A

Scientific Production

The research work presented in this thesis was reported to the scientific community through paper submissions to renowned conferences and journals. The process of doing research, submitting paper, gathering feedback, and improving the work helped to achieve the maturity hereby presented. The published paper to date is presented below.

- "A Semantic and Knowledge-Based Approach for Handover Management," published in Sensor MDPI, 2021.
- "NetSel-RF: A Model for Network Selection Based on Multi-Criteria and Supervised Learning," published in Applied Sciences, 2020.
- "ZTHM-5G: Zero-Touch Handover Management in 5G," Submitted.
- "MEC IoT: Monitorización de estructuras civiles en el contexto IoT," published in the proceedings of the 2017 IEEE Colombian Conference on Communications and Computing (COLCOM), 2017.

The published paper is available in the next pages.