FACULTY OF ENGINEERING
AND ARCHITECTURE

Techno-Economic Analysis of 5G Networks: Enabler Technologies and Deployment Strategies

Asma Chiha Ep Harbi

Doctoral dissertation submitted to obtain the academic degree of
Doctor of Engineering

Supervisors
Prof. Sofie Verbrugge, PhD - Prof. Didier Colle, PhD
Department of Information Technology
Faculty of Engineering and Architecture, Ghent University

August 2023

GHENT
UNIVERSITY

ISBN 978-94-6355-742-9
NUR 986
Wettelijk depot: D/2023/10.500/74

Members of the Examination Board

Chair

Prof. Filip De Turck, PhD, Ghent University

Other members entitled to vote

Prof. Sidharta Gautama, PhD, Ghent University
Prof. Jeroen Hoebeke, PhD, Ghent University
Bart Lannoo, PhD, Be Mobile
Prof. Jason Whalley, PhD, Northumbria University, United Kingdom

Supervisors

Prof. Sofie Verbrugge, PhD, Ghent University
Prof. Didier Colle, PhD, Ghent University

Acknowledgements

In his book Almuqadimah (the introduction to social science), Ibn Khaldun, the famous sociologist, states that “**Man is social by nature**”, which implies that human beings possess an inherent inclination to live among others, establish meaningful social relationships and engage in dialogues that contribute to their personal and collective growth.

Therefore, I humbly acknowledge that my achievements thus far would not have been possible without the support and encouragement of the TE team, my dear friends, my loving family, and my devoted husband and children. And, as the prophet Mohammed peace be upon him said: “He who does not thank people does not thank ALLAH (God)” [Al Adab al-Mufrad, 218]: **"من لا يشكر الناس لا يشكر الله"**

I want to thank all those without whom this dissertation would not have been possible:

To my beloved parents, **Hamadi** and **Samia**, who have always been my unwavering pillars of support throughout my academic journey. Your constant love, encouragement, and sacrifices have been the driving force behind my success. This book is a testament to your unwavering faith in me and the countless hours of hard work you have put in to ensure my academic and personal growth. I dedicate this book to you with all my heart and gratitude.

الى أمي و أبي, أهديكما هذا العمل و أقدر كل جهد بذلتماه من أجلي, أحبكما كثيرا!

To my esteemed supervisors **Sofie** and **Didier**, and mentors **Marlies** and **Frederic**, who have guided me with their unparalleled expertise, wisdom, and patience. Your valuable insights, constructive criticism, unwavering support, and encouragement have played a pivotal role in molding my research and academic endeavours. My heartfelt gratitude is beyond measure . I dedicate this book to you as a token of my admiration and gratitude.

To my loving siblings, **Souhail**, **Ines**, **Emna**, **Aicha**, **Ilyes**, and **Mohammed**, who have been my constant source of inspiration, love, and support. Your unwavering faith in my abilities and your unwavering support have been the cornerstone of my academic success. This book is dedicated to you with all my love and appreciation.

I extend my deepest appreciation to my mother-in-law, **Zeinab**. Your continuous belief in my abilities, along with your words of wisdom and encouragement, have been a source of inspiration that I am truly grateful for. I would also like to remember and honor the memory of my late father-in-law, **Moncef.** Though he is no longer with us, his spirit of resilience and dedication to education continues to inspire me. His unwavering belief in the pursuit of knowledge has left an indelible mark on my academic journey, and I am committed to carrying his legacy forward.

To my dear friends, **Hajer**, **Malek**, **Najeh**, **Sana**, and all the others, who have been my pillars of strength, confidants, and cheerleaders. Your support, encouragement, and unwavering belief in me have made this academic journey a memorable one. I am deeply grateful for your friendship, and I dedicate this book to you as a token of my gratitude.

شكرا على كل شيئ, شكرا على الضحكات و القهقهات, شكرا على دعمكن غير المحدود, شكرا على الطرود العابرة للقارات, و شكرا على عبارة "أنت لها", أحبكن كثيرا!

To my wonderful children, **Alaa**, **Mariem**, and **Omar**, who have been my constant source of joy, inspiration, and motivation. Your boundless love, patience, and understanding have been my anchor during these challenging times. This book is dedicated to you with all my heart and love.

*Aan mijn geweldige kinderen, **Alaa, Mariem** en **Omar**, die mijn constante bron van vreugde, inspiratie en motivatie zijn geweest. Jullie grenzeloze liefde, geduld en begrip zijn mijn anker geweest tijdens moeilijke tijden. Dit boek is aan jullie opgedragen met heel mijn hart en liefde.*

أحبكم كثيرا!

Finally, to my loving husband, **Oussema**, who has been my rock, partner, and soulmate. Your unwavering love, support, and encouragement have been the driving force behind my success. This book is dedicated to you with all my love and gratitude for being my constant companion in this journey called life.

شكرا على كل شيئ!

Table of Contents

List of Figures

List of Tables

List of Acronyms

1G	First Generation of cellular networks
2G	Second Generation of cellular networks
3G	Third Generation of cellular networks
3GPP	3rd Generation Partnership Project
4G	Fourth Generation of cellular networks
5G	Fifth Generation of cellular networks
5GC	5G Core Network
6G	Sixth Generation of cellular networks

A

ACPU	Average Cost Per User
ADSL	Asymmetric Digital Subscriber Line
AF	Application Function
AI	Artificial Intelligence
AMF	Access and Mobility Management Function
AMPS	Advanced Mobile Phone System
AN	Access Network
AS	Application Server
APN	Access Point Name
AR	Augmented Reality
ARPU	Average Revenue Per User
AUSF	Authentication Server Function

B

BBUs	Baseband Units
BEREC	Body of European Regulators for Electronic Communications
BIPT	Belgian Institute for Postal Services and Telecommunication
BOM	Bill Of Materials
BSC	Base Station Controller
BTS	Base Transceiver Station

C

CAPEX	Capital expenditure

CCAM	Cooperative, Connected, and Automated Mobility
CCK	Complementary Coded Keying
CDMA	Code Division Multiple Access
CDMA2000	Code Division Multiple Access 2000
CLM	Cooperative Lane Merging
CN	Core Network
CO	Central Office
CP	Control Plane
CS	Circuit Switching
C-RAN	Cloud-RAN
CU	Centralized Unit

D

DAE	Digital Agenda for Europe
DESI	Digital Economy and Society Index
DOCSIS	Data Over Cable Service Interface Specification
DSL	Digital Subscriber Line
DSLAM	Digital Subscriber Line Access Multiplexer
DSM	Digital Single Market
DU	Distributed Unit

E

EC	European Commission
EDGE	Enhanced Data rates for GSM Evolution
EEA	European Economic Area
eMBB	enhanced Mobile Broadband
eNB/eNodeB	Evolved Node B
EPC	Evolved Packet Core
ETSI	European Telecommunications Standards Institute
EU	European Union
E-UTRAN	Evolved UMTS (Universal Mobile Telecommunications System) Terrestrial Radio Access Network
E2E	End-to-End

F

FCC	Federal Communications Commission
FDMA	Frequency Division Multiple Access
FDD	Frequency Division Duplex
FTTB	Fiber-to-the-Building
FTTC	Fiber-to-the-Cabinet
FTTH	Fiber-to-the-Home

FTTdp	Fiber to the distribution point
FWA	Fixed Wireless Access

G

GB	Gigabyte
GEO	Geostationary Earth Orbit
g.fast	Giga-bit fast access subscriber technology
GHz	Gigahertz
GGSN	Gateway GPRS Support Node
gNB	Next Generation NodeB
GPRS	General Packet Radio Service
GS	Group Specification
GSM	Global System for Mobile Communications
GSMA	Global System for Mobile Communications Association,

H

HetNet	Heterogeneous Network
HLR/VLR	Home/Visitor Location Register
HSPA/+	High-Speed Packet Access/Plus
HSS	Home Subscriber Server

I

ICT	Information and Communication Technology
IEEE	Institute of Electrical and Electronics Engineers
IMT	International Mobile Telecommunications
IoT	Internet of Things
IS-54	Interim Standard 54
IS-95	Interim Standard-95
ISG	Industry Specification Group
ISP	Internet Service Provider
IP	Internet Protocol
ITU	International Telecommunication Union

K

Kbps	Kilobits per second
KPIs	Key Performance Indicators

L

LAN	Local Area Network
LEO	Low Earth Orbit
LTE/-A	Long-Term Evolution/-Advanced

M

MB	Megabyte
Mbps	Megabits Per Second
MEC	Multi-access Edge Computing
MEO	Medium Earth Orbit
MHz	Megahertz
MIMO	Multiple-Input Multiple-Output
MNOs	Mobile Network Operators
MME	Mobility Management Entity
mMIMO	Massive Multiple-Input Multiple-Output
mMTC	Massive Machine Type Communications
mm-wave	Millimeter Wave
ms	Milliseconds
MVNO	Mobile Virtual Network Operator
MTSO	Mobile Telephone Switching Office
MSC	Mobile Switching Center

N

NAS	Non-Access Stratum
NEF	Network Exposure Function
NF	Network Function
NFV	Network Function Virtualization
NIC	Network Interface Card
NGN	Next Generation Networks
NG-RAN	Next Generation Radio Access Network
NMT	Nordic Mobile Telephone
NO	Network Operator
NR	New Radio
NRF	NF Repository Function
NSA	Non-Stand Alone
NS	Network Slicing
NSSF	Network Slice Selection Function
NSSAI	Network Slice Selection Assistance Information

O

OEM	Original Equipment Manufacturer

OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
ONU	Optical Network Unit
OPEX	Operational expenditure
OSS	Operation and Support Subsystem
OTT	Over The Top

P

PCEF	Policy Control Enforcement Function
PCF	Policy Control Function
PCRF	Policy Control and Charging Rules Function
PDN	Packet Data Network
P-GW	Packet Data Network Gateway
PN	Pseudo-Noise
PON	Passive Optical Network
PPP	Public Private Partnership
PS	Packet Switching
PSTN	Public Switched Telephone Network

Q

QoS	Quality of Service
QoE	Quality of Experience

R

RAN	Radio Access Network
RNC	Radio Network Controller
RO	Road Operator
ROI	Return on Investment
RRH	Remote Radio Head
RSU	RoadSide Unit

S

SA	Stand Alone
SDN	Software Defined Networking
SIM	Subscriber Identity Module
SGSN	Serving GPRS Support Node
S-GW	Serving Gateway
SLA	Service-level Agreement
SMF	Session Management Function

SMS	Short Message Service
SP	Service Provider

T

TACS	Total Access Communication System
TCO	Total Cost of Ownership
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
Tbps	Terabits per second

U

UDM	Unified Data Management
UDN	Ultra-Dense Networks
UE	User Equipment
UMTS	Universal Mobile Telephone System
UP	User Plane
UPF	User Plane Function
URLLC	Ultra-Reliable and Low Latency Communications
UTRAN	UMTS (Universal Mobile Telecommunications System) Terrestrial Radio Access Network

V

V2I	Vehicle to Infrastructure
V2N	Vehicle to Network
V2X	Vehicle to Everything
VDSL/VDSL2	Very-high-bit-rate Digital Subscriber Line
VLAN	Virtual Local Area Network
VM	Virtual Machine
VOD	Video On Demand
VoIP	Voice over IP
VR	Virtual Reality

W

WAN	Wide Area Network
WAP	Wireless Access Point
WCDMA	Wideband Code Division Multiple Access
Wi-Fi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access

WMANs	Wireless Metropolitan Area Networks

Samenvatting
– Summary in Dutch –

Dankzij veelbelovende technologieën zoals netwerkvirtualisatie, netwerkslicing en edge computing, kunnen 5G netwerken potentieel altijd en overal snelle en betrouwbare mobiele breedbandconnectiviteit bieden. Om echter het succes van de uitrol van 5G-netwerken te garanderen, moeten we niet alleen de technische uitdagingen die met deze technologieën gepaard aanpakken, maar ook hun economische implicaties.

In dit thesisboek hebben we een uitgebreide techno-economische analyse uitgevoerd van het uitrollen van 5G-netwerken, met als doel de haalbaarheid ervan te beoordelen en strategische aanbevelingen te doen aan mobiele netwerkoperators (MNO's). Om dit overkoepelende onderzoeksdoel te behalen, zijn vijf zaken onderzocht: de haalbaarheid van 5G-uitrol voor verschillende soorten regio's en diensten, de besparingen (kwantitatief) die verschillende 5G-technologieën met zich meebrengen, nieuwe verdeelsleutels voor kosten, potentiële business- en samenwerkingsmodellen, en de optimale strategie voor de uitrol van 5G betreffende technologische keuzes en timingbeslissingen.

Om deze vijf onderzoeksdoelen aan te pakken, voerden we een techno-economische analyse uit van verschillende case studies. We hebben verschillende techno-economische modellen ontwikkeld die in dit onderzoek worden gebruikt, waaronder scenario-analyse van relevante scenarios, kosten-batenanalyse, sensitiviteitsanalyese, analyse van business modellen en verdeelsleutels voor kosten.

Wat betreft de haalbaarheid van 5G-netwerken voor verschillende soorten regio's werden twee soorten diensten geanalyseerd in drie verschillende regio's: enhanced Mobile Broadband (eMBB) en Ultra-Reliable and Low Latency Communication (URLLC) in landelijk gebied, voor het luchtruim en op de snelweg. Voor het eMBB-scenario hebben we de rendabiliteit geëvalueerd van geïntegreerde satelliet-4G- en 5G-netwerken voor connectiviteit in, respectievelijk, landelijke gebieden en vliegtuigen. De onderzochte oplossing bleek rendabel te zijn voor de noden van eindgebruikers (2 tot 5 Mbps), maar duur voor snelle breedband. Voor het URLLC-scenario hebben we CCAM-diensten op de snelweg in grensgebieden bestudeerd. Deze analyse biedt inzicht in de Total Cost of Ownership (TCO) die gepaard gaan met de verschillende scenarios. We stelden verschillende prijszettingen voor; de op kosten gebaseerde prijszetting bleek de meest effectieve aanpak.

Wat betreft de potentiële besparingen van 5G-technologieën hebben we voornamelijk gefocust op twee technologieën, netwerkvirtualisatie en edge

computing. Ons onderzoek toonde aan hoe het cachen van populaire inhoud op Mobiele Edge Computing (MEC) zowel OPEX-kosten, als de Average Cost Per User (ACPU) kan verlagen. Er moet hier echter wel een afweging gemaakt worden tussen de mate van personalisatie en cachingsnelheid. Daarnaast werd ook de impact van Network Function virtualization (NFV) op implementatiekosten bekenen, waarbij we een kostenbesparing van 12 tot 45% van de kosten van het kernnetwerk aantoonden als gevolg van het gebruik van NFV en netwerkslicing.

We wilden nieuwe verdeelsleutels ontwikkelen om kosten te delen tussen aanbieders van diensten en stakeholders, die beter geschikt zijn voor de unieke kenmerken van 5G-technologieën. Daarvoor hebben we twee kostenmodellen ontwikkeld: één voor de toewijzing van kosten aan dienstverleners in gesegmenteerde netwerken en de tweede om kosten te delen tussen stakeholders. We hebben een nieuwe verdeelsleutel voorgesteld voor gesegmenteerde 5G-netwerken.De resultaten van het model toonden aan dat eMBB-video een aanzienlijk deel van de netwerkkosten draagt in vergelijking met spraak, ten gevolge van de hogere bandbreedtevereisten. Ook hebben we een verdeelsleutel ontwikkeld om kosten te verdelen onder stakeholders die 5G-gebaseerde mobiliteitsdiensten aanbieden. We hebben belangrijke kostfactoren geïdentificeerd, waaronder het type connectiviteit en de mate van samenwerking tussen actoren, en hebben aangetoond dat het implementeren van een collaboratief businessmodel, waarbij infrastructuur gedeeld wordt, de investeringskosten voor betrokken operators aanzienlijk kan verminderen.

We hebben onderzocht welke potentiële business- en samenwerkingsmodellen een succesvolle implementatie van 5G mogelijk maken, om 5G-connectiviteit aan te bieden in vliegtuigen, en vervolgens in wagens. Voor de 5G connectiviteit in vliegtuigen, hebben we verschillende businessmodellen geïdentificeerd, afhankelijk van wie de service aan de eindgebruiker zal leveren, namelijk wholesale versus retail modellen, waarbij de keuze gebaseerd is op het niveau van controle en risico dat de luchtvaartmaatschappij bereid is te nemen. Het retailmodel is voordeliger voor klanten omdat het een kortere waardeketen biedt en een betere kwaliteit garandeert. Voor low-cost maatschappijen wordt het sponsormodel voorgesteld. De freemium wholesale en retail modellen maken branding en gratis diensten mogelijk voor zowel MNO's/luchtvaartmaatschappijen als passagiers, maar vereisen een voldoende hoog conversiepercentage van gratis naar premium om rendabel te zijn. Het complementaire servicemodel is beter geschikt voor luchtvaartmaatschappijen die inzetten op klanten die betalen voor de premium service, inclusief connectiviteit. Verder hebben we verschillende prijsmodellen voorgesteld voor elk businessmodel om maximale winstgevendheid te garanderen. Voor de 5G connectiviteit in wagens, hebben we drie businessmodellen voorgesteld die afhankelijk zijn van opkomende 5G-technologieën, specifiek netwerk slicing en MEC, en rekening houden met de gefaseerde uitrol van 5G-netwerken (non-standalone en standalone implementaties). Van de verschillende geïdentificeerde modellen toont het collaboratieve businessmodel het meeste potentieel op basis van MNO-feedback, maar dit model vereist een herziening van huidige regelgeving.

Om technologische keuzes en timingbeslissingen voor 5G-implementatiestrategieën te optimisaliseren, hebben we verschillende scenario's bestudeerd voor twee soorten connectiviteit: Vehicle to Infrastructure (V2I) en Vehicle to Network (V2N), met verschillende concentraties van RSU's en macrocellen in het onderzochte gebied. Voor V2I-connectiviteit concludeerden we dat het installeren van de eerste wegkantunits aanzienlijke kosten met zich meebrengt, terwijl het toevoegen van nieuwe wegkantunits geen significante extra uitgaven vraagt. Voor V2N-connectiviteit toonden onze resultaten aan dat het toevoegen van extra macrocellen resulteerde in hoge cumulatieve kosten, waarbij OPEX de dominante kostenpost is vanwege de sitehuurkosten die nodig zijn om de nieuwe macrocellen te hosten. Daarnaast hebben we verschillende MEC-implementatieopties bestudeerd, variërend van het installeren van MEC in een gecentraliseerd model (in het datacenter) tot een volledig gedistribueerd model waarbij MEC-nodes naast elke microceltoren worden gehost. De keuze tussen implementaties heeft aanzienlijke implicaties voor zowel end-to-end (E2E) latency als implementatiekosten. We hebben drie MEC-clusteringmodellen voorgesteld die kunnen helpen bij het maken van deze afweging tussen latency en kosten. We hebben deze modellen geëvalueerd vanuit zowel kosten- als E2E-latency perspectieven, waarbij bleek dat MEC-clusteringmodellen beter presteren dan het niet-geclusterende alternatief op het gebied van kosteneffectiviteit en efficiëntie. Met MEC-clusteringmodellen kan er tot 71% bespaard worden op de TCO, en kan de E2E-latency met 1-6 ms verminderd worden. Bepaalde modellen zijn kosteneffectiever dan andere, afhankelijk van de beschikbaarheid van bestaande netwerkinfrastructuur. Netwerkdeling, zowel in actieve als passieve vorm, kan de kosten van MEC-implementaties aanzienlijk verminderen met respectievelijk 31% en 25%. Bovendien werd de impact van het penetratieniveau van wagens op de TCO geëvalueerd. Zoals verwacht, hoe hoger het aantal tegelijkertijd verbonden auto's (hoge penetratiepercentages), hoe hoger de TCO van de vereiste implementatie moet zijn om aan de gedefinieerde KPI's te voldoen. Resultaten toonden ook aan hoe duur het is om in een greenfield van de snelweg, waar geen bestaande glasvezelkabel of elektriciteitsvoorzieningen zijn, een netwerk op te zetten. Greenfield-implementatie is driemaal duurder dan wanneer deze infrastructuur reeds aanwezig is.

Hoewel de techno-economische modellen in deze doctoraatsthesis ontwikkeld zijn om de uitrol van 5G-netwerken te beoordelen, geloven we dat ze gemakkelijk kunnen worden aangepast om andere soorten technologieën te evalueren, rekening houdend met de specifieke kenmerken van die technologieën.

Toekomstig onderzoek kan het uitbreiden van de analyse naar andere 5G-gebruiksscenario's zoals mMTC omvatten, het opnemen van dynamisch verkeer in de verdeelsleutel voor netwerkslicing en het toekennen van beschikbare midellen, en het evalueren van de levering van 5G-connectiviteit aan andere bewegende opervlakken en aan andere soorten regio's zoals stedelijke en dichtbevolkte gebieden. We willen ook het belang benadrukken van het evalueren van de sociaal-economische voordelen van 5G en de mogelijke impact ervan op het behalen van de Sustainable Development Goals (SDGs) om tot een allesomvattend begrip van de mogelijke impact te komen.

Summary

5G networks have the potential to offer fast and reliable mobile broadband connectivity anytime, anywhere, and to anything thanks to promising technologies like network virtualization, network slicing, and edge computing. However, to ensure the success of 5G networks rollout, not only the technical challenges associated with these technologies need to be addressed, but also their economic implications.

In this dissertation, we conducted a comprehensive techno-economic analysis of the 5G network deployment aiming at assessing its feasibility and providing strategic recommendations to mobile network operators (MNOs). To achieve this overarching research objective, five research challenges have been tackled, including assessing the feasibility of 5G deployment for different types of regions and services, quantifying the cost savings associated with various 5G technologies, developing new cost allocation models, identifying potential business and collaboration models, and identifying technology choices and timing decisions for optimized 5G deployment strategies.

These five research challenges were addressed by conducting a techno-economic analysis of several case studies. Therefore, we developed several techno-economic models to be used in this research including scenario analysis, cost-benefit analysis, sensitivity analysis, business modelling analysis, and cost allocation models.

Regarding the feasibility of deploying 5G networks for different types of regions and services, two types of services were analyzed in three different types of regions: enhanced Mobile Broadband (eMBB) and Ultra-Reliable and Low Latency communication (URLLC) in rural, in the sky, and on the highway. For the eMBB use case, we evaluated the economic viability of an integrated satellite-4G and 5G networks to provide connectivity to rural and to airplanes respectively. The studied solution was found to be economically viable for good throughput per user (2 to 5 Mbps) service but expensive for high-speed broadband. For the URLLC use case, we studied the provision of CCAM services on the highway in cross-border environments, analysis results aim at providing insights into the Total Cost of Ownership (TCO) associated with various deployment scenarios. Several revenue models were proposed and the cost-based revenue model was found to be the most effective approach.

Regarding the potential cost savings of 5G technologies, we studied mainly the impact of two technologies namely network virtualization and edge computing. Our research showed how caching popular content on mobile edge computing (MEC can significantly save on OPEX costs and decrease the required Average Cost Per User (ACPU). However, there is a trade-off between the level of content personalization available and the caching rate. The impact of Network Function

virtualization (NFV) on deployment cost has also been assessed, demonstrating a cost saving of 12 to 45% of the core network costs resulting from the use of NFV and network slicing.

Regarding developing new cost allocation models that are better suited to the unique characteristics of 5G technologies to share costs among services and stakeholders, we developed two cost allocation models one for allocation of cost to services in sliced networks and the second one to share costs among stakeholders. For the former case, we proposed a novel cost allocation model for 5G-sliced networks, and two other models were developed to provide input to the proposed model namely resource allocation and 5G VNF dimension models. The results of the model showed that eMBB video bears a significant portion of the network costs due to its important requirement in terms of throughput compared to voice. For the latter case, we developed a cost allocation model to distribute costs among stakeholders involved in the provision of 5G-based connected mobility services. We identified key cost allocation drivers including connectivity type and the degree of collaboration between actors and showed that implementing a collaborative business model using infrastructure sharing can significantly reduce investment costs for involved operators.

Regarding identifying potential business and collaboration models that can enable successful 5G deployment, we identified several business models to provide 5G connectivity first to airplanes and second to connected cars. For the first case study, we identified different business models depending on who will provide the service to the end user, namely wholesale versus retail models, with the choice based on the level of control and risk the airline company is willing to take. The retail model is more beneficial to customers as it offers a shorter value chain and higher quality assurance. For low-cost structures, the sponsorship model is suggested. The freemium wholesale and retail models offer branding opportunities and free services to both MNOs/airline companies and passengers but require a high enough conversion rate from free to premium to work effectively. Premium airline companies are better suited to the complementary service model, where customers pay for the full premium service that includes connectivity. Furthermore, we proposed various pricing schemes for each business model to ensure maximum profitability. For the second case study, we proposed three business models that rely on emerging 5G technologies, specifically network slicing and MEC, and account for the phased rollout of 5G networks (non-standalone and standalone deployments). Among the various models identified, the collaborative business model shows the most promise based on MNO feedback, but it requires regulatory review.

Regarding identifying technology choices and timing decisions for optimized 5G deployment strategies, we studied different deployment scenarios for two types of connectivity: Vehicle to Infrastructure (V2I) and Vehicle to Network (V2N) with varying densifications of roadside units and macrocells in the studied area respectively. For the V2I connectivity, it was concluded that installing the first roadside units incurs a significant cost while adding new roadside units will not have a significant capital expenditure. For the V2N connectivity, our results showed that adding additional macrocells resulted in a high cumulative cost, where OPEX is the dominant cost due mainly to the site rental cost required to host the new macrocell.

Additionally, we studied different MEC deployment options, ranging from installing MEC in a centralized model (in the data center) to a fully distributed model where MEC nodes are hosted next to each microcell tower. The choice between deployment models carries significant implications for both end-to-end (E2E) latency and deployment costs. As a result, it becomes necessary to strike a balance between these factors, and to this end, we proposed three MEC clustering models that can help achieve this trade-off between latency and cost. We evaluated those models from both cost and E2E latency perspectives, revealing that deploying MEC clustering models outperforms non-clustering deployment in terms of both cost-effectiveness and efficiency. With MEC clustering models, you can save up to 71% on the TCO and reduce E2E traffic latency by 1-6 ms. Certain models are more cost-effective than others depending on the availability of prior network infrastructure, and network sharing in both its active and passive forms can significantly reduce the cost of MEC deployments by 31% and 25%, respectively. Moreover, the impact of the penetration rate of connected cars on the TCO was evaluated. Results showed that as expected, the higher the number of simultaneously connected cars (high penetration rates), the higher the TCO of the required deployment needs to be to meet the defined KPIs. Results demonstrated also how expensive it is to deploy in a green field of the highway where no existing fibre cable or electricity facilities. The green field deployment is three times more costly than the pre-equipped field deployment.

Although the developed techno-economic models have been used to assess the 5G networks rollout, we believe that they can be easily adapted to evaluate other types of technologies with respecting of course the specificities of those technologies.

Finally, future work for the completed research can include extending the analysis to other 5G use cases such as mMTC, incorporating dynamic traffic into network slicing cost allocation and resource allocation models, and evaluating the provision of 5G connectivity to other moving platforms and to other different types of regions like urban and dense urban areas. We want also to highlight the importance of evaluating the socio-economic benefits of 5G and its potential impact on achieving the Sustainable Development Goals (SDGs) for a more holistic understanding of its potential impact.

1
Introduction

"Don't you think dreams and the Internet are similar? They are both areas where the repressed conscious mind vents." - Yasutaka Tsutsui

In contemporary society, our daily routines, including work and social interactions, heavily rely on Internet connectivity. The consequences of even a single hour without internet access on the global economy would be significant. According to [1], a 60-minute interruption in World Wide Web connectivity would result in a loss of €1.7 billion to the world economy. This figure would increase to €17 billion after ten hours and to €41.7 billion after a full day. In modern times, fast and reliable internet access is no longer considered a luxury, but rather a necessity. The continual evolution of smartphone technology has further heightened the need for fast and reliable internet connectivity anytime and anywhere. Consequently, the need for broadband networks has become increasingly apparent in recent decades. Broadband connectivity plays a pivotal role in communication, education, commerce, and other vital services. Furthermore, the coronavirus crisis highlighted the critical role that broadband networks played in enabling remote work and telecommuting, increasing flexibility for both employees and employers and preventing the economic catastrophes associated with the pandemic. In the Broadband Commission report [2], broadband was dubbed the "Hidden Hero" of the pandemic.

This chapter provides the necessary background to comprehend the research presented in this dissertation. Initially, an overview of broadband networks is given, comprising wired and wireless network types. The focus then shifts to explaining the fifth generation (5G) of the cellular network in greater detail, including its technology, use cases, and the enabler key technologies.

1.1 Broadband networks

In telecommunications, broadband refers to a permanent, high-speed internet connection with a wide bandwidth for data transmission. It surpasses the older narrowband dial-up internet access in terms of speed and reliability. Broadband technology enables us to communicate information through voice, video, and data by utilizing a wide range of devices and network systems.

1.1.1 Definition, network architecture, and economic considerations

Initially, "broadband" was a technical term but later transformed into a marketing phrase for any Internet access technology with relatively high speed. The 802.16-2004 standard [3] specifies that broadband must have instantaneous bandwidths exceeding 1 MHz and support data rates above 1.5 Mbit/s. Recently, the Federal Communications Commission (FCC) revised the definition to include download speeds of at least 25 Mbit/s and upload speeds of at least 3 Mbit/s [4]. For the European Commission (EC), a satisfactory broadband internet access service, at a fixed location, should comprise a download sync speed of at least 30 Mbps, an upload sync speed of at least 1.5 Mbps, and low latency to enable effective voice and video calls, and unrestricted data usage [5].
The generic architecture of a broadband network is captured in Figure 1-1.

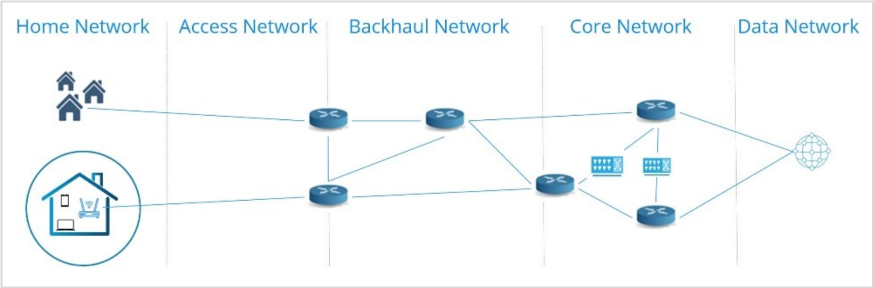


Figure 1-1: High-level architecture of a broadband network

The high-level network architecture consists of mainly three parts: the core network, the backhaul network, and the access network. These network parts link the end-user or the household/business to the data network (e.g., internet). The household/business is served by the home network also known as Local Area Network (LAN).
Broadband technology comprises both wired and wireless high-speed transmission technologies. The following section provides a definition for each of the three primary components of a network, accompanied by examples from both wired and wireless networks.

Access Network: The access network is the part of the network that connects the end user's premises or the user equipment (UE) to the service provider's distribution point or base station in the wired and wireless network, respectively. Examples of access network technologies in the wired world are DSL, cable modem, and fiber-optic cable, while examples in the wireless world include 2G/3G/4G/5G cellular technologies, and satellite [6]and [7].

Backhaul Network: The backhaul network is the part of the network that connects the access network to the core network in both the wired and wireless worlds. It aggregates traffic from the access network and transmits it to the core network for further processing and routing. Examples of backhaul network technologies include copper, fiber-optic cable, microwave links, and satellite [6]and [7].

Core Network: The core network is the central part of the network that provides connectivity between different access networks and/or other networks (e.g., data network) and is responsible for routing traffic between different parts of the network and providing various services such as security, quality of service (QoS), and mobility management in both the wired and wireless worlds. Examples of core network technologies in the wireless world include Evolved Packet Core (EPC) and 5G Core (5GC) [6] and [8].

It's important to note that the definitions of these terms may vary depending on the context and the specific technology being used. However, the definitions provided by the international telecommunication union (ITU) and the 3rd Generation Partnership Project (3GPP) are widely used in the wired network and cellular industry respectively and provide a general understanding of these concepts.

Economic considerations

Deploying new broadband networks or upgrading existing ones involves a complex evaluation of various factors that can significantly impact the feasibility and success of the deployment. To ensure a viable return on investment, it's crucial to carefully assess the economic factors associated with the broadband network. Apart from selecting the appropriate broadband technology to meet the service requirements in specific geographic conditions, which will be discussed in the next section, the cost of deploying and maintaining the network infrastructure is a critical factor to consider. In what follows, a detailed look at the challenges associated with each element of the broadband network is given.

Firstly, the access network is often considered the most challenging element of broadband networks due to its direct interface with end-users and their devices, introducing various complexities and constraints in terms of coverage, capacity, reliability, and quality of service. Additionally, the access network may have to operate in different environments, each with its own set of challenges and requirements. Even after selecting the suitable broadband technology, estimating the cost of deploying the access network is the most challenging factor. A study by the Global System for Mobile Communications Association (GSMA) Intelligence [9] found that the Radio Access Network (RAN) accounts for the largest share of the total cost of a mobile network deployment, followed by the backhaul network and the core network. An estimated 60-80% of the capital expenditure (CAPEX) and 40-60% of the operating expenditure (OPEX) are going towards the RAN. The backhaul network accounts for about 20-30% of the CAPEX and 20-30% of the OPEX, while the core network accounts for 5-15% of the CAPEX and 10-20% of the OPEX

Secondly, the backhaul network connects the core network to the access network and requires high-speed connections to the core network, which can be expensive to deploy, especially in rural or remote areas with low population density. The backhaul network may also require complex routing and traffic management systems, which can add to the cost and complexity of deploying and maintaining this part of the network. According to [10], "the key challenge of backhaul is in providing cost-effective capacity at the lowest possible total cost of ownership (TCO)".

Finally, building and maintaining a cost-effective core network for broadband networks is a critical challenge. Several factors contribute to the cost of a core network, such as equipment, installation, maintenance, and upgrades. The challenge is to design a network that is scalable, reliable, and secure while keeping costs under control [11].

In conclusion, deploying new broadband networks or upgrading existing ones requires careful evaluation of the economic factors associated with each element of the network. By estimating the cost of deploying and maintaining the network infrastructure while comparing different technological options, a viable return on investment can be achieved.

In the next subsection, some of the aforementioned technologies are classified into two primary categories: wired and wireless broadband networks. The link between each technology and the specific part of the high-level network architecture (Figure 1-1) it impacts is also highlighted. While most technologies primarily impact the access network part, hence, the term broadband access networks is widely used to refer to broadband networks, some technologies can span more than one network part of the architecture presented in Figure 1-1.

1.1.2 Wired broadband networks

As explained previously, several technologies are classified as wired broadband connectivity. In the following, each technology will be described apart.

Cable TV connectivity or Cable modem

Cable TV connectivity or also commonly referred to as cable modem facilitates the provision of broadband by cable operators, utilizing the same coaxial cables that transmit visual and auditory signals to your television set. Typically, cable modems are external gadgets featuring two ports - one that connects to the cable wall socket and another linking to a computer as illustrated in Figure 1-2. These modems can transmit data at a speed of at least 1.5 Mbps.

Many cable TV companies provide high-speed Internet access by dedicating one or more of their 6-MHz channels to this service, utilizing modulation techniques to offer various levels of connectivity to their customers. The customers receive a cable modem, which links to the cable alongside the television, providing Ethernet connectivity to their computer. To enable wireless connections to the Internet for PCs and laptops, a Wi-Fi access point is frequently attached to the cable modem [12]. Additionally, TV can still be watched simultaneously. The actual transmission speed depends on factors such as the type of cable modem, cable network, and volume of traffic. The architecture of the cable modem network spans across the three previously explained network parts namely the radio access, backhaul and core network.

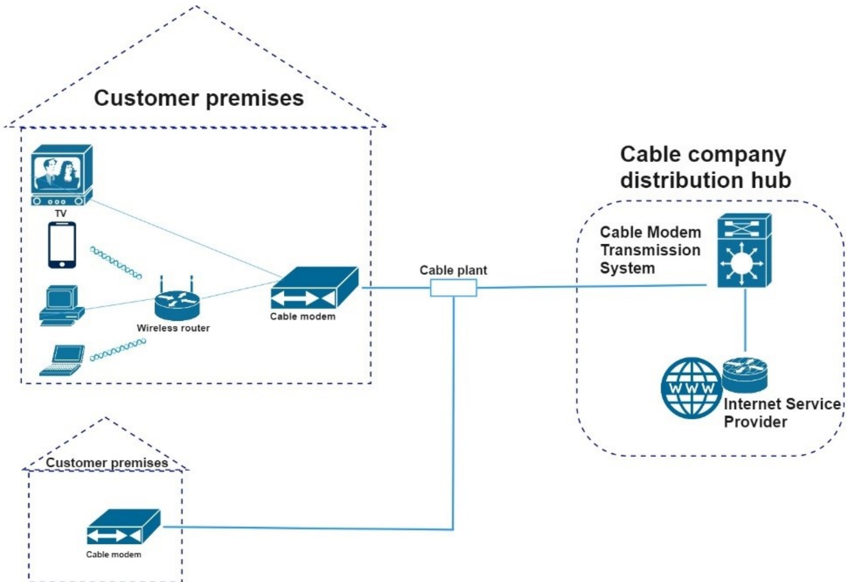


Figure 1-2 *An illustration of the Cable modem technology: based on* [13]

Digital Subscriber Line (DSL)

DSL is a technology that allows for faster transmission of data over traditional copper telephone lines. The speed of DSL service can range from several hundred Kbps to several Mbps, depending on the distance between the home or business and the nearest telephone company facility. DSL utilizes the standard telephone lines found in most homes. Despite these lines being designed only for voice signals, advanced techniques have been developed to enable fast data transmissions. A DSL modem is then used to connect to the subscribers' computers and home telephone lines. The data rate of a DSL line is heavily dependent on the distance between the subscriber and the central office. As the distance increases, the data rate decreases. However, even for the longest runs of up to 6 km, data rates of up to approximately 2 Mbps are achievable.

DSL technology has undergone significant advancements, giving rise to newer versions like ADSL, VDSL, VDSL2, and G.fast. VDSL can offer data rates of up to 52 Mbps, while VDSL2 can reach up to 200 Mbps over shorter distances. This remarkable progress is due to the use of fiber in the backhaul network part, replacing copper as depicted in *Figure 1-4*. The deliverable rate is dependent on the distance between the subscriber and the digital subscriber line access multiplexer (DSLAM). The latest DSL technology, G.fast, can deliver combined download and upload speeds of up to 1 Gbps over existing telephone wiring. Achieving this requires the distribution point unit to be located closer to the home. The evolution of DSL technology is captured in *Figure 1-3*.

As shown in *Figure 1-4*, the ADSL technology spans across the access and backhaul network parts, yet, the new advancements of DSL technology cover the access network part only.

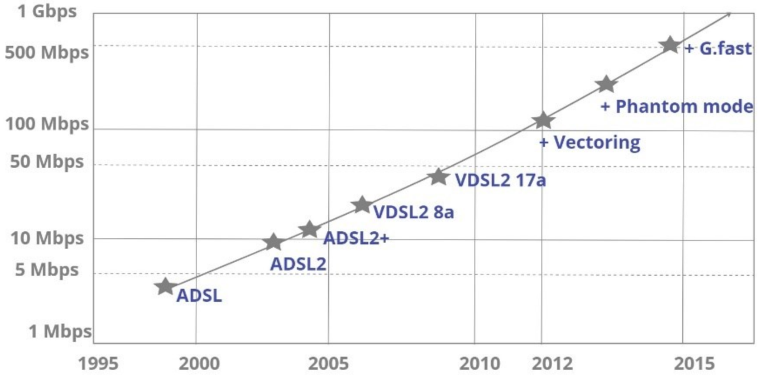


Figure 1-3 Evolution of DSL technology over time. The x-axis represents years, while the y-axis represents data. The graph shows how DSL technology has advanced over time, with increasing data rates enabling faster internet speeds for users: based on [14].

Fiber

Fiber optic technology utilizes transparent glass fibers, which are as thin as human hair, to transmit light instead of electrical signals for data transmission. This enables data to be transmitted at much higher speeds than traditional DSL or cable modem services, often reaching tens to hundreds of Mbps. The actual speed experienced by the user may depend on factors such as the proximity of the fiber to the computer and the service provider's configuration, including bandwidth usage. Fiber can also deliver other services such as voice (VoIP) and video-on-demand simultaneously. Although fiber broadband is currently only available in limited areas, telecommunications providers have plans to expand their fiber networks and offer bundled voice, Internet, and video services. The technology can be installed all the way to the customer's home or business and is known as Fiber to The Home (FTTH), to the curb outside also called Fiber to The Curb/Cabinet (FTTC), or to a location in between the provider's facilities and the customer also known as Fiber to the distribution point (FTTdp). The different fiber access options and their possible combination with DSL are depicted in *Figure 1-4*. Fiber technology is used for both access and backhaul network elements.

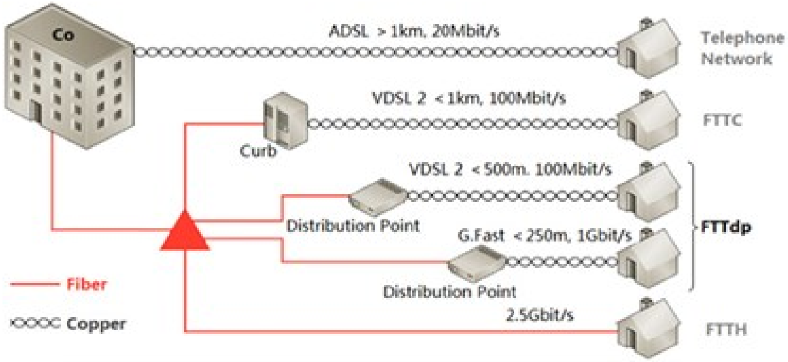


Figure 1-4 Overview of the different Fiber access types with comparison and combination with DSL access technologies, based on [15].

The European market share of the different wired broadband technologies as of 2019 is illustrated in *Figure 1-5*.
The data presented in *Figure 1-5* highlights the dominance of DSL in several European countries, including France, Italy, and the UK. However, it is worth noting that certain countries such as Spain, Latvia, Lithuania, and Romania have made significant strides in the fiber market and have a relatively larger market share compared to other fixed broadband technologies.

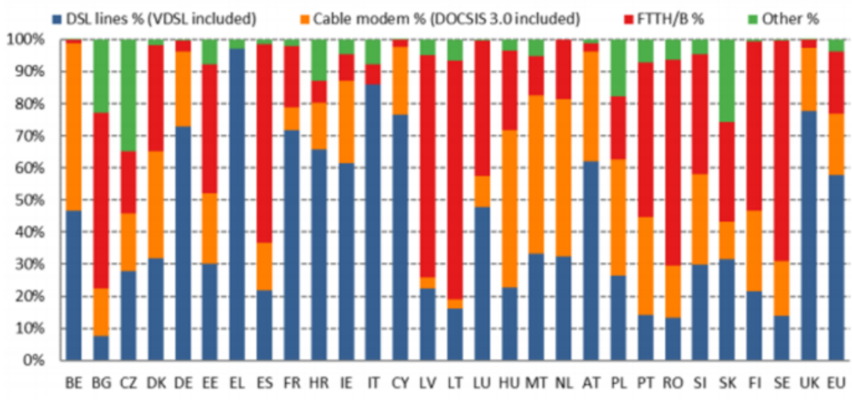


Figure 1-5 Fixed broadband subscriptions – technology market shares in the EU (% of subscriptions), July 2019 [16] showing the dominance of DSL technology in several European countries.

On average, DSL technology has the highest market share in Europe with 72% in 2013 and 58% in 2019 followed by modem cable. During the same period, the market share of DSL decreased significantly, while fiber gained traction, increasing its market share from a mere 5% in 2013 to almost 20% in 2019, as depicted in *Figure 1-6*.

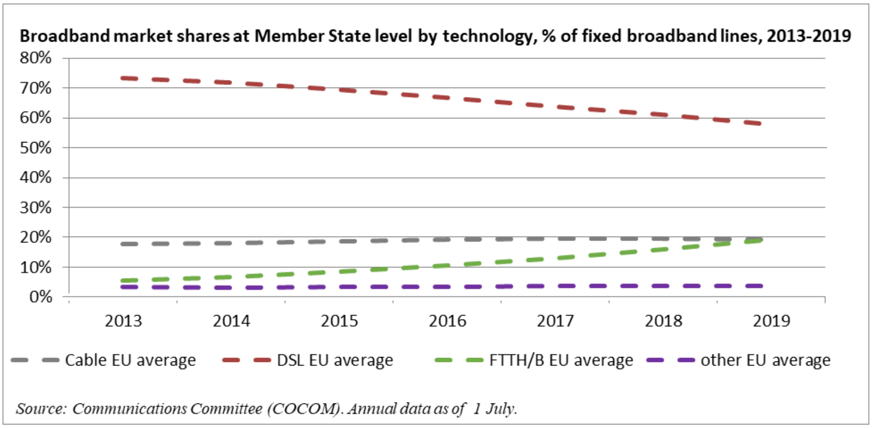


Figure 1-6 The European average of technology market shares for fixed broadband subscriptions, July 2013- 2019 [16] showing the decline of DSL market share in favor of fibre.

1.1.3 **Wireless broadband networks**

To comprehend the various categories of wireless broadband networks, we should first grasp the concept of wireless communication or transmission.
Wireless communication involves transmitting data, voice, and video without physical cables or wires by using electromagnetic signals that travel from transmitting devices to end-user equipment. The initial wireless transmitters used to transmit coded signals, such as Morse code [17], but advancements in modulation led to the transmission of voice and music, which became known as radio. Wireless transmissions rely on electromagnetic waves to carry signals, data, voice, and video over a communication path.
The foundation for modern wireless networking was established in the early 1970s through the launch of ALOHAnet in Hawaii, a wide area network (WAN) that utilized ultra-high frequency signals to broadcast data across the islands. The technology behind ALOHAnet played a critical role in the creation of Ethernet in 1973 and contributed significantly to the development of the first wireless standard, 802.11 [18].

History of wireless technology

Wireless technology has been in existence for over a century, but it wasn't until the last 15 years, especially after the ratification of the 802.11ac [19] and 4G [20] standards, that it advanced enough to allow the development of comprehensive applications and services for widespread adoption by consumers and enterprises. The evolution of wireless features has progressed from simple data transfers with speeds of just several Kbps to operations that require gigabits of data for completion.

With each new generation of wireless technology, more sophisticated capabilities are created, providing users with greater flexibility in accessing information and services from almost anywhere. Simultaneously, mobile devices have become more potent and versatile, allowing users to perform complex tasks. The progress in performance, capacity, and coverage will continue, enabling people to stay connected wherever they go.
The history of the evolution of wireless technology is illustrated in *Figure 1-7*.

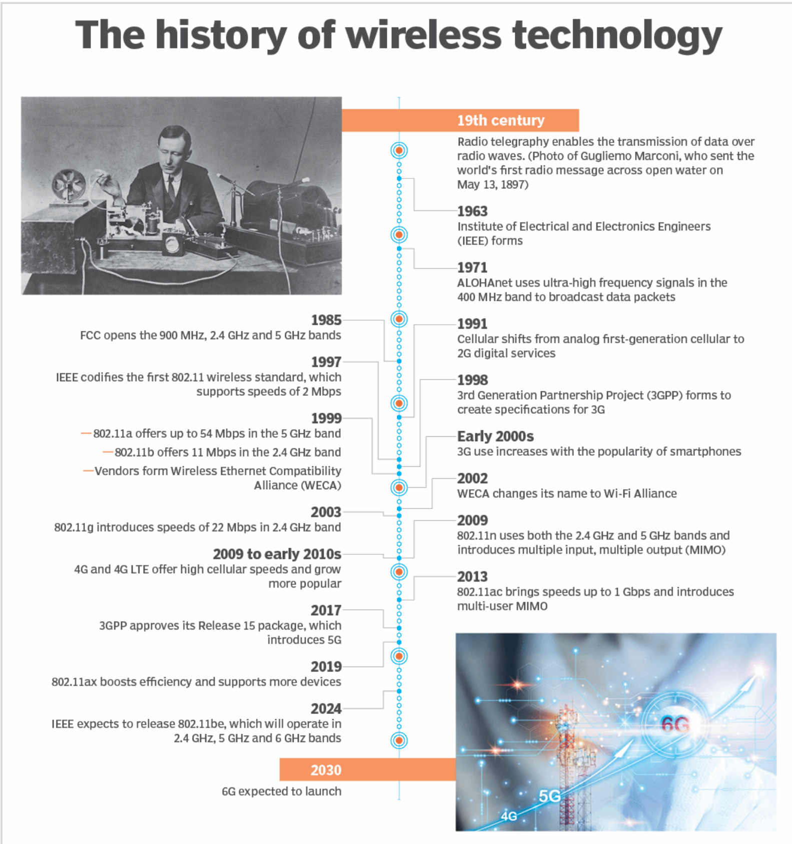


Figure 1-7 The evolution of wireless technology, photo by TechTarget [21]

So, in technical words, what is a wireless network?

A wireless network is a grouping of devices that communicate through radio frequencies for sending and receiving data. Unlike wired networks that rely on physical cables to establish connections, wireless networks enable the connection of endpoint computing devices such as smartphones, laptops, and tablets to peripheral devices without using wired cabling.
Wireless networks rely on radio transmission to broadcast and receive signals within a specific range of electromagnetic radiation spectrum or "spectrum." Data

is transmitted over wireless networks through small, embedded hardware antennas within devices.

Different wireless networks utilize different frequency ranges within the spectrum, and various channels help reduce congestion within each frequency range. These frequencies are regulated by government agencies to avoid interference between different wireless networks and other electronic devices that use the same frequencies. Different wireless networks may use different wireless technologies, such as Wi-Fi, Bluetooth, or cellular networks, depending on the application and requirements [12].

Figure 1-8 shows a comparison between the various types of wireless networking technologies in use today based on the range and data rates. These technologies will be detailed separately in the following sub-sections.

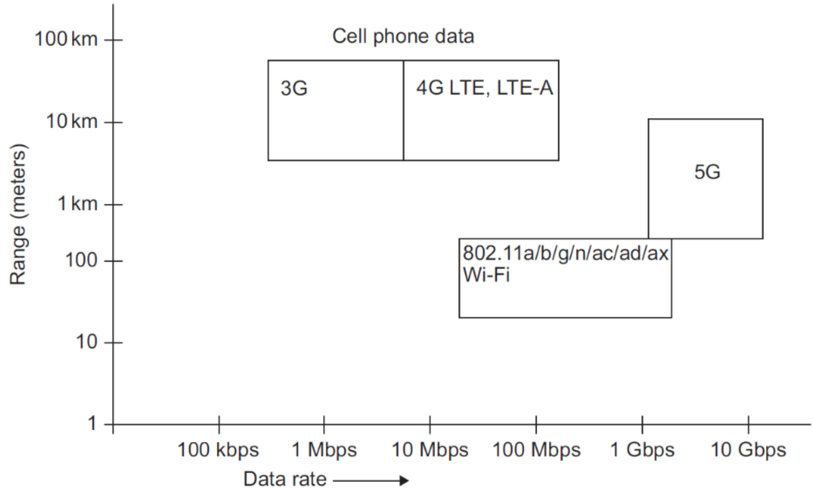


Figure 1-8 Summary of wireless networking technologies shown as data rate versus range, adopted from [12].

Figure 1-1 illustrates the three main components of a generic network architecture: the core, the backhaul, and the access network. Wireless technologies are classified based on their impact on specific parts of the network architecture. The access network can be further divided into two parts: the shared access network utilized by all users and the private Local Area Network (LAN) used inside limited areas such as residences, schools, or office buildings. To categorize different wireless technologies, we follow an inside-to-outside approach, starting from the LAN to the backhaul network.

Wireless technologies for the Local Area Network (LAN)

In the high-level network architecture presented in Figure 1-1, wireless networks for LAN relates to the home network part. The famous wireless access technology used for LAN is Wi-Fi. Wi-Fi, short for "wireless fidelity," is a wireless networking technology that has gained widespread popularity. It is based on the IEEE's wireless Ethernet standard, 802.11, and was initially designed to extend Ethernet LANs wirelessly. This technology has allowed laptops, personal computers, and even cell phones to connect to an organization's or company's LAN without the need for a wired connection. This has simplified the process by eliminating the need for new wiring and provided users with greater mobility, especially since laptops are now more commonly used than fixed PCs. In today's world, Wi-Fi has become the primary wireless link for smartphones and tablets to access the Internet [].

Figure 1-9 shows how Wi-Fi network works. The data to be sent over the network is encoded into electromagnetic waves that are propagated through the air, which are received and decoded by the wireless network interface card (NIC) installed in the receiving device. Wireless networks use a wireless access point (WAP) to act as a central hub to manage communication between devices. The WAP serves as a connecting link between the wireless and wired networks, enabling wireless devices to utilize resources like servers and printers available on the wired network.

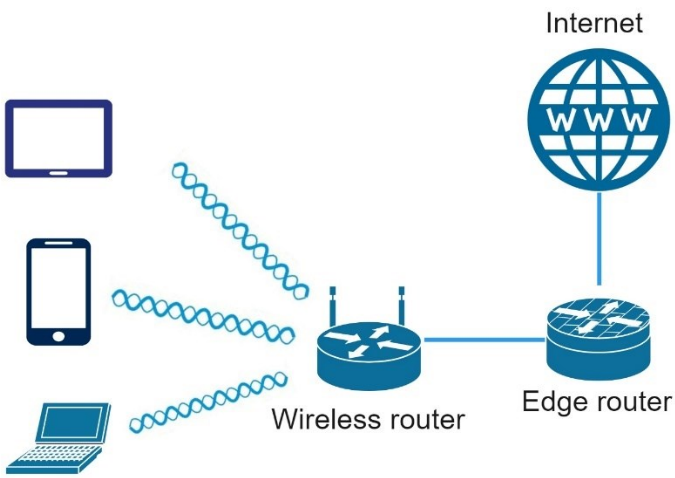


Figure 1-9 Illustration of a simple Wi-Fi network

Wi-Fi's popularity can be attributed to its low cost, high data rate, and ease of deployment. The IEEE 802.11 standard includes different versions such as 802.11a, 802.11b, 802.11g, and 802.11n, each with varying data rates, with a maximum of 54/11/54/600 Mbps for 802.11a/b/g/n respectively, and a range of 100-200m [22]. 802.11a operates at the unlicensed 5GHz band, while 802.11b/g operates at the unlicensed 2.4GHz band. Despite 802.11a providing a greater number of frequency channels and higher throughput, it experiences more multi-path propagation issues compared to 802.11b/g. The 802.11b standard employs the Complementary Code Keying (CCK) technique with a 25 MHz bandwidth, while 802.11a/g utilize Orthogonal Frequency Division Multiplexing (OFDM) modulation with a 20 MHz bandwidth [23].

In *Figure 1-10*, the evolution of Wi-Fi standards with their respective use case is illustrated.

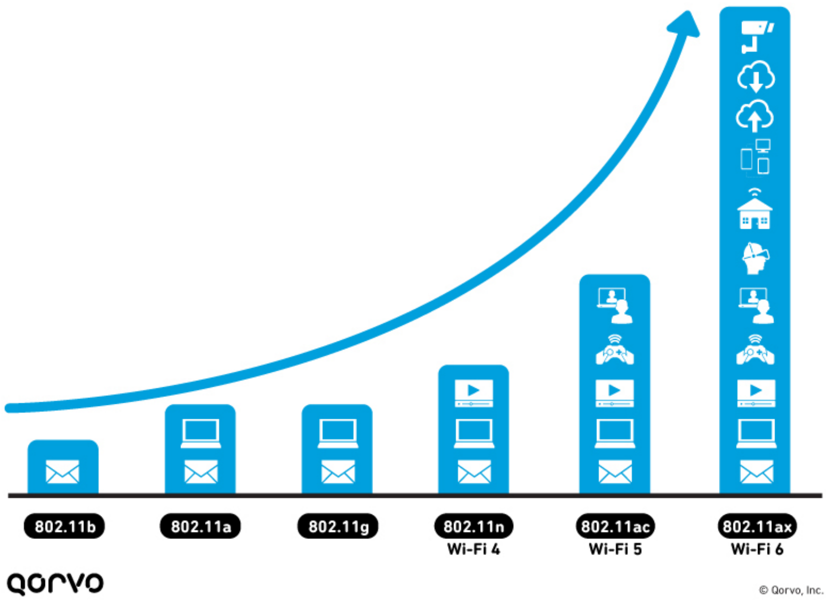


Figure 1-10 The evolution of Wi-Fi standards and their respective use cases, adopted from [24]

An additional wireless broadband technology within the IEEE 802 standards is 802.16, commonly referred to as WiMAX. This technology was specifically developed to cater to the requirements of Wireless Metropolitan Area Networks (WMANs). To learn more about WiMAX, we refer the reader to this article [23].
While Wi-Fi is more an indoor connectivity solution, users need an outdoor mobile broadband to use with their smartphones and tablets while moving outside the Wi-Fi range.

Cellular networks: evolution from 1G to 6G: 3GPP standards

Cellular network technology spans across the access and home network parts illustrated in Figure 1-1. A cellular network represents a wireless network enabling mobile devices, including smartphones and tablets, to communicate with one another and access the internet via a network of interconnected cells. Each cell is typically served by a base station, which uses radio waves to transmit and receive data between the mobile device and the network [25]. The evolution of cellular networks has been driven by several technical drivers that have enabled each new generation of networks to provide increased capacity, data rates, and coverage. Below is a summary of the evolution of cellular networks from 1G onwards, highlighting the main technical drivers. Although 1G and 2G networks are not classified as wireless broadband networks, we have included them in this

overview to provide a comprehensive understanding of the evolution of cellular networks. However, it is only with the advent of 3G networks and beyond that cellular networks are commonly referred to as broadband networks.
The evolution of the different cellular communication is illustrated in *Figure 1-11* and each generation will be detailed further in the following sub-sections.

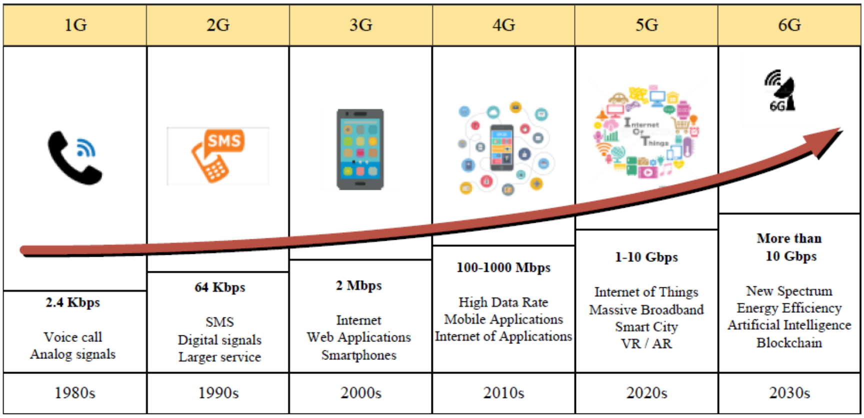


Figure 1-11 The evolution of cellular communication [28]

1G (First Generation)
The first generation of cellular networks, 1G, was introduced in the 1980s and provided analog voice services where each call used a separated narrowband frequency channel. The main technical driver for 1G was the development of the cellular network architecture, which allowed mobile phones to be used over a wider area than was previously possible. The first 1G network was launched in Japan in 1979 by NTT, and subsequently, other countries followed. Multiple cellular standards, including Nordic Mobile Telephone (NMT), Advanced Mobile Phone System (AMPS), Total Access Communications System (TACS), and others, were adopted across different regions [26] and [27]. These networks operated on the 800 MHz (USA) and 900MHz (EU) frequency band and used frequency modulation technology for voice communication. The data rates were slow, and the networks supported only voice calls [26].

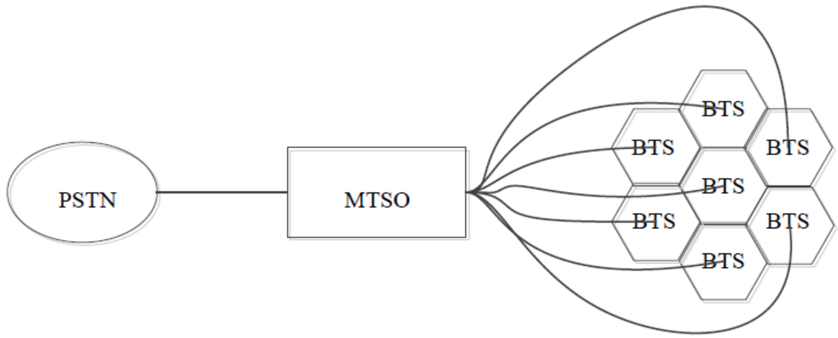


Figure 1-12 The 1G network architecture, adopted from [28].

The network architecture of 1G networks is presented in Figure 1-12. In cellular networks, cells are represented by a hexagonal form. In order to enhance capacity, these cells can be subdivided into smaller units called sectors. The Mobile Telephone Switching Office (MTSO) establishes connections between Base Transceiver Stations (BTS) and the Public Switched Telephone Network (PSTN). The MTSO manages handoffs, call routing, registration, authentication, and other functions in a circuit-switched[1] (CS) [29] network. Operators obtain licensed spectrum from regulatory organizations to provide voice services.

2G (Second Generation)

Second-generation cellular networks were introduced in the 1990s and used digital technology, which offered several advantages over 1G. The primary driver behind 2G was the need for improved voice quality and capacity. 2G networks employed digital modulation methods, including Time-Division Multiple Access (TDMA) and Code-Division Multiple Access (CDMA), enabling a higher number of users to share the same frequency band and introducing circuit-switched data services [26]. The most popular 2G standard was GSM (Global System for Mobile Communications). GSM was introduced in Europe in 1991 and quickly gained worldwide acceptance. 2G also enabled the introduction of short message service (SMS), which was a significant milestone in the history of mobile communication. 2G networks operated in the 900 MHz and 1800 MHz frequency bands. The European Telecommunications Standards Institute (ETSI) Special Mobile Group (SMG) played a key role in the development of the GSM standard. This group chose TDMA as the standard, which was subsequently followed by other digital standards adopted in different countries. Alternatively, CDMA was first initiated within the IS-95 standard (later called cdmaOne) [26]. GSM primarily offers voice and circuit-switched data services, attaining data rates of up to 9.6 kbps. [30].

[1] Circuit switching establishes a dedicated physical path between two communication devices for the duration of a communication session. This path is reserved exclusively for the two devices to exchange data until the session is terminated. During this time, the path is unavailable for other communications [32].

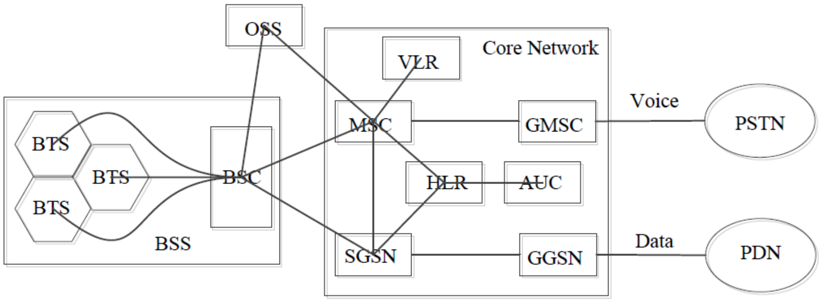


Figure 1-13 The network architecture of 2G GSM networks, adopted from [28].

Figure 1-13 presents the network architecture of 2G.

The Base Station Subsystem (BSS) is composed of two key components: the Base Transceiver Stations (BTSs) and the Base Station Controller (BSC). The entire GSM network is monitored and controlled by the Operation and Support Subsystem (OSS). The Mobile Switching Center (MSC) provides services such as registration, authentication, call routing, call location, and other features. The Home/Visitor Location Register (HLR/VLR) stores subscriber information. The HLR provides subscriber data to the Gateway Mobile Switching Center (GMSC) for call routing to the appropriate MSC. A Serving GPRS Support Node (SGSN) manages mobility and packet routing, while the Gateway GPRS Support Node (GGSN) connects the GPRS network to external packet-switched[2] networks [28].

2.5G and 2.75G

The lower data rate of GSM has prompted cellular systems to develop new technologies, such as General Packet Radio Service (GPRS) [31], [33], and [34]. GPRS, viewed as a 2.5G technology, employs both circuit and packet switching similar to GSM. It can boost data rates up to 50 kbps, using transmission and modulation methods akin to those in GSM.

GPRS serves as a crucial stepping stone towards an enhanced GSM environment, supporting higher data rates through EDGE. Predating 3G, EDGE is a radio technology that allows data transmission and reception at speeds up to 200 kbps. Building upon the GSM standard, EDGE utilizes the same transmission mechanism and bandwidth as GSM, but with a different modulation technique. While 8PSK offers higher data rates within a smaller area, GMSK ensures reliable wide coverage. EDGE was designed to improve packet-switching services and enable high-speed data applications, such as multimedia [33] and [34].

[2] Packet switching is a method of sending data over a network in which the data is divided into small packets and sent individually across the network. Each packet is sent independently and may take a different path through the network to reach its destination. When the packets arrive to their destination, they are reassembled into their original form [32].

3G (Third Generation)

The third generation (3G) of cellular networks was introduced in the early 2000s and marked a significant improvement over its predecessor, 2G. 3G networks enabled faster data transfer rates and supported multimedia applications such as video calling and mobile internet browsing. The main technical driver behind 3G was the use of a new radio access technology called Wideband Code Division Multiple Access (WCDMA) and high-speed packet access (HSPA) technologies, which allowed data rates of up to several megabits per second. 3G networks operated on a variety of frequency bands, including 850 MHz, 900 MHz, 1700 MHz, and 2100 MHz, depending on the region. 3G networks supported data rates of up to 2 Mbps, which made it possible to stream video and use more advanced data services like video calling.

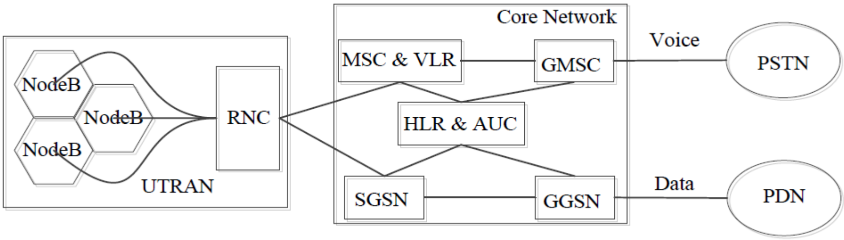


Figure 1-14 The 3G WCDMA network architecture, adopted from [28].

The network architecture of 3G networks is presented in Figure 1-14. The BSC functions in 2G were replaced by the Radio Network Controller (RNC), while the Base Transceiver Station (BTS) functions were taken over by the NodeB. WCDMA, also known as Universal Mobile Telephone System (UMTS), is composed of NodeB and RNC groups, which collectively make up the UMTS Terrestrial Radio Access Network (UTRAN). The two cellular protocols used in 3G were WCDMA (UMTS) and CDMA2000 [28].

HSPA+ (Evolved High-Speed Packet Access) is an advanced 3rd Generation Partnership Project (3GPP) standard introduced in late 2008 and widely adopted by 2010. On the other hand, 3.9G Long-Term Evolution (LTE) boasts capabilities that exceed those of conventional 3G mobile communications [33] and [34]. Both ITU and 3GPP have reached the consensus that LTE can be rightfully called 4G technology [35].

4G (Fourth Generation)

The fourth generation of cellular networks, 4G, was introduced in the late 2000s and provided even higher data rates and improved network efficiency. 4G is a technology that exclusively enables packet-switched based networking. The main technical drivers for 4G were the introduction of long-term evolution (LTE) technology, and the use of multiple-input and multiple-output (MIMO) antenna systems. 4G networks operated on a variety of frequency bands, including 700 MHz, 850 MHz, 1800 MHz, 2100 MHz, and 2600 MHz, depending on the region. 4G networks supported data rates of up to 100 Mbps, which made it possible to stream high-definition video and use more advanced data services like mobile gaming and virtual reality.

4G utilizes orthogonal frequency division multiplexing (OFDM) to enable advanced multi-antenna transmission and expandable transmission bandwidths up to 20 MHz. One of the critical components of the system technology is MIMO, which allows for multi-stream transmission and higher data rates. This results in improved spectral efficiency, connection quality, and the ability to adjust radiation patterns for signal gain and mitigation. Adaptive beamforming in antennas is used to generate an interference array, which in turn improves the signal quality.

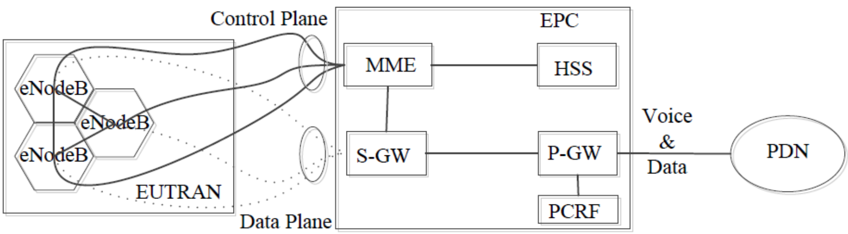


Figure 1-15 The 4G LTE network architecture, adopted from [28].

The network architecture of 4G networks is presented in Figure 1-15. 4G is a globally standardized cellular technology that utilizes an evolved packet core (EPC) to replace the core network functions, and an eNodeB to replace the NodeB. Together, the group of eNodeBs form the evolved UTRAN (EUTRAN). The EUTRAN to EPC connection utilizes both control plane and user plane signals, allowing for flexible evolution rates and network deployment options.

The Home Subscriber Server (HSS) element, taken from 3G and 2G, is a central database that stores information about all network subscribers. The packet data network gateway (P-GW) serves as a gateway that connects packet data networks (PDN) to the outside world and is accessible via an access point name (APN). The serving gateway (S-GW) routes data from the base station to the PDN gateway, which serves the same purpose as the GGSN and SGSN in UMTS and GSM.

The mobility management entity (MME) utilizes signaling messages and the HSS to govern high-level mobile functioning. The policy control and charging rules function (PCRF), as part of the P-GW's policy control enforcement function (PCEF), regulates charging features and makes policy control decisions that are flow-based. The control plane is used by nodes to exchange signaling messages, while the user plane is used for data transmission between the UE and the PDN. Signaling messages may be used by nodes to start, stop, or change a UE's data session [28].

With the adoption of 4G, data speeds have significantly increased to up to 100 Mbps). As demand for mobile broadband communication capacity continues to grow year after year, the wireless technology has evolved to include LTE Advanced (LTE-A). LTE-A is designed to achieve even higher peak throughput rates than LTE, with a theoretical maximum rate exceeding 1 gigabit per second (Gbps) [36].

5G (Fifth generation)

The fifth generation of cellular networks, 5G network are the latest generation of cellular networks and were introduced in the late 2010s. The standardization process for 5G communication is now complete and the implementation of this standard is taking place globally [37]. The International Telecommunication Union (ITU) has defined the vision and specifications for 5G with the goal of meeting eight Key Performance Indicators (KPIs) and three common use cases namely enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communication (URLLC), and massive Machine Type Communications (mMTC) [38]. These KPIs include achieving a data rate of Gigabits per second (Gb/s) for eMBB, air interface delay of a few milliseconds (ms) for URLLC, and Indicator type communication for mMTC, with a requirement for millions of connections per square km (1 M/km2) [38].

To achieve these KPIs, several enabling technologies have been developed and considered during the standardization process. Technical trials are currently being conducted to ensure the successful implementation of these technologies. Examples of these technologies include the Dual Connectivity Architecture, Massive Multiple-Input Multiple-Output (mMIMO), mm-wave communications, flexible frame structure among other wireless technologies [39].

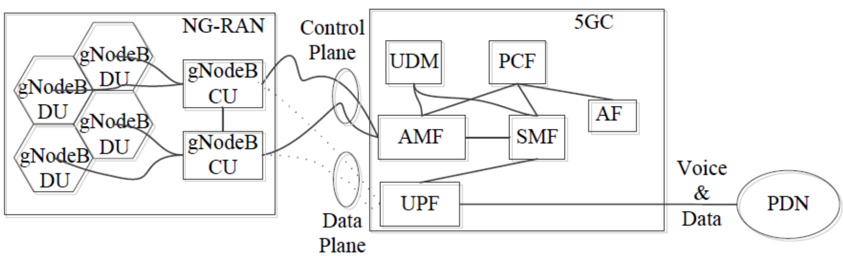


Figure 1-16 The 5G NR network architecture, adopted from [28].

The network architecture of 5G networks is presented in Figure 1-16. The 5G New Radio (NR) network is built around the "next generation NodeB," or gNB, which serves as the base station (BS). Unlike previous generations of cellular networks, the NG-RAN consists of two separate units: the centralized unit (CU) and distributed unit (DU). The gNB replaces the eNodeB, while the 5G core (5GC) supplants the EPC.

Within the 5G network, various functions are responsible for specific tasks. The access and mobility management function (AMF) is accountable for mobility management among other tasks. The session management function (SMF) allocates UE IP addresses, performs UPF selection and control etc.. The unified data management (UDM) function handles subscription management, registration, user data, and mobility management. The policy control function (PCF) enforces policy rules for control plane (CP) functions.

The user plane function (UPF) plays a critical role in connecting external data networks to the 5G network, as well as managing UP QoS. The application function (AF) manages the policy framework and other policy control tasks [28]. Further explanation regarding all these 5GC network functions is provided in section 1.2.2.

Table 1-1 Comparison between different generation of cellular networks, adopted from [28]

Cellular Generations	Standards	Applications	Multiple Access Techniques	Physical Resources	Duplex Methods	Switching Techniques
1G	AMPS, NMT, TACS	Voice call with analog signals	FDMA	Frequency	FDD	CS
2G	GSM, IS-54, IS-95	Voice services with digital signals, SMS. From 2.5G, email and web-browsing.	TDMA	Time slots	FDD	CS From 2.5G CS&PS
3G	WCDMA, CDMA2000	Mobile TV, video telephony, and video conference	CDMA	Time slots/ PN codes	FDD/TDD	CS&PS
4G	LTE	High data rate applications, e.g. HD TV, cloud computing, video gaming, etc	OFDMA	Time/ Frequency	FDD/TDD	PS
5G	NR	IoT, enhanced broadband, smart city, VR, AR etc.	OFDMA	Time/ Frequency	FDD/TDD	PS

The remarkable surge in connected devices has resulted in a significant uptick in data traffic. This, along with the emergence of new applications such as autonomous vehicles and Virtual Reality/Augmented Reality (VR/AR), has created a demand for data rates and latency beyond what 5G networks can deliver. These challenges are fuelling the development of 6G communication systems, anticipated to be deployed by 2030 [40].

6G (Sixth generation)

To fully realize the potential of 6G networks and address the challenges presented by emerging technologies, researchers from academia and industry are working to define the main qualifying technologies that will differentiate 6G from its

predecessor. While 6G standard has not yet been finalized, here below we list some of the 6G expected capabilities
The development of 6G wireless communication networks aims to enhance data throughput, system capacity, spectrum efficiency, latency reduction, and coverage expansion up to 100 times more than 5G. The focus is to enable faster and more efficient communication, particularly for the Internet of Everything (IoE), and support the growth of an industrial omnipresent intelligent mobile society. 6G will operate at higher frequencies like mm-wave, terahertz, and visible light, allowing data rate reaching Tbps, and a user experience data rate of 10 Gbps. It will also integrate artificial intelligence (AI) technology, prioritize endogenous security solutions, and leverage new technologies like edge computing, cloud computing, AI, and blockchain [36].

Hybrid solution: Fixed wireless access (FWA)
Fixed wireless access (FWA) is a wireless technology used to provide broadband internet access to fixed locations like homes and businesses. FWA utilizes radio waves to transmit data between a wireless access point or base station and a receiver located at the customer's premises, typically a router or modem that connects to the customer's devices via Ethernet or Wi-Fi. While FWA mainly relies on cellular networks for access, it can also use other technologies such as satellite for backhaul, as shown in *Figure 1-17*.
FWA is often used in rural or remote areas where laying physical infrastructure like cable or DSL is not feasible or cost-effective. Additionally, FWA can be used to provide internet access to temporary locations like events or construction sites. The FWA technology spans across the access and home network parts of the high-level network architecture depicted in Figure 1-1, with backhaul networks included when technologies other than cellular networks are used.

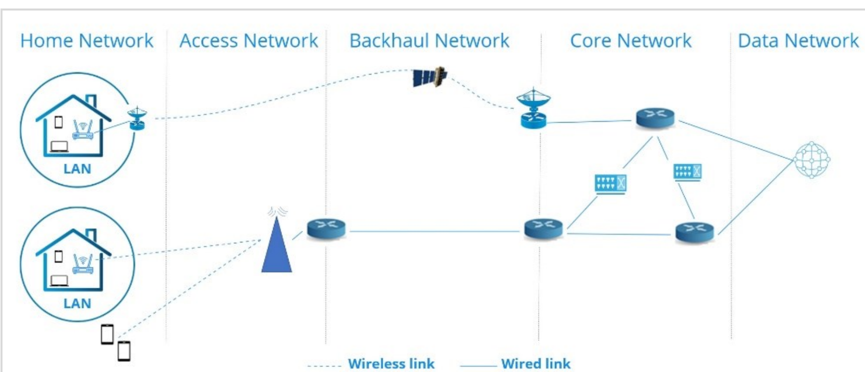


Figure 1-17 Fixed wireless access technology using cellular networks (shown in the bottom part of the figure) and satellite communication (shown in the upper part of the figure) for backhauling. This figure shows the impacted network element of Figure 1-1 namley the access network and the backhaul network.

Satellite

Satellite communications cover broadcast satellites which are unidirectional satellite for television use and broadband satellites which are bidirectional satellites for Internet access. In this section we describe the latter one. Broadband satellite technology uses communication satellites to provide internet access to remote areas that are not served by traditional wired broadband networks. The orbit is the trajectory of the satellite. There are three types of satellite orbits used for broadband internet services: Geostationary Earth Orbit (GEO), Medium Earth Orbit (MEO), and Low Earth Orbit (LEO) [41]. Choosing a satellite system involves selecting an appropriate orbit that meets the key performance indicators (KPIs) for the intended services.

- Geostationary Earth Orbit (GEO): Satellites in GEO are positioned approximately 35,800 km above the equator, allowing them to stay in a fixed position relative to the Earth's surface. This means that a user can maintain a constant connection with a satellite in GEO, as the satellite appears to remain stationary in the sky. However, the distance to the satellite creates a delay in data transmission known as latency, which can impact some applications, such as online gaming or video conferencing. GEO satellites are often used for broadcasting and delivering broadband services to remote regions [42].
- Medium Earth Orbit (MEO): Satellites in MEO are positioned at an altitude of around 5,000 to 12,000 km above the Earth's surface. They are closer to the Earth than GEO satellites, which reduces latency and allows for faster data transfer rates. However, since MEO satellites move relative to the Earth's surface, they require a larger number of satellites in a constellation to ensure continuous coverage. MEO satellite constellations are often used for providing broadband services to airplanes, ships, and land-based users [42].
- Low Earth Orbit (LEO): Satellites in LEO are positioned at an altitude of approximately 500 to 2,000 km above the Earth's surface. They are much closer to the Earth than GEO or MEO satellites, which minimizes latency and allows for very high data transfer rates. LEO satellite constellations typically consist of hundreds or even thousands of satellites, which can provide continuous coverage and enable seamless handoffs between satellites as they orbit the Earth. LEO satellite constellations are being used to deliver broadband internet services to consumers and businesses, and some companies are planning to deploy global broadband networks using LEO satellite constellations [42] and [43].

Satellite broadband works by transmitting data from the user's premises to a satellite using a satellite terminal. The satellite then relays the data to a ground station, which is connected to the internet backbone. The data is then sent back to the satellite and transmitted back to the user's premises. Satellite can also be used only for backhauling other broadband access technologies as shown in *Figure 1-18*, hence, satellite impacts two parts of the network architecture presented in Figure 1-1, namely the backhaul and the access network.

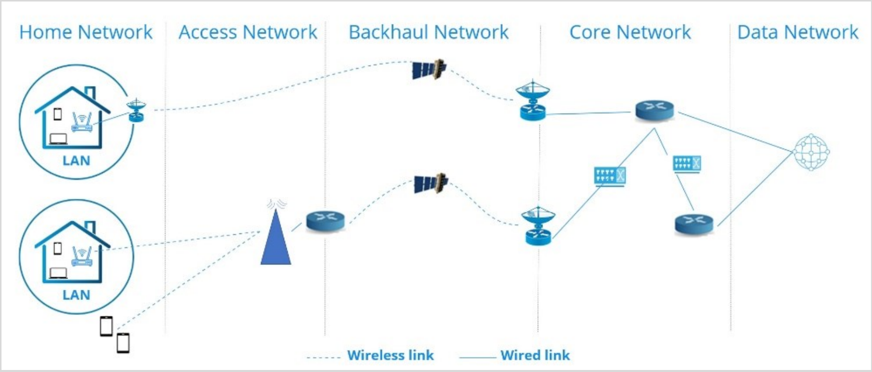


Figure 1-18 Satellite communication as backhaul, and access network technology (shown in the upper part of the figure) and satellite communication used for backhauling other access technology (shown in the bottom part of the figure). This figure shows the impacted network element of Figure 1-1 namely the access network and the backhaul network.

1.2 5G networks

In the previous section, we provided a high-level introduction to 5G networks, placing them in the context of the evolution of cellular networks through the different generations. However, as this dissertation aims to study 5G networks from a techno-economic perspective, it is important to delve deeper into the details of this latest generation of mobile networks. This section will provide a more comprehensive overview of 5G networks, including the famous 5G use cases, the spectrum that will be used, the 5G network architecture, and the enabling technologies that make 5G possible

1.2.1 5G use cases

Based on the IMT vision 2020 and beyond that defined the basic guidelines for future broadband networks, 3GPP has adopted the following three main use cases as the main use cases of 5G networks [44]:
The IMT vision 2020 laid out fundamental guidelines for future broadband networks and has identified three primary use cases. Following this vision, 3GPP has adopted the same three use cases for 5G networks. They are as follows:
Enhanced Mobile Broadband (eMBB)
The demand for mobile broadband is expected to increase, leading to the development of Enhanced Mobile Broadband. This new usage scenario will require additional features and application areas to ensure better performance and a more seamless user experience. The usage scenario encompasses various cases, such as wide-area coverage and hotspot, each with unique requirements. For example, the hotspot scenario requires extremely high traffic capacity to support a dense user population, while mobility needs are minimal, and the user data rate surpasses that of wide area coverage. In contrast, for wide area coverage, users

seek continuous coverage, medium to high mobility, and a significantly enhanced user data rate compared to existing rates. However, the data rate requirement might be less stringent than in the hotspot case [44]. eMBB will enable a range of applications such as virtual/ augmented reality, and ultra-high-definition video streaming.

Ultra-Reliable and Low Latency Communications (URLLC)

This use case is designed for applications that require highly reliable and low latency connectivity such as autonomous vehicles, remote surgery, and industrial automation. With URLLC, 5G networks can achieve latencies as low as 1 millisecond and provide the reliability of up to 99.999% [45].

Massive Machine Type Communications (mMTC)

This use case is aimed at supporting a substantial number of connected devices, such as sensors and IoT devices requiring both low data rates and power consumption. With mMTC, 5G networks can support up to 1 million devices per square km, which is a significant improvement compared to the 4G networks [45]. These three use cases are presented in Figure 1-19.

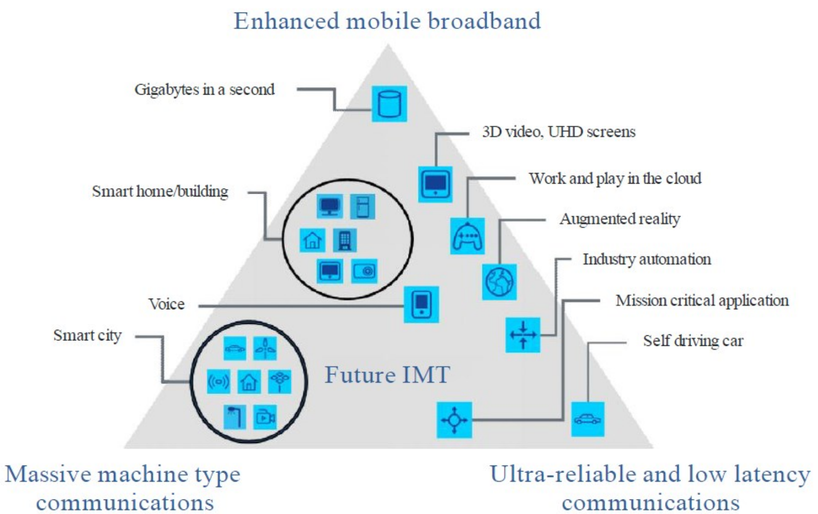


Figure 1-19 5G use cases according to the IMT 2020 vision [44].

The 5G use cases have different network requirements in terms of latency, peak data-rate, mobility, etc. The radar chart in Figure 1-20 illustrates the network requirement of the three 5G uses cases. As can be seen in the figure, eMBB requires mostly high data rate, high traffic capacity per area, and high mobility, whereas mMTC requires high connection density and medium network energy efficiency and URLLC is demanding in terms of mobility and latency.

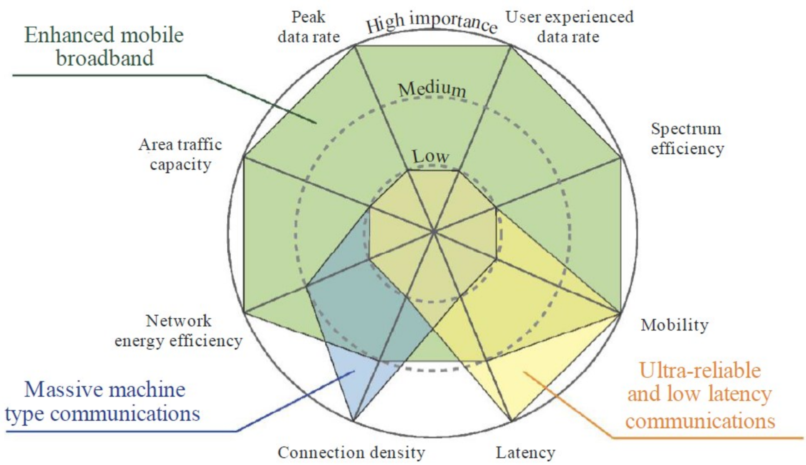


Figure 1-20 5G use cases requirements [44].

1.2.2 **5G spectrum**

In order to support the diverse use cases outlined in the previous section, it is necessary for 5G networks to access a new spectrum, as the current spectrum allocated to mobile communication networks is insufficient to meet the requirements of the identified use cases [46]. As a result, 5G networks will utilize a range of frequency bands, including low-band, mid-band, and high-band frequencies [47] and [48]. These three bands are detailed in the following section and presented in Figure 1-21.

Low-band frequencies (sub-1 GHz) have better coverage and can penetrate walls and other obstacles, but they have limited capacity and slower data speeds. They will be used primarily for wide-area coverage and indoor use.

Mid-band frequencies (1 GHz to 6 GHz), a.k.a sub-6 GHz, provide a balance between coverage and capacity, making them ideal for providing both wide-area coverage and high-speed data transfer. They will be used for the 5G networks' backbone and to deliver high-speed data services in urban and suburban areas.

High-band frequencies, also referred to as millimeter waves (mmWave), these waves offer high capacity and speed, although their coverage area is limited and can be obstructed by obstacles such as buildings and trees. They will be used for providing high-speed data services in densely populated urban areas, stadiums, and other venues with high data usage.

In addition to these three frequency bands, 5G networks may also use unlicensed spectrum and shared spectrum allowing several users to use the same frequency band. Overall, the use of different spectrum types and bands is essential for the success of 5G networks, as it allows for the optimal balance between coverage, capacity, and speed.

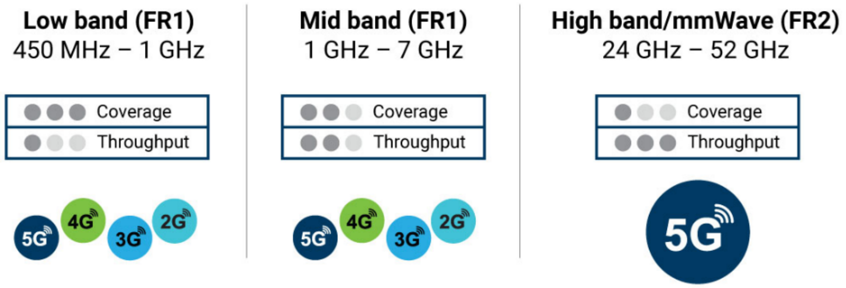


Figure 1-21 5G frequency bands compared to the previous cellular generations, [49].

It is worth emphasizing that the frequency bands utilized by older mobile generations have been incorporated and are still being used in the new generation, as can be seen in Figure 1-21.

1.2.3 5G network architecture

The main objective of 5G communication networks is to provide services that cater to the various demands of highly mobile systems, including high throughput, low latency, and a large number of connections as described under the framework of the three 5G use cases (section 1.2.1). Additionally, the goal is to achieve a completely connected society while ensuring human-centric coexistence with machine-type applications. The acquisition of new spectrum bands (as described in section 1.2.2) alone is not sufficient to meet these diverse requirements. Therefore, the 5G framework needs to employ a dynamic, logical, and flexible architecture supported by various advanced technologies that can support different applications [50].

In contrast to earlier mobile communication networks, 5G embraces a smarter architecture that is not constrained by the proximity of base stations or complex infrastructure limitations. It employs a versatile deployment approach, integrating innovative technologies such as network slicing, software-defined networking, and network function virtualization [50]. This approach has led to a comprehensive reassessment of the entire architectural system encompassing both the Radio Access Network (RAN) and the Core Network, as well as the functional division between them, in accordance with the 3GPP New Radio (NR) standards [51].

3GPP defined the 5G system architecture in [55], outlining the necessary features for an operational 5G system. Usually, a mobile system architecture comprises two primary components: the Core Network (CN) and the Access Network (AN). The 3GPP's 5G system architecture entails interaction between the user equipment (UE) and an endpoint, which could be another UE or a server like the Application Server (AS). This 3GPP system includes the AS, 5G Core Network (5GC), Next Generation Radio Access Network (NG-RAN), and UE. It facilitates communication between the data network and UE through the AN and CN [50].

Figure 1-22 5G system end-to-end architecture based on [50].

In the following, we discuss the main technical advancements in the two main 5G architecture elements namely the NG-RAN or a.k.a the 5G new radio (5G-NR) and 5GC.

NG-RAN or 5G-NR

The 3GPP New Radio (NR) is a versatile and unified air interface that enables the three primary categories for 5G communications as described previously in section 1.2.1, namely, eMBB, URLLC, and mMTC among other 5G vertical applications. It operates within frequency bands of up to 52.6 GHz by enabling network deployment across frequency spectrums below 6 GHz, covering mmWave and centimeter wave (cmWave) frequency bands [52]. It supports massive MIMO on a large scale, boosting network coverage and capacity performance, thereby catering to the substantial demand for data services. The NR air interface is based on an OFDM-based signal modulation like LTE [50] and [52].

5GC

The mobile core network in general is a group of functions typically housed within one or multiple devices that serve multiple objectives. These functions include authenticating devices before connecting them to the network, delivering internet (IP) connectivity for voice and data services, ensuring that the promised quality of service (QoS) requirements are met, monitoring user mobility to ensure uninterrupted service, and keeping track of subscriber usage for billing and charging purposes. The 3GPP has maintained the concept of a flat architecture with a separation between the Control Plane (CP) and User Plane (UP) to meet service requirements. This allows for a scalable and adaptable network, as operators can utilize this functional split for dimensioning, deployment, and customization to fit their specific needs [53]. In addition, 3GPP identifies two representations to be used in the 5GC architecture [54], [55], and [53]:

1. The point-to-point reference presentation: this architecture has been used for decades in the previous mobile network generations. The reference point representation, which emphasizes the interactions between pairs of network functions through point-to-point reference points (e.g., N7) connecting any two network functions (e.g., SMF and PCF), is employed when there is a need to depict such interactions between functions (Figure 1-23).

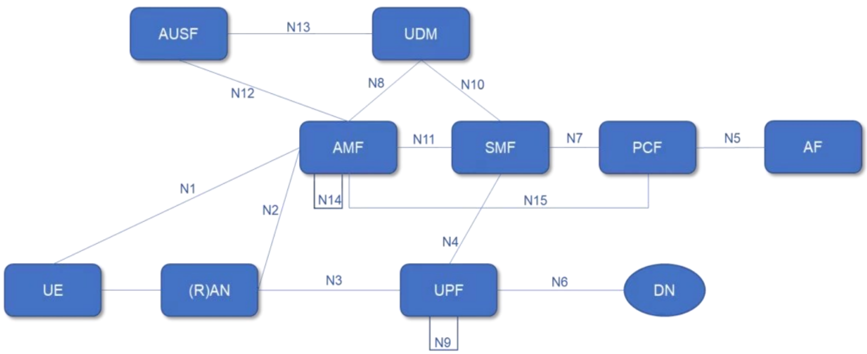


Figure 1-23 Non-Roaming 5G System Architecture in reference point representation [56].

2. Service-Based Architecture (SBA): service-based architecture, on the other hand, is a more flexible and modular approach to network design. In this architecture, the network functions are divided into smaller, independent building blocks called services. Each service provides specific functionality that can be accessed and used by other services. This enables the creation of more complex services by combining smaller, independent services, without the need to modify or replace the entire network. This illustration also incorporates point-to-point reference points connecting NFs wherever needed (Figure 1-24).

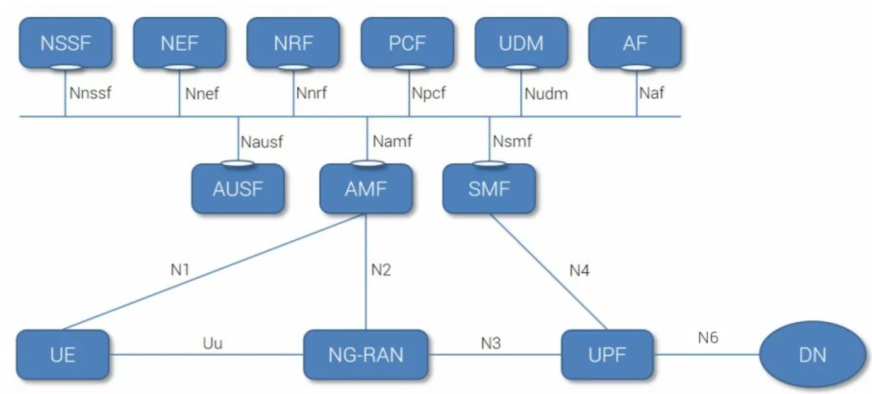


Figure 1-24 5G System Service-based architecture [57].

As can be seen in Figure 1-23 and Figure 1-24, each network function (NF) has a name and executes certain functionalities, in the following we delve into the various NFs in the 5G core network. It is important to note that not all functionalities of each function are described but the famous ones. The major 5G NFs are described in the following and the comparison with the 4G NFs is highlighted [57]:

“**Access and Mobility Management Function (AMF)** supports a range of functionalities such as Termination of Non-Access Stratum (NAS) signaling, NAS ciphering & integrity protection, registration management, connection management, mobility management, access authentication and authorization, and security context management. AMF also includes some of the MME functionality from the EPC world (i.e., the 4G/LTE core network).
Session Management Function (SMF) supports session management tasks such as session establishment, modification, release, UE IP address allocation & management, and traffic steering configuration for UPF for proper traffic routing. SMF also incorporates some of the MME and packet gateway (PGW) functionality from the EPC world.
User plane function (UPF) carries out functions such as packet routing & forwarding, packet inspection, and QoS handling, and acts as an external PDU session point for interconnecting to Data Network (DN). UPF also includes some of the serving gateway (SGW) and packet gateway (PGW) functionality from the EPC world.
Policy Control Function (PCF) provides a unified policy framework, policy rules for control plane functions, and access to subscription information for policy decisions. PCF also includes some of the PCRF functionality from the EPC world.
Authentication Server Function (AUSF) acts as an authentication server and incorporates some of the HSS functionality from the EPC world.
Unified Data Management (UDM) is responsible for generating Authentication and Key Agreement (AKA) credentials, user identification handling, access authorization, and subscription management. UDM also incorporates some of the HSS functionality from the EPC world.
Application Function (AF) facilitates application influence on traffic routing, accessing NEF, and interaction with a policy framework for policy control. AF has the same functionalities as AF in the EPC world.
Network Exposure Function (NEF) enables the exposure of capabilities and events, secures the provision of information from external applications to the 3GPP network, and translation of internal/external information. NEF is not present in the EPC world.
NF Repository function (NRF) facilitates the service discovery function and maintains the NF profile and available NF instances. NRF is not present in the EPC world.
Network Slice Selection Function (NSSF) helps in selecting the Network Slice instances to serve the UE, determining the allowed NSSAI, and determining the AMF set to be used to serve the UE. NSSF is not present in the EPC world.”

5G non-standalone and standalone architecture

After understanding the NG-RAN and the 5GC, the question raises is how these two 5G main elements are interconnected and deployed. In the following section, we address this question.

5G networks can be deployed based on two different options, namely the Non-Stand Alone (NSA) and Stand Alone (SA) architectures [58]. Figure 1-25 illustrates how the NSA architecture links the NG-RAN to both the 5GC and 4G (LTE) core network, while the SA architecture, depicted in Figure 1-26, solely connects the 5GC and NG-RAN.

The NSA architecture utilizes a combination of the NG-RAN with the current 4G access network (E-UTRAN) and its core network (EPC) to enable the use of 5G NR technology without the need for a network replacement. It is an intermediary step towards the full 5G network deployment [59]. However, while the 5G NR can provide benefits such as lower latency and higher capacity, only 4G services will be available at its capacity. The Next Generation NodeB (gNB), acting as the logical node for 5G NR base stations (BS), links to the 4G/LTE BS, known as the enhanced NodeB (eNB), via the X2 interface. The gNB is responsible for establishing connections between the User Equipment (UE) and the 5GC interface. The eNB is connected to the EPC, more specifically to the Mobility Management Entity/Serving Gateway components (MME/S-GW), through the S1 interface. On the other hand, the S1-U interface links the gNB to the EPC. Furthermore, the gNB can establish links with other gNBs through the X2-U interface [60]. The E-UTRAN obtains information about the UE from the MME and local E-UTRAN to determine whether to utilize dual connectivity. The 4G components of the NSA manage the control plane for early 5G networks. Thus, the QoS implementation for 5G SA is reliant on the heavy deployment of 4G infrastructure [61].

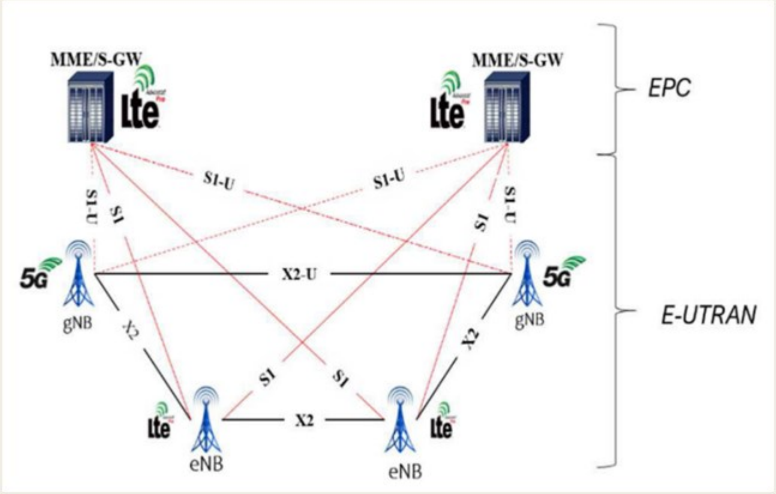


Figure 1-25 The 5G NSA architecture [50].

The SA architecture represents the complete 5G architecture where the NG-RAN provides both LTE and NR radio access. The eNB delivers control and user plane functions for LTE/E-UTRAN towards the UE, while the gNB provides the same functions for 5G NR. The Xn interface connects the gNBs and eNBs, and the NG interfaces connect them to the 5GC, which is linked to the AMF via the NG-C interface and the UP function through the NG-U interface [61].

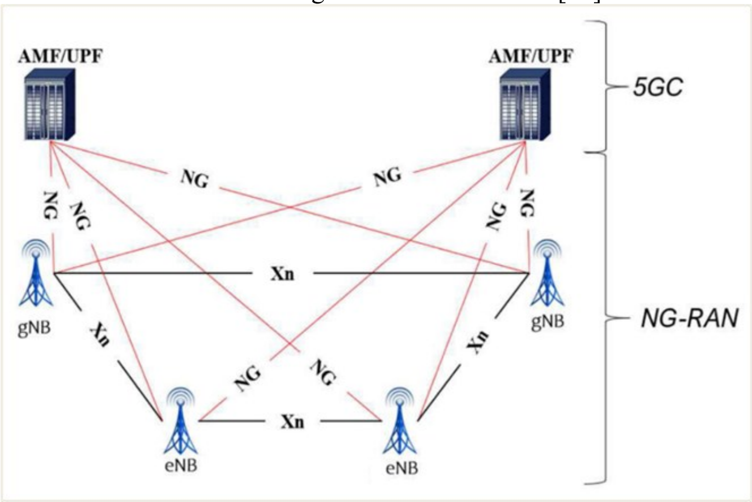


Figure 1-26 The 5G SA architecture [50].

1.2.4 **5G key technologies**

In this section, we will cover several key technologies that are integral to 5G networks, including SDN, NFV, network slicing, MEC, HetNet, Massive MIMO, and mmWave. For further information on additional 5G technologies, we recommend referring to [62].

SDN, NFV, and network slicing

Traditional network architectures are often inflexible, complex, and costly to build, maintain, and operate. In contrast, emerging concepts such as Network Function Virtualization (NFV) and Software Defined Networking (SDN) offer more dynamic, programmatically efficient network configuration, while network slicing allows for the creation of logical end-to-end networks that can be dynamically created to meet the specific requirements of different services [63].

SDN is an approach to network architecture that separates the control plane from the data plane, allowing for centralized control of the network through an SDN controller [64]. This enables dynamic and programmatically efficient network configuration, making it easier to troubleshoot and maintain the network. In contrast, **NFV** involves moving network functionalities from specialized hardware to software, allowing them to run on standard-based hardware and be easily moved to any location within the network. This approach reduces the need to add new equipment and enables more flexibility to scale up or down, experiment, and deploy new services with lower risk [63].

When combined, SDN and NFV are considered the key enablers of the network slicing concept. **Network slicing** refers to the creation of logical end-to-end networks that can be dynamically created to meet the specific requirements of different services. According to 3GPP [65]: "A network slice is viewed as a logical end-to-end network that can be dynamically created. A given User Equipment (UE) may access to multiple slices over the same Access Network (e.g. over the same radio interface). Each slice may serve a particular service type with an agreed upon Service-level Agreement (SLA)". Network slicing allows different service providers with disparate traffic requirements to share the same infrastructure resources. This does not mean deploying Virtual Local Area Networks (VLANs) to isolate traffic flows, but it means splitting physical network resources and network functions to deploy an exclusive end-to-end (E2E) implementation for each application, e.g. eMBB, URLLC. Network slicing allows Mobile Network Operators (MNOs) to open their virtualized networks to vertical segments such as the healthcare sector, car manufacturers, intelligent transport providers, and service providers that lack physical network infrastructure.

These technologies provide the flexibility to tailor services to the specific needs of customers, resulting in more efficient use of resources and a better overall customer experience [63].

A network slice refers to a chain of customized logical network functions (NFs) that are designed to support specific use cases or business models' requirements. Each slice is designed to support a set of KPIs such as capacity, latency, reliability, etc. Figure 1-27 presents two examples of network slices.

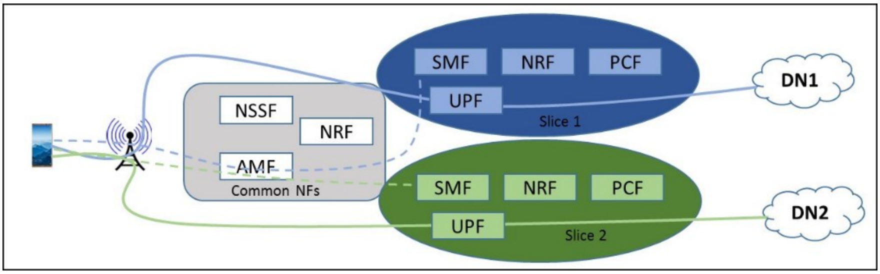


Figure 1-27 Examples of network slices [65].

The diagram presented in Figure 1-27 depicts the sharing of control plane network functions (NFs) among various network slices. Notably, certain NFs, namely NSSF, NRF, and AMF, are shared among slices, serving distinct purposes such as selecting network slice instances for the UE, maintaining the NF profile and available NF instances, and handling UE mobility, respectively. Conversely, NFs like SMF, PCF, and UPF are dedicated to specific slices. It is worth noting that NRF can serve as both a shared and dedicated NF.

Multi-access Edge Computing

Multi-access Edge Computing (MEC) technology is gaining more and more recognition as an important factor in enabling advanced applications that require low latency and high quality. MEC is considered crucial, alongside other 5G technologies, for the success of applications such as Vehicle to Everything (V2X) and Augmented Reality (AR)/Virtual Reality (VR) games as outlined in ETSI GS MEC [67]. MEC helps reduce latency by hosting services close to users and data generators [68]. In 2014, MEC was introduced to provide cloud-computing capabilities to the RAN close to mobile subscribers [69]. The MEC Industry Specification Group (ISG) of the European Telecommunications Standards Institute (ETSI) developed the MEC architecture specification based on specific use cases [67] An example of MEC-based network architecture for different use cases is illustrated in Figure 1-28.

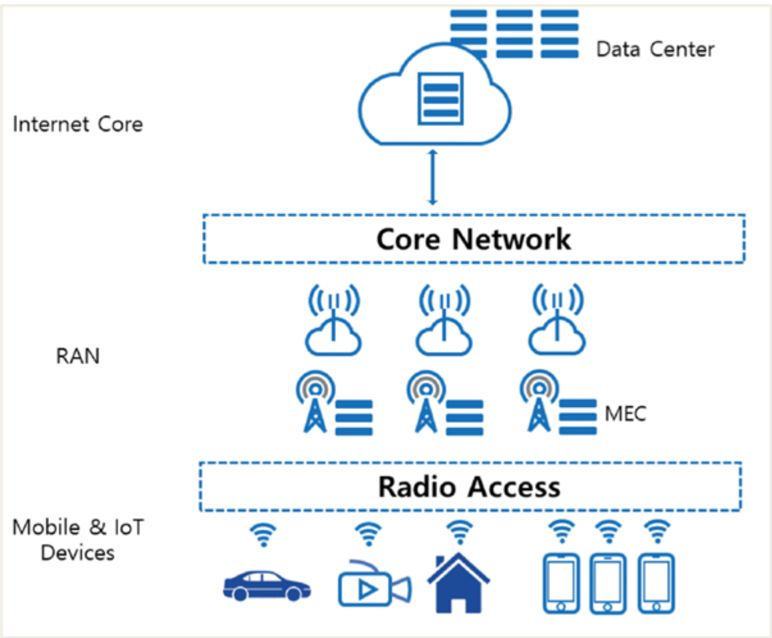


Figure 1-28 Generic MEC network architecture based on [70].

HetNet

One key technique to improve network capacity, coverage performance, and energy efficiency is network densification through the deployment of various cell types, such as macro-, micro-, pico-, and femtocells. This approach has been widely used in current wireless cellular networks, such as 3G and 4G LTE,

resulting in a multitier cellular HetNet (Heterogeneous Network) [71]. Additionally, wireless HetNets may include remote radio heads (RRHs) and wireless relays, further enhancing network performance. Unlike the current LTE system, where multi-hop communications were deemed supplementary, relaying and multi-hop communications are anticipated to be fundamental components of the 5G wireless architecture [71] and [72].
HetNet provides a more efficient and flexible wireless communication infrastructure. This allows for better coverage, higher capacity, and improved quality of service for users, particularly in areas with high traffic demand.

Massive MIMO

Massive MIMO is an advanced physical-layer technology for future wireless access, considered the most compelling technology currently available for sub-6 GHz frequencies. The technology involves using large antenna arrays at base stations to simultaneously serve multiple terminals with smart processing to exploit the unique propagation signatures of each terminal, resulting in superior capacity as illustrated in Figure 1-29. Massive MIMO offers two major benefits: First, it provides excellent spectral efficiency by multiplexing many terminals at the same time-frequency resource, which is possible due to the sufficiently different channels to different terminals in diverse propagation environments. Second, Massive MIMO achieves superior energy efficiency due to the array gain, allowing for a reduction of radiated power. Additionally, its ability to perform well with low-accuracy signals and linear processing can lead to considerable savings. Notably, Massive MIMO performs equally well in line-of-sight and rich scattering propagation environments [73].

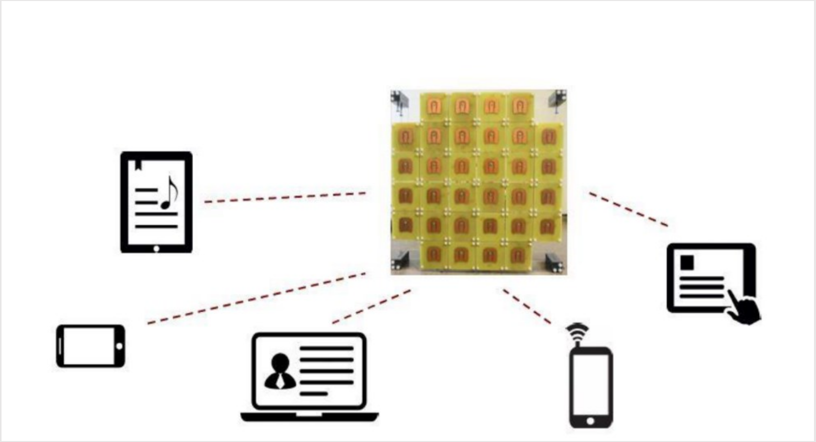


Figure 1-29 Massive MIMO uses large antenna arrays to spatially multiplex many terminals [73].

mmWave

The mmWave frequencies have promising applications, including high-resolution video streaming and outdoor point-to-point backhaul links. While mmWave technologies have already been standardized in IEEE 802.11ad for short-range services, their use in cellular applications is relatively uncommon, likely due to high propagation loss, penetration loss, and rain fading [74]. Nonetheless, with the expected proliferation of small cells such as pico- and femtocells in future 5G HetNets, short-range mmWave technologies could prove highly advantageous. As a result, mmWave frequencies are among the potential technologies that could satisfy the requirements of the 5G network [71].

1.3 5G Network rollout

In the preceding section, we delved into the realm of emerging 5G technologies and their potential to enhance network capabilities, while also highlighting their promising cost-saving attributes. In this subsequent section, our focus shifts towards a comprehensive exploration of the global 5G rollout landscape, examining the progress across three distinct continents. Through this lens, we will dissect the impact of policies embraced in these regions and their influence on shaping the trajectory of 5G development.
Reports published prior to 2020 [98], [99] and [99] extensively highlighted the fact that despite the widespread enthusiasm for 5G, the economic benefits in terms of GDP growth and employment resulting from new services and products will take time to materialize globally. It's becoming increasingly evident, especially in Asia, that achieving perfected 5G technology will necessitate a longer period, potentially spanning a decade [98]. This realization has been acknowledged by both China and Japan, with China Telecom noting the comparatively higher costs of 5G compared to preceding generations. Consequently, the complete rollout of 5G might proceed at a more deliberate pace than initially foreseen within the industry [99]. In fact, in the light of market saturation, revenues for Mobile Network Operators (MNOs) have experienced decline [100]. Yet, the implementation of 5G broadband in densely populated urban areas could rejuvenate their prospects. By amplifying smartphone activity, especially within online services like social networking, entertainment video streaming, and gaming, 5G could energize the Business-to-Consumer (B2C) market. Additionally, it could enable MNOs to generate new revenue streams through Internet of Things (IoT) applications and innovative use cases targeting industrial users in the Business-to-Business (B2B) market. However, commercially less attractive regions, such as rural and remote areas, pose challenges for MNOs to achieve a wide coverage of 5G [100]. Indeed, the uncertainty surrounding financial returns, encompassing technology and network capital costs, has contributed to reduced enthusiasm among operators. This hesitation to invest without a well-established business case and proven revenue potential further adds

to the complexity. Globally, MNOs are assessing the consequences of heightened demand for new infrastructure and its associated funding. Initial strategies include delaying investments in new 5G infrastructure or integrating existing core networks with upgraded 5G radio access networks to optimize backhaul capital investments. As a last resort, exploring alternative investment sources becomes essential, with the government emerging as a key entity with sufficient remaining capital. Consequently, the imperative of the 5G campaign is to convince governments that the substantial costs are justified by the substantial social and economic benefits. This endeavour draws support from a multitude of studies that project trillions of Euros in economic value and the creation of millions of jobs [98].

Yet, policymakers embarking on the journey of 5G development must anticipate and **tackle core rollout challenges** that could impede progress if left unattended according to [102].

A **key challenge** pertains to **the economic viability of the 5G business case** and its broader ecosystem. Beyond consumer entertainment streaming, use cases like IoT applications for smart cities and autonomous vehicles demand high levels of reliability, resilience, and security. These requirements should be meticulously integrated into the 5G network's design. Effective 5G systems design is crucial to overcoming this economic vulnerability [103]. This dissertation will thoroughly address this challenge across various chapters, notably in chapters 2, 3 and 5. Furthermore, it stands as the central research focus within the initial stages of this study.

Another **significant hurdle linked to 5G technology is its novelty**. Initial deployment and operations depend heavily on suppliers, given the unfamiliarity of the technology to the workforce. A substantial cost component revolves around workforce training, spanning network engineering, operations, and instruction for installers dealing with a dense concentration of small cells. For instance, China's 5G rollout encountered the challenge of establishing an extensive installer training program, demanding meticulous training materials [102].

In densely urbanized settings, a third major difficulty emerges – **the identification of sufficient outdoor sites with proper backhaul, and potentially, fronthaul connections**. Sites must adhere to Inter-Site Distance (ISD) requirements, based on propagation distance, for the specified density. They must also offer power supplies, grounding, and high-speed broadband connectivity to core networks and edge servers. Deploying 5G base stations may involve diverse mounting options like walls, roofs, or lampposts. However, ensuring the physical protection of local servers, securing roadside units, managing appearance concerns, and navigating complex permitting processes pose additional challenges [104].

The fourth challenge, though seemingly remote, **pertains to future cybersecurity**. 5G networks, due to their high data rates, become prime targets for aggressive cyberattacks and malware infiltration, especially for critical infrastructure applications. Although security was retroactively integrated into the standardization by 3GPP committees, the level of security in major deployments remains uncertain [105].

Furthermore, **there are concerns about the health effects associated with 5G-related RF radiation.** A portion of the population across several countries worries about these potential effects. The World Health Organization (WHO) has established a program to assess the health implications of exposure to electromagnetic fields encompassing 5G frequencies. Governments must address these concerns transparently and responsibly, taking into account public sentiment without ignoring protests [106].

However, recent reports from 2022 onward, such as [101] and [107], have demonstrated the positive impact of effective policies implemented by various countries to address some of the aforementioned challenges. These policies have not only encouraged MNOs in these nations to make significant strides in their 5G rollout but have also led to notable progress in 5G penetration. This advancement spans both the B2C sector, measured through the number of 5G subscribers, and the B2B market, quantified by the number of trials and projects focusing on industrial use cases.

Moving forward, the subsequent sections delve into detailing the current status of 5G deployment in specific countries. Additionally, we explore how the 5G promotion policies adopted in these nations have played a crucial role in fostering the successful implementation and adoption of 5G technology. Lastly, we draw valuable lessons from these experiences and offer recommendations that could be applied to other countries and the European context.

1.3.1 Global 5G advancements

The global adoption of 5G is rapidly increasing, as presented in Figure 1-30, with over 30% of regions launching 5G networks. Early-stage B2C 5G applications focus on extended reality (XR), high-definition (HD), and immersive experiences, particularly in entertainment through live streaming, virtual reality (VR), and augmented reality (AR). B2B 5G applications are undergoing extensive verification and demonstration, with implementations in sectors like industry, healthcare, and autonomous driving. However, scaling up these applications requires further investigation.

Major countries and regions are recognizing the economic and social potential of 5G and are actively creating favourable environments for 5G innovation. Governments are initiating strategic plans, projects, and policies to drive 5G application ecosystems, transform the economy, and promote vertical industry integration. Countries like China, South Korea, the U.S., and Europe have each devised strategic plans to accelerate 5G adoption. South Korea and the U.S. are focusing on strategic development and technical advancement respectively, while Europe is implementing policies and projects for 5G integration. China is actively promoting the convergence of 5G with advanced technologies to achieve its vision of being a worldwide leader in 5G and encouraging large-scale 5G deployment.

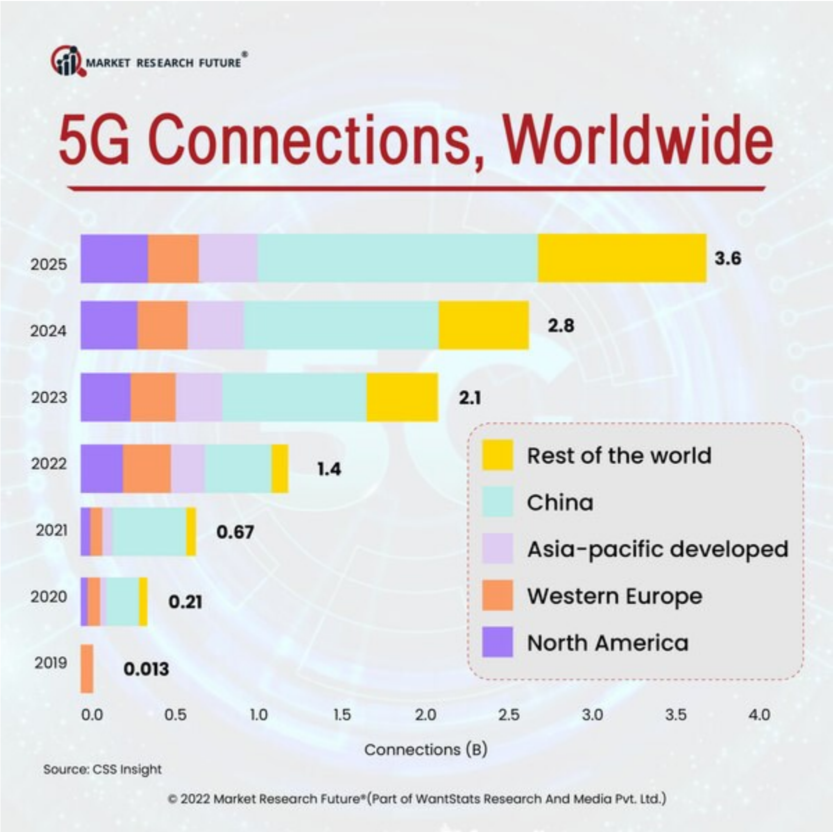


Figure 1-30 Worldwide 5G connections between 2019 and 2025, from [108]

In the following, we delve into the cases of China, the USA, and Europe, examining their unique contexts and approaches.

1.3.1.a 5G in China: current status and adopted policies

5G rollout status

China has been at the forefront of 5G commercial adoption since issuing 5G commercial licenses in June 2019. Over the past three years, 5G network coverage has expanded from cities to rural areas, , with over 1.037 million 5G base stations deployed by the end of August 2021 and with 70% of new phones being 5G capable and over 25% of users adopting 5G, as presented in Figure 1-30. This shift is reflected in the financial reports of telecom companies, demonstrating the economic benefits of 5G implementation []. Notably, China's **5G commercialization began with the business sector (B2B)**. After four years of development, **5G to business** has gained significant traction in various industries, forming a supportive framework for continued growth.

In fact, the economic benefits of 5G commercialization have become more evident, significantly boosting telecom operators' revenues. In the period from

January to August 2021, telecom services yielded a total revenue of CNY991.9 billion (around 125 billion Euro), a remarkable 8.4% year-on-year growth, marking the highest level since 2014.

Furthermore, the surge in 5G user penetration has positively impacted mobile data services, elevating the average revenue per user (ARPU) for mobile services. In the first half of 2021, China Mobile's 5G ARPU reached CNY88.9 (11 Euro), significantly surpassing its overall mobile ARPU of CNY52.2 (7 Euro).

China has achieved significant milestones in building high-quality and targeted 5G networks. The adoption of standalone (SA) networking is becoming predominant, offering advantages in supporting vertical industries and network capabilities like slicing and low latency. **Collaboration in network co-building and sharing** between telecom operators **is now the norm,** leading to substantial savings in construction costs and energy consumption. For instance, as of September 2021, China Telecom and China Unicom successfully **co-built and shared** over 500,000 5G base stations, resulting in substantial savings of up to CNY100 billion (around 13 billion Euro) in network construction costs. Moreover, this collaboration has led to an annual conservation of over 12 billion kWh of electricity

In conclusion, China's 5G development journey has been marked by significant achievements in both technological advancement and market leadership, reshaping the telecommunications landscape on a global scale. These achievements are the direct result of well-implemented and impactful policies [109].

Adopted policies

During the 14^{th} Five-Year Plan period (2021-2025), China's focus on building a modern socialist nation coincides with a pivotal phase for extensive 5G network deployment. China's ambition includes speeding up the rollout of 5G networks, aiming for a 56% user penetration rate within this period. This vision encompasses 10 key sectors for 5G for B2B applications, spanning areas like transportation, energy, manufacturing, agriculture, education, healthcare, and cultural tourism, with the aim of developing 5G-based use cases and industry ecosystems.

Comprehensive initiatives are in motion to implement the government's 5G development principles. The Ministry of Industry and Information Technology (MIIT) has placed priority on 5G development in its planning, manifesting in the "1+2+9" strategy. The MIIT's efforts extend to policy guidance, network construction, use case creation, and industrial growth, as outlined in documents such as the "Notice on Accelerating the Development of 5G," and the "Dual Gigabit Network Coordinated Development Action Plan (2021–2023) [110].

Further, the MIIT, in collaboration with other central government bodies, introduced the 5G Application "Set Sail" Action Plan (2021–2023), outlining 32 tasks across eight initiatives to guide 5G adoption in consumer, business, and public sectors. Fifteen key areas for B2B applications were identified, including healthcare, education, energy, transportation, and more. This plan promotes knowledge sharing and support mobilization through regular summaries and nationwide promotion. Local governments have also taken part in supporting 5G

development by issuing various policies and setting objectives aligned with their economic conditions. By August 2021, 569 5G policies had been released, indicating robust government support.
These different policies have focused on several aspects among which we list [110]:
Strategic Importance: Recognizing the fundamental importance of 5G to economic and social development.
Leadership and Investment: China's commitment to becoming a leader in 5G coverage, technology, and adoption, is backed by significant investments.
Economic Benefits: Focusing on the economic benefits of 5G commercialization, driving revenue growth for telecom operators.
Standalone Networking: Emphasizing standalone (SA) networking with a 5G core network, providing superior capabilities for vertical industries.
Network Sharing: Encouraging collaboration in network co-building and sharing among telecom operators to improve cost efficiency.
Policy Implementation: Effectively implementing policies through close coordination with both central government agencies and local authorities.
Public Engagement: Addressing public concerns about health effects and electromagnetic fields, promoting transparent and responsible communication.
These policies collectively underscore China's strategic approach to 5G development, which has demonstrated remarkable efficiency in fostering the rollout of 5G technology up to the present moment.

1.3.1.b 5G in the USA: current status and adopted policies

5G rollout status
In the early phases of its introduction in the United States, 5G networks exhibited extensive population coverage but a limited user adoption rate. As of June 2021, over 80% of the population had access to 5G networks; however, the rate of user penetration remained under 10%. During the initial stage, telecom operators primarily rolled out mmWave 5G networks due to the absence of mid-band spectrum. Yet, these high-frequency networks offered only constrained coverage and incurred substantial deployment expenses. Post-2019, a shift occurred with the integration of low-band spectrums to establish 5G networks, enabling broader coverage. This expansion was facilitated by the widespread implementation of spectrum-sharing technologies to offer nationwide 5G services. In March 2021, a significant milestone was reached when the Federal Communications Commission (FCC) concluded the auction of C-band (3.7–3.98 GHz) spectra. Following this, major telecom operators embarked on planning for the deployment of C-band networks, which will take precedence in their forthcoming 5G network development endeavours [101].

Adopted policies
The United States government has designated the advancement of the 5G industry as a paramount national priority. This commitment is evidenced through the formulation of strategic plans, facilitation of 5G technology research and

development, allocation of crucial 5G spectrums, and the reinforcement of robust cybersecurity measures.

In October 2018, the FCC introduced the 5G FAST Plan [101], a comprehensive strategy aiming to advance 5G network development and enhance the United States' position in this domain. The plan includes key components:

Spectrum Allocation: This involves auctioning mmWave spectra, assigning mid-band spectra, optimizing low-band spectra utilization, and fostering new Wi-Fi opportunities in license-free frequency bands.

Infrastructure Policy Update: A focus on easing 5G development barriers at all levels, particularly through policies addressing small cells, commonly used in 5G networks. These new policies aim to expedite small cell site approval processes.

Regulation Amendments: Changes include eliminating network neutrality policies, streamlining approval for new network device connections to existing poles, encouraging 5G backhaul network deployment with reduced costs, promoting investment in advanced networks and services, encouraging optical fiber network investment by lifting regulations on enterprise data service tariffs, and safeguarding the network supply chain by restricting purchases from entities posing national security risks [101].

The profitability of deploying 5G networks in densely populated urban areas is very likely, **while the economic feasibility of such deployment in rural regions is challenging**. **To bridge the digital gap and extend economic prospects to rural areas, the FCC established the 5G Fund for Rural America**. This initiative will allocate $9 billion over the next decade in two phases to introduce 5G wireless broadband services to rural regions. The initial phase, spanning ten years, **involves $8 billion for 5G deployment in areas where telecom operators lack incentives for independent deployment.** The subsequent phase, with at least $1 billion and any leftover funds from the first phase, **will support 5G deployment for precision agriculture initiatives** [101].

1.3.1.c 5G in Europe: current status and adopted policies

5G rollout status

Since its inaugural commercial launch in 2019, 5G has swiftly gained traction across Europe, with the majority of countries introducing commercial services. By June 2021, 81 operators in 30 European nations had initiated commercial 5G offerings [101]. Notably, prominent national mobile network operators like those in the UK, Germany, France, and Italy have established operational 5G networks. Initially, the mid-band spectrum was predominantly employed for network deployment. However, from the latter part of 2020, some European operators embraced low-band frequencies and dynamic spectrum sharing (DSS) technologies, augmenting network accessibility. For instance, Deutsche Telekom harnessed DSS to cover 80% of Germany's populace. High-band 5G networks have proven valuable in addressing escalating demand for high-speed data. Spain and Finland are employing mmWave networks for targeted coverage in hotspot areas, enabling specialized 5G applications. By June 2021, according to the Global

System for Mobile Communications Association (GSMA), 5G networks enveloped 28.51% of Europe's population and 51.48% of Western Europe [107].

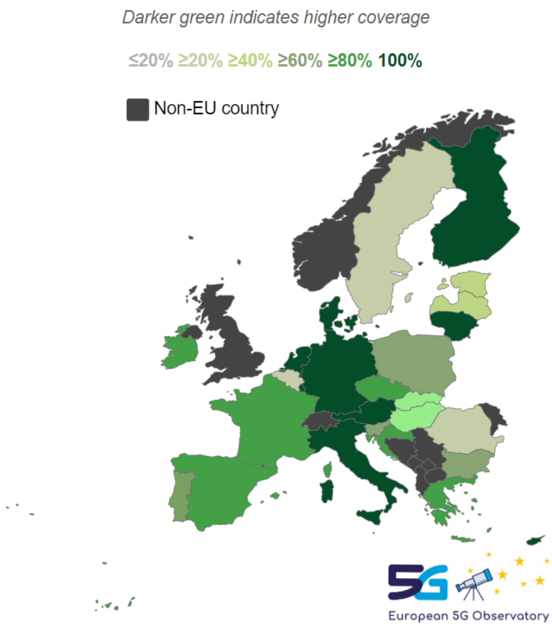


Figure 1-31 5G coverage Map across Europe in 2022, source [111]

In Figure 1-31, the 5G coverage of the populated areas is presented. According to the 5G observatory, the source of this figure, the data was last updated in July 2022, serving as a broad reference point without implying specific quality of service metrics. Encompassing all 5G coverage, including that facilitated by Dynamic Spectrum Sharing (DSS), this information pertains to locations covered by at least one operator. According to the same source [111], and based on this coverage map, 81% of the populated areas in Europe are covered by 5G networks. And here, raise several questions, what about less populated areas, like rural areas and low-traffic corridors? And what policies were put in place by the EU to foster the deployment and adoption of 5G?

Adopted policies

Elevating competitiveness in the digital economy and establishing a single digital market stand as paramount objectives within the European Union (EU). Positioned as a cornerstone for digital economy advancement and pivotal to the creation of a cohesive digital market, 5G has gained significant attention. Since 2012, the EU has been promoting 5G network deployment, commercial adoption, and innovative industry applications. This has been achieved through strategic initiatives such as crafting a comprehensive 5G development roadmap, fostering research collaboration among member states, initiating government-led research projects, and other concerted efforts.

In 2016, the EU released the 5G Action Plan, designed to foster investment in 5G networks, catalyze innovative ecosystems, and enhance Europe's collective 5G competitiveness. This plan furnishes EU member states with directives for formulating 5G development strategies and roadmaps. Its objectives encompass coordinating network deployment, spectrum allocation, cross-border service continuity, and standardization across member states, engendering a pivotal mass for 5G innovation within the EU's single market [112]. The EU's 5G ambitions target initial network rollout by 2018, widespread commercial adoption by the end of 2020, and comprehensive coverage for urban areas and major transportation routes by 2025. In addition to the 5G Action Plan, the European Commission (EC) adopted a new strategy towards its broadband program called "Connectivity for a European Gigabit Society 2025" in September 2016 [113]. The main targets (for 2025) announced in this agenda were:

1 Provide gigabit connectivity for all of the main socio-economic drivers.

2 Guarantee uninterrupted 5G coverage for all urban areas and major terrestrial transport paths.

3 Access to connectivity offering at least 100 Mbps for all European households.

In December 2018, the EU formally introduced the European Electronic Communications Code (EECC), establishing the foundational framework for the existing regulatory model [114]. The EECC outlined objectives including ensuring the availability of pioneering 5G spectrum in the EU by the close of 2020 and establishing a streamlined authorization structure for small-area wireless access points (SAWAPs). By December 21, 2020, the EU mandated its member states to integrate the EECC into national legislation, aiming to bolster the advancement of 5G networks across EU nations.

On March 9, 2021, the European Commission underscored in its publication "2030 Digital Compass: The European Way for the Digital Decade" that 5G plays a pivotal role in establishing secure, high-performance, and sustainable digital infrastructure [115]. The Digital Decade, as presented in Figure 1-32, outlines ambitions for universal Gigabit network coverage for all European households by 2030, complete 5G coverage across populated European areas, and the deployment of at least 10,000 climate-neutral and highly secure edge cloud nodes.

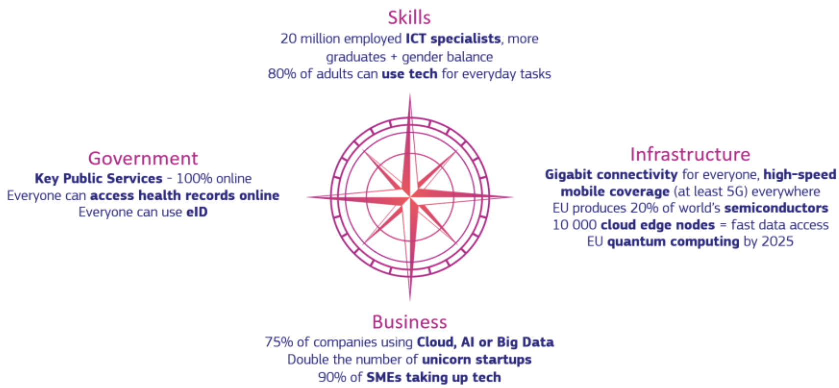


Figure 1-32 The Digital Decade Compass: objectives and targets, source [116]

The EU has prioritized 5G as a catalyst for industrial transformation. Starting in 2016, the EU has introduced policies and initiatives to engage European stakeholders in 5G trials and adoption across industries. These efforts span various domains, from studying 5G's viability in key sectors to advancing breakthroughs in hardware and core technologies. The 5G Infrastructure Public-Private Partnership (5G PPP) projects have progressed through phases, focusing on integrating 5G into vertical industries and establishing test platforms. The EU has invested over EUR400 million, fostering private investment of over EUR1 billion, to solidify its global standing in 5G applications. European countries like the UK and Germany have also launched initiatives to support 5G adoption in industries, such as the UK's 5G Testbeds and Trials Programme and Germany's focus on Industry 4.0. Collaborative efforts across operators, industries, and research institutions have resulted in extensive testing and innovation in areas like autonomous driving, healthcare, and more. Clusters and platforms have formed to encourage replicable, commercialized 5G applications across various sectors within different countries [101].

1.3.1.d Comparison of the studied policies

European policies are centred around fostering 5G deployment through research projects that showcase its integration into industrial use cases. These efforts also explore the feasibility of economically implementing 5G in challenging regions like rural and cross-border areas—areas that often hold less appeal for Mobile Network Operators (MNOs), as previously discussed[3]. In contrast, the American approach includes direct allocation of deployment funds to rural areas, exemplified by the "5G Fund for Rural America" ($10 billion). Furthermore, a significant aspect of the American strategy involves the removal of regulations

[3] The research presented in this dissertation is conducted within the context of European funded projects aimed at tackling the challenges of 5G rollout in areas that are less attractive to MNOs.

that hinder 5G rollout progress, such as Net neutrality, while also simplifying barriers to 5G development across all levels.

Conversely, China's approach has centred on facilitating network sharing among MNOs and adopting the 5G standalone network early in the deployment phase, with a key emphasis on B2B as the primary revenue generator for 5G deployment.

Considering that this research study was developed within the context of European-funded research projects like SAT5G, 5G-CARMEN, and 5G-Blueprint [117], the insights gained and recommendations provided to policymakers in section 8.2.1 hold relevance not only within the European context but also for other nations embracing analogous policies. These insights and recommendations stem from a foundation of best practices observed in the policies of various countries, as elucidated in the preceding section. They have been juxtaposed against the prevailing European landscape, highlighting opportunities for enhancement.

1.4 Research challenges

As highlighted in the previous section, the need for fast and reliable mobile broadband connectivity is critical for both Business to Customer (B2C) and Business to Business (B2B) markets. 5G networks have the potential to meet the essential KPIs required for providing connectivity-based services for these markets. From a user perspective, the primary drivers for 5G networks are high-speed connections, anywhere, anytime, and to anything. However, the deployment of 5G infrastructure, using the previously discussed key enabler technologies such as mmWave and massive MIMO, can be expensive. On the other hand, technologies like NFV, network slicing, and beamforming can help reduce rollout costs. Hence, it is crucial to investigate the economic considerations of deploying the 5G network, including identifying the factors driving the high deployment cost, determining which cost-saving technologies to utilize, and selecting the appropriate technologies to meet service requirements while minimizing deployment costs. As highlighted in section 1.1 and in section 1.3, understanding the economic implications of the 5G network rollout is of utmost importance. Particularly in challenging environments such as remote, rural, and cross-border regions, there exist obstacles that impede the rapid realization of investment benefits, as discussed in section 1.3. These hindrances arise from the significant costs involved and the sparsely distributed populace across these areas. Acknowledging these complexities, we have intentionally selected these regions as the primary focus of our study. The goal is to assess the economic viability of deploying 5G, illuminating a path forward amidst these unique challenges to inform concerned stakeholders such as MNOs and policymakers.

Therefore, this research aims to address the challenges and opportunities associated with the 5G network rollout from a techno-economic perspective in the selected use cases.

Techno-economic analysis (TEA) is a field that combines engineering and economic principles to evaluate the feasibility, sustainability, and profitability of technological solutions. In today's society, where sustainability and economic efficiency are increasingly emphasized, TEA provides valuable insights that can help policymakers and investors make informed decisions.

TEA can help to answer critical questions such as: What are the costs associated with implementing new technologies, and how can we optimize the design and operation of these systems to maximize their efficiency and profitability? What impact will the implementation of these technologies have on end-users in specific fields? What are the potential economic benefits and costs associated with adopting these new technologies, and how can we ensure that their deployment is both sustainable and viable? Moreover, TEA can be applied to identify emerging business models, especially as new technologies disrupt traditional industries. Section 1.5 of this book describes the techno-economic models used to accomplish this goal.

Therefore, the research objective of this dissertation is to provide a comprehensive techno-economic evaluation of the 5G network rollout and its implications in terms of network deployment strategies, the feasibility of the network deployment, and the viability of potential business models and we explore the interconnections between 5G technologies and these factors in detail to shed light on the opportunities and challenges that arise from these technological advancements. This overarching research objective can be achieved by tackling five research questions, which we will explore in greater detail in the following subsections.

1.4.1 RQ1: Feasibility of 5G deployment for different types of regions and services

The deployment of broadband networks, particularly 5G, relies on two significant factors: the characteristics of the targeted region and the specific services or Key Performance Indicators (KPIs) required. The characteristics of a region include its geography, population density, and household distribution, among others, which can classify regions as remote, rural, urban, or dense urban. Service KPIs refer to requirements such as the high capacity for broadband services like video streaming or ultra-reliable, low-latency internet access for connected mobility services.

For example, deploying terrestrial broadband networks in rural and remote areas is a challenging and costly endeavour due to geographical constraints and low population density, particularly for the backhaul network, as explained in section 1.3. In such situations, satellite communication can serve as an alternative solution to backhaul 5G networks, enabling coverage in hard-to-reach locations at reasonable costs. However, it's important to thoroughly evaluate the economic viability of such an integrated solution. Yet, there is a lack of techno-economic analysis of integrated 5G and satellite solutions in the literature according to a survey of 150 techno-economic analysis papers that assess 5G networks [75]. Hence, it is crucial to address this gap. In contrast, providing reliable and low-latency internet access for services such as connected mobility is a challenge in urban and densely populated areas. Here, 5G technology shows promise as a

potential solution. Despite this potential, the literature has not extensively addressed the economic feasibility of deploying 5G-based connected mobility services [75].
To address these gaps, this research question focuses on both factors, namely the type of region to be served and the services provided, in assessing the feasibility of the 5G rollout.

1.4.2 RQ2: Quantification of 5G technologies cost savings

The potential cost savings that 5G technologies, such as network virtualization, content caching using MEC, and network slicing, offer in deployment have been widely discussed [76]. These technologies are designed to help reduce capital and operational expenses associated with deploying, maintaining, and operating a 5G network, as explained in section 1.2.4. However, many papers in the literature, including white papers and reports issued by standardization bodies [76], fail to provide quantitative analysis to support their claims regarding the potential cost savings of various technologies, except for SDN technology, which has been thoroughly addressed in [77]. Therefore, there is a need for a comprehensive and quantitative assessment of the potential cost savings of 5G key enabler technologies.
Hence, this research question aims to quantify the cost savings that can be achieved through the adoption of these technologies and to identify which technology should be used under specific conditions to provide a clear understanding of the economic benefits of these technologies.

1.4.3 RQ3: Cost allocation of deployment cost to services and stakeholders

Once the deployment cost of a 5G network has been determined for a specific region with particular service KPIs, it is crucial to allocate this cost among the various services being offered. This allows service providers to understand which services are more expensive and demanding in terms of network resources. Furthermore, deployment costs must be shared between all stakeholders involved within specific business models. However, due to the novelty of technologies used in the 5G network and the complexity of its ecosystem, new allocation models are necessary. For instance, network slicing, as discussed in section 1.2.2, enables services to be separated from one another in a logical sub-network, customized according to their specific KPIs. Yet, a significant challenge that arises with network slicing is how to allocate the network's cost to the running network slices and services. Since the network has been dimensioned to support a particular set of services, each network slice or service must bear a proportionate share of the total network cost based on the resources it consumes. Properly allocating costs is essential for pricing the different end-to-end services according to their required resources.

Therefore, the focus of this research question is to identify new allocation models that address the unique aspects of 5G's emerging technologies while respecting the complexity of its ecosystem.

1.4.4 RQ4: Identification of potential collaboration and business models to save on the investment cost

The emergence of the 5G networks presents a unique opportunity for businesses to innovate and develop new business models. With its promising technologies like network slicing and edge computing, businesses can explore new applications and services that were previously impractical. 5G's faster data transfer rates, lower latency, and increased network capacity make it possible for businesses to unlock new potentials. However, the complexity of the 5G ecosystem poses a significant challenge in establishing sustainable and viable business models. With various stakeholders from diverse backgrounds involved, they must collaborate to leverage the full potential of 5G technology. For instance, network infrastructure sharing is a collaboration model enabled by the introduction of network virtualization [78]. In this model, mobile network operators (MNOs), in the case of 5G networks, collaborate in the network deployment by sharing some of their network infrastructure in passive and/or active mode [86]. Network sharing can offer significant cost savings and improve network coverage [79] and [80]. However, it also presents business challenges, including the loss of competitive advantage, revenue sharing, and the ceding of control and ownership. MNOs must coordinate and share their network infrastructure with their partners and comply with various regulatory requirements [80]. Consequently, successful network-sharing collaboration models require careful assessment of these challenges.

Therefore, this research question focuses on exploring and evaluating sustainable business and collaboration models that can help involved stakeholders to capitalize on the introduction of 5G technologies.

1.4.5 RQ5: Technology choices and timing decisions for optimized 5G deployment Strategies

The first research question aims to tackle the challenge of assessing the impact of various regions and services on the feasibility of deploying 5G technology. However, it's important to note that the choice of 5G technology can have a significant impact on deployment costs, particularly when there are alternative technologies available. For instance, when deploying a 5G network for B2B customers, it's crucial to determine whether it's more cost-effective to utilize a slice of the public 5G network or to deploy a private 5G network. Furthermore, in addition to technology choice, deployment timing also has a significant impact on the feasibility of the deployment. For instance, deciding when to deploy 5G networks for connected mobility given a specific level of adoption for connected cars is critical to ensuring its viability.

Therefore, this research question aims to identify the optimal deployment strategies based on the selected technology. To achieve this goal, several sub-research questions must be addressed, including what technology to adopt, how to combine certain technologies, when to adopt them, and for which use cases. The answers to these questions will provide valuable strategic recommendations for mobile network operators (MNOs), policymakers, and regulators. By following the identified recommendations, MNOs can optimize their network resources, reduce investment costs, and provide better user experiences. Additionally, regulators and policymakers can use these findings to make informed decisions about 5G rollout and investment policies, promoting the growth of the digital economy.

1.5 Research approach: modelling the techno-economic impact

The techno-economic analysis (TEA) is an interdisciplinary domain that draws on concepts from both engineering and economics to evaluate the feasibility, sustainability, and profitability of technological solutions. TEA can be used to address various research challenges related to technology, such as selecting the appropriate technology, dimensioning the deployment, making pricing decisions, and conducting cost allocation studies. By integrating knowledge from multiple fields, TEA provides a comprehensive approach to assessing the economic and technical feasibility of new technologies. This can help decision-makers and innovators in various industries make informed decisions and foster innovation.

Therefore, given the nature of the research challenges we want to address in this dissertation (discussed in section 1.4), and in light of the challenging nature of the selected use cases (further detailed in section 1.5.6), we adopted a research approach that relies on a generic techno-economic methodology. Our developed TEA methodology, presented in Figure 1-33, involves five main iterative steps. First, we identify the value network and business models. This step involves defining the different business roles and stakeholders involved in the 5G ecosystem of the specific case study, as well as the interactions between them. We consider various ways in which stakeholders interact, which can lead to different value network configurations and corresponding business models. Then, by analyzing the different business models together with the case study specifications, e.g., region to cover, service KPIs, possible technologies, etc. we identify different deployment scenarios that can be analyzed in the later steps of the methodology. Next, we conduct a viability study of the business case by developing a Total Cost of Ownership (TCO) model along with revenue models. The fourth step involves analyzing the impact of new technologies on the business case. Specifically, we investigate the effects of innovative technologies for generating potential cost savings and optimizing the network deployment scenario. Finally, we conduct sensitivity analysis to assess the degree of uncertainty linking the different models outputs to the inputs. This step is important for understanding the range of possible outcomes and identifying key drivers of the business case.

By following these five steps, we provide a structured approach for analyzing the potential of the 5G networks rollout and evaluating the corresponding business cases. This methodology can help stakeholders make informed decisions about whether and how to invest in 5G networks, taking into account the complexities of the value network and the potential impact of new technologies.

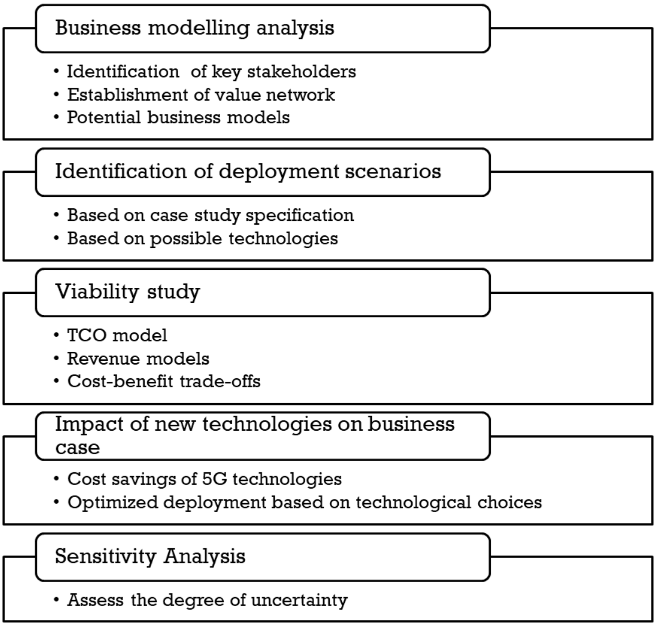


Figure 1-33 Techno-economic methodology

We developed a generic TEA framework that translates the adopted TEA methodology into interconnected TEA research models in order to generate the required results and insights. The developed TEA framework is presented in Figure 1-34, where the 5G case study specifications together with technological choices and insights from the business modelling analysis feed the scenario analysis model identification. The latter generates a list of possible deployment scenarios to be evaluated and compared using the rest of the TEA models presented in the diagram namely the economic viability study models and the cost allocation models. To ensure the accuracy and reliability of the results, sensitivity analysis is also conducted to refine the findings, assess the uncertainty of certain assumptions, and derive cost drivers for each deployment scenario.

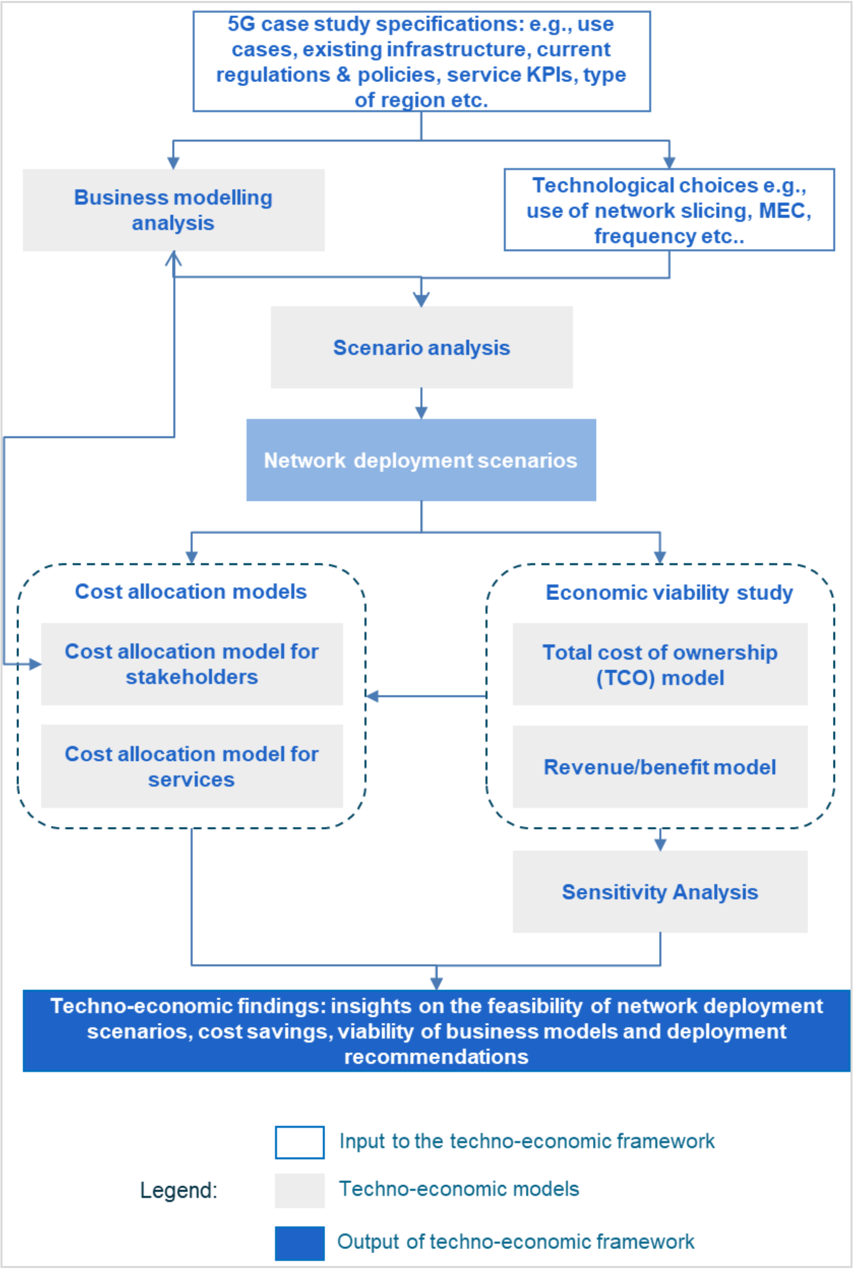


Figure 1-34 Generic techno-economic analysis framework

Throughout the analysis, a variety of techno-economic models have been employed, falling into two main categories: qualitative and quantitative models. For instance, within the realm of business uropa io analysis, the Value Network Analysis represents a qualitative approach. Conversely, quantitative models include the Total Cost of Ownership (TCO) model, the cost allocation model, and sensitivity analysis. In the following, we define each of the TEA models presented in the developed framework in Figure 1-34.

1.5.1 Business modelling analysis

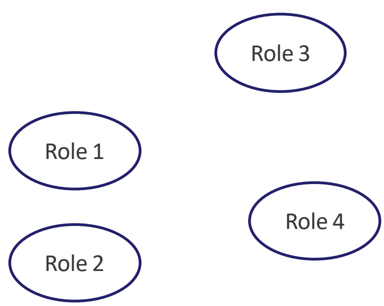
To effectively evaluate the viability of a 5G network roll-out, it is necessary to consider not only the technical aspects and challenges but also the disruptive impact it will have on current business practices, market dynamics, and stakeholders.

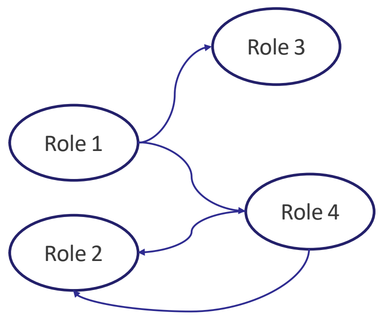
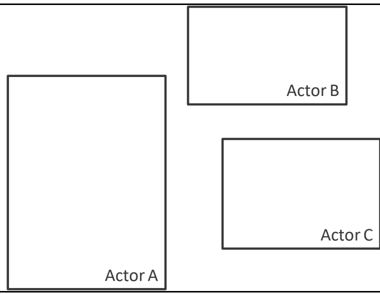
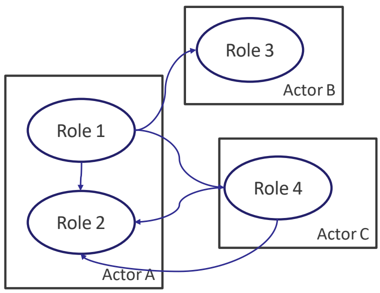
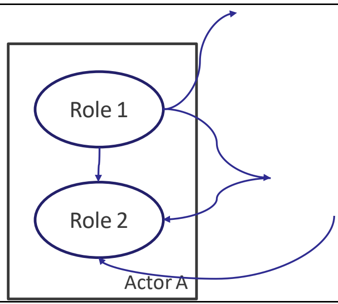
In this section, the focus will be on defining key business modelling terms, including stakeholders and business roles, as well as introducing the value network representation. This section will also explain the selected business modelling framework, which was chosen after a thorough comparison of commonly used frameworks. Finally, the business modelling methodology utilized in this work will be briefly outlined.

Definition of a stakeholder, roles, and the value networks

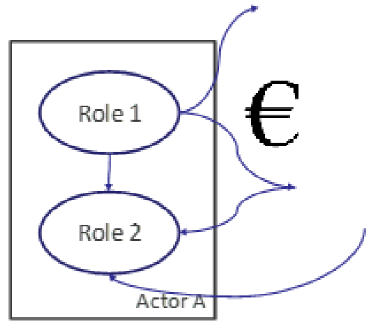
Given the varied definitions and interpretations of terms like value networks, business models, and business roles in the literature, the purpose of this section is to offer unambiguous definitions of these terms, along with a detailed explanation of the qualitative and quantitative analysis steps necessary to evaluate a particular process.

Table 1-2 Definition of a stakeholder, roles, and value networks [89]

Identification of the business roles. Business roles signify the specific tasks that need to be accomplished during the deployment and operation of a network, essentially defining "what is being done." These roles involve distinct responsibilities such as base station and tower deployment, equipment maintenance, and service provisioning. It is worth noting that the level of detail required for a given analysis may vary, but each business role should be viewed as an indivisible task.	

Value network A value network illustrates how different roles interact through the exchange of value streams, which can be either tangible or intangible. Tangibles refer to goods, services, or revenues that are exchanged through a contractual transaction, while intangibles relate to knowledge or benefits that support the core product but are not necessarily part of the contractual agreement. Each value stream represents a transfer of value, and a unique value network can be used to represent each process.	
Identification of business actors Business actors denote "by whom" the tasks are being carried out. This step involves identifying actors based on real-world examples.	
Value network configuration Actors may assume various roles, and as such, they can be assigned to these roles to indicate each actor's responsibilities in a given value network configuration. It is important to note that each process has only one value network, representing the interaction of roles, but there can be multiple value network configurations in which different actors assume different roles.	
Business model A business model illustrates the inbound and outbound value streams of a single actor within a particular value network configuration.	

Business case
A business case involves a quantitative assessment of a business model, combining incoming revenues with outgoing costs in an economic analysis that includes measures such as Net Present Value, Internal Rate of Return, and so on. It is worth noting that the business case for each actor will depend on the chosen value network configuration, given that the roles an actor takes on may vary.



In this dissertation, the term ecosystem has been also used to refer to a complex network of interconnected organizations, entities, and individuals that collaborate, interact, and influence each other within a particular industry or market [90]. In a business context, an ecosystem model considers not just a single company or organization in isolation but takes into account the broader context in which it operates. This broader context includes suppliers, customers, competitors, partners, regulators, and other stakeholders that impact and are impacted by the activities of the focal organization.

Overview of different value network options
After conducting a thorough literature review, various frameworks have been identified that represent how businesses generate added value and profit. Three of the most used frameworks are the Business Model Canvas (BMC) [91], Value Network Analysis (VNA) [92], and E3 Value Model (E3) [93]. While the main differences between these frameworks are reflected upon in this section, it is worth noting that BMC and VNA cannot be compared one-to-one as they operate on different layers. BMC is a tool that helps to understand how added value is created and how it leads to revenue and costs at an individual business actor level. However, for this work, the aim is to analyze how different types of stakeholders or actors will interact with each other to achieve the 5G roll-out. Therefore, a broader perspective is necessary, and Value Network Analysis (VNA) is the framework that will be used to identify these interactions between the business roles and the responsible actors.
The most significant advantage of Allee's approach, as compared to E3, is the ability to convert intangible assets such as human knowledge, internal structures, ways of working, reputation, and business relationships into negotiable forms of value. VNA provides a way to model, analyze, evaluate, and improve the capability of a business to convert both tangible and intangible assets into other forms of negotiable value, thereby realizing greater value for itself [94] .

Business modelling methodology

A comprehensive methodology has been developed with a clear focus on identifying viable business models for providing 5G connectivity in different circumstances such as considering various use cases, areas to be covered, and service KPIs. The methodology involves several steps, beginning with the identification of the ecosystem. This involves defining the various business roles and stakeholders involved, as well as understanding their interactions. The second step is to identify the value network [81]. This involves analysing the different ways stakeholders interact with each other, resulting in different configurations of the value network. These configurations lead to the development of individual business models that are best suited for each stakeholder. This methodology is presented in *Figure 1-35*.

Figure 1-35 Business modelling methodology

In Chapter 3, we employed the outlined methodology to pinpoint potential business models pertinent to 5G-based inflight connectivity. In our pursuit of analyzing and drawing comparisons among these identified business models, we leveraged the value network intelligence framework proposed by V. Allee [92] and [94] as our foundational framework. However, we tailored this framework to accommodate the unique aspects of the use cases and eliminate any extraneous elements unsuitable for our specific use case. For instance, in the "Brand & Relationship" category, we combined components such as perceived value and reciprocity under the overarching concept of branding potential. This allowed us to explore the beneficiaries of enhanced perception and associated advantages. Furthermore, we introduced new categories like technology within the "Business Performance segment" to address the technology prerequisites necessary to uropa customer demands and ensure QoS. In alignment with our objectives, we opted to exclude certain evaluation metrics, including "Flow Path Optimization". This decision aligns with our framework's primary objective: firstly, to meticulously analyze the proposed business models, and secondly, to conduct a comparative assessment based on diverse metrics. Flow Path Optimization, initially introduced by Verna Allee, entails scrutinizing a specific business model with the intent of enhancing it. For instance, by dissecting the links within the Value Network, Allee aimed to identify areas ripe for improvement and optimization..

Additionally, in Chapter 5, we investigated possible business models for connected mobility services using the same approach. For a more comprehensive understanding of the various steps involved in the business modelling methodology, we recommend referring to these two chapters.

1.5.2 Scenario analysis

We adopt the definition of scenario analysis as defined in [3], where scenario analysis is a technique used in strategic planning to examine the potential outcomes of different sets of defined input parameters on a quantitative model. This can provide valuable insights for decision-making, particularly when there is limited scientific information available to accurately predict future outcomes. In the context of cellular technology, such as 5G, scenario analyses can be useful for assessing the demand for networks and informing deployment decisions. In 3GPP literature, scenarios often refer to different deployment situations, such as urban or rural areas, and are used to evaluate various 5G deployment options.
Throughout Chapters 2, 3, 5, and 7, we utilized the scenario analysis model to identify potential deployment scenarios for analysis and comparison. This process helps to identify optimal deployment strategies for the given situations.

1.5.3 Economic viability study

The economic viability study serves as a mechanism to assess the economic feasibility of a project or policy. This is accomplished by identifying all the costs and revenues and comparing them to determine if the project or policy is worth pursuing. The goal of a viability study is to determine whether the benefits of a project or policy outweigh its costs, and if so, by how much [81].
In the subsequent sections, we introduce the chosen cost model, specifically the Total Cost of Ownership (TCO) model. We define its essence, draw comparisons with alternative models, elucidate the rationale behind its selection, and subsequently detail its seamless integration into the overarching generic cost model adopted for this study. Additionally, we elaborate on the benefits estimation model employed in this dissertation.

Generic cost model

To enhance the comprehension of the cost model's description, we clarify first the used economic terminologies before delving into the specifics of the cost model.

Definition

We have adopted [82]'s definition of CAPEX and OPEX.
Capital expenditures, commonly referred to as CAPEX, involve investing in fixed infrastructure that is gradually depreciated over time. For network operators, CAPEX includes expenses such as acquiring land, network infrastructure like base stations and routers, and software for network management systems. It's worth noting that equipment purchases are always considered CAPEX, regardless of whether the payment is made in a lump sum or spread out over time.
In contrast, operational expenditures, also known as OPEX, do not contribute to the infrastructure itself but rather encompass ongoing expenses necessary to maintain the company's operations. These costs typically consist of technical and commercial operation expenses, as well as administrative expenses. For network operators, OPEX primarily comprises rental and leasing fees for infrastructure such as land, buildings, and network equipment, as well as personnel wages.

In most cases, the overhead cost is defined as the cost of marketing, helpdesk, human resources, finance, etc.

Model description

The cost model for 5G networks typically consists of two main components: **the network dimensioning process module** and **the cost calculation module**. The network dimensioning process module uses the use case specifications, as well as information about the existing infrastructure, the area being studied, the number of connected devices, and the KPIs of the use cases. This information is used to determine the necessary network infrastructure. The cost calculation module is used to determine the total cost of ownership (TCO) of the 5G network. This is done by using the bill of materials (BOM) generated by the dimensioning process, along with unit costs for network elements and distinctions between capital expenditures (CAPEX) and operational expenditures (OPEX). The TCO is then calculated by summing the CAPEX, OPEX costs, and overhead costs.
It's important to note that the specific details of the cost model may vary depending on factors such as the type of network being deployed, the location, and the specific use cases being supported, and it will be detailed accordingly in each chapter.

The Total Cost of Ownership (TCO)

The TCO model is a tool for assessing the financial viability of a project, particularly in the context of technology and asset-intensive projects. TCO focuses on evaluating all costs associated with owning, implementing, and maintaining an asset or technology solution over its entire lifecycle [95]. This includes not only the initial acquisition costs or CAPEX but also ongoing operational (OPEX), maintenance, and disposal costs. The TCO model is typically used to provide a comprehensive understanding of the true costs associated with a project, which can be especially useful for comparing different technology options or evaluating the long-term impact of various decisions.

Comparing TCO to alternative models

While the Total Cost of Ownership (TCO) model is a widely used approach for assessing the costs associated with owning and maintaining an asset or technology solution, there are several alternative models that focus on specific aspects of financial analysis and decision-making. Here are some alternative models to consider:
Total Cost of Acquisition (TCA) or CAPEX Analysis: This model focuses specifically on the costs associated with acquiring an asset or technology solution. It includes expenses such as purchase price, installation costs, and any immediate setup expenses. Unlike TCO, it does not consider ongoing operational or maintenance costs [96].
Return on Investment (ROI): ROI measures the financial return generated by a project relative to its cost. It compares the net profit generated by the project to the

initial investment. While it doesn't capture the full lifecycle costs like TCO, it provides a simple way to assess profitability [97].
Cost-Benefit Analysis (CBA): Similar to ROI, CBA compares the costs of a project with its benefits, both in monetary and non-monetary terms. While it provides insight into the overall value of a project, it may not address ongoing operational costs comprehensively [81].
Operational Expenditure (OPEX) Analysis: This focuses solely on the ongoing operational costs of a project, excluding capital expenditure. It's particularly relevant for projects where ongoing expenses are a primary concern.

After careful consideration of these comparisons, the TCO model was chosen due to its numerous advantages. In fact, the TCO model presents a range of strengths. It provides a thorough examination of costs related to owning and running technology or systems. This includes direct costs like upfront expenses, operating and maintenance costs, as well as indirect costs such as energy and personnel expenses. TCO considers the entire lifecycle of a project, encompassing acquisition, operation, maintenance, and disposal, aiding in recognizing long-term costs and potential savings. Moreover, TCO supports decision-making by giving a comprehensive view of total expenses, aiding in comparing choices and informed economic decisions [95].
However, the TCO model has certain weaknesses. Estimating costs for different TCO components can be subjective and challenging. It might not account for intangible benefits, like improved quality or customer satisfaction, which are hard to quantify. TCO models might lack adaptability for different contexts or projects, demanding extra effort for customization. They often overlook external factors, such as environmental effects or social costs, leading to incomplete economic analysis. Additionally, TCO models simplify assumptions, like projected lifespan or inflation rates, which can introduce uncertainties and inaccuracies in cost calculations.

To overcome the shortcomings of the TCO model, our research approach has incorporated several strategies:

- **Cost Estimation and data collection strategy:** to reduce the potential impact of subjectivity in the data used, which could undermine result accuracy and reliability, we have implemented a robust data collection strategy. Since our research was conducted within the framework of European projects, we sourced data from diverse partners spanning various countries in Europe. Additionally, we conducted thorough desk research to compile and benchmark data. By leveraging multiple data sources, we have enhanced the validity of our findings.
- **Generic cost model and scenario analysis**: to ensure the versatility of our model for application in various use cases, we constructed a generic cost model. By employing scenario analysis, we generated a range of scenarios. These scenarios were then subjected to the generic model, which was further tailored to align with the unique aspects of the studied scenarios.

This approach allows for the adaptability of the model while accommodating the specific requirements of each scenario..

- **Sensitivity analysis**: To tackle the challenge posed by simplifying inputs and assumptions, we conducted a sensitivity analysis. This process evaluated the influence of various assumptions on the outcomes of the model, providing a more nuanced understanding of the potential impacts. The sensitivity analysis model is described in detail in 1.5.5.

However, **it is crucial to recognize a noteworthy factor**: the inherent limitation of the TCO approach in adequately incorporating indirect benefits. Consequently, TCO could be integrated into a more comprehensive quantitative analysis, encompassing not only economic factors but also societal and environmental metrics. Alternatively, it could serve as a complement to qualitative assessments.

Benefit estimation model

In this dissertation, two methods have been used for estimating the total benefits or revenues of deploying 5G-based solutions namely the willingness to pay (WTP) and the average revenue per user (ARPU):

WTP is the maximum amount of money that a customer or end-user is willing to pay for a particular product or service. It is a measure of the perceived value of the product or service to the customer. WTP is often determined through market research and surveys, and it can vary based on factors such as the customer's income, preferences, and the availability of alternative options [83].

On the other hand, ARPU is a financial metric used in telecommunications and other subscription-based businesses to measure the average revenue generated per user or customer. It is calculated by dividing the total revenue generated by the number of customers or subscribers. ARPU can be used to evaluate the profitability of a business and its ability to monetize its customer base [84].

In this work, estimating the total revenue that will be generated by 5G for a specific deployment (i.e., certain use cases in a specific region) was challenging. Therefore, the ARPU was calculated based on the average cost per user (ACPU), with a profit margin added on top.

In Chapters 2, 3, 5, 6, and 7, we utilized the economic viability study to assess the feasibility of the deployment scenarios or use cases under study. For more information about the mathematical formulations of the cost and revenue models, as well as the various steps involved in the analysis, we refer readers to these chapters.

1.5.4 Cost allocation model

To fully comprehend the concept of cost allocation, it is important to first define some relevant economic terminology. This will provide a solid foundation for understanding the definition of a cost allocation model.

First direct versus indirect cost: direct costs are expenditures that can be attributed to a particular product or service, whereas indirect costs are expenses associated with managing and operating the company providing this product or service [85]. In the context of mobile networks, direct costs refer to expenses that can be directly attributed to the provision of mobile network services, such as network equipment, and salaries of network engineers that work exclusively on these services. Indirect costs, on the other hand, refer to expenses that are necessary for the operation of a mobile network but cannot be directly attributed to the provision of mobile services, such as rent, utilities, and administrative expenses.

A cost allocation model is a method used to distribute or allocate the costs of a particular project, service, or asset among various stakeholders or departments. The purpose of a cost allocation model is to ensure that all parties responsible for a project or service contribute to its cost based on their level of involvement or benefit [87]. Generally, the cost that needs to be allocated is the indirect cost, as the direct costs can be easily linked to stakeholders, products or services.

In the context of 5G deployment, a cost allocation model can be used to allocate the cost of deploying 5G infrastructure among various stakeholders involved in the project, such as telecommunication companies, government agencies, and other entities. Furthermore, a cost allocation model can also be used to allocate the cost of deploying 5G infrastructure to the specific services or applications that are utilizing the 5G network. This can be done for example by analyzing the usage and traffic data of each service and allocating the cost of deploying 5G infrastructure accordingly. For example, a high-bandwidth service such as video streaming may be allocated a higher portion of the 5G deployment cost than a low-bandwidth service such as email. In general, a cost allocation model can help ensure a fair distribution of the cost of 5G deployment among various stakeholders or services. In this dissertation, both cost allocation models have been used namely the cost allocation to stakeholders and services/slices.

In Chapter 4, we utilized the cost allocation model to allocate the costs of the 5G network to the services being offered. In Chapter 5, we applied the cost allocation concept to distribute the 5G deployment costs among the stakeholders involved. To better understand the distinction between these two applications of the cost allocation model, we encourage readers to refer to these chapters.

1.5.5 **Sensitivity analysis**

One of the famous definitions of sensitivity analysis is: "*The study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input*" [88]. Every model is based on a set of assumptions, which introduces uncertainty into the input parameters. Consequently, sensitivity analysis aims to assess the influence of varying parameters on the model outputs. Once these parameters are identified, those with a negligible impact are usually disregarded, while greater attention is devoted to the important ones [81].

Throughout Chapters 2, 3, 5, 6, and 7, we utilized the sensitivity analysis model. For a more detailed understanding of the use and application of this model, we suggest referring to these chapters.

1.5.6 Choice of studied use cases and data collection

Since our primary objective is to create a comprehensive Techno-Economic Analysis (TEA) of 5G rollout, and as the deployment of 5G networks in urban environments has been extensively explored in various published techno-economic studies, as highlighted in this survey [75], we have opted to complement existing research by concentrating on more challenging areas. These areas encompass rural, remote, and cross-border regions. As previously discussed, these regions tend to be less attractive to MNOs to deploy 5G due to uncertainties surrounding the investment's profitability, attributed to their lower population densities. Given this direction, the research challenges outlined in section 1.4 can be thoroughly examined, drawing on the distinct attributes of the chosen use cases. To feed the techno-economic models, to study the selected use cases, with the required data, and since our research was conducted within the framework of European projects, we sourced data from diverse partners spanning various countries in Europe including British Telecom (BT), Telecom Italia (TIM), Deutsche Telekom, SES, Airbus, Thales among others. Additionally, we conducted thorough desk research to compile and benchmark data as explained in the data collection strategy in section 1.5.3. By leveraging multiple data sources, we have enhanced the validity of our findings.

In the following section, a detailed description of the studied use cases and the specific research questions investigated within each case will be provided.

1.6 Outline and research contributions

This PhD dissertation consists of a series of publications that were produced during the research. The original publications were not modified except for formatting changes to ensure consistency. In this section, we provide an overview of the remaining chapters of the dissertation, explain how they are interconnected with the identified research questions in section 1.4 (in Figure 1-36), and highlight their main contributions. A comprehensive list of all publications written during the research period, including those not included in the dissertation text, is provided in Section 1.7.

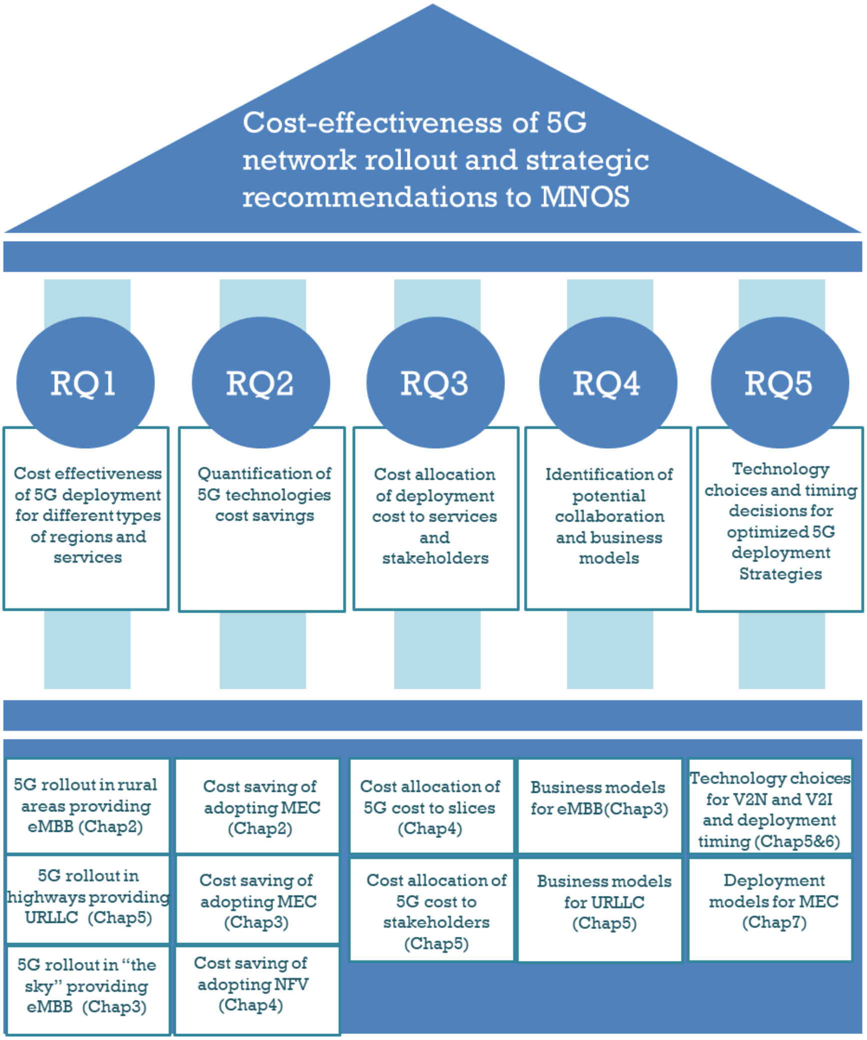


Figure 1-36: Overview of how the different contributions in Chapters 2 to 7 are linked to the different research questions and the overarching research objective. Under each research question, we put the related chapters that answer the research question or its sub-parts by highlighting the research contributions done in those chapters.

Chapter 2 describes how an integrated satellite-4G network solution can bridge the digital divide in Europe by providing broadband connectivity to rural and remote areas where deploying terrestrial networks is challenging and expensive. While our analysis is based on a 4G solution, we adopted several technologies from the 5G paradigm since the 5G standard was not finalized at the time of writing.

Satellite communication holds great promise in providing full coverage to remote and rural areas, but it is important to acknowledge the high cost of satellite capacity if we want to deliver broadband connectivity. Therefore, the question remains: is it economically feasible to use satellites for broadband connectivity? Additionally, how can the 5G network be leveraged to make connectivity more affordable for end users?

This study brings the following contributions related to the research questions previously formulated in section 1.4: Firstly, we explore **the economic feasibility of deploying the integrated satellite-4G solution in rural areas to provide eMBB** (RQ1[4]). Secondly, we indicate how **MEC technology can reduce the Average Cost Per User (ACPU)** to make the solution more affordable for rural area residents (RQ2). Finally, we show **how the ACPU is affected by the population density of these rural areas**, i.e., how the ACPU varies across a broader range of areas (RQ1).

Chapter 2 discusses the provision of broadband connectivity (eMBB) to remote and rural areas. On the other hand, **Chapter 3** focuses on inflight eMBB connectivity and presents a techno-economic analysis of an integrated satellite-5G network solution. As described previously, while satellite communication offers the potential to reach users anywhere, the high costs of broadband connectivity using this technology are a major hurdle. The study answers several research questions: First, **is "internet in the sky" feasible from both technical and economic perspectives using the integrated satellite-5G solution** (RQ1)? And, **how can the use of MEC technology help reduce costs** (RQ2)? Finally, **what are the potential business and pricing models for delivering inflight broadband connectivity** (RQ4)? By addressing these research questions, the study provides insights into the viability of the satellite-5G solution and its potential impact on the aviation industry.

In the previous chapters, we examined the feasibility of deploying the 5G network for specific service types across various regions. The next question is, how should we distribute the deployment cost among the services to be offered?

In **Chapter 4 we propose an allocation model to answer this research question for slice-based 5G network** (RQ3). This chapter presents a model for distributing the cost of network deployment among the various deployed slices, which can subsequently be utilized to determine the prices of different E2E services. To provide the required inputs to the proposed model, a resource allocation algorithm and a 5G network function (NF) dimensioning model are also introduced. The chapter also explores the outcomes of the different models, in addition to **the cost savings on the core network part that stems from utilizing NFV** (RQ2).

Besides the eMBB service investigated in the previous chapters, **Chapter 5** also investigates the provision of URLLC services. This chapter delves into the provision of cooperative, connected, and automated mobility (CCAM) services, which fall under the category of URLLC services. Connected mobility services

[4] RQ1-5 refer to the research questions discussed in section 1.3.

require a reliable and low-latency network, and the 5G technology has been identified as the ideal provider for CCAM requirements. However, providing CCAM services involves multiple stakeholders from various fields, such as MNOS, OEMs, and regulators. This chapter **investigates the potential business and collaboration models that take benefit from emerging 5G technologies** (RQ4). To evaluate the viability of the identified business models, **a cost-benefit analysis is conducted** (RQ1) **and has been used also to assess the impact of the time of the deployment on the total cost** (RQ5). Additionally**, a cost allocation model is developed to share the 5G deployment costs among the involved business actors** (RQ3). In Chapter 5, we encountered the challenge of the lack of data on the penetration rates of connected cars, which required us to simulate different deployment scenarios assuming the same year as the reference for their deployment. To address this issue and account for the uncertainty in predicting the penetration rate of connected cars, we proposed a scenario-based approach with three distinct penetration rate scenarios: low, medium, and high (RQ5). This approach allowed us to conduct a more precise techno-economic analysis, which is detailed in **Chapter 6**.

In the realm of providing CCAM services, the MEC paradigm plays a vital role in the 5G network architecture by reducing E2E latency. However, to ensure optimal deployment of MEC in terms of E2E latency and deployment cost, MNOs must carefully consider different deployment models. In **Chapter 7, we propose various MEC deployment models and evaluate them from both E2E latency and cost perspectives**. Our findings can inform the decision-making process of MNOs as they determine their deployment strategies for MEC technology (RQ5).

The final chapter of this dissertation, **Chapter 8**, presents a comprehensive conclusion on the techno-economic analysis of deploying a 5G network. It effectively summarizes the key findings from previous chapters and utilizes them to provide strategic guidelines for 5G network deployment, with a specific emphasis on the role of the key 5G technologies. Furthermore, the book concludes with an insightful overview of potential areas for future research, highlighting avenues for further exploration in this field.

1.7 Publication List

1.7.1 Publications in international journals (listed in the Science Citation Index)

1. Chiha, A., Van der Wee, M., Colle, D., & Verbrugge, S. (2020). Techno-economic viability of integrating satellite communication in 4G networks to bridge the broadband digital divide. Telecommunications Policy, 44(3), 101874.
2. Chiha, A., Van der Wee, M., Goratti, L., & Colle, D. (2021). Techno-economic analysis of inflight connectivity using an integrated satellite-5G network. International Journal of Satellite Communications and Networking, 39(4), 322-338.

3. Chiha, A., Van der Wee, M., Colle, D., & Verbrugge, S. (2020). Network slicing cost allocation model. Journal of Network and Systems Management, 28, 627-659.
4. Chiha, A., Denis, B., Verbrugge, S., & Colle, D. (2023). Techno-Economic Analysis of MEC Clustering Models for Seamless CCAM Service Provision. IEEE Communications Magazine, 61(2), 32-37.
5. Chiha, A., Vannieuwenborg, F., Denis, B., Colle, D., & Verbrugge, S. (2023). Cooperative, Connected and automated mobility (CCAM) services provisioning in cross-border settings: Techno-economic analysis in the light of technical challenges. Transport Policy, 140.

1.7.2 Publications in international conferences (listed in the Science Citation Index) (p1)

6. Chiha Ep Harbi, A., Poli, F., Denis, B., Vannieuwenborg, F., García-Roger, D., Monserrat, J. F., & Aigner, W. (2021, September). Techno-economic and Simulation Study of a V2I-based Cooperative Manoeuvring Case in a Cross-border Scenario. In 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall) (pp. 1-7). IEEE.
7. Chiha, A., Van der Wee, M., Briggs, K., & Colle, D. (2020, June). Techno-economic evaluation of a brokerage role in the context of integrated satellite-5G networks. In 2020 6th IEEE Conference on Network Softwarization (NetSoft) (pp. 1-5). IEEE.

1.7.3 Publications in other international conferences (c1)

8. Chiha Ep Harbi, A., Van der Wee, M., Verbrugge, S., & Colle, D. (2018). Techno-economic viability of integrating satellite communication in 4G networks to bridge the broadband digital divide. In ITS2018, the 29th European Conference of the International Telecommunications Society (pp. 1-11).
9. Chiha Ep Harbi, A., Van der Wee, M., Fitch, M., Briggs, K., Watts, S., Goratti, L., & Jou, C. T. (2019). Managing integrated satellite-5G networks: techno-economic evaluation of a brokerage role. In EuCNC2019, the European Conference on Networks and Communications.
10. Chiha Ep Harbi, A., Camps-Arago, P., Vannieuwenborg, F., Deckers, L., Kenis, E., Vandenberghe, W., … & Colle, D. (2021). The importance of business and governance model blueprints for teleoperated transport in cross-border settings. In 27th ITS World (pp. 1-10).
11. Chiha, A., Degrande, T., Verbrugge, S., Colle, D., Avdikos, G., Aigner, W., ... & Garcia-Roger, D. (2022, October). Business Case Evaluation of Cooperative, Connected and Automated Mobility Service Provision in cross-border settings [INVITED PAPER–Project 5G-CARMEN]. In 2022 IEEE Future Networks World Forum (FNWF) (pp. 209-214). IEEE.

References

[1] Ian Wright (2022), The Cost of Shutting Down the Internet. Retrieved, February 10, 2023, from https://merchantmachine.co.uk/the-cost-of-shutting-down-the-internet/

[2] Broadband Commission, (2020). COVID-19 crisis Broadband Commission Agenda for Action For faster and better recovery. Retrieved, February 10, 2023, from https://broadbandcommission.org/wp-content/uploads/2021/02/Agenda4Action_English2020.pdf

[3] Theory, I.M. (2007). IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems Draft Amendment: Management Information Base Extensions. doi: 10.1109/IEEESTD.2004.226664.

[4] Micah Singleton (2015). The FCC has changed the definition of broadband. Retrieved, February 12, 2023, from https://www.theverge.com/2015/1/29/7932653/fcc-changed-definition-broadband-25mbps

[5] Berec (2020). Broadband as a Universal Service: Ensuring the availability of an adequate broadband internet access service, including the underlying connection, at a fixed location. Retrieved, February 12, 2023, from https://isportal.berec.europa.eu/view-doc/broadband-as-a-universal-service-ensuring-the-availability-of-an-adequate-broadband-internet-access

[6] 3GPP. (2021). TS 23.002: Network Architecture. Retrieved February 13, 2023, from https://www.3gpp.org/ftp/Specs/html-info/23002.htm

[7] ITU-T. (2017). G.9700/Y.2705: Broadband Integrated Services Digital Network (B-ISDN) Digital Subscriber Signalling System No. 2 (DSS2) User-Network Interface (UNI) Layer 1 Specification.

[8] ITU. (2012). ITU Handbook on Telecommunications: Access Networks. Retrieved February 13, 2023, from https://www.itu.int/dms_pub/itu-r/opb/hdb/R-HDB-62-2022-PDF-E.pdf

[9] GSMA Intelligence. (2018). The Mobile Economy 2018. Retrieved February 13, 2023, from https://data.gsmaintelligence.com/api-web/v2/research-file-download?id=28999769&file=The%20Mobile%20Economy%202018.pdf

[10] Chundury, R. (2008). Mobile broadband backhaul: Addressing the challenge. Ericsson Review, 3, 4-9.

[11] Van Giang, N., & Kim, Y. H. (2014). Slicing the next mobile packet core network. In 2014 [1]1th International Symposium on Wireless Communications Systems (ISWCS) (pp. 763-767). IEEE.

[12] Frenzel, L. E., Jr. (2018). Networking: Wired and Wirele–s - All Devices Talking to One Another. In Electronics Explained: Fundamentals for Engineers, Technicians, and Makers ⁽2nd ed., Chapter 9).

[13] Nguyen, T. T., Berscheid, B., Nguyen, H. H., & Salt, J. E. (2017). A Novel Iterative OFDMA Channel Estimation Technique for DOCSIS 3.1 Uplink Channels. IEEE Transactions on Broadcasting, 63(2), 361-375. https://doi.org/10.1109/TBC.2016.2636738

[14] Bai, T., Zhang, H., Zhang, R., Yang, L. L., Al Rawi, A. F., Zhang, J., & Hanzo, L. (2017). Discrete multi-tone digital subscriber loop performance in the face of impulsive noise. IEEE Access, 5, 10478-10495.

[15] Router-switch.com. (n.d.). FTTD Solution: Cost-effective Fiber Access Solution for SMBs. Router-switch.com. https://www.router-switch.com/solution/fttd.html

[16] European Commission. (2020). Digital Economy and Society Index (DESI) 2020, Telecom country chapters. Commission Staff Working Document. Retrieved from https://www.europarl.europa.eu/RegData/docs_autres_institutions/commission_europeenne/swd/2020/0111/COM_SWD%282020%290111%28PAR04%29_EN.pdf

[17] Nassa, V. K. (2011). Wireless Communications: Past, Present and Future. Dronacharya Research Journal, 50.

[18] Goldsmith, A. (2005). Wireless communications. Cambridge University Press. https://doi.org/10.1017/CBO9780511841224

[19] IEEE. (2013). IEEE 802.11ac™-2013. Retrieved February 15, 2023, from https://standards.ieee.org/standard/802_11ac-2013.html

[20] 3GPP. (2011). 4G Evolution for Mobile Operators. Retrieved February 15, 2023, from https://www.3gpp.org/news-events/partner-news/4g-evolution-for-mobile-operators

[21] Chuck Moozakis. (n.d). What is wireless communications? Everything you need to know. Retrieved February 15, 2023, from https://www.techtarget.com/searchmobilecomputing/definition/wireless

[22] Perahia, E., & Stacey, R. (2008). Next Generation Wireless LANs: Throughput, Robustness, and Reliability in 802.11n. Cambridge University Press.

[23] R.Q. Shaddad, A.B. Mohammad, S.A. Al-Gailani, A.M. Al-hetar, M.A. Elmagzoub,. (2014). A survey on access technologies for broadband optical and wireless networks, Journal of Network and Computer Applications, Volume 41, Pages 459-472, ISSN 1084-8045, https://doi.org/10.1016/j.jnca.2014.01.004.

[24] Sharma, J. (2020, March). The Wi-Fi Evolution. White paper. Wireless Communication Business Unit.

[25] Sesia, S., Toufik, I., & Baker, M. (Eds.). (2011). L–E - The UMTS long term evolution: from theory to practice (2nd ed.). John Wiley & Sons.

[26] Farley, T. (2005). Mobile telephone history. Telektronikk, 101(3-4), 22-34.

[27] Hillebrand, F. (2013). The creation of standards for global mobile communication: GSM and UMTS standardization from 1982 to 2000. IEEE Wireless Communications, 20(5), 24-33.

[28] Shah, A. F. M. S. (2022). A survey from 1G to 5G including the advent of 6G: Architectures, multiple access techniques, and emerging technologies. In 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC) (pp. 1-6). IEEE.

[29] Abramov, V., Li, S., Wang, M., Wong, E. W. M., & Zukerman, M. (2012). Computation of blocking probability for large circuit-switched networks. IEEE Communications Letters, 16(11), 1892-1895.

[30] Halonen, T., Romero, J., & Melero, J. (2004). GSM, GPRS and EDGE Performance: Evolution Towards 3G/UMTS. Chichester, UK: Wiley.

[31] Rappaport, T. S., et al. (2013). Millimeter wave mobile communications for 5G cellular: it will work! IEEE Access, 1, 335-349. https://doi.org/10.1109/ACCESS.2013.2260813

[32] Tanenbaum, A. S., & Wetherall, D. (2011). Computer Networks. Pearson Education.

[33] Abrol, A., & Jha, R. K. (2016). Power optimization in 5G network: A step towards GrEEn communication. IEEE Access, 4, 1355-1374. https://doi.org/10.1109/ACCESS.2016.2549641

[34] Gupta, A., & Jha, R. K. (2015). A survey of 5G network: architecture and emerging technologies. IEEE Access, 3, 1206-1232. https://doi.org/10.1109/ACCESS.2015.2461602

[35] Akyildiz, I. F., Gutierrez-Estevez, D. M., & Reyes, E. C. (2010). The evolution to 4G cellular systems: LTE-Advanced. Physical Communication, 3(4), 217-244. doi: 10.1016/j.phycom.2010.08.001.

[36] Solyman, A. A. A., & Yahya, K. (2022). Evolution of wireless communication networks: From 1G to 6G and future perspective. International Journal of Electrical and Computer Engineering, 12(4), 3943.

[37] Alsharif, M. H., Kelechi, A. H., Albreem, M. A., Chaudhry, S. A., Zia, M. S., & Kim, S. (2020). Sixth generation (6G) wireless networks: vision, research activities, challenges and potential solutions. Symmetry, 12(4), 6676. doi: 10.3390/sym12040676.

[38] David, K., & Berndt, H. (2018). 6G vision and requirements: Is there any need for beyond 5G?. IEEE Vehicular Technology Magazine, 13(3), 72-80. doi: 10.1109/MVT.2018.2848498.

[39] Chen, S., Liang, Y. C., Sun, S., Kang, S., Cheng, W., & Peng, M. (2020). Vision, requirements, and technology trend of 6G: how to tackle the challenges of system coverage, capacity, user data-rate and movement speed. IEEE Wireless Communications, 27(2), 218-228. doi: 10.1109/MWC.001.1900333.

[40] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini. (2020). What should 6G be?. Nature Electronics, vol. 3, no. 1, pp. 20–29, doi: 10.1038/s41928-019-0355-6.

[41] Misra, D., Misra, D. K., & Tripathi, S. P. (2013). Satellite communication advancement, issues, challenges and applications. International Journal of Advanced Research in Computer and Communication Engineering, 2(4), 1681-1686.

[42] Media & Entertainment Delivery over Hetnet with Optimized Satellite Architecture project, Technical note 2, issue 2.2. (2016).
[43] Farserotu, J., & Prasad, R. (2000). A survey of future broadband multimedia satellite systems, issues and trends. IEEE Communications Magazine, 38(6), 128-133. doi: 10.1109/35.846084.
[44] ITU-R. (2015). IMT Vision_Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond, document Rec. ITU-R M.2083-0.
[45] 5GPP. (n.d.). 5G KPIs. Retrieved April, 08, 2018 from https://5g-ppp.eu/kpis/.
[46] Erunkulu, O. O., Zungeru, A. M., Lebekwe, C. K., & Chuma, J. M. (2020). Cellular communications coverage prediction techniques: A survey and comparison. IEEE Access, 8, 113052-113077.
[47] Tripathi, P. S. M., & Prasad, R. (2018). Spectrum for 5G services. Wireless Pers. Commun., 100(2), 539-555.
[48] IEEE 5G. (2017). IEEE 5G and beyond technology roadmap White Paper. Institute of Electrical and Electronics Engineers. Retrieved July 23, 2020, from https://5g.ieee.org/images/_les/pdf/ieee-5g-roadmap-white-paper.pdf
[49] Sleiman, D. (2021). RF and 5G new radio: Top 5 questions answered. Retrieved March 10, 2023, from https://www.exfo.com/es/recursos/blog/rf-5g-new-radio-top-5-questions/
[50] Erunkulu, O. O., Zungeru, A. M., Lebekwe, C. K., & Chuma, J. M. (2021). 5G mobile communication applications: A survey and comparison of use cases. IEEE Access, 9, 97251-97295.
[51] 5G; System Architecture for the 5G System (3GPP TS 23.501 Version 15.2.0 Release 15). (2018). ETSI TS 123.501, 15, 4-220.
[52] Vook, F. W., Ghosh, A., Diarte, E., & Murphy, M. (2018). 5G New Radio: Overview and Performance. In 2018 [5]2nd Asilomar Conference on Signals, Systems, and Computers (pp. 1247-1251). Pacific Grove, CA, USA: IEEe. doi: 10.1109/ACSSC.2018.8645228
[53] 3GPP. (2016). Study on Architecture for Next Generation System (Release 14), document 3GPP TR 23.799, V14.0.0, 21-24.
[54] 3GPP TS 23.503, V15.6.0. (2019). Policy and Charging Control Framework for the 5G System (5GS); Stage 2 (Release 15).
[55] 3GPP TS 23.502, V16.4.0. (2020). Procedures for the 5G System (5GS); Stage 2 (Release 16).
[56] 3GPP. (2017-04). 3GPP TS 23.501 V0.4.0. System Architecture for the 5G System
[57] 3GPP TS 23.501 V15.0.0 (2017-12) System Architecture for the 5G System (Stage 2)
[58] Rommer, S., Hedman, P., Olsson, M., Frid, L., Sultana, S., & Mulligan, C. (2019). 5G Core Networks: Powering Digitalization (1st ed.). Academic.

[59] Lei, W., Zhang, Q., Song, L., Gao, X., & Chen, X. (2020). 5G procedure, RAN architecture, and protocol. In 5G System Design (pp. 227-271). Springer.

[60] Cisco. (2018). 5G Non Standalone Solution Guide, StarOS Release 21.5. San Jose, CA, USA.

[61] ETSI TS 123.501. (2018). 5G; System Architecture for the 5G System (3GPP TS 23.501 Version 15.2.0 Release 15), pp. 4-220, Vol. 15.

[62] Wong, V. W. S., Li, Z., Niu, Z., Lu, J., & Li, G. Y. (Eds.). (2017). Key technologies for 5G wireless systems. Cambridge University Press.

[63] Matias, J., Taleb, T., Flinck, H., & Dutta, S. (2015). Toward an SDN-enabled NFV architecture. IEEE Communications Magazine, 53(4), 187-193. doi: 10.1109/MCOM.2015.7060494

[64] Wood, T., Ramakrishnan, K. K., Hwang, J., Liu, G., & Zhang, W. (2015). Toward a software-based network: integrating software defined networking and network function virtualization. IEEE Network, 29(3), 36-41.

[65] IEEE Standards Association. (2017, December). Network Slicing and 3GPP Service and Systems Aspects (SA) Standard. Retrieved May 15, 2019, from https://sdn.ieee.org/newsletter/december-2017/network-slicing-and-3gpp-service-and-systems-aspects-sa-standard

[66] IEEE SDN. (2017, December). Network slicing and 3GPP service and systems aspects (SA) standard. Retrieved from https://sdn.ieee.org/newsletter/december-2017/network-slicing-and-3gpp-service-and-systems-aspects-sa-standard

[67] ETSI. (2022). ETSI GS MEC 002 V2.2.1 (2022-01“: "Multi-access Edge Computing (MEC); Phase 2: Use Cases and Requiremen”s." Retrieved from https://www.etsi.org/deliver/etsi_gs/MEC/001_099/002/02.02.01_60/gs_MEC002v020201p.pdf

[68] Soua, R., Turcanu, I., Adamsky, F., Führer, D., & Engel, T. (2018). Multi-Access Edge Computing for Vehicular Networks: A Position Paper. In 2018 IEEE Globecom Workshops (GC Wkshps) (pp. 1-6). IEEe. doi: 10.1109/GLOCOMW.2018.8644392

[69] ETSI. (2014). Mobile-edge-computing- Introductory technical white paper. ETSI, White Paper, Retrieved April 2018, from https://portal.etsi.org.

[70] Huh, J.-H., & Seo, Y.-S. (2019). Understanding Edge Computing: Engineering Evolution With Artificial Intelligence. IEEE Access, 7, 164229-164245. doi: 10.1109/ACCESS.2019.2945338.

[71] Bogale, T. E., & Le, L. B. (2016). Massive MIMO and mmWave for 5G Wireless HetNet: Potential Benefits and Challenges. IEEE Vehicular Technology Magazine, 11(1), 64-75. doi: 10.1109/MVT.2015.2496240.

[72] Osseiran, A., Boccardi, F., Braun, V., Kusume, K., Marsch, P., Maternia, M., et al. (2014). Scenarios for 5G mobile and wireless communications: The vision of the METIS project. IEEE Communications Magazine, 52(5), 26-35.

[73] Larsson, E.G. & Van der Perre, L. (2017). Massive MIMO for 5G.

[74] Rahimian, A., & Mehran, F. (2011). RF link budget analysis in urban propagation microcell environment for mobile radio communication systems link planning. In Proc. Int. Conf. Wireless Communications Signal Processing (pp. 1-5).

[75] Oughton, E. J., & Lehr, W. (2022). Surveying 5G Techno-Economic Research to Inform the Evaluation of 6G Wireless Technologies. IEEE Access, 10, 25237-25257. doi: 10.1109/ACCESS.2022.3153046.

[76] Blanco, B., García-Sánchez, A. J., Melendi, D., Díaz-Díaz, R., Fernández-Palacios, J. P., & García-Haro, J. (2017). Technology pillars in the architecture of future 5G mobile networks: NFV, MEC and SDN. Computer Standards & Interfaces, 54, 216-228.

[77] Bram Naudts. (2017). Techno-economic Analysis of Software-Defined Telecommunications Networks.

[78] Liang, C., & Yu, F. R. (2014). Wireless network virtualization: A survey, some research issues and challenges. IEEE Communications Surveys & Tutorials, 17(1), 358-380.

[79] ITU. (2016). Sharing networks, driving growth. Geneva: International Telecommunication Union.

[80] Khan, A., Kellerer, W., Kozu, K., & Yabusaki, M. (2011). Network sharing in the next mobile network: TCO reduction, management flexibility, and operational independence. IEEE Communications Magazine, 49(10), 134-142. doi: 10.1109/MCOM.2011.6035827.

[81] Verbrugge, S., Casier, K., Van Ooteghem, J., & Lannoo, B. (2009). Practical steps in techno-economic evaluation of network deployment planning.

[82] Verbrugge, S., Colle, D., Pickavet, M., Demeester, P., Pasqualini, S., Iselt, A., Kirstädter, A., Hülsermann, R., Westphal, F.-J., & Jäger, M. (2006). Methodology and input availability parameters for calculating OpEx and CapEx costs for realistic network scenarios. Journal of Optical Networking, 5, 509-520.

[83] Rappoport, P., & Alleman, J. (2004). Modeling demand for telecom services using surveys. Geneva, Switzerland.

[84] Suryanegara, M. (2018). The Economics of 5G: Shifting from Revenue-per-User to Revenue-per-Machine. In 2018 [1]8th International Symposium on Communications and Information Technologies (ISCIT) (pp. 191-194). Bangkok, Thailand: IEEe. doi: 10.1109/ISCIT.2018.8588006.

[85] Spielman, E. (2023). Direct Costs vs. Indirect Costs: What Are They, and How Are They Different? Retrieved April 20, 2023, from https://www.businessnewsdaily.com/5498-direct-costs-indirect-costs.html

[86] BEREC. (2018). BEREC Report on Infrastructure Sharing

[87] Young, H. Peyton. (1994). Cost allocation. Handbook of game theory with economic applications 2 1193-1235.

[88] Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models. John Wiley and Sons.

[89] OASE deliverable D6.1 (2010). Overview of Tools and Methods and Identification of Value Networks.

[90] Rong, K., Lin, Y., Li, B., Burström, T., Butel, L., & Yu, J. (2018). Business ecosystem research agenda: more dynamic, more embedded, and more internationalized. Asian Business & Management, 17, 167-182.

[91] Osterwalder, A., & Pigneur, Y. (2009). Business Model Canvas: nine business model building blocks. Self published.

[92] V. Allee (2010). Presentation: What is Value Network Analysis? Accessed: 12 04 2019. [Online]. Available: https://www.slideshare.net/baba2729/what-isvaluenetworkanalysis.

[93] E3 model, Accessed: 12 04 2019. [Online] Available:https://www.e3value.com/

[94] V. Allee, "Value network analysis and value conversion of tangible and intangible assets," J. Intellect. Cap, vol. 9, no. 1, p. 5–24, 2008.

[95] Ellram, L. (1994). A taxonomy of total cost of ownership models. Journal of business logistics, 15(1), 171.

[96] Ellram, L. M. (1995). Total cost of ownership: an analysis approach for purchasing. International Journal of Physical Distribution & Logistics Management, 25(8), 4-23.

[97] Phillips, J. J. (1994). Measuring return on investment (Vol. 2). American Society for Training and Development.

[98] Blackman, C., & Forge, S. (2019). 5G Deployment State of Play in Europe, USA and Asia. Policy Department for Economic, Scientific and Quality of Life Policies, Directorate-General for Internal Policies. PE 631.060 – April 2019.

[99] Streaming Media (2019), MWC 19: Operators promise the moon but won't shoulder the cost, 27 February, https://www.streamingmedia.com/Articles/ReadArticle.aspx?/ArticleID=130139.

[100] Fildes, Nic (2019b), "Europe's 5G Struggle", Financial Times, 28 January.

[101] Sun, P. (2023). 5G Development Around the Globe. In A Guidebook for 5GtoB and 6G Vision for Deep Convergence (Chapter 2). Posts & Telecom Press. https://doi.org/10.1007/978-981-99-4024-0_2

[102] Forge, S., & Vu, K. (2020). Forming a 5G strategy for developing countries: A note for policy makers. Telecommunications Policy, 44(7), 101975.

[103] European Commission. (2018). Fixed and mobile Convergence in Europe: Quality Measurements for 5G and network densification.

[104] Forge, S., Horvitz, R., & Bohlin, E. C. (2019). Light deployment Regime for small-area wireless access points (SAWAPs), final report. Retrieved from https://op.europa.eu/en/publication-detail/-/publication/463e2d3d-1d8f-11ea-95ab-01aa75ed71a1/language-en/format-PDF/source-112125706

[105] Uher, J. J., Harper, J. R., Mennecke, R. G., Patton, P. M., & Farroha, S. (2018). The Next Wave: Investigating end to end security in 5G capabilities and IoT extensions (Vol. 21, No. 4). National Security Agency (NSA), USA.

[106] Electromagnetic fields and public health – Base stations and wireless technologies," World Health Organization (WHO), Factsheet, 8 October 2014.

Available at https://www.who.int/en/news-room/fact-sheets/detail/electromagnetic-fields-and-public-health-mobile-phones

[107] GSMA. (2022). The Mobile Economy Europe 2022 [Report].

[108] Market Research Future. (2023). Worldwide 5G Connections in 2023. Retrieved from https://www.marketresearchfuture.com/news/worldwide-5g-connections-in-2023.

[109] Chen, S. Z., & Kang, S. L. (2018). A tutorial on 5G and the progress in China. Frontiers of Information Technology & Electronic Engineering, 19(3), 309-321.

[110] Lee, C. K., & Yu, L. (2022). A multi-level perspective on 5G transition: The China case. Technological Forecasting and Social Change, 182, 121812.

[111] 5G observatory (2023) Retrieved from: https://5gobservatory.eu/observatory-overview/interactive-5g-scoreboard/#EU-scoreboard

[112] European Commission. (2016) 5G Action Plan. Retrieved from: https://digital-strategy.ec.europa.eu/en/policies/5g-action-plan

[113] European Commission. (2016). Broadband For Europe. Gigabit Society 2025. Retrieved from https://ec.europa.eu/digital-single-market/en/broadband-europe

[114] European Commission. (2018)., Retrieved from: https://digital-strategy.ec.europa.eu/en/library/proposed-directive-establishing-european-electronic-communications-code

[115] European Commission. (2021). 2030 digital compass: your digital decade. Retrieved from https://futurium.ec.europa.eu/en/digital-compass/discussion-forum/welcome-message-dg-connect-building-common-digital-future-empowers-everyone-0

[116] European Commission. Retrieved from https://digital-strategy.ec.europa.eu/en/policies/europes-digital-decade

[117] 5GPPP, third phase projects. Retrieved from https://5g-ppp.eu/5g-ppp-phase-3-projects/

2

Techno-economic viability of integrating satellite communication in 4G networks to bridge the broadband digital divide

Asma Chiha, Marlies Van der Wee, Didier Colle, and Sofie Verbrugge

Published in Telecommunications Policy

This chapter presents an analysis that was conducted before the 5G standards had been finalized, and therefore, it is based on a 4G solution. However, even at that time, several technologies from the 5G paradigm, such as MEC technology, were already being adopted. The chapter proposes integrating satellite communication into 4G networks to provide enhanced broadband connectivity to rural areas in Europe and bridge the digital divide. To evaluate the economic viability of this integration, we use a techno-economic framework based on the Total Cost of Ownership (TCO) model. We simulate two scenarios to assess the viability of the integrated satellite-4G solution and to evaluate the impact of adopting the MEC technology, one without MEC and the other with MEC. Moreover, we evaluate the economic viability of the proposed solution across a wider range of areas based on various population densities.

***

Abstract— Bridging the broadband digital divide between urban and rural areas in Europe is one of the main targets of the Digital Agenda for Europe. Though many technological options are proposed in literature, satellite communication has been identified as the only possible solution for the most rural areas, due to its global coverage. However, deploying an end-to-end satellite solution might, in some cases, not be cost-effective. The aim of this study is to give insights into the economic effectiveness of integrating satellite communications into 4G networks in order to connect the most rural areas (also referred to as white areas) in Europe. To this end, this paper proposes a converged solution that combines satellite communication as a backhaul network with 4G as a fronthaul network to bring enhanced broadband connectivity to European rural areas, along with a techno-economic model to analyze the economic viability of this integration. The model is based on a Total Cost of Ownership (TCO) model for 5 years, taking into account both capital and operational expenditures, and aims to calculate the TCO as well as the Average Cost Per User (ACPU) for the studied scenarios. We evaluate the suggested model by simulating a hypothetical use case for two scenarios. The first scenario is based on a radio access network connecting to the 4G core network via a satellite link. Results for this scenario show high operational costs. In order to reduce these costs, we propose a second scenario, consisting of caching the popular content on the edge to reduce the traffic carried over the satellite link. This scenario demonstrates a significant operational cost decrease (more than 60%), which also means a significant ACPU decrease. We evaluate the robustness of the results by simulating for a range of population densities, hereby also providing an indication of the economic viability of our proposed solution across a wider range of areas.

Keywords— Techno-economics, broadband digital divide, satellite, 4G

2.1 Introduction

While the ICT revolution continues at a fast pace bringing new technologies such as Internet of Things (IoT), machine learning and block chain, half of the people on earth are still without a fast and reliable Internet connection. However, broadband Internet is no longer seen as a comfort, but as a basic service: "Like electricity a century ago, broadband is a foundation for economic growth, job creation, global competitiveness, and a better way of life" [35]. The disconnected users are mostly located in rural areas. In [62], broadband connectivity is seen as the essential protector against the depopulation of these rural areas, thanks to the wide range of available services such as public e-services, e-learning, telemedicine and online social networking.

From a policy perspective, providing high-speed broadband connectivity for rural users is one of the main challenges facing the European countries to satisfy targets announced in the Digital Agenda for Europe [55]. Therefore, the public-policy focus on the digital divide is shifting towards broadband Internet access. In

Europe, only 53% of rural homes are covered by next generation access compared to 80% of overall EU households [31].

Based on [45], important global reasons for not having broadband Internet access are[5]:

a) Internet service is not available;

b) Internet service is available but does not correspond to household needs;

c) Cost of service is too high.

It is clear that when it comes to bridging the broadband digital divide, one should look at both technical and economic parameters. Many technical solutions have been proposed to tackle the absence of broadband infrastructure. A full wireless-based solution with the combination of WIMAX and Wi-Fi technologies was proposed by [50]. This paper investigates the use of WIMAX to provide backhauling support, while Wi-Fi is used to provide access to the end user. This solution cannot provide global coverage of the areas under study as Wi-Fi access points are only installed where houses are clustered. In another study, a fiber-based solution was developed to connect rural areas in India [17]. Using a geo-intelligent algorithm authors aim to identify the optimized GPON tree to reach rural areas from the nearest provider's central office. The algorithm considers the terrain constraints such as difficult hilly terrain, deserts, and coastal regions, and demonstrates an incremental fiber saving compared to the anticipated planning approach of 25.6% to 80%, depending on the terrain characteristics, but is still expensive. Dean Segel in [16] proposed to use a satellite network to augment or substitute the terrestrial connection in rural areas with limited connection. A hybrid-terrestrial solution was proposed where the communication unit is QoS aware, allowing to forward traffic over different backhauling systems depending upon the latency-sensitivity of the traffic. Within the same technology, i.e. satellite, Liolis et al. in [49] have identified several use cases of integrating satellite into 5G networks. One of these use cases deploys satellite as a backhaul network to reach unserved rural areas, while 5G will be deployed as an access network.

Though relevant from a technological perspective, the studies mentioned above only focus on bringing broadband connectivity to the unserved/underserved areas without taking into account the cost of the service, the third cause of not having broadband access.

A variety of studies address both the lack of the network infrastructure and the cost of the broadband service. Van der Wee et al in [63], for example, applied a cost-benefit model on specific scenarios, which are different in term of area type, demand uptake, and revenue scheme. Results prove that deploying a FTTH is only viable in a dense urban area with an aggressive take-up. Frías et al. in [37] introduced a cost model to compare a fixed broadband network (FTTH) with a wireless one (LTE) and applied it to rural Spain. Furthermore, a techno-economic modelling method for choosing the adequate rural broadband access solution was

[5] Note that the ITU mentions four reasons in total, the fourth being: knowledge or skills needed to use the Internet are lacking. This has been addressed by several papers (Preston et al., 2007), (Abrardi & Cambini, 2019), and (Moghaddam et al., 2013), but will not be tackled in this paper.

proposed by Krizanovic Cik, V et al. [60]. A second techno-economic model that applies geo-based multi-objective optimization to find areas with the highest concentration of unserved/underserved users in the USA at the lowest cost to service providers was developed by Durairajan et al. in [19]. A converged optical-wireless architecture for interconnecting rural areas in Europe proposed by Polykalas et al. in [55], was consolidated by a techno-economic study for connecting the rural areas and islands complex in Greece. This study proved that, only by using 30% of public funding and 30 % of cost sharing (i.e. infrastructure sharing between telco and grid operator), the entire Greek islands complex can be covered by a high-speed fixed-mobile network. Recently, Ioannou et al. in [43] carried a comparative techno-economic study between FTTdp (Fiber To The distribution point) and LTE networks in order to estimate the needed ARPU for rural Greece, finding that both solutions are too expensive for the citizens of rural areas (ARPU for FTTdp is 332 euro and for LTE is 66 euro for 30 Mbps download speed). On a larger scale, Feijóo et al. in [36] used a techno-economic framework to calculate the needed total investments to fill the broadband gap for urban, suburban and rural geo-types in EU-28. This framework proved first, that in term of broadband investment, the main gaps are occurring in rural areas, and second that around 50 billion euro is needed to bridge these gaps. Though relevant, these models did not look at the willingness to pay of the end user. We should bear in mind that in most rural areas, inhabitants have a lower income and the price of the service is crucial for its adoption.

From the discussed literature, we can deduct that rural areas are still underserved and that, while technological solutions exist, they frequently are too expensive for market players to invest in. This paper therefore combines technical aspects with techno-economic evaluation that takes into account Willingness to Pay, in order to propose a viable solution for connecting remote rural areas, specifically the European white areas. The proposed solution is a hybrid satellite-wireless solution in which a geostationary satellite (GEO satellite) network is to be deployed as a backhaul network. The GEO satellite is chosen thanks to its global coverage and high capacity. For the access network (fronthaul), we select the 4G technology, as it is not only one of the Next Generation Network (NGN) technologies but also provides up to approximately 100 Mbps as a peak data rate. In order to evaluate the economic viability of the proposed solution we use a Total Cost of Ownership (TCO) model and compare the Average Cost Per User (ACPU) to the Willingness To Pay (WTP) of the end users derived from desk research.

The article proceeds by presenting the actual conditions of broadband connectivity in Europe in section 2.2. In section 2.3, the proposed converged satellite-4G solution is presented. Section 2.4 highlights the cost model, a simulation of the model based on a hypothetical use case is presented in section 2.5. Section 2.6 evaluates the robustness of the results by simulating for a range of population densities and by assessing the impact of the caching data rate. The last section, section 2.7, summarizes the results and discusses future work.

2.2 Broadband connectivity in Europe

There is no universal consensus on the minimum download and upload speeds for a service to be classified as a broadband. The European Commission mentions both 144 kbps and 3 Mbps, while the FCC (Federal Communication Commission) in the United States defines broadband at speeds of minimum 25 Mbps [18]. However, a consensus exists on the key role that broadband access plays in job creation, social and digital inclusion, society's well-being and the economic growth. On the European level, there is a current debate on broadening the concept of the Universal Service Obligation (USO) to include broadband access [18], [8] and [67].

The idea behind a USO is to ensure all citizens have the right to the same services independently of their location. In Europe, the universal service from a telecommunication perspective is outlined by the directive 2002/22/EC "Universal service and users' rights relating to electronic communications networks and services (Universal Service Directive)" [22]. In this Directive, the European Commission (EC) states that "The concept of universal service should evolve to reflect advances in technology, market developments and changes in user demand". To reflect this dynamicity, the EC states in the same directive (under Article 15) that the Universal Service Directive (USD) should be reviewed in the light of not only the technological development but also economic and social ones. Following this decision, the EC has reviewed the directive three times since 2005. The main points included in these three consultations are:

- In the 2005 consultation on broadening the scope of the Universal Service Directive (USD) to include broadband, the EC concluded that only a minority of the European citizens are using broadband services (e.g. penetration rates of around 25% in countries such as the Netherlands and Denmark [2]. As a result of the low penetration rate, the criteria to include broadband services in the USO were not satisfied.
- The results of the 2010 consultation were different from the one of 2005 giving the significant increase of the broadband penetration (24% at EU 25 in 2009 compared to only 8% at EU 25 in 2004 [32]). However, the EC concluded that: "At this stage, it would not be appropriate to include mobility or mandate broadband at a specific data rate at EU level" [23], mainly due to the huge investment required to ensure the availability and affordability of broadband to the whole Europe. On the other hand, in order to minimize the possible market distortion resulting from broadening the scope of USO, the EC advices Member States to use public intervention tools to guarantee broadband availability outside the framework of the USO.
- In the 2016 consultation, [29] and [18] estimated the cost of providing the primary basket (consisting of 9.6Mbps connectivity) by 2020 for all the EU citizens to be 13.7 billion Euro. In addition to that, the cost of the social tariffs of broadband (considered at the same level of social telephony tariffs) would be around 147 million Euro annually. The study concluded that the most

efficient mechanism of financing the broadband universal service would be public funds. They argue this choice by the fact that the entire society would benefit from the availability of broadband on both the economic and social sides. Currently, the USF (Universal Service Funding) is financed by the telecommunication sectors together with public funds. Therefore, it seems unfair to provide broadband connectivity using USFs, because other stakeholders will also benefit from a wider broadband access like the online service providers etc…. Hence, other means of funding have to be investigated.

As a conclusion of these reviews on the scope of the universal service, the EC, Member States, telecommunication operators and all the broadband contributors agreed that broadband connections should be made available at an affordable price to all the EU citizens, however without including an exact minimum bitrate in the USO. Nevertheless, the EC still refers in its reports e.g. [27], to areas served by less than 2 Mbps as areas without broadband services[6].
Therefore, the EC has set a clear policy towards the deployment of high-speed broadband giving its socio-economic benefits. In 2010, the EC announced the Digital Agenda for Europe (DAE) which consist of 3 main targets: (1) availability of broadband for all Europeans in 2013, (2) deployment of 30 Mbps broadband capability to all Europeans by 2020, and (3) adoption of 100 Mbps broadband by 50% of European households by 2020[24]. Following the announcement of the Digital Agenda for Europe 2020, the European Commission (EC) has adopted a new strategy towards its broadband program called "Connectivity for a European Gigabit Society 2025" in September 2016 [28]. The main targets (for 2025) announced in this agenda were:

4 Provide gigabit connectivity for all of the main socio-economic drivers.
5 Guarantee an uninterrupted 5G coverage for all urban areas and major terrestrial transport paths.
6 Access to connectivity offering at least 100 Mbps for all European households.

The last ambitious target of this agenda aims to the global coverage of Europe by at least 100 Mbps before 2025. As mentioned above, the rural EU average coverage for NGA (Next Generation Access) technologies at 39.2%, continues to be significantly lower than the total NGA coverage (76.0%) [31]. However, as urban areas start to reach the saturation stage of NGA coverage, deployment should shift towards rural areas to bridge the gap. As mentioned before, billions of euro should be invested for deploying network infrastructure. The most important parts of these investments are foreseen to come from the private sector, especially from the broadband providers. However, the private sector considers the rural

[6] In this paper, we consider basic or universal broadband services to have at least 2Mbps as download speed.

areas as a non-viable market because the cost of deploying a network outside the urban and suburban areas is on average 80 % higher [59]. Moreover, a study carried by the European investment bank showed that 41% of the entire broadband investment should be allocated to rural areas in order to offer high-speed broadband to all European citizens. A similar study by Point Topic even mentions 63% instead of 41% [54].

Therefore, the EC has encouraged Member States to use public financing in line with the European Union competition and State Aid rules to achieve the speed, coverage and growth targets of the Europe 2025 agenda. The EC has defined a set of conditions that should be satisfied in order to acquire subsidies for the broadband rollout. These conditions are built on the classification of areas into white, grey and black. White areas are where there currently is no broadband provider and unlikely to be in the near future; grey ones are areas where there is only one broadband provider and black areas are covered by more than one broadband provider [25]. Hence, white and grey areas are the only eligible areas to receive the State Aid funding.

Several Member State proposals have been approved by the EC concerning the public funding of broadband deployment. More specifically, 60 proposals were approved during the last three years by the EC [26], most of them detailed how they aim to deploy new broadband infrastructure while assuming the possible rate of the service demand.

However, there will still be a large need for private funds in reaching the target.

On the other hand, in order to reach the 100% coverage in rural areas, not only the network rollout and its cost should be considered but also the price of the service and the willingness to pay (WTP) by the end user, as the ability of the operators to invest is driven by the end-users' willingness to pay [47]. Ovando, C. et al. in [52] proved that for rural areas only a very high take-up could make the deployment of broadband networks feasible, while the demand of broadband services is very sensitive to price based on the socio-economic specifications of those areas. Moreover, not only the price of the service can affect the adoption of broadband but also other factors such as the age of end users as proven in [46]. Finally, Glass, V. et al. in [39] believe that marginal increases in speed do not attract customers anymore, but that the inclusion of new services such as video could lead to an increase in take-up and willingness to pay. It is exactly this diversification of services and the offering of new services such as video streaming, video gaming and virtual reality VR applications, that requires a high-speed broadband.

Recently, Liu et al. in [68] proposed a model to estimate the willingness to pay for broadband Internet speed in the United States. Results of this study reveals a concave households' valuation of the broadband bandwidth. For example, in order to increase the bandwidth from 4Mbps to 10Mbps, customers are willing to pay 2.34$ per Mbps (14$ in total). Yet, to increase the bandwidth from 100Mbps to 1Gbps they only willing to pay 0.2$ per Mbps (19$ in total). In addition to that, a noteworthy result found in this research is that costumers value a decrease in latency from 300-600 milliseconds to 10 milliseconds by around 8$/month, which is equivalent from moving from Internet over satellite service to 4G services (we

will come to this later on this paper: section 2.5.3). It is important to mention here that this data is coming from two surveys that took place in the United States. Therefore, extrapolating these results to the European level should take into account the difference in term of willingness to pay between the American and the European citizens, but at least, they give an order of magnitude of the correlation between the WTP and the broadband speed.
According to [30], the average European price of a mobile broadband service at 2 Mbps speed (which results in 1.6 GB data per month assuming that users are active only 9 hours per day and consume in average 80% of the bandwidth) is between 23 to 31 euro per month. This price should be valid for both urban and rural areas. However, if we aim for a good service adoption, this price has to be decreased by 15% to be adopted by rural inhabitants as well as non-adopters in urban areas according to a wide survey carried out in the USA [11]. We should hence account for an average WTP of 19 euro per month for a decent broadband connection in a rural area.

In the light of this literature overview, we can infer that solutions that optimize for both technical (broadband technology) and economic (demand and willingness to pay) requirements for network deployment in unserved areas are needed and, to the best of the authors' knowledge, not readily available in literature. In the next section, we present our solution that covers both requirements.

2.3 Converged Satellite-Wireless architecture for interconnecting rural areas

We propose a converged satellite-wireless solution that takes into account (1) the technical requirements of high-speed broadband and (2) the economic constraints, being the cost of the network deployment and the willingness to pay of the rural inhabitants. Therefore, we investigate in the first subsection (section 2.3.1) the status of satellite broadband. In section 2.3.2, similar proposed solutions in literature are discussed. We finalize this section by detailing the technical description of our solution.

2.3.1 Status of satellite broadband

Using satellite for broadband connectivity has several advantages: it is the only readily available technology that has a worldwide coverage, while also being capable to provide high-speed capacity. For that reason, the EC has considered its first target of the DAE, which is 30Mbps broadband speed for all citizens by 2013 as achieved [24] relying on the global coverage of satellite.
However, according to [45] there are two main challenges facing the deployment of satellite solutions in rural areas:

1 The unawareness of the end user about the availability of the Internet satellite providers, which justify why broadband satellite in Europe has a market share below 1% compared to other technologies [40]. As a solution, the EC together with the satellite providers have developed an online tool (www.broadbandforall.eu) in order to establish a connection between users and the local satellite suppliers.
2 The insufficient political support for the satellite broadband. In fact, the State Aid documentation refers for the next generation solutions as wired or mobile solutions and satellite solutions have been marginalized. On the other hand, there is a general consensus among satellite operators that a 30 Mbps Service Level Agreement (SLA) cannot be achieved at any reasonable and affordable price without some intervention.

In fact, an end-to-end satellite solution may not be viable, for many reasons. First, the end users need to buy a satellite receiver (well known as VSAT) which is around 800 euro without counting the price of the installation [64]. In addition to that, the price of the broadband service itself is very expensive, about 99 euro for a volume of 2GB per month, with maximum speeds of only 10Mbps/2Mbps (download/upload) [64]. Giving that, we can think what the price for 30 Mbps would look like.

As a conclusion, satellite operators are not able to expand their networks to meet the goals of the EC agenda, given the present funding policies. On the other hand, the EC will not achieve the goal of 30Mbps coverage for all Europe without the help of satellite operators.

As a possible solution for this “dilemma”, we suggest in this paper to combine the 4G networks as fronthaul network with the satellite networks as a backhaul. The use of 4G networks aims to decrease the cost of the radio access network. In addition to that, satellite networks from a technical perspective are the most eligible to serve as a backhaul networks for rural areas given their advantages as presented above.

2.3.2 State of the art of Satellite-4G network solutions

Section 2.1 introduced some papers that support the use of satellite as a backhaul network to serve rural areas [16] and [49]. They motivate the choice of using satellite technology as a backhaul network by the expensive cost of upgrading the terrestrial backhaul route, which could cost millions of dollars, yet the potential new revenue cannot cover this cost.

From a technical perspective, several sources studied the compatibility between 4G and satellite networks [21], [44], [61] and [69]. Some technical challenges generated by the combination of satellite and 4G networks have been addressed as well. Zangar et al. in [70] addressed in their work the “key challenge” in the 4G-Satellite integration which is the radio resource allocation policy. The main challenging aspect according to [1] in order to ensure the smooth combination of 4G and satellite networks is the management and the handling of the TCP connection. The same authors recently analysed the bandwidth aggregator techniques for hybrid satellite-xDSL links [1].

2.3.3 Proposed Satellite-4G solution

This paper presents a solution that takes advantage of the satellite communication's global coverage and uses it as a backhaul network. In addition, the solution relies on a 4G access network, as one of the NGA technologies, to be deployed as a fronthaul network. The proposed solution aims to provide broadband connectivity to white areas where there is no broadband provider and it is unlikely to be in the near future. However, in areas where there are already basic broadband services (but not NGN), satellite solutions can complement the existing terrestrial solution as described in [49]., but this case is out of scope of this paper.
The complete network solution is composed of:

1. A 4G core network that treats and processes the offered services,
2. A satellite gateway connected to the 4G core network via a fiber connection, which is responsible for forwarding the traffic from the core network to the radio access network (RAN) via a satellite link,
3. A satellite terminal installed near the RAN that receives the traffic from the satellite gateway via the satellite link and sends it to the RAN and vice versa,
4. Finally, a RAN which consists of eNodeBs (evolved Node B, i.e. mobile base stations), responsible of carrying out the traffic from the 4G core to the end user and vice versa.

The network architecture is presented as follows:

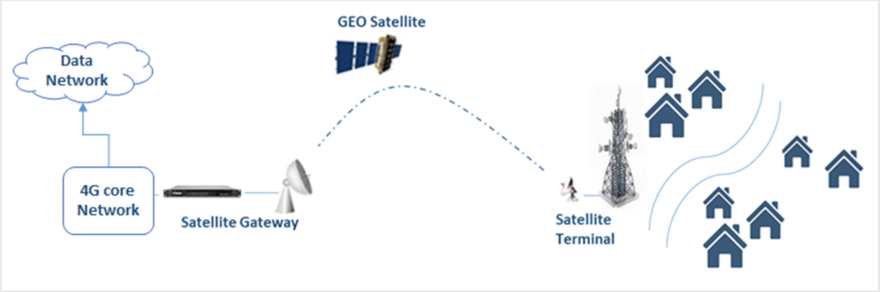


Figure 2-1 Network architecture of the integrated satellite-terrestrial solution

To study the viability of the integration of satellite communication within the 4G mobile network, a cost model is defined and discussed in the next section.2.4

2.4 Cost Model

The proposed model takes into account both the Capital (CAPEX) and the Operational Expenditures (OPEX), designed for converged networks, and considers a planning horizon of 5 years. This model aims to calculate the Total Cost of Ownership (TCO) as well as the Average Revenue Per User (ACPU) for the studied scenario. This model aims to study the economic feasibility of the 4G-satellite solution in rural Europe. In section 2.4.1, we will discuss in more detail the structure of the model. The mathematical formulation of the model is presented in section 2.4.2.

2.4.1 Model structure

The main inputs of the model are the bill of materials (BOM), the number of users, the minimum bitrate per user, the average margin of profit of telecoms operators and the time horizon of the project. Those inputs feed into a cost model that consists of four sub-models in alignment with the network architecture components presented in the previous section. The first sub-model is designed for the edge site. It incorporates the CAPEX of the radio access network (RAN), the capex of the satellite terminal, the common capex and the OPEX of all edge components. The second one models the satellite network used for the backhaul; both CAPEX and OPEX are taken into account. The third block builds the model of the costs for the 4G core network and the last block in the diagram englobes all overhead costs. After the calculation of the CAPEX and OPEX for all these blocks, the TCO can be derived. Hence, given the TCO as well as the number of users, the ACPU can be derived as an output of the model. The structure of the model is presented in Figure 2-2.

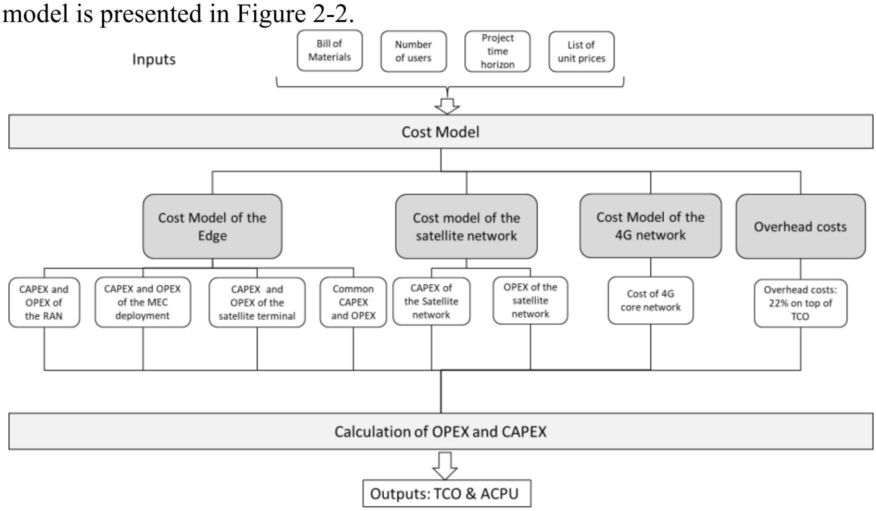


Figure 2-2 Cost model structure.

2.4.2 Mathematical formulation of the model

The previous section has presented, from a high-level perspective, the structure of the proposed cost model. In this section, we will detail each sub-model formula used to calculate both CAPEX and OPEX parts. An overview of abbreviations can be found in Table 2-1.

Cost model of the edge site

The cost of the edge incorporates the CAPEX of the RAN, the CAPEX of the satellite terminal, the common CAPEX and the OPEX of all the edge components. The number of equipment needed depends on the dimensioning process of the site, which in turn depends on the bitrate to be provided in the site (Equation 1). The CAPEX of the RAN (Equation 3) depends on the number of eNodeBs to be deployed, which is the maximum of the number calculated based on the bitrate and the coverage area of the served site (Equation 2).

Equation 2-1

$$Br_s = N_u \times Br_u$$

Equation 2-2

$$N_{enb} = MAX({}^{Cov_a}/_{Cov_{enb}}, {}^{Br_s}/_{Bw_{enb}})$$

Equation 2-3:

$$Capex_{RAN} = N_{enb} \times C_{enb} + C_T + C_{Inst}$$

Equation 2-4:

$$Opex_{RAN} = Site_{rental} + (P_{eNBs} + P_{ST}) \times C_{watt} + M$$

Furthermore, the CAPEX of the satellite terminal consists of the satellite terminal equipment, which is calculated based on its link capacity with the satellite. The common CAPEX incorporates all common capital costs needed to build the edge infrastructure. Finally, the OPEX of the edge (Equation 4) consists of the cost of the power consumption of all the equipment, the cost of the site rental per year and the maintenance costs.

Cost model of the satellite network

The cost model of the satellite network consists of two main parts. First, the CAPEX of the satellite gateway, which is the cost of the equipment and the building required to deploy the satellite gateway. Second the OPEX of the satellite network (Equation 6), which is the cost of the satellite capacity (Equation 5), in addition to the cost of the maintenance and the power consumption.

Equation 2-5

$$C_{satCapS} = (Br_S + Br_T) \times 12 \times C_{satMbps}$$

Equation 2-6

$$Opex_{sat} = C_{satCapS} + P_{satGat} \times C_{watt} + M$$

Cost model of the 4G core network
For modelling the 4G core network, an estimation of the cost of the core network per user should be made. Given this estimation in addition to the number of users, the cost of the 4G core network can be calculated. This model uses results found by Bouras et al. in [10], which calculate the cost of 4G core network deployed to serve 100,000 users (corresponding to about 100 base stations).
Additional costs
We have built a realistic cost model using the following assumptions for the additional costs:

- Hardware installation cost is 15% of the hardware costs [12]. The cost of the hardware installation is part of the CAPEX costs and it is expressed by the following formula:

Equation 2-7

$$C_{Inst} = 15\% \times C_{Hw}$$

- Maintenance cost is 10 % of the CAPEX costs [12]. Maintenance costs are counted in the OPEX costs.

Equation 2-8

$$M = 10\% \times CAPEX$$

- In most cases, the overhead cost is defined as the cost of marketing, helpdesk, human resources, finance etc. According to [6], it is around 22% on top of the sum of the CAPEX and OPEX costs.

Equation 2-9

$$Ovhd_C = 22\% \times (CAPEX + OPEX)$$

Total Cost of Ownership (TCO)
The Total Cost of Ownership (TCO) of the proposed solution is counted as the sum of the CAPEX, the OPEX of 5 years and the overhead costs of 5 years.

Equation 2-10

$$TCO = CAPEX + \sum_{t=1}^{5} OPEX(t) + Ovhd_c(t)$$

Table 2-1 Nomenclature

Nomenclature	***Designation***	**Nomenclature**	**Designation**
Br_u	bitrate per user	H_u	average active hours per user
Br_S	bitrate per site	N_u	number of users
Cov_a	coverage area	$C_{satMbps}$	cost of 1 Mbps per month via satellite link
Cov_{enb}	eNB coverage	$C_{satCapS}$	cost of satellite capacity for a site S
Bw_{enb}	eNB bandwidth	Br_T	bitrate per site for traffic control and overhead
N_{enb}	number of eNB	**M**	maintenance costs
C_{enb}	cost of eNB	P_{satGat}	power consumption satellite gateway per year
C_T	cost of the tower	C_{watt}	cost of energy
N_s	number of servers	N_{stg}	number of storage
Mec_{mng}	cost of MEC software management	C_{stg}	cost of storage
C_s	cost of server	**N**	number of sites served by the satellite gateway
C_{Inst}	cost of hardware installation	C_{Hw}	cost of the hardware
$Ovhd_C$	overhead costs	P_{eNBs}	power consumption of the eNBs per year
$Site_{rental}$	Cost of renting the site for deploying the eNBs and the satellite terminal per year	P_{ST}	power consumption of the satellite terminals per year
$C_{satCapS\_Caching}$	cost of satellite capacity for a site S with caching data on the edge		

2.5 Simulation: hypothetical use case

To evaluate the proposed solution from a techno-economic point of view, we define a hypothetical use case, which will first be described. Inputs used to run the cost model are listed in the second section and finally results are analyzed in the third section.

2.5.1 Use case description

The case consists of a satellite backhaul connected to a cell tower located in a rural area in the EU covering two villages about 5km apart connected via a rural main road. The villages are home to 350 families, with an average of 3 users per home. The predominant traffic on the cell is eMBB (enhanced Mobile Broadband).

From the mathematical formulation of the model presented previously in section 2.4.2, we can conclude that the number of users as well as the minimum bitrate required per user are the main cost drivers of our model. First, they affect the dimensioning process; see (Equation 2-1) and Equation 2-2). In addition, the OPEX costs (Equation 2-5) are directly driven by the bitrate per site.

In order to generate realistic results taking into account the bitrate per user cost driver, there are two ways to proceed. The first one is to forecast the average consumed mobile data traffic per user (i.e. monthly download volume) in the considered timeframe (2020-2024) and then calculate the bitrate per user that generates this amount of mobile data traffic. The second option is to set an initial bitrate per user according to the DAE 2020 target, namely 30 Mbps. We can assume that the first case corresponds to a conservative scenario and the second one to an aggressive scenario, and hereafter their description:

Conservative scenario

In order to have an idea about the future mobile data traffic, we refer to the well-known Cisco VNI report [14]. It forecasts that in 2022, the monthly mobile data traffic for Western Europe will be 8.8GB per user. To know the user bitrate needed to generate this amount of data we rely on Analysis Mason data used in the BATS project report [7]. The calculation takes into account the number of hours during which the user is active, then it results that each 1GB per month corresponds to an average busy hour data rate of 7.8 kbps. Hence, 8.8 GB per month results in 68.64 kbps per user as an average busy hour data rate. However, this bitrate does not meet the definition of the broadband speed set by the EC. Thus, we consider the minimal speed in this scenario the one of the universal broadband service, being 2 Mbps. Hence, in this scenario we assume that the bitrate per user (3 users per family, 350 families) is 2 Mbps, thus taking into account the active user rate of 80% and the number of active hours of 9 per day, the resulted average busy hour data rate is 630 Mbps per site.

Aggressive scenario

In order to guarantee a good Quality of Experience (QoE) for the end user, the EC requires a bitrate of 30 Mbps per user for high-speed broadband services, as explained above. Hence, in this scenario we assume that the bitrate per user (3

users per family, 350 families) is 30 Mbps, thus taking into account the active user rate of 80% and the number of active hours of 9 per day, the resulted average busy hour data rate is 9450 Mbps per site.
For all the defined scenarios, on top of the site bitrate, we expect 10% traffic and control plane overheads.

2.5.2 **Model inputs**

Inputs used to run the cost model are presented in the following tables. General inputs of the model are described in Table 2-2. Parameters used for modelling the edge site, which consists of two parts: the RAN and the satellite terminal, as well as the satellite network and the 4G core network are presented in Table 2-3.

- **General inputs**

Table 2-2 General inputs

Parameter	Conservative scenario	Aggressive scenario
Study period	5 years (2020-2024)[7]	
Area	78.5 km^2	
Average busy hour data rate per site	630 Mbps	9450 Mbps
Average active hours per user	9 [8]	
Active user rate (%)	80 [2]	
Cost of 1 kWh (€)	0.114 [33]	
Conversion rate ($ to €)	0.9	
Time Value of Money	5%	

[7] Use case definition
[8] These assumptions are derived based on Internet research, so may be realistic but not precise.

- **Network inputs**

Table 2-3 Network inputs

Parameter	Value
Edge inputs	
RAN	
Macro cell: 3 antennas, 1BBU, Software upgrades and maintenance	25-30K €
Macro cell bitrate	Maximal deployment is 5 * 3 sectors per eNB, (5*3*140= 840Mbps) [58].
building, rigging and materials (tower 10m)	10k $ [20].
Power consumption	18 144 kWh per year [58].
Satellite terminal (ST)	
Cost of ST	4K $ [38].
Capacity of the satellite link/ST (Mbps)	150 [38].
Satellite terminal power consumption	438 kWh per year
Air conditioning	~500 € [5].
Common power consumption: cooling etc...	30835 kWh per year
Edge maintenance	10% of CAPEX [12].
Satellite Network inputs	
Satellite capacity cost ($/Mbps/month)	15-7 (2020-2025) [20].
Satellite gateway infrastructure (€)	Cost of satellite gateway is included in the cost of satellite capacity
Maintenance	10% of CAPEX [12].
4G core network inputs	
Cost of 4G core network (€)	2200K: deployed to serve 100 BS and each BS

	serve 1000 users [10].
Cost of 4G core per user (€)	22

2.5.3 Results and interpretation

Results for the baseline scenario

As a first step, we simulate the proposed architecture for the scenarios discussed previously: the conservative and the aggressive scenario, based on collected inputs presented above and assumptions discussed in section 2.4.2.
Results of the TCO for the three defined scenarios are presented in Figure 2-3.

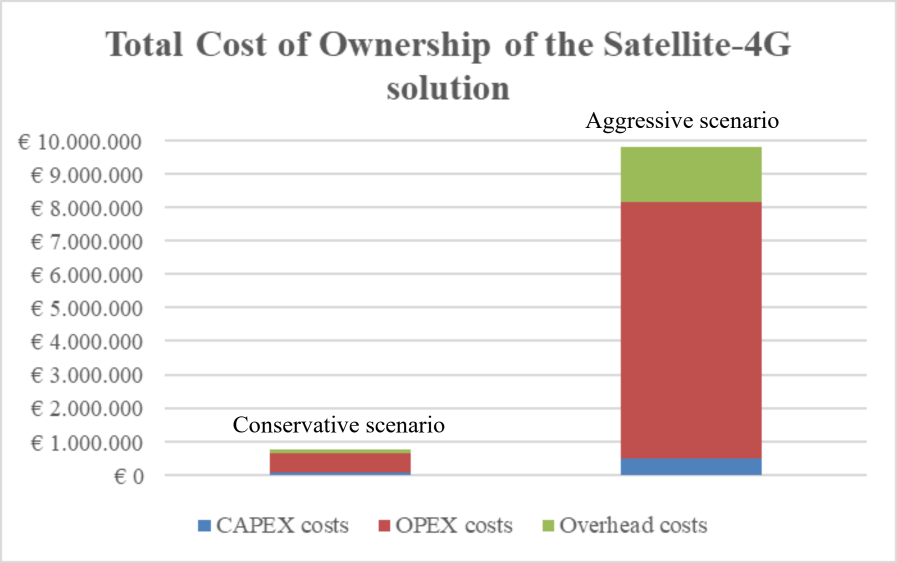


Figure 2-3 Comparison of the Total Cost of Ownership for the considered scenarios

The monthly ACPU for the conservative and aggressive scenarios can as such be derived: 12 and 155 euro respectively. The ACPU for the aggressive scenario (2Mbps) is not expensive comparing to the willingness to pay of the rural inhabitants (19 euro as discussed in section 2.2). Hence, the satellite-4G solution is viable for rural areas with a modest bitrate per user. However, if we opt for a good QoE (i.e. the aggressive scenario), the ACPU results in 155.56 euro, which is way higher than the WTP (especially if we take into account that the ACPU calculated here does not yet include a profit margin for the telecom operators or service providers).

In the view of optimizing these results, we examine the resulted cost components in more detail, we find that the OPEX are very high, as clearly observed in Figure 2-3. This is due to the satellite capacity (backhaul) that should be paid monthly. One of the solutions proposed to decrease these OPEX is to cache a percentage of the popular content on the edge and by doing so we decrease the amount of traffic that needs to be carried out via satellite link, in turn decreasing the operational costs [66].

Simulation with caching data on the edge to decrease OPEX

1. Caching technology

Tremendous growth of the video content request is foreseen for the next years. According to the Cisco VNI Forecast Highlights Tool, Internet video traffic will be 80% of all consumer Internet traffic by 2022 comparing to 73% in 2017, which corresponds to 29.5 EB (Exabyte) per month by 2022 compared to only 8.9 EB per month in 2017 [15]. This growth is driven by the expansion of VoD (Video on Demand) libraries. According to the same forecasting tool, consumer IP VoD traffic will reach 3.8 EB per month by 2022, yet, it was only 2.6 EB per month in 2017.

The traditional centralized networks cannot mitigate this exponential growth of user demand giving the heavy load on the backhaul links and the long latency [65]. In order to accommodate the huge bandwidth requirements resulting from the massive demand of the high-definition video content, new network architectures based on caching strategies have been proposed. Essentially, caching techniques aim to store a duplicated copy of the most popular content in the network edge. In this way, the end user does not have to download it from a central location in the network. Hence caching techniques transform the bandwidth requirement into a storage requirement [9]. Several papers in literature optimize content caching techniques and algorithms [66], [65] and [57].

The main question that should be addressed when a caching-based solution is chosen, is how much data we have to store in the edge. Cha et al. in [13] proved that the Pareto law is valid for video content consumption in Internet. They found that 10% of the most popular videos on YouTube account for approximately 80% of the views. On the other hand, the 90% of the remaining video content has only 20% of views (the same results were found for Daum (a video service in Korea) as well).

2. Caching scenario specifications

Two reasons are behind the use of the caching technology in this scenario. The first one is to decrease the amount of the traffic that needs to be sent via satellite link, and by doing so, decrease the cost of the satellite capacity and hence the OPEX. Furthermore, caching popular content on the edge will not only solve the issue of increasing demand for bandwidth, but will also decrease the latency of the video traffic, as the content is placed closer to the end users. Therefore, our second reason is to decrease the latency for VOD service, and as discussed previously in section 2.2, costumers significantly value the decrease of latency.

In the network architecture presented in Figure 2-1, we need to add Multi Edge Computing (MEC) infrastructure on the edge site [71]. The role of the MEC is to cache a percentage of the popular content locally on storages on the edge and to communicate with the base station to receive users' requests and send back the corresponding content. An intelligent algorithm is there to update the cached data according to the frequency of usage and downloads of new content. We assume that the Pareto law is applicable here (as argued in the previous section). Hence 20% of popular content will be stored and 80% of user video requests are served from cached data. Therefore, 80% of the user's traffic is video (as argued previously in section 12.5.1), and 80% of the video requests can be served from the cached data (which is 20% of the popular content), resulting in 64% of the user traffic to be served from the local cache. As this assumption directly affects the OPEX, a sensitivity analysis on the caching rate is given in section 2.6. The new network architecture is presented in the following figure:

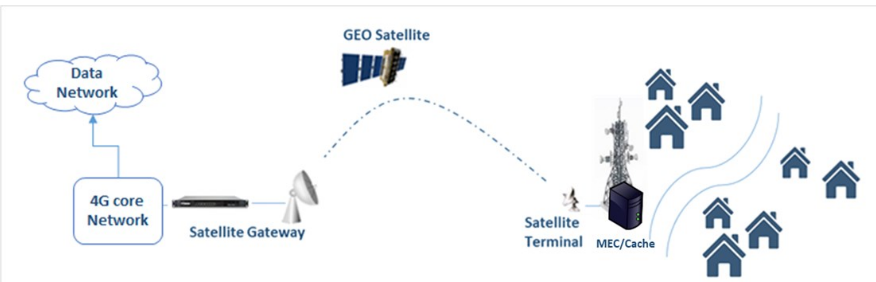


Figure 2-4 Network architecture of the proposed solution with the MEC deployment

In addition to changes in the network architecture of the proposed solution, there are also changes in the cost model structure (Figure 2-2). We need to take into account in the new model both the CAPEX and OPEX of the MEC deployment. Based on how much data can be cached on the edge, the number of storage equipment and servers required is calculated, and then the cost of the entire MEC infrastructure is derived based on the following equations:

Equation 2-11

$$Cost_{MEC} = CAPEX_{MEC} + OPEX_{MEC}$$

Equation 2-12

$$CAPEX_{MEC} = N_S \times C_S + N_{stg} \times C_{stg} + \frac{Mec_{mng}}{N}$$

The main difference between the two scenarios (without and with caching data on the edge) is the cost of the satellite capacity. Based on the number of user requests that will be served from the cached data R_{cd}, Equation 2-4(Equation 2-4) becomes the following equation:

Equation 2-13

$$C_{satCapS} = ((1 - R_{cd})Br_S + Br_T) \times 12 \times C_{satMbps}$$

The model inputs related to this deployment are recapitulated in the table below:

Table 2-4 MEC inputs

Parameter	Value
Data Caching rate on the edge	20% - 80%: 20% of popular content will be stored and 80% of video user requests are served from cached data.
Popular content volume	YouTube catalogue: 10ˆ9 MB as a total volume of popular content, which is based on 10ˆ8 MB YouTube movies, each having a size of 10MB [57].
MEC infrastructure costs: (Lenovo, 2018) • Server • TruDDR4 memory: 64 GB of RAM • Physical storage: 10 TB of SAS disks on 14 disks • 8 vCPU at 2Ghz • License cost	~700€ ~800€ 2800€ Free for VMWare 12k€
Management software: • 8 vCPU at 2,6Ghz, • 32 GB of	60 k€ to cover 50 sites[9]

[9] These assumptions are derived based on Internet research, so may be realistic but not precise

RAM, • 600 GB of disk, with the license	
MEC infrastructure power consumption	3022.2 kWh per year

When we re-run the model, while assuming that popular content is cached on the edge, results on the TCO for the conservative and the aggressive scenarios for the baseline and caching scenario are presented in Figure 2-5. A quick comparison between the TCO of the solution with and without caching, shows that the TCO with caching is approximately one third of the one without caching. This comparison shows an increase in the CAPEX costs in the case of using the cache due to the cost of the MEC infrastructure, but a much larger decrease in OPEX for the satellite backhaul. As a conclusion, the use of caching popular content on the edge with the assumed rate of 64% as explained previously, proves a reduction of 60.5% in OPEX costs, resulting in an important decrease (57.2 %) in terms of ACPU. A comparison between the obtained ACPU for the aggressive and conservative scenarios for both cases with and without caching is recapitalized in Table 2-5.

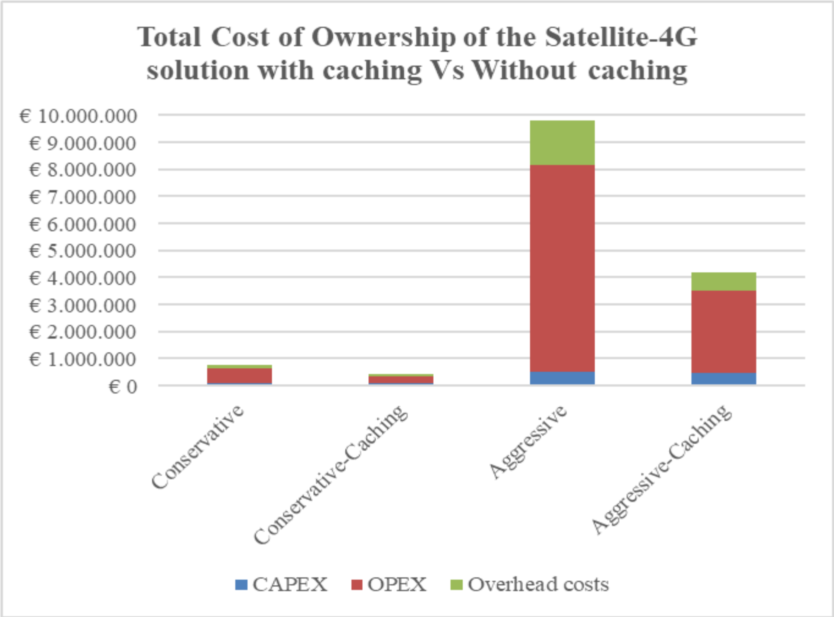


Figure 2-5 Total expenditures of satellite-4G solution with caching data on the edge Vs the baseline scenario

Table 2-5 ACPU with and without caching

ACPU	Conservative scenario	Aggressive scenario
Satellite-4G without caching	12.1 €	155.6 €
Satellite-4G with caching	6.7 €	66.6 €
ACPU reduction rate	44.6 %	57.2 %

From the table above, we can extract several conclusions:

- The ACPU reduction rate resulting from the use of the caching technology is more than 55% (for the aggressive scenario).
- The ACPU reduction rate increases when the bitrate required per site increases, this is explained by the fact that the bigger data rates we offer, the more requests are served from the cached data.
- The ACPU for the conservative scenario without caching is economically viable (compared to the WTP discussed previously). Yet the ACPU for this scenario with the use of caching concept is very cheap. Still, one needs to take into account the conversion of ACPU to actual ARPU

(prices), using profit margins for the telecom operator and service providers.
- The deployment of the solution without using the caching concept is not cost-effective in the case of the aggressive scenario.
- A good speed thus a good QoE can be offered to the inhabitants of white areas thanks to the satellite-4G solution, in case a caching of the popular content is deployed in the edge.

These results can give insights to the network operators on the deployment and pricing strategies that they can follow. For example, it might be that operators would start as a first step with a modest bitrate per user in order to provide the universal broadband connectivity to these unserved areas (white areas) in the EU. The ACPU generated in this case is around 6.7 euro, yet the willingness to pay of the end user is 19 euro so operators have a good margin to define their pricing strategy. By choosing the right margin of profit, operators can move forward smoothly to the aggressive scenario. In this way, they can cross-subsidize the offer of high-speed broadband by their first offer, which is the basic broadband service. On the other hand, the combination of the mobile and the satellite networks could be eligible to receive public funding according the EC criteria presented previously, as satellite networks were somehow marginalized from receiving public funding.
To validate our results, we compare them to the findings in [43]. In this latter, in order to cover rural areas in Greece with a population density of 33 inhabitants per km^2, the required ARPU for the FTTdp solution is 332 euro for. To make a fair comparison, we will compare our generated ARPU for the aggressive scenario (i.e. a data rate of 30 Mbps) for the same population density. The generated ARPU for the Satellite-4G is only 67 euro, taking into account the average margin profit of 11% on top of the ACPU (Investopedia, 2018). We can hence prove that the combination of backhaul satellite and fronthaul 4G networks are more cost effective than the FTTdp technology.
Moreover, a quick back of the envelope calculation of the fiber-4G solution presented previously in the introduction section, shows that the CAPEX to deploy such a solution are more than 5M euro, which results on an ACPU of 67.7 euro per month for a timeframe of 5 years, this without counting OPEX. Hence, the combination of the fiber and 4G mobile networks is not the right broadband solution to cover rural areas (a table summarizing this back of the envelope calculation is drafted in the appendix).
On the other hand, the price of the satellite-only broadband solution presented in [64] is 99 euro for 10 to 20 Mbps as a download speed for a total 2 GB data per month. Making the extrapolation, this would result in a price of almost 200 euro per month for providing 30 Mbps. In addition, we have to count the cost of the satellite terminal that the end user has to buy and its installation and maintenance

costs. Hence, the total price per month is also more expensive than the 4G-satellite solution, even if considering the baseline scenario without caching.
The ACPU of the conservative scenario could motivate operators to offer more speed to the end user while bearing in mind that offering broadband connectivity in rural areas will come up with different services, e.g. tech-agriculture, farm business, e-services etc. This means new sources of benefits for the operators, which could make the use case more cost-effective than providing only the universal Internet service. Offering diverse services could be more profitable with the deployment of network slicing within the upcoming 5G networks due to the virtualization of network functions [41], but investigating this further is out of scope of this paper.
However, the price itself is not the only contributing factor as explained previously in the introduction. Hence, operators in rural areas should look at the specific demand by the targeted users (e.g. new services targeted to their life).

2.6 Sensitivity analysis

In the simulation phase of the proposed cost model, the bitrate per user has been considered as a cost driver. Based on this cost driver, we generated results for two different scenarios, namely the conservative with 630 Mbps per site and the aggressive with 9450 Mbps per site. Besides the bitrate per user, there are two important factors also affecting the required bitrate per site, which are the number of users and the caching rate. Therefore, we study these two parameters in more detail by running a sensitivity analysis for the model. We studied first the impact of varying the number of users and hence the population density on the ACPU for the conservative scenario (in 2.6.1) and second the impact of the caching rate on the CAPEX and OPEX costs for the aggressive scenario (in 2.6.2).
The input variables used in the sensitivity analysis are described in Table 2-6 below. Note that for this sensitivity analysis, we ran the cost model for the range in input variables as described in the table.

Table 2-6 Input variables for sensitivity analysis

Input variable	Distribution	range	Simulated scenario	Related formula in the model
Population density: number of users	Uniform	13-400 inhabitants /km^2	Conservative	Equation 2-1
Caching rate	Uniform	0%-90%	Aggressive	Equation 2-13

2.6.1 Impact of the population density on the result

In this section, we study the variation of the ACPU with and without caching data on the edge for the conservative scenario (as an output) according to the variation of the population density per km^2 (as an input). Results of this simulation are presented in Figure 2-6:

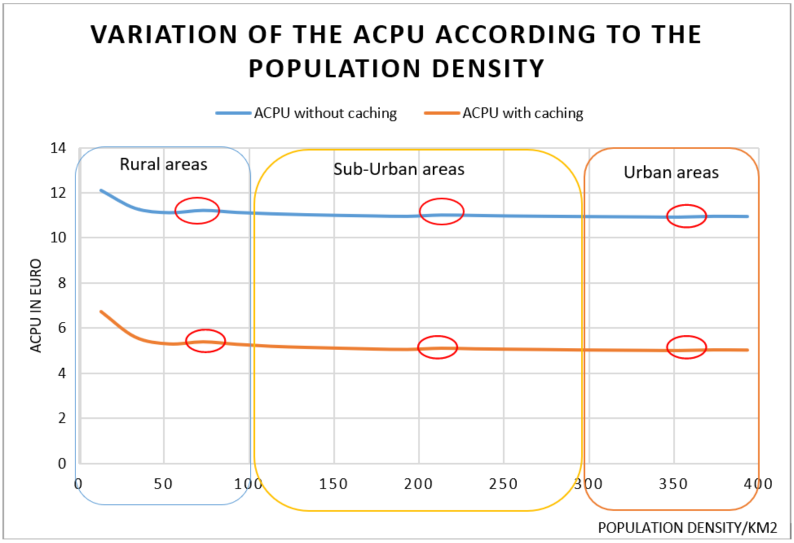


Figure 2-6 Variation of the ACPU for the conservative scenario according to the variation of the population.

Based on the European criteria of the classification of regions [34], we classified in the figure above three main regions based on the population density. Rural areas presented with the blue rectangle (density under 100 inhabitants per km^2), sub-urban areas presented with the yellow rectangle (density between 100 and 300 inhabitants per km^2) and urban areas presented in the orange rectangle (density over 300 inhabitants per km^2). Several deductions can be extracted from the results of this simulation:

- A significant drop of the ACPU is seen in rural areas comparing to a slight decrease in sub-urban and urban areas. This is due to the non-need of adding new base stations within a specific margin of population density. For example, moving from the density of 13 inhabitants per km^2 to 26 does not require the installation of new base station which means the same cost of the infrastructure will be divided by the double number of users.
- A saturation point is reached starting from the population density of 200 inhabitants per km^2. This is because increasing the number of users extensively induces the increase of the required data rate per site thus increasing the number of base stations and satellite terminals that should be installed.
- Transition points between different densities with high ACPU can be observed in Figure 2-6, indicated with red circles. This high ACPU is due

to the need to deploy new base stations to satisfy the additional bandwidth requirement, yet the number of users are not enough to cover this additional cost (the cost of the new installed base station).

- Similar to other broadband solutions, the cost per user of deploying the satellite-4G solution in urban areas is much cheaper than in rural areas.

The low cost of the proposed solution in sub-urban and urban areas can give insights to the operators about the suitable use of a similar solution to improve their services at a low cost. Two use cases for both the sub-urban and urban areas might be of interest.

The first use case for sub-urban areas consists on the consolidation of the "poor" terrestrial connection with the satellite connection to improve the speed and thus the QoE. This use case is mainly relevant to homes and small office home office (SOHO) premises located in underserved areas of developed countries, which are served with terrestrial telecommunication network infrastructure (xDSL or cellular access) of poor bandwidth performance (e.g., users are located far from the DSLAM or far from the 4G cell tower). In such underserved areas of developed countries, the use of satellite to complement the existing terrestrial broadband access link can lead to a hybrid satellite/terrestrial multi-play scenario, which can be envisaged in order to benefit from low-latency of terrestrial networks and high-bandwidth of satellite networks [49].

The second use case concerns the use of the satellite solution in urban areas. This use case involves the delivery and offload of content (multimedia, network software updates etc.) in the mobile network via satellite alone or a combination of satellite and terrestrial links for local caching. Onward delivery is by fetching from the cache and has the advantage of lower latency and potential improved QoE. Satellites are well suited to provide such broadcast/multicast resources over wide areas to aggregate the largest audience possible and hence to reduce the global delivery cost. Combining satellite broadcast/multicast resources with the terrestrial unicast resources is a powerful way to optimize the content delivery costs and improve scalability [49]. The feasibility of these two uses cases from a techno-economic point of view is out of scope of this paper, but can be tackled in future studies.

2.6.2 Impact of the caching rate on the result

We argued previously the assumption on the caching rate in section 2.5.1. However, this assumption has a big impact on the results, as can be seen from the mathematical formulation of the model (section 2.4.2). To this end, we study in this section the variation of the CAPEX and OPEX in function of the caching rate. Results of this simulation in presented in the figure below:

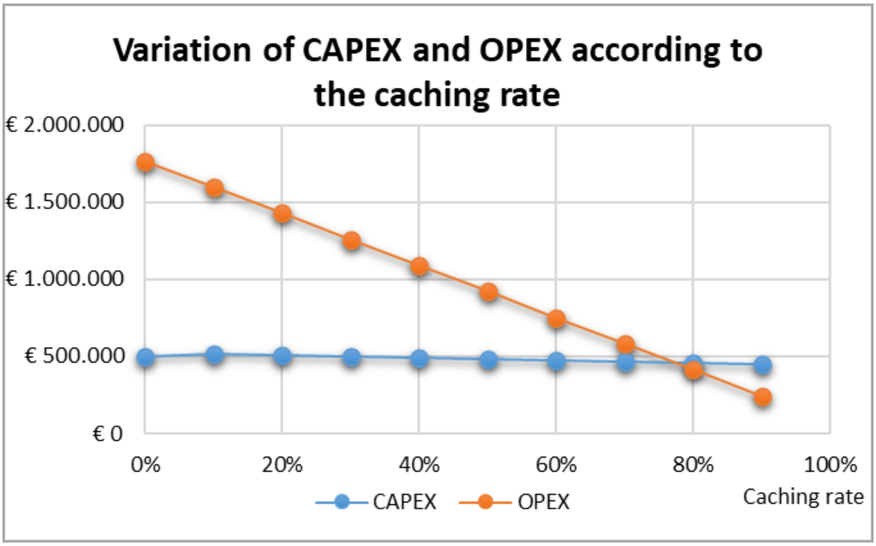


Figure 2-7 variation of the CAPEX and OPEX for the aggressive scenario in function of the variation of the caching rate

Several interpretations can be extracted from the Figure 2-7 related to CAPEX and OPEX. On the one hand, the CAPEX increases with the use of caching due to the cost of the MEC deployment, this can be observed in the Figure 2-7 at the 10% caching rate point. On the other hand, the CAPEX decreases slightly (can be seen starting from point 20%) with the increase of the caching rate due to the decrease of the traffic that needs to be sent via satellite, thus the decrease of the number of satellite terminals needed to handle the traffic. On the operational costs side (OPEX), there is a significant decrease in function of the caching rate increase. The more data is cached, the less traffic is sent via the satellite link, thus the less the satellite capacity cost is paid. One could think, based on these results, to use a high rate of caching to decrease the operational costs, however an in-depth analysis of the consumer traffic and forecasting and optimization algorithms for popular content are needed to determine the reasonable range of the caching rate as elaborated in [13].

To conclude, the main observations from the impact of varying the key inputs of the model on the outcomes are summarized in Table 2-7 below.

Table 2-7 Main observations of the variation of the key parameters of the cost model.

Input variable	ACPU	CAPEX	OPEX
Population	The ACPU is very sensitive to the	The CAPEX increases with	The OPEX increases with the increase of

density: number of users	population density variation hence to the number of users and it is very low in dense urban areas.	the increase of the number of users because new sites needs to be deployed to cover the need of the new users.	number of users because of the increase of the total bitrate required for the new users.
Caching rate	The ACPU decreases significantly due to the decrease of the OPEX while increasing the caching rate.	Slightly affected by the variation of caching rate.	Significantly affected by the variation of the caching rate because we save on the satellite capacity costs when we cache more data on the edge.

2.7 Summary and future work

In this paper, we proposed a new broadband solution to provide the broadband connectivity to white rural areas in Europe, which is the combination of backhaul satellite communication with 4G fronthaul networks. A cost model to study the feasibility of integrating satellite communications within the 4G mobile network to bridge the broadband gap in rural Europe was proposed. Results from this techno-economic model show an Average Cost Per User (ACPU) of 12 to 155 euro, depending on the required bitrate per user, and the amount of popular content that can be cached. Therefore, for the universal broadband service (2Mbps), results prove that the proposed solution is economically viable based on the comparison of the ACPU to the WTP of the end user. In addition, the deployment of the solution without using the caching concept is not cost-effective in the case of the high data rates are required (case of aggressive scenario). However, results show that caching popular content on the edge can significantly save on OPEX costs (60%) and thus decrease the required Average Cost Per User (ACPU). Yet, the resulted ACPU of providing high-speed broadband as required by the DAE 2020 (30 Mbps) is still expensive for rural inhabitants, even with caching (66 euro for 64% caching rate and 44 euro for 80%). In addition, we proposed in this paper a sensitivity analysis to study the impact of the most important inputs of our model. Results of this sensitivity shows that indeed both the population density and the caching rate has a significant impact to a certain extend on the resulted ACPU.

Results of this modelling approach are able to give both the mobile and the satellite operators an overview of the feasibility of the proposed solution, and provide insights into the most cost-effective solution. Results for conservative and aggressive scenarios can guide operators to fix good strategies for the deployment of broadband services in rural areas in such way to guarantee both the service adoption by the rural inhabitants and to generate good revenues.

As future work, we aim to elaborate a socio cost-benefit model to be able to analyse the proposed solution on the long run from different perspectives. Furthermore, the model suggested in this paper restricts on the cost part for 5 years, our intention in the future is to expand it to a cost-benefit model for 10 years to give insights to both mobile and satellite operators about the revenues of this solution. Last but not least, the sensitivity analysis elaborated in this paper has investigated only two parameters, namely the caching data rate and the number of users. Yet still several key parameters and assumptions that should be studied carefully within this analysis, such the uptake of the service and the geographical location of the rural areas (i.e. Western Europe, Asia, Africa etc.).

Acknowledgment

The authors would like to acknowledge all partners in the SaT5G project. This work was conducted within the framework of the SaT5G project, funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 761413. This work has also partially been supported by the COST-action CA15127 (RECODIS).

Appendix

Table 2-8 Back of the envelope calculation for the Fiber-4G solution

Input	Value
Assumptions	• Area of 100 km2, • 100 km far from the nearest operator central office; • 1300 users; • price of m of fiber 50 euro; • max coverage of a macro cell is 5^2*Pi=78.5 km^2;
Price of the Fiber	5000000
Number of base stations	2
Price of base station	25000
Price of BSs	50000
Length of fiber between BSs	5000
Price of fiber to connect BSs	250000
Total price:CAPEX	5275000
ACPU/month for 5 years	67.7 euro

References

[1] Abdelsalam, A., Luglio, M., Roseti, C., & Zampognaro, F. (2017, July). TCP connection management through combined use of terrestrial and satellite IP-Based links. In 2017 40th International Conference on Telecommunications and Signal Processing (TSP) (pp. 37-42). IEEE.

[2] Picot, A., & Wernick, C. (2007). The role of government in broadband access. Telecommunications Policy, 31(10-11), 660-674.

[3] Abdelsalam, A., Luglio, M., Roseti, C., & Zampognaro, F. (2019). Analysis of bandwidth aggregation techniques for combined use of satellite and x DSL broadband links. International Journal of Satellite Communications and Networking, 37(2), 76-90.

[4] Abrardi, L., & Cambini, C. (2019). Ultra-fast broadband investment and adoption: A survey. Telecommunications Policy.

[5] Airconco. (2018). Accessed 8 March 2018. Retrieved from https://www.airconco.com/air-conditioning-price-calculator/

[6] Atkearney. (2018). Accessed 7 March 2018. Retrieved from https://www.atkearney.com/documents/20152/434204/Telco_Fitness.pdf/427dcc76-15f9-b9a3-0fd5-5c2e157918d6

[7] BATS, Deliverable D5.2. (2016). Accessed 12 February 2018. Retrieved from https://cordis.europa.eu/project/rcn/106027/reporting/en

[8] Bohlin, E., & Teppayayon, O. (2010, August). Broadband universal service in Europe: A review of policy consultations 2005-2010. TPRC.

[9] Borst, S., Gupta, V., & Walid, A. (2010, March). Distributed caching algorithms for content distribution networks. In 2010 Proceedings IEEE INFOCOM (pp. 1-9). IEEE.

[10] Bouras, C., Ntarzanos, P., & Papazois, A. (2016, October). Cost modeling for SDN/NFV based mobile 5G networks. In 2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT) (pp. 56-61). IEEE.

[11] Carare, O., McGovern, C., Noriega, R., & Schwarz, J. (2015). The willingness to pay for broadband of non-adopters in the US: Estimates from a multi-state survey. Information Economics and Policy, 30, 19-35.

[12] Casier, K. (2010). Techno-economics of FTTH deployment in the presence of competition (Doctoral dissertation).

[13] Cha, M., Kwak, H., Rodriguez, P., Ahn, Y. Y., & Moon, S. (2009). Analyzing the video popularity characteristics of large-scale user generated content systems. IEEE/ACM Transactions on networking, 17(5), 1357-1370.

[14] Cisco. (2017). Cisco VNI reports. Accessed 10 March 2018. Retrieved from https://www.cisco.com/c/m/en_us/solutions/service-provider/vni-forecast-highlights.html

[15] Cisco. (2019). Cisco VNI Forecast Highlights Tool. Accessed 20 January 2019.Retrieved from https://www.cisco.com/c/m/en_us/solutions/service-provider/vni-forecast-highlights.html#

[16] Dean Segel, J. (2009). Rural broadband hybrid satellite/terrestrial solution. USA. Accessed 10 January 2019. Retrieved from

https://patentimages.storage.googleapis.com/d8/1e/c1/ee36d8d93f709e/US7477597.pdf
[17] Dalela, P. K., Basu, S., Majumdar, S., Yadav, A., Singh, A., Nagpal, N., & Tyagi, V. (2014, December). Geo-intelligence based terrain-aware mobile backhaul planning for rural India. In 2014 IEEE International Conference on Advanced Networks and Telecommuncations Systems (ANTS) (pp. 1-6). IEEE.
[18] Davies, R. (2016). Broadband as a universal service. Briefing, Members' Research Service PE 581.977, EPRS: European Parliamentary Research Service.
[19] Durairajan, R., & Barford, P. (2017). A techno-economic approach for broadband deployment in underserved areas. ACM SIGCOMM Computer Communication Review, 47(2), 13-18.
[20] ESA. (2016). ESA SPECSI study: Internal and confidential document. Accessed 20 October 2017.
[21] ETSI. (2002). TR 101 865: V1.2.1. Satellite Earth stations and systems (SES); satellite component of UMTS/IMT-2000. Accessed 10 March 2018. Retrieved from https://www.etsi.org/deliver/etsi_tr/101800_101899/101865/01.02.01_60/tr_101865v010201p.pdf
[22] European Commission. (2002). Accessed 10 January 2019. Retrieved, from https://eur-lex.europa.eu/legal-content/GA/TXT/?uri=celex:32002L0022
[23] European Commission. (2011). Third review of the universal service. Brussels.
[24] European Commission. (2013a). Accessed 15 March 2018. Retrieved from https://ec.europa.eu/digital-single-market/en/content/broadband-all.
[25] European commission. (2013b). Accessed 15 March 2018. The Broadband State Aid Rules. Retrieved from https://ec.europa.eu/digital-single-market/en/state-aid
[26] European Commission. (2014). Accessed 15 March 2018. Retrieved from http://ec.europa.eu/competition/sectors/telecommunications/broadband_decisions.pdf
[27] European Commission. (2015). Mapping of Broadband and Infrastructure Study - SMART 2012/0022. Brussels.
[28] European Commission. (2016a). Accessed 15 March 2018. Broadband For Europe. Gigabit Society 2025. Retrieved from https://ec.europa.eu/digital-single-market/en/broadband-europe
[29] European commission. Directorate-General for Communications Networks, C. a. (2016b). Review of the scope of universal service. Brussels.
[30] European Commission. (2017). Mobile Broadband Prices in Europe. Executive Summary. Brussels.
[31] European Commission. (2018). Broadband Coverage in Europe 2017. European Commission. Accessed 15 March 2018. Retrieved from https://ec.europa.eu/digital-single-market/en/news/study-broadband-coverage-europe-2017

[32] Eurostat. (2009). Accessed 10 August 2019. Retrieved from https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/digital_competitiveness_report_2009.pdf
[33] Eurostat. (2017). Accessed 12 February 2018. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Electricity_price_statistics
[34] Eurostat. (2018). Accessed 12 February 2018. Retrieved from https://ec.europa.eu/eurostat/web/rural-development/methodology
[35] FCC. (2010). Accessed 2 February 2018. National Broadband Plan report. The U.S. Federal Communications Commission. Retrieved from https://www.fcc.gov/general/national-broadband-plan
[36] Feijóo, C., Ramos, S., Armuña, C., Arenal, A., & Gómez-Barroso, J. L. (2018). A study on the deployment of high-speed broadband networks in NUTS3 regions within the framework of digital agenda for Europe. Telecommunications Policy, 42(9), 682-699.
[37] Frías, Z., González-Valderrama, C., & Martínez, J. P. (2015). Keys And Challenges To Close The Rural Broadband Gap. The Role of LTE Networks In Spain.
[38] Gilat. (2015). White paper: Satellite Backhaul vs Terrestrial Backhaul: A Cost Comparison.
[39] Glass, V., & Stefanova, S. K. (2010). An empirical study of broadband diffusion in rural America. Journal of Regulatory Economics, 38(1), 70-85
[40] Gómez, B., et al., D.-G. f. (2016). Reforming EU telecoms rules to create a digital union. Accessed 12 February 2018. Retrieved from https://publications.europa.eu/en/publication-detail/-/publication/9d8aa519-ec17-11e5-8a81-01aa75ed71a1/language-en/format-PDF/source-86976357
[41] Han, B., Tayade, S., & Schotten, H. D. (2017, July). Modeling profit of sliced 5G networks for advanced network resource management and slice implementation. In 2017 IEEE Symposium on Computers and Communications (ISCC) (pp. 576-581). IEEE.
[42] Investopedia. (2018). average-profit-margin-company-telecommunications. Accessed 10 January 2018. Retrieved from https://www.investopedia.com/ask/answers/060215/what-average-profit-margin-company-telecommunications-sector.asp
[43] Ioannou, N., Katsianis, D., & Varoutas, D. (2018). Comparative Techno-Economic Evaluation of LTE Fixed Wireless Access and FTTdp G. fast network deployment for providing broadband services of at least 30Mbps in rural areas.
[44] ITU report. (2010). RM2176: Vision and requirements for the satellite radio interface(s) of IMT-Advanced.
[45] ITU report. (2015). Measuring the Information Society Report 2015. Accessed 15 January 2018. Retrieved from https://www.itu.int/en/ITU-D/Statistics/Pages/publications/mis2015.aspx

[46] LaRose, R., DeMaagd, K., Chew, H. E., Tsai, H. Y. S., Steinfield, C., Wildman, S. S., & Bauer, J. M. (2012). Broadband Adoption| Measuring Sustainable Broadband Adoption: An Innovative Approach to Understanding Broadband Adoption and Use. International Journal of Communication, 6, 25.
[47] Lemstra, W. (2016). Broadband developments in Europe: A retrospective review of the determinants of supply and demand. IBEI Research Paper, November.
[48] Lenovo. (2018). Accessed 15 January 2018; Retrieved from https://www.senetic.fr/lenovo/serveurs_lenovo/
[49] Liolis, K., Geurtz, A., Sperber, R., Schulz, D., Watts, S., Poziopoulou, G., & Fitch, M. (2019). Use cases and scenarios of 5G integrated satellite-terrestrial networks for enhanced mobile broadband: The SaT5G approach. International Journal of Satellite Communications and Networking, 37(2), 91-112.
[50] Lo, K., Kwong, K., Chieng, D., Ting, A., Choong, K., & Abbas, M. (2012). Bridging digital divide in Malaysia using low cost WiMAXWiFi multi-hop network.
[51] Moghaddam, B. K., & Khatoon-Abadi, A. (2013). Factors affecting ICT adoption among rural users: A case study of ICT Center in Iran. Telecommunications policy, 37(11), 1083-1094.
[52] Ovando, C., Pérez, J., & Moral, A. (2015). LTE techno-economic assessment: The case of rural areas in Spain. Telecommunications Policy, 39(3-4), 269-283.
[53] Point Topic. (2015). Europe to fall short for next generation broadband coverage. Accessed 10 February 2018. Retrieved from http://point-topic.com/free-analysis/europe-to-fall-short-for-next-generation-broadband-coverage/
[54] Point Topic. (2013). Europe' s broadband investment needs Quantifying the investment needed to deliver superfast broadband to Europe. London. Accessed 10 February 2018. Retrieved from http://pointtopic.com/free-analysis/europes-broadband-investment-needs/
[55] Polykalas, S. E., Vlachos, K., & Ellinas, G. (2017, May). Technologies and architectures for broadband Digital Divide elimination. In 2017 24th International Conference on Telecommunications (ICT) (pp. 1-6). IEEE.
[56] Preston, P., Cawley, A., & Metykova, M. (2007). Broadband and rural areas in the EU: From technology to applications and use. Telecommunications Policy, 31(6-7), 389-400.
[57] Rossini, G., & Rossi, D. (2012, September). A dive into the caching performance of content centric networking. In 2012 IEEE 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD) (pp. 105-109). IEEE.
[58] SAT5G. (2017). Internal and confidential document. Accessed 25 October 2017.
[59] Schneir, J. R., & Xiong, Y. (2016). A cost study of fixed broadband access networks for rural areas. Telecommunications Policy, 40(8), 755-773.

[60] Krizanovic Cik, V et al. (2016). Broadband ecosystem elements in techno-economic modelling and analysing of broadband access solutions for rural areas. Technical Gazette, 23(1), 279-290.
[61] Siyang, L., Fei, Q., Zhen, G., Yuan, Z., & Yizhou, H. (2013). LTE-satellite: Chinese proposal for satellite component of IMT-Advanced system. China Communications, 10(10), 47-64.
[62] Townsend, L., Sathiaseelan, A., Fairhurst, G., & Wallace, C. (2013). Enhanced broadband access as a solution to the social and economic problems of the rural digital divide. Local Economy, 28(6), 580-595.
[63] Van der Wee, M., Verbrugge, S., Tahon, M., Colle, D., & Pickavet, M. (2014). Evaluation of the techno-economic viability of point-to-point dark fiber access infrastructure in Europe. IEEE/OSA Journal of Optical Communications and Networking, 6(3), 238-249.
[64] VSAT. (2018). Remote business and residential. Accessed 10 March 2018. Retrieved from http://www.vsat-systems.com/End-User-Service/Plans/Remote-businesses-and-residential/
[65] Wang, S., Zhang, X., Zhang, Y., Wang, L., Yang, J., & Wang, W. (2017). A survey on mobile edge networks: Convergence of computing, caching and communications. IEEE Access, 5, 6757-6779.
[66] Wang, X., Chen, M., Taleb, T., Ksentini, A., & Leung, V. C. (2014). Cache in the air: Exploiting content caching and delivery techniques for 5G systems. IEEE Communications Magazine, 52(2), 131-139.
[67] Xavier, P. (2003). Should broadband be part of universal service obligations?. info, 5(1), 8-25.
[68] Liu, Y. H., Prince, J., & Wallsten, S. (2018). Distinguishing bandwidth and latency in households' willingness-to-pay for broadband internet speed. Information Economics and Policy, 45, 1-15.
[69] Zangar, N., Abdennebi, M., Tabbane, S., & Tohmé, S. (2009, September). Design, mobility management and analysis of a hybrid WiMAX/satellite network. In 2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications (pp. 1133-1137). IEEE.
[70] Zangar, N., & Hendaoui, S. (2019). Leveraging multiuser diversity for adaptive hybrid satellite-LTE downlink scheduler (H-MUDoS) in emerging 5G-satellite network. International Journal of Satellite Communications and Networking, 37(2), 113-125.
[71] Hu, Y. C., Patel, M., Sabella, D., Sprecher, N., & Young, V. (2015). Mobile edge computing—A key technology towards 5G. ETSI white paper, 11(11), 1-16.

3

Techno-economic analysis of inflight connectivity using an integrated satellite -5G network

Asma Chiha, Marlies Van der Wee, Leonardo Goratti [10], Didier Colle

Published in the International Journal of Satellite Communications and Networking

While the 5G network helps fulfil the promised KPIs of the three famous use cases, eMBB, uRLLC, and mMTC, there are still areas where broadband connectivity is non-existent, interrupted, or slow, such as onboard planes. Inflight connectivity is now viewed as a "must-have" service, and demand and uptake are significantly increasing over the last few years. In this chapter, we propose an integrated solution using both satellite and 5G networks to provide inflight connectivity and assess the economic feasibility of offering broadband services to passengers on commercial flights. To this end, we developed a techno-economic framework that takes into account both capital and operational expenses in computing the Total Cost of Ownership (TCO), as well as the average cost per user and Megabyte of data.

***

[10] Safran Passenger Innovations, Germany

Abstract—The demand for mobile broadband services is increasing exponentially alongside with user expectations regarding the reachability of these services and their prices. This paper presents an integrated satellite and 5G network for providing inflight connectivity and evaluates the economic viability of offering broadband connectivity to passengers on commercial airplanes by the development of a techno-economic framework, which considers both capital and operational expenditures to compute the Total Cost of Ownership (TCO), but also an average cost per user and per Megabyte. Results show high operational costs mainly due to satellite bandwidth usage. Therefore, caching popular content on the network onboard is beneficial to reduce the traffic carried over the satellite link, thus lowering the operational costs as well. Furthermore, the framework is used to compare the identified business models for Inflight Entertainment and Connectivity offerings and their pricing strategies, alongside a benchmark against the current inflight connectivity pricing. Finally, a sensitivity analysis is elaborated in order to mitigate the uncertainty of inputs (e.g. rate of caching) used to feed the TCO model. Following concrete recommendations are the main result of this research: (1) Providing inflight broadband services with a 2 to 5 Mbps throughput per user is feasible if satellite communications is integrated into the 5G network. (2) Caching popular data reduces the operational costs and hence the average cost per user (from 25 to 32% depending on the caching rate adopted). (3) This framework allows to provide recommendations on the best suited business models and related pricing schemes.

3.1 Introduction

5G network drivers are centred around providing high-speed connections anytime, anywhere and to anything. There are however still areas where broadband connectivity is non-existent, interrupted or slow. One of these cases is onboard planes; this inflight connectivity is no longer seen as a comfort from passengers' perspective but as a "must-have" service for which the demand and uptake is significantly increasing over the last years [1].

Inflight Entertainment and Connectivity (IFEC) onboard airplanes was introduced nearly two decades ago with the Terrestrial Flight Telecommunication System (TFTS) [2], which allowed passengers to do voice calls on the plane using handsets connected to their seats. The radio link between the aircraft and the fixed telephony network was established via macro cells that can cover up to 240 km^2 and up to 10 km of altitude [3]. This connection is known as an Air-To-Ground (ATG) connectivity system. In its first introduction and given the high investments and maintenance costs, this service did not get a good market adoption compared to present figures. A second attempt to provide connectivity to airplanes was made by Boeing in 2006, offering internet connectivity. The service was used by a limited number of airlines but quickly removed due to a low market demand [4]. Unlike the previous ATG solution, the Boeing solution used a satellite link to connect the Wireless Local Area Network (WLAN) on the plane with the network on the ground.

After these two first attempts, many solutions were proposed by different companies and adopted by airlines, both offered by specific vendors of aircraft connectivity solutions such as Panasonic Avionics (adopted by Air Lingus [5] and Air France [6]), as well as by connectivity providers such as OnAir (operates in Emirates airline aircrafts [7] and Hong Kong Airlines aircrafts [8]). Over time, the user demand for more speed and lower prices has dramatically increased, which pushes airlines to look for advanced IFEC solutions to cope with passengers' expectations [1].

Integrating satellite communication into 5G network enables a broad range of use cases where the benefits of satellite can be leveraged [1], such as complementing and extending dense terrestrial cells; providing connectivity to moving platforms (ships, airplanes, vehicles and trains); or offloading temporarily congested networks among other benefits.

In this paper, we propose an integrated satellite-5G (Sat5G) solution aiming at providing broadband access for passengers on commercial airliners, that will allow providing innovative media services onboard. This solution provides broadband access to fulfil the private demand of passengers independently of the airplane's high priority functionalities. Passengers can use their own personal device(s), or PEDs - as they do on the ground - or can rely on the seat screens to access local media content (i.e. from the entertainment system onboard), while the network requests are individual and specific to each passenger's activity. To the best knowledge of the authors, there is no existing integrated solution that combines 5G technology with satcom to provide IFEC services. Though, the advanced 5G technology and the innovations in satellite systems can catalyse new opportunities in the IFEC market to deliver broadband access to airplane passengers with unprecedented quality of experience, specifically relying on the O3b satellite constellation for medium earth orbit. This latter can be achieved by the adoption of new technologies such as caching techniques and 5G network slicing. This work is the result of collaboration between different partners in the European Sat5G research project [9].

The rest of the paper unfolds as follows: a detailed description of the integrated satellite and 5G network solution is presented in section 3.2. Section 3.3 first details the involved stakeholders, their interactions and the generic value network and then identifies the potential business models and analyses them using a suggested framework. A viability study of providing inflight connectivity using the proposed solution is elaborated in section 3.4. The main takeaways from using the integrated satellite and 5G network to provide inflight connectivity, as well as future research steps, are summarized in the last section (section 3.5).

3.2 Integrated satellite and 5G network for inflight connectivity

In addition to the several benefits that satellite communication can leverage to 5G network, satellite technology was chosen to be adopted here, because of its global

coverage and reliable connection. Using the new High Throughput Satellite (HTS) technology, a significant increase in the offered backhaul capacity can be achieved compared to the old generation of satellites [10]. An integrated 5G system combines this HTS backhaul with mobile network connectivity onboard aircrafts, complementing pre-existing WLAN to provide broadband services to passengers.
A seminal work regarding architecture options for an integrated satellite-terrestrial network within 5G can be found in the ETSI TR 103611 report [11], which identified five different architecture scenarios, with the key innovation being the introduction of a Non-Terrestrial Network (NTN) segment, which allows mapping the satellite system onto mobile network equivalents:

i. Direct 3GPP Access; Non-Terrestrial (NT) 5G User Equipment (UE) connects to the 5G Core Network (5GC) directly using a 5G compliant NTN New Radio Access Network (NTN 5G RAN); in this case the NT UE is assumed a device capable of using NTN New Radio (NTN NR)-based access directly connecting to the 5G core through the NTN 5G RAN,

ii. Indirect 3GPP Access; the end-to-end connectivity between a 5G UE (e.g. smartphone) is established indirectly by the presence of a gNB combined with an NTN NR-based NTN NT UE, which acts as a relay toward the NTN 5G RAN and thus connects to the 5GC,

iii. Indirect Mixed 3GPP NTN Access; 5G UEs are served by a gNB combined with an NTN-enabled NT UE that is served by a trusted mixed 3GPP NTN access network to connect to the 5GC; the mixed case arises from non-3GPP Layer 2 and non-3GPP Layer 1,

iv. Indirect Access via NTN Transport Network; the 5G UEs are served by a gNB connected to an NTN terminal that is linked to the 5GC via a trusted NTN transport network, where the trust comes from the agreement between operators, respectively the mobile network operator and the NTN operator,

v. Indirect Untrusted Access; a similar situation to the fourth one is envisaged but in this case the NTN access network is untrusted in 3GPP terms and additional 3GPP functionalities (e.g. non-3GPP interworking function) are required.

In the particular case of an aircraft scenario, we envisage a progressive adoption of solutions fuelled by the 5G system based on technology maturation. The first opportunity to adopt a solution based on 5G technology is provided by the architecture in the Indirect Access via NTN Transport Network scenario described above, which is referred to as IA-NTN in the remaining of this paper. Such a scenario requires a less tight level of integration between mobile and satellite networks, though it requires creating the necessary interfaces and modules to ensure a trusted end-to-end network. In addition, the IA-NTN architecture

scenario maintains sufficient modularity from the technology standpoint, as well as on the business side. The modularity directly implies a faster and less complex design phase, followed by a slimmer certification process to comply with the stringent aviation standards. On the other hand, from the market viewpoint, different types of liaisons can be set in place between the stakeholders, and new joint ventures might rise. In a more accrued end-to-end 5G system for aircraft connectivity, the architecture options described in the second and third scenarios can be adopted, although this requires further maturation and integration of the satellite technology in the 5G technology mainstream.
In this paper we shall focus the techno-economic analysis on the architecture option described in the IA-NTN scenario and shown in Figure 3-1.

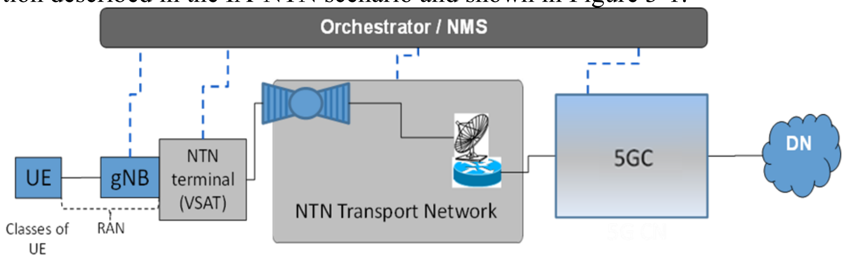


Figure 3-1: Architecture option of indirect access via an NTN transport network [11]

The existing IFEC, as the name says, is composed of an Inflight Entertainment (IFE) part and a connectivity (C) part. As shown in Figure 3-2, an IFEC system is composed of a network segment for onboard connectivity, namely for passengers and cabin crew, whereas a second network segment is dedicated to connecting to a ground infrastructure. The IFE part of the system is focused on delivering different type of media contents to passengers, including a catalogue of movies, TV series, music and newspapers stored in a server onboard. Typically, passengers consume content either through the embedded seat screens or with their personal devices (smartphones, tablets, etc.). In the latter case, passengers need to download an App on their personal devices from the onboard store in order to access the media resources. Commonly, this is done by connecting to the onboard wireless network, which is based on different generations of Wi-Fi (e.g. Wi-Fi 5). On the other hand, we register limited penetration of mobile technology for onboard connectivity, still. Traditionally, the reasons for a little penetration of mobile network technology for onboard connectivity can be found in a higher complexity of the connectivity system, including at spectrum level, if compared to Wi-Fi. It is worth noticing that the possibility to revisit the onboard connectivity taking advantage of 5G and network virtualization, as well as of Multi-access Edge Computing (MEC) gives rise to adopt new models both at technology and business level. For example, the use of unlicensed spectrum to deliver mobile network connectivity can also provide an innovative new opportunity [12].

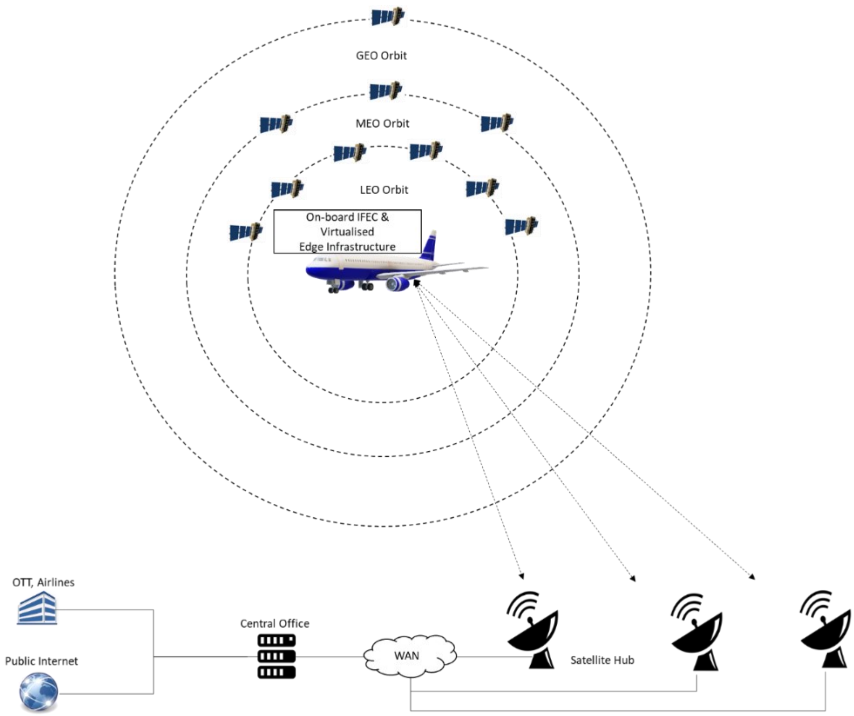


Figure 3-2: Future aircraft connectivity including multi-constellation backhaul and edge computing platform onboard

Existing connectivity solutions that airline companies purchase for connecting to a ground infrastructure rely mostly on a satellite backhaul, though other activities can be found also regarding ATG links. Satellite systems of different operators in the Geostationary Orbit (GEO) offer the advantage of global coverage though at the penalty of higher end-to-end latency due to the altitude of the space segment (i.e. 500ms two-ways) and lower capacity, a trend that is changing with the increasing adoption of HTS satellites that make use of multiple spot beams. As shown in Figure 3-2, satellite connectivity to aircrafts is currently enriched with newer opportunities that come from Non-Geostationary Satellite orbits (NGSO) like the Medium Earth Orbit (MEO) and the Low Earth orbit (LEO), as well as from more efficient use of Ku and ka bands. Although for aircraft connectivity, GEO constellations remain the de facto situation, the market is currently opening also to NGSO constellations, which offer significantly lower latency and higher throughput per satellite beam. For example, note the activities undertaken by OneWeb and SpaceX along the future release of mega-constellations. As a matter of fact, it must be considered that by changing the orbit of the satellite system other challenges arise, including the remarkably larger number of satellites, the complexity and costs to launch and manage such fleet and the necessary mechanisms to ensure seamless handover as in the case of an aircraft that travels over geographical distances.

Another challenge that arises with NGEO satellite is the high Doppler shifts (e.g. up to 480 kHz Doppler shift for LEO in Ka band for the whole satellite coverage) [13], which can affect the QoE/QoS of end users by reducing the provided capacity. However, many techniques have been implemented to counter its effects, such as built-in Doppler compensate per beam mechanism [13] and others are under developing in the research community [14] and [15].

As already mentioned above, the techno-economic analysis presented in this work is based on the future evolution of the IFEC platform in which 5G technology is adopted and the whole aircraft may become a flying edge infrastructure complying with the principles of the MEC paradigm. As shown in Figure 3-2, a central office facility that may belong to different stakeholders in the IFEC value chain provides overall resource coordination, services orchestration and connectivity to different data networks.

3.3 Key stakeholders and value network

As discussed in the previous section, the considered scenario is providing broadband access for passengers on commercial airliners, alongside additional individual media requests. This scenario proposes a bi-directional broadband access for each passenger for private use which is transparent to the moving platform. The network requests are therefore individual and proper to each passenger's activities. Typically, the 5G network ecosystem is already complex due to the involvement of many stakeholders; in addition to that, we add here another actor which is the satellite network operator. This section presents first the key stakeholders involved in providing the inflight connectivity using the integrated sat5G network (in 3.3.1) and the interactions between them which represents the value network (in 3.3.2). Different business models will be analysed in the light of how value can be exchanged between these stakeholders in section 3.3.3.

3.3.1 Identification of key involved stakeholders in Sat5G inflight connectivity

The market of IFEC typically includes different stakeholders, following the architecture requirements (Figure 3-2). The platform includes the IFE and the connectivity system onboard, as well as the onboard satellite terminal and the satellite antenna. The satellite equipment is typically based on components supplied by satellite equipment vendors. The IFE requires one or more media servers, wireless access points that in the future can embed 5G connectivity, and seat screens. Airlines are the main customers of the IFEC provider and purchase a preferred IFEC package that includes specific equipment. Aircraft manufacturers (e.g. Airbus) request the IFEC provider for equipment that has to be installed on linefit (new airplanes) but also retrofit on older aircraft models. A satellite network operator (SNO) provides satellite connectivity to the aircrafts of an airline and an onboard operator (or the airline in some business models) offers connectivity service to passengers relying on the equipment supplied by the IFEC company. In our case, the onboard operator is a mobile network operator (MNO) who provides the 5G network inside the aircraft to the airline, because the onboard operator should have the right to operate 5G network spectrum and only MNOs holding licences can take up this role. The end user in this case is the passenger that makes use of the inflight connectivity. It is worth noticing that all the equipment installed onboard is subject to a rigorous and complex compliance certification process that follows regulations set by regulators and requires skilled expertise.

3.3.2 Generic Value network

Following the identification of the stakeholders, this section describes the generic value network which depicts the interaction between the involved stakeholders. We adopt here the value network as proposed by Allee who defines it as “a web of relationships generating tangible or intangible value by dynamic exchanges between two or more individuals, groups, or organizations” [16]. The term generic is used since it provides a general overview and can be different when stakeholders take up multiple roles and relate with different interactions. Relationships between the different actors are shown in Figure 3-3. The focus of the business model analysis will be on the interactions between the stakeholders coloured in orange, while the stakeholders coloured in blue are not core, but complete the value network ecosystem.

The MNO is placed in the center of the network because it is the main end provider of the service. It offers the service, equipment and insights for the service in exchange for a fee paid by the subscriber, which is the airline in this case. In order to provide these values, the MNO has to buy equipment from an equipment vendor and capacity from a satellite network operator. We assume that the MNO owns and operates the core network on the ground. Intangible arcs include product suggestions, expertise and trends.

For both SNOs and MNOs, regulations are very important, since spectrum licenses are sold via regulators but are also to be used in different countries with different law systems. The study of the impact of the complexity of these regulations on the business model is out of the scope of this paper.

The airline links the end user (the passengers), content provider and MNO by providing the moving platform. As the end user demands content in exchange for a fee (tangible value such as direct payment or intangible value such as loyalty), the content provider plays an important role in finding which content should be cached. A content distributor provides the content to the content provider, as well as industry expertise in exchange for a fee, feedback and trends. It becomes clear that this use case involves many actors, leading to many possible value network configurations which can be then translated into different business models. The next section details these different configurations, focusing mainly on the stakeholders that are marked in orange in Figure 3-3.

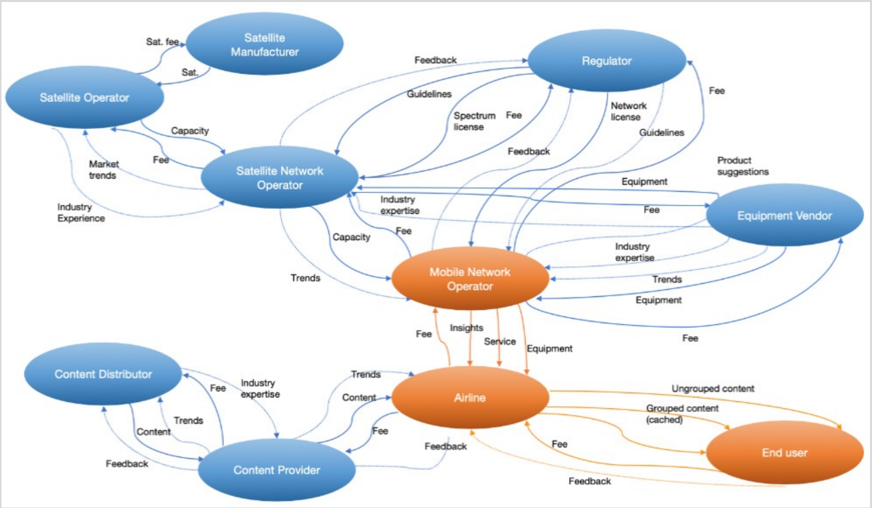


Figure 3-3: Generic Value Network of providing inflight connectivity using integrated satellite 5G network

3.3.3 **Identification and qualitative comparison of feasible business models**

In order to analyze the potential business models for this use case and compare between them, a customized framework is developed. To build this framework, the value network intelligence framework, developed by Allee [17], is used as a starting point, non-relevant links in the framework were removed, but the main structure is kept. The description of the suggested framework is presented in detail in Table 3-1.

Table 3-1: Explanation of the developed framework

TOPIC	DESCRIPTION
Value optimization	
Value creation	How is the value created in the network, by whom and for whom?
Financial impact	What is the financial impact for the stakeholders? Who makes revenues and are there any missed opportunities?
Intangible asset management	What are the intangible assets involved in the value network?
Business performance	
Technology	What technology is required to meet the customer requirements?
Channel management	Which channels are used to bring the service to the end user and how efficient is the channel structure?

QoS	What is the required Quality of Service (QoS) for the end users and are there implications on the performance of the proposed solution?
Branding & relationships	
Branding potential	Who can gain a better perception and what are the advantages that are linked to that?
Differentiation opportunities	Which additional opportunities are identified in the business case that stakeholders can benefit from in terms of relationships and branding?
Network vitality	
Structural dependency	Who depends on whom and are there threats in the structural dependency of the stakeholders?
Specialist knowledge	Which stakeholders require specialist knowledge of the inflight connectivity?
Financial risk	Who takes the highest financial risk in the investment? Linked to this is of course also the revenue potential. A trade-off has to be made.

The summary of the identified business models is presented in Table 3-2.
In order to make a qualitative comparison between the identified business models, the suggested framework will be applied to each business model separately first, after which the business models are compared.

Table 3-2: Summary of identified business models

Model	Short description
Wholesale	The airline buys the project from an MNO and sells the service to end users.
Retail	The MNO installs the project on the platform provided by the airline and sells the service directly to end users.
Sponsorship	The airline buys the project from an MNO and generates revenues by selling sponsorship projects. End users are offered the service for free.
Freemium wholesale	The airline buys the project from an MNO and sells the service to premium end users. A limited version of the service is offered to for free.
Freemium retail	The MNO installs the project on the platform provided by the airline and sells the service directly to premium end users. A limited version of the service is offered to for free.
Complementary service	The airline buys the project from an MNO and offers the full service for free to all end users.

Wholesale model with connectivity for paying users

As described in the Table 3-2, the airline company buys the project from the MNOs which includes network equipment, hardware installation, connectivity over satellite etc. and offers the inflight connectivity to the passengers. The airline collects all revenues and has control over everything inside the airplane and has thus more independency setting up prices (e.g. price for weekend versus working days). In [18], Panasonic aviation adopts this business model. The application of the suggested analysis framework is summarized in Table 3-3.

Table 3-3: Application of the suggested framework to the wholesale model

Value optimization	
Value creation	Value for end user: connectivity Value for airline: increased ticket sales and revenues for connectivity service
Financial impact MNO	MNO makes profits by selling the project, but can miss out on revenues by not selling directly to the end users
Financial impact airline	Increased revenues for airline Threat of hidden costs for airline
Financial impact end user	End user pays a price – different pricing options available
Intangible asset management	End user feedback can be incorporated in service
Business performance	
Technology	Sat5G integration required to reach high performance
Channel management	Business performance might encounter delays in case of problems due to indirect link end user – MNO
QoS	Available bandwidth is split between paid users
Branding & relationships	
Branding potential	Increased perception of airline that can lead to higher loyalty
Differentiation opportunities	Airline can brand itself as an onboard internet provider and grow its name Opportunity to serve as platform for other connectivity-dependent revenues: sales – promotions
Network vitality	
Structural dependency	Quality requirements and delivery times MNO may not be aligned with airline needs
Specialist knowledge	Airline needs internal experts / managers
Financial risk	Airline takes the financial risk

Retail model with connectivity for paying users

In a retail model, the MNO installs, migrates and operates the full inflight connectivity project onboard, using the aircraft platform from the airline, and provides the service to the end user. This allows the MNO to manage the customer experience and optimize the pricing model. On the other hand, it implies that the airline is not directly involved in the revenue stream from the end user. Yet, as it serves as the platform to provide the service, the airline can have a share in revenues, agreed upon and contracted beforehand. An example of an MNO that works via this business model is Gogo. Examples of airlines working together with Gogo are Alaska Airlines, American Airlines, Vietnam Airlines, Virgin America, etc. [19]. The application of the suggested analysis framework is summarized in Table 3-4.

Table 3-4: Application of the suggested framework to the retail model

Value optimization	
Value creation	Value for end user: connectivity Value for airline: increased ticket sales
Financial impact MNO	MNO has higher revenue potential by selling directly to end users
Financial impact airline	Overhead costs / Share in revenues possible
Financial impact end user	End users pays a price – different pricing options available
Intangible asset management	End user feedback can be incorporated in service
Business performance	
Technology	Sat5G required to reach high performance MNO has the opportunity to use multicasting to its advantage
Channel management	Business performance efficiently handled thanks to direct link between end user and MNO
QoS	Available bandwidth is split between paid users, depending on the pricing model
Branding & relationships	
Branding potential	Less branding potential for airline, although still significant
Differentiation opportunities	Airline can brand itself as an onboard internet company and grow its name
Network vitality	
Structural dependency	Quality requirements from end user are directly linked to MNO MNO's business is dependent on airline
Specialist knowledge	Airline does not need internal experts / managers
Financial risk	MNO takes the financial risk

Sponsorship model with free connectivity for all passengers

In a sponsorship model, the airline buys the inflight connectivity project from an MNO and sells sponsorship deals to sponsors to offset the costs of the project partly or completely. In this model, third parties could reach a captive audience and generate ancillary revenues by spreading a promotional message or improve their company brand. In return, they have to pay the airline for this advertising opportunity and the required connectivity to do so. The sponsors have to evaluate the project based on the potential increase in sales. An interesting example of this model is the partnership between Delta Airlines and LinkedIn at the end of 2018, in which passengers were offered free internet in addition to free online courses on topics such as 'Leadership' on the platform LinkedIn Learning [20]. The application of the suggested analysis framework is summarized in Table 3-5.

Table 3-5: Application of the suggested framework to the sponsorship model

Value optimization	
Value creation	Value for end user: connectivity Value for end user: frictionless service (passengers do not have to worry about payment onboard or in advance) Value for airline: increased ticket sales Value for sponsorship partner: end user credentials, data and the opportunity to advertise and promote
Financial impact MNO	MNO makes profits by selling the project to the airline
Financial impact airline	Airline has an opportunity to create extra revenues if sponsorship demand is high, resulting in a higher revenue potential Threat of hidden costs for airline
Financial impact end user	Service is free
Intangible asset management	End user feedback can be incorporated in service
Business performance	
Technology	Sat5G required to reach high performance, although air-to-ground can also be used depending on end user needs and sponsor needs
Channel management	Business performance might encounter significant delays in case of problems due to indirect link end user - MNO
QoS	Available bandwidth is split between all users
Branding & relationships	
Branding potential	High branding potential for sponsor via log-in page, promotions, advertisements, sales Less branding potential for airline, although still significant
Differentiation opportunities	Opportunity to differentiate from other airline companies providing paid connectivity

Network vitality	
Structural dependency	Quality requirements and delivery times of MNO may not be aligned with airline needs
Specialist knowledge	Airline needs internal experts / managers / sales Sponsor needs internal managers
Financial risk	Airline takes the financial risk of finding adequate sponsors

Freemium wholesale model and Freemium retail model

The freemium wholesale model and retail model are based respectively on the wholesale model and retail model explained earlier, with a major difference in customer segments and revenue streams. This difference is based on the fact that customers can be divided into two segments: the free users and the premium users. The free users will have access to broadband internet until a certain limit is reached (e.g. 50 MB) or at a much lower speed than premium users [21] with possible ads. The premium users have unlimited access to broadband or a good bitrate compared to free users. They often form a smaller percentage of the customers, although they are the only customer segment generating direct revenue streams for the airline (wholesale model) or the MNO (retail model). In the application of the framework to these two models the difference is in the value creation where the part offering the free connectivity can has additional value by managing end user credentials, collecting data and the opportunity to advertise and promote.

Complementary service model with free connectivity for all passengers

The complementary service model is similar to the wholesale model with the main difference is that the airline offers the connectivity to all end users for free. Free inflight Wi-Fi can be a key differentiator for airlines that want to obtain a higher market share. A Wi-Fi connection without charge can drive preference and serve as a platform for connectivity-dependent revenues such as taxi service, other flights, promotions, and food or drink orders [19]. Loyalty and perception by the current customers can be improved and new customers can be attracted by introducing a frictionless service. The price of inflight connectivity can be reflected in the ticket price for premium customers and be part of their amenity package, or broadband could be traded for loyalty program points. The latter could be very effective as frequent travelers are often travelling for business – but not always in business class, which will make them value a stable Internet connection more than other rewards [21]. The application of the framework to the complementary service model is summarized in Table 3-6.

Table 3-6: Application of the framework to the complementary service model

Value optimization	
Value creation	Value for end user: connectivity Value for end user: frictionless service (passengers do not

	have to worry about payment onboard or in advance) Value for airline: end user credentials, data and the opportunity to advertise and promote
Financial impact MNO	MNO makes profits by selling the project to the airline
Financial impact airline	Airline has an opportunity to create extra revenues if ticket demand rises, resulting in a higher revenue potential Threat of hidden costs for airline
Financial impact end user	Service is free
Intangible asset management	End user feedback can be incorporated in service
Business performance	
Technology	Sat5G required to reach high performance for passengers
Channel management	Business performance efficiently handled thanks to direct link between end user and MNO
QoS	Risk of passengers consuming large amounts of data, reducing performance for other passengers
Branding & relationships	
Branding potential	Increased perception of airline that can lead to higher loyalty
Differentiation opportunities	Airline can brand itself as a free onboard internet provider (including internet for business needs) and grow its name Opportunity to differentiate from other airline companies Opportunity to serve as platform for other connectivity-dependent revenues: sales - promotions
Network vitality	
Structural dependency	Quality requirements and delivery times MNO may not be aligned with airline needs
Specialist knowledge	Airline needs internal experts / managers / sales
Financial risk	Airline takes the financial risk (incorporating connectivity service in ticket price may lead to lower ticket demand)

By applying the suggested framework, presented earlier in Table 3-1, to the identified business models, several conclusions were drawn. In cases where the airline company wants full control over pricing, the wholesale model is suggested. The retail model is beneficial for airline companies who are more risk-adverse and do not want to drift away from their core business. The customer will also benefit from the retail model, as the value chain is shorter, and the quality is assured. The sponsorship model is suggested when the need or interest for marketing of the company is rather low. This could be because tickets already sell out fast, or that a low-cost structure is the main marketing strategy. The freemium wholesale and retail models are very interesting for both MNOs/airline companies and passengers, because they bear many branding opportunities, and customers are attracted by free services. Combining this with a good quality service and the option to generate extra connectivity-dependent revenues, can make up a very viable business model. Key to making the freemium business model work is having a conversion rate from free to premium that is high enough. The complementary service model will be preferred for premium airline companies, as customers from these companies prefer to pay for the full premium service, in which connectivity is already included.

Currently, 4 out of 5 airline operators prefer the wholesale model over the retail model, although some airlines offer the service for free [6]. The inflight connectivity market is expected to become a $30 billion market by 2035, compared to a 1$ billion market in 2018 [22]. The fact that this market will become large implies that, if the conversion rate is high enough, a freemium business model should certainly be considered in the rest of the analysis.

3.4 Techno-economic analysis and business models

Following the qualitative analysis of potential business models, this section will present a quantitative analysis of the economic viability of the IFEC using integrated satellite and 5G network. First, the evaluation of the Total Cost of Ownership (TCO) of the proposed solution is detailed in section 3.4.1. In section 3.4.2, revenue models for different business models and benchmarking with real life examples are described. Section 3.4.3 presents the evaluation of the uncertainty of the input using sensitivity analysis.

3.4.1 Quantitative business models evaluation: TCO model

We use a TCO model to evaluate the economic viability of the proposed solution for the specific use case. The cost model structure adapted in this analysis is presented in Figure 3-4.

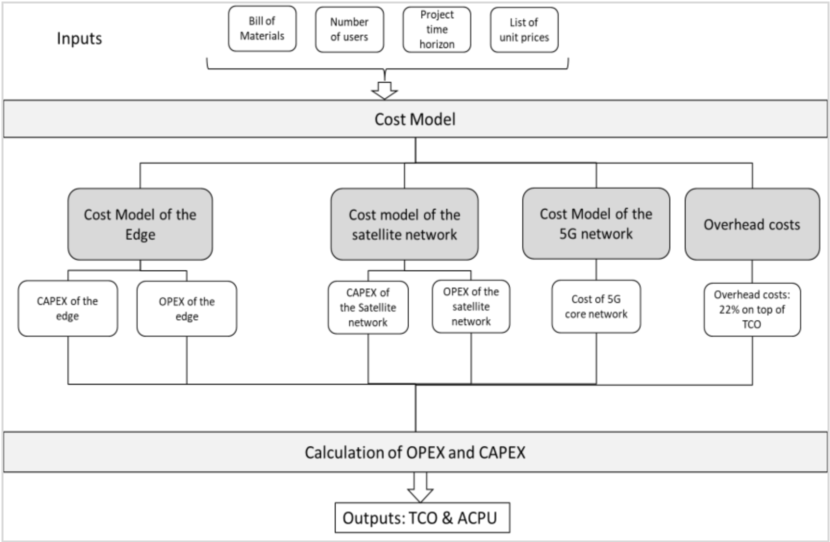


Figure 3-4: TCO model structure

The main inputs of the model are the bill of materials (BOM), the number of passengers/users, the minimum bitrate per user, the flight duration, number of flights per year and the time horizon of the project. Those inputs feed into a cost model that consists of four sub-models in alignment with the network architecture components presented in Figure 3-5. After the calculation of the CAPEX and OPEX for all blocks, the TCO can be derived. Finally, given the TCO as well as the number of users, the ACPU (Average Cost Per User) can be derived as an output of the model.

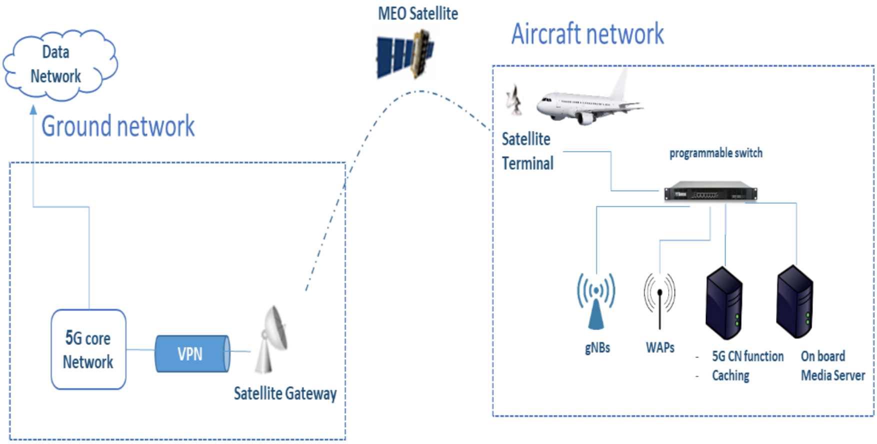


Figure 3-5: Network architecture of the proposed Sat5G integrated solution

Hereafter, we will detail the formulas used to calculate both CAPEX and OPEX for each sub-network following the structure of the cost model.
The cost of the aircraft network incorporates the CAPEX of the 5G gNB, Wi-Fi Access Points (WAPs), the media server and the CAPEX of the satellite terminal and the OPEX of all the network components. The CAPEX and OPEX of the aircraft network are calculated using the following formulas:

Equation 3-1

$$CAPEX_{NetAircraft} = Nb_{gNB} \times CAPEX_{gNB} + Nb_{WAP} \times CAPEX_{WAP} + CAPEX_{ST} + CAPEX_{MediaServ}$$

Equation 3-2

$$OPEX_{NetAircraft} = (Nb_{gNB} \times P_{gNB} + Nb_{WAP} \times P_{WAP} + P_{ST} + P_{MediaServ}) \times C_{watt} + M$$

The cost of the satellite link depends a lot on how much satellite capacity is needed to serve the aircraft users. The total satellite capacity needed is:

Equation 3-3

$$Sat_{Cap} = N_u \times Br_u$$

The OPEX of the satellite network is calculated using the following formulas:

Equation 3-4

$$C_{satCapS} = Sat_{Cap}\,(1 - C_{rate}) \times 12 \times C_{satMbps}$$

Equation 3-5

$$Opex_{sat} = C_{satCapS} + P_{satGat} \times C_{watt} + M$$

The cost of the airline data center is mainly the CAPEX of the servers used and the OPEX of them which is the power consumption and the maintenance costs.
For the core network or operator central office the CAPEX and OPEX are calculated like below:

Equation 3-6

$$CAPEX_{core} = C_{HW_{Mbps}} \times Sat_{Cap} \times Rate_{ctrl\_traffic}$$

Equation 3-7

$$OPEX_{core} = C_{SW_{Mbps}} \times Sat_{Cap} \times Rate_{ctrl\_traffic}$$

We have built the cost model using the following assumptions for the additional costs:

- *Hardware installation cost is 15% of the hardware costs [23]. The cost of the hardware installation is part of the CAPEX costs and it is expressed by the following formula:*

Equation 3-8

$$C_{Inst} = 15\% \times C_{Hw}$$

- *Maintenance cost is 10 % of the CAPEX costs [23]. Maintenance costs are counted in the OPEX costs.*

Equation 3-9

$$M = 10\% \times CAPEX$$

- *In most cases, the overhead cost is defined as the cost of marketing, helpdesk, human resources, finance etc. According to [24], it is around 22% on top of the sum of the CAPEX and OPEX costs.*

Equation 3-10

$$Ovhd_C = 22\% \times (CAPEX + OPEX)$$

The Total Cost of Ownership (TCO) of the proposed solution is counted as the sum of the CAPEX, the OPEX of 5 years and the overhead costs of 5 years.

Equation 3-11

$$TCO = CAPEX + \sum_{t=1}^{5} (OPEX(t) + Ovhd_c(t))$$

With:

Br_u	bitrate per user	*Nu*	number of users
Nb_{gNB}	number of gNB installed in the aircraft	*CsatMbps*	cost of 1 Mbps per month via satellite link
C_{Inst}	cost of hardware installation	*M*	maintenance costs
C_{enb}	cost of gNB	*NbWAP*	number of WAPs installed in the aircraft
C_{ST}	cost of satellite terminal	*Cwatt*	cost of energy
$C_{Mediaserv}$	cost of the media server	*Ratectrl_traffic*	percentage of the control traffic from the overall traffic
C_{rate}	caching rate	*CHW_Mbps*	cost per Mbps of control traffic of the hardware of the 5G core network
C_{SW_Mbps}	cost per Mbps of control traffic of the software of the 5G core network	*P(i)*	power consumption of (i) device: i:{gNB; WAP; Media server}

Inputs used to run the cost model are presented in Table 3-7. In addition, the cost model is based on several assumptions that should be made to have as realistic results as possible:

1. We assume the use of MEO satellites within the O3B constellation;
2. Satellite handovers within the O3B constellation are pure satellite procedures and they do not affect the 5G network, hence we consider only one satellite gateway to communicate with the 5G core network;
3. We assume that the airplane flies within the coverage of MEO satellites;
4. When the ground gNB that the UE is attached to hands his traffic to the gNB located on board, different business entities are involved. This might need a different process comparing to the normal handover within 5G network (because the new gNB is backhauled by satellite connectivity, it might be different because the satellite link carries less capacity and more latency). However, for sake of simplicity, we will model this case as a normal handover within the 5G network context;
5. We use a combination of small cells and WAPs. WAPs are mainly used for the local content, but they can carry traffic to the 5G core network.

The scenario considered here is a flight from Munich to Portugal. We assume that the airplane flies within the coverage of MEO satellites. Since there are specific locations for satellite gateways within the O3B constellation e.g. Nemea (Greece), Sintra (Portugal) etc. and in this scenario, we chose to work with the satellite gateway located in Sintra in Portugal. The virtualized 5G network core is located in the airline data center in Munich.

Table 3-7: Cost model input

Aircraft Network	Price
Wi-Fi AP: Aero-adapted [25]	5 000 – 10 000 €
5G Small Cell: Aero-adapted gNB HW: Ettus B210 SDR, Intel Nuc Server with 4 cores CPU, 16 GB RAM and 500 GB Disk SW: OAI 4G CN srsLTE [25]	50 000 – 100 000€
Number of gNBs	1
Number of WAPs	3
Media Server HW: COTS SW: Open Source / Proprietary [25]	30 000 – 50 000€
Number of media servers	1 for caching and 1 for media server
Caching license [25]	12 000€
5G Satellite network	Price
VSAT	200 000 – 500 000€

Aircraft Network	Price
SES AvL antenna + radome [25]	
5G Core Network	
Power usage: 5G	14 907 kWh / year
Aircraft Network	Quantity
Power usage: Servers	298 kWh / year
Power usage: gNB	0,7 kWh / year
5G Satellite network	Quantity
Power usage: VSAT	447 kWh / year
Satellite capacity [25]	75 -> 35 between 2020-2025 $ / Mbps / user
VPN: BT [25]	8087 €: 20K euro for the VPN divided by the number of flights

Assuming a caching rate of 40%, the results of the studied scenarios are presented in Figure 3-6.

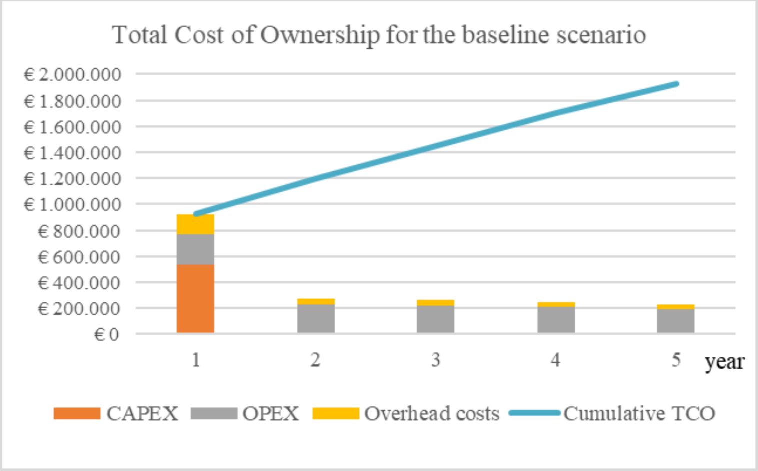


Figure 3-6: TCO and other cost components of the SaT5G proposed solution

The proposed solution with a caching rate of 40% results in 3.5 euro per user as an ACPU and 0.1 euro per 100 MB of data consumption. A sensitivity analysis will be elaborated on the caching rate in section 3.4.3.

3.4.2 Revenue models for different business models and benchmarking

Access to the onboard wireless network can be purchased by a passenger either online while buying the flight ticket, or directly onboard, and both using different models. For instance, a minimum amount of time or unlimited navigation with reduced download speed can also be made available depending on the airline. Not every revenue model is feasible for every business model. Feasible pricing models for the identified business models described in section 3.3.3 are discussed in the Table 3-8.

Table 3-8: Feasible pricing models per business model

Business model	Feasible considered pricing models
Wholesale	Data-based, subscription-based, tiered-bandwidth, per-flight, time-based pricing
Retail	Data-based, subscription-based, tiered-bandwidth, per-flight, time-based pricing
Sponsorship	Average sponsorship deal size per month
Freemium wholesale	Subscription-based, per-flight pricing
Freemium retail	Subscription-based, per-flight pricing
Complementary service	Indirect increase in ticket price

In order to validate the findings of this use case, a benchmarking with the existing inflight connectivity service is made and presented in Table 3-9 (assuming a profit margin of 30% on top of the calculated ACPU).

Table 3-9: Comparison between SaT5G prices and existing solutions

Airline	Provider	Business model(s)	End-user price	Pricing model(s)	Average price per MB	Comparison with SaT5G price (30% as profit margin)
Aer Lingus [26]	Panasonic Avionics, Aeromobile, Deutsche Telekom	Retail, Complementary service	Economy class: €6,95 / 50 MB; €13,95 / 120 MB; €29,95 / 270 MB Business class: free	Data-based pricing	€ 0,122	**74 times** more expensive than the Sat5G costs.
Air Canada [27]	Gogo	Retail	$49,95 / 1 month; $5 / 1 hour; $14 / 24 hours	Subscription-based pricing, time-based pricing	€ 0,017	**12 times** more expensive than the Sat5G costs.
Air France [6]	Panasonic Avionics	Unclear	€5 / 20 MB; €10 / 50 MB; €30 / 200 MB	Data-based pricing	€ 0,200	**130 times** more expensive than the Sat5G costs.

From Table 3-9, we can conclude that under the taken assumptions, the Sat5G integrated solution is cheaper comparing to other inflight internet providers. However, it is noteworthy that in the cost model of the Sat5G solution, the cost of the 5G mobile network spectrum is not included and it might be that the existing services count the spectrum costs in the price of the service in order to pay off this high cost.

3.4.3 Sensitivity analysis

As discussed previously, the assumption on the caching rate has a significant impact on the results since it affects the required satellite capacity to serve the aircraft. Therefore, in this section we study the impact of the caching rate variation on the ACPU.

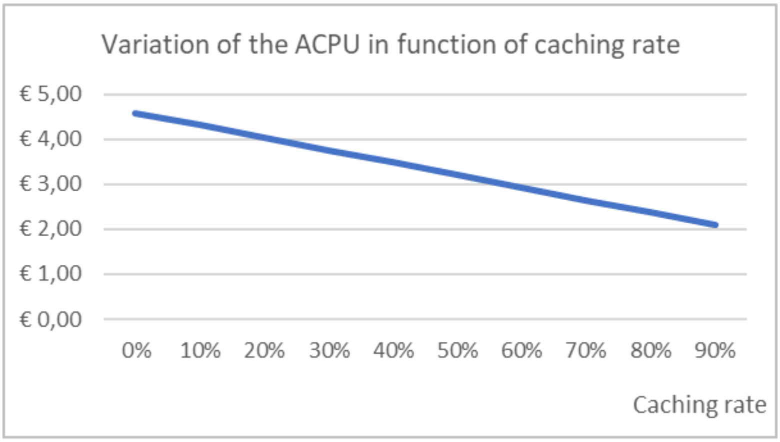


Figure 3-7: Impact of the caching rate on the ACPU

It is clear from Figure 3-7 that the ACPU is, as predicted, strongly related and affected by the caching rate assumed. The ACPU decreases when the caching rate increases, since we serve the users more from the cached data on the media server located on the aircraft than from the traffic sent via the satellite link. Thus, we use less satellite capacity.

In addition to the caching rate that significantly impacts the ACPU, there is also the satellite capacity cost that is considered as an important OPEX component. Therefore, in Figure 3-8, the ACPU is drawn in function of the variation of the satellite capacity cost.

Results show that there is a linear relation between the satellite capacity cost and the ACPU and the average cost per MB of data. The ACPU is very sensitive to the satellite capacity cost but still not expensive even with 105 $/Mbps/month, comparing the inflight connectivity prices as discussed in the previous section.

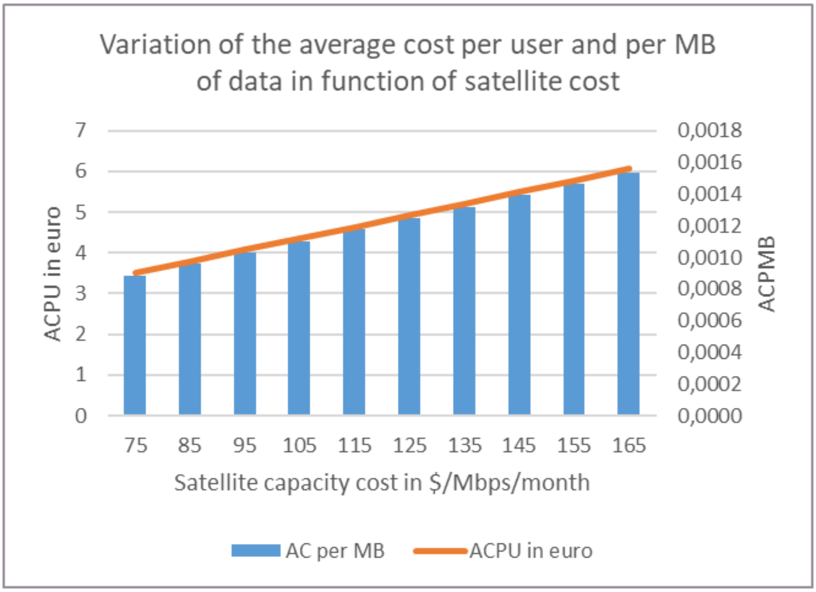


Figure 3-8: *impact of the satellite capacity cost on the ACPU and ACPMB*

3.5 Conclusion and recommendations

The integration of satellite and 5G network provides a highly promising solution for providing connectivity to remote or harder to reach locations. Following a network architecture building on indirect access via a Non-Terrestrial Network, this paper investigated the cost and the possible business models for providing broadband service to onboard users. Different business models were identified for the proposed solution based on who will provide the service to the end user. Depending on the level of control and risk the airline company wants to take, the choice for a wholesale versus a retail model can be made. The customer will benefit more from the retail model, as the value chain is shorter, and the quality is assured. In case the main goal is a low-cost structure, the sponsorship model is suggested. The freemium wholesale and retail models are very interesting for both MNOs/airline companies and passengers, because they bear many branding opportunities, and customers are attracted by free services. Key to making the freemium business model work is having a conversion rate from free to premium that is high enough. The complementary service model will be preferred for premium airline companies, as customers from these companies prefer to pay for the full premium service, in which connectivity is already included. Different pricing schemes were proposed and matched to the identified business models. In addition, a Total Cost of Ownership (TCO) model was built to investigate the economic viability of providing inflight connectivity using the integrated networks. The main results of this model showed that providing inflight broadband service with a good throughput per user (2 to 5 Mbps) with reasonable prices is feasible thanks to the integration of satellite communications into the 5G network. Moreover, the impact of new technology such as caching techniques on the business case has been investigated and it proved that caching popular data reduces the OPEX and hence the ACPU (from 25 to 32% depending on the caching rate used). However, the higher the caching rate, the lower the personalization of the content on the aircraft, so a trade-off needs to be made. The extension of this analysis to other moving platforms such as trains and ships would be an interesting piece of follow-up work.

References

[1] future travel experience, "Airlines face big decisions as inflight connectivity uptake continues to rise," 2018. [Online]. Available: https://www.futuretravelexperience.com/2018/02/airlines-inflight-connectivity-uptake-continues-to-rise/. [Accessed 19 11 2019].

[2] ETSI, "ETS 300 321 “Radio Equipment and Systems (RES); Terrestrial Flight," European Telecommunications Standards Institute, Sophia Antipolis - Valbonne - FRANCE, October 1996.

[3] ETSI, "White Paper No. 4: GSM operation onboard aircraft," ETSI , France, 2007.

[4] Wikipedia, "Connexion by Boeing," [Online]. Available: https://en.wikipedia.org/wiki/Connexion_by_Boeing.

[5] A. Lingus, "Wi-Fi and connectivity," [Online]. Available: https://www.aerlingus.com/experience-aer-lingus/onboard-features/wifi-connectivity/.

[6] Air France, "Air France launches its new in-flight Wi-Fi offer," [Online]. Available: https://corporate.airfrance.com/en/news/air-france-launches-its-new-flight-wi-fi-offer .

[7] emirates. [Online]. Available: https://www.emirates.com/nl/dutch/experience/inflight-entertainment/onboard-wifi. [Accessed 2019].

[8] edreams. [Online]. Available: https://www.edreams.com/blog/in-flight-wifi/.

[9] SaT5G, "SaT5G," [Online]. Available: https://www.sat5g-project.eu/. [Accessed 12 05 2020].

[10] SES, "four-reasons-high-throughput-satellite-will-be-game-changer," [Online]. Available: https://www.ses.com/four-reasons-high-throughput-satellite-will-be-game-changer. [Accessed 25 04 2020].

[11] ETSI, "ETSI TR 103 611 V0.0.6: Satellite Earth Stations and Systems (SES); Seamless Integration of Satellite and/or HAPS (High Altitude Platform Station) Systems into 5G and related Architecture Options," European Telecommunications Standard Institute, 2020.

[12] anixter, "the unlicensed spectrum," [Online]. Available: https://www.anixter.com/en_us/resources/literature/technology-application-guides/lte-u-and-vowifi-mobile-on-the-unlicensed-spectrum.html. [Accessed 12 05 2020].

[13] SaT5G, "D4.4 Harmonisation of SATCOM with 5G Control and User planes," Sat5G consortuim, 2019.

[14] O. G. A. &. V.-C. A. Kodheli, "Integration of Satellites in 5G through LEO Constellations," in *GLOBECOM 2017 IEEE Global Communications Conference*, 2017.

[15] A. A. V.-C. M. C. S. A. S. C. N. M. B. E. e. a. Guidotti, " Architectures and key technical challenges for 5G systems incorporating satellites," *EEE Transactions on Vehicular Technology,* vol. 68, no. 3, pp. 2624-2639, 2019.

[16] V. Allee, "Value network analysis and value conversion of tangible and intangible assets," *J. Intellect. Cap,* vol. 9, no. 1, p. 5–24, 2008.

[17] V. Allee, "Presentation: What is Value Network Analysis?," 12 04 2019. [Online]. Available: https://www.slideshare.net/baba2729/what-isvaluenetworkanalysis.

[18] E. D. T. Monitor, "Inflight entertainment and communication: Technologies and market," 2017.

[19] F. C. Reinhart, "The anatomy of inflight connectivity," Gogo.

[20] Forbes, "Delta Aspires To Offer Free In-Flight Wi-Fi For All Passengers," [Online]. Available: https://www.forbes.com/sites/bishopjordan/2018/10/02/delta-free-in-flight-wi-fi/#4d795d331243. [Accessed 11 04 2019].

[21] I. Aviation, "Making inflight broadband pay," [Online]. Available: https://www.inmarsataviation.com/en/benefits/revenue-opportunities/making-inflight-broadband-pay.html. [Accessed 01 04 2019].

[22] I. Aviation, "Broadband in the sky will be a $130 billion market by 2035," [Online]. Available: https://www.inmarsat.com/news/34067/. [Accessed 13 04 2019].

[23] K. Casier, Techno-economics of FTTH deployment in the presence of competition (Doctoral dissertation), 2010.

[24] "atkearney," [Online]. Available: https://www.atkearney.com/documents/20152/434204/Telco_Fitness.pdf/427dcc76-15f9-b9a3-0fd5-5c2e157918d6.

4

Network slicing cost allocation model

Asma Chiha, Marlies Van der Wee, Didier Colle, and Sofie Verbrugge

Published in the Journal of Network And Systems Management

The previous chapters explored the possibility of deploying 5G networks for specific services in different regions. The next issue to address is how to fairly distribute the cost of deploying the network among the services to be provided. Therefore, this chapter offers a solution to this research question by presenting an allocation model specifically designed for slice-based 5G networks. This model outlines how to distribute the costs of network deployment among the different slices, which can then be used to determine the prices of various end-to-end slices (services). To provide the necessary inputs for this proposed model, the chapter introduces a resource allocation algorithm and a 5G network function dimensioning model. Additionally, the chapter examines the various outcomes of the different models and the cost savings that arise from utilizing NFV in the core network.

***

Abstract—Within the upcoming fifth generation (5G) mobile networks, a lot of emerging technologies, such as Software Defined Network (SDN), Network Function Virtualization (NFV) and network slicing are proposed in order to leverage more flexibility, agility and cost-efficient deployment. These new networking paradigms are shaping not only the network architectures but will also affect the market structure and business case of the stakeholders involved. Due to its capability of splitting the physical network infrastructure into several isolated logical sub-networks, network slicing opens the network resources to vertical segments aiming at providing customized and more efficient end-to-end (E2E) services. While many standardization efforts within the 3GPP body have been made regarding the system architectural and functional features for the implementation of network slicing in 5G networks, techno-economic analysis of this concept is still at a very incipient stage. This paper initiates this techno-economic work by proposing a model that allocates the network cost to the different deployed slices, which can then later be used to price the different E2E services. This allocation is made from a network infrastructure provider perspective. To feed the proposed model with the required inputs, a resource allocation algorithm together with a 5G network function (NF) dimensioning model are also proposed. Results of the different models as well as the cost saving on the core network part resulting from the use of NFV are discussed as well.
Keywords: NFV, SDN, resource allocation, cost model, 5G, satellite, network function resource modelling.

4.1 Introduction

The exponential increase of the Internet speed together with emerging services and applications are considered obvious from a customer point of view. However, telco operators face serious investment challenges to meet these increasing expectations of the end user while keeping a reasonable Average Revenue Per User (ARPU) to not experience a customer churn.

Currently, traditional network architectures are static, decentralized and complex. In addition, network infrastructures are designed to fulfil the requirements of dedicated, specific services using specialized network hardware (known as hardware appliances or middleboxes). These services, however, do not make full use of the available resources all the time. This inefficiency leads to high capital costs for building, maintaining and operating the networking infrastructure, leading to long return on investment cycles. Furthermore, the optimization of the network infrastructure for a specific service prevents the addition of new features and services, slowing down or even preventing innovations. Technical improvements and new technological trends, such as SDN, NFV and network slicing, have to be investigated in order to make the network architecture more flexible and agile.

For the sake of flexibility and easy troubleshooting of the network, a centralized network intelligence component is needed. Nevertheless, we cannot achieve a centralized control of the network without disassociating the routing process (control plane) from the forwarding process of network packets (data plane). The

SDN paradigm is an approach that enables dynamic and programmatically efficient network configuration by splitting the control plane and the data plane and centralizes the control of the network by using an SDN controller [1].
On the other hand, Network Function Virtualization (NFV) moves the network functionalities from hardware into software such that they can run on a range of standard-based hardware which may easily be moved to any location within the network if needed. The focus of the NFV concept is on reducing the need to add new equipment. This enables more flexibility to scale up or scale down, and more opportunity to innovate, experiment and deploy new services with lower risk [2].
These new emerging technologies being SDN and NFV are jointly beneficial and complementary to each other since they both share the same feature of encouraging innovation, creativity and competitiveness [1], [3]. Therefore, NFV and SDN together are considered as the key enablers of the network slicing concept. According to [4]: "A network slice is viewed as a logical end-to-end network that can be dynamically created. A given User Equipment (UE) may access to multiple slices over the same Access Network (e.g. over the same radio interface). Each slice may serve a particular service type with agreed upon Service-level Agreement (SLA)". Network slicing allows different service providers with disparate traffic requirements to share the same infrastructure resources. This does not mean deploying Virtual Local Area Networks (VLANs) to isolate traffic flows, but it means splitting physical network resources and network functions to deploy an exclusive end-to-end (E2E) implementation for each application, e.g. enhanced Mobile Broadband (eMBB), ultra-Low Latency Communication (uRLLC).
The deployment of multiple network slices requires the network to reserve enough resources for each specific slice instance according to its requirements. For the radio access network (RAN) part, the network needs to allocate physical radio resources to each specific network slice based on the required bandwidth. On the core network part, an isolated set of logical network functions may be used which are provisioned for this specific network slice combined with logical and physical network functions which are shared among multiple network slice instances. Network slicing gives Mobile Network Operators (MNOs) the opportunity to open their virtualized networks to vertical segments such as the healthcare sector, car manufacturers, intelligent transport providers etc., as well as to service providers that lack physical network infrastructure. MNOs need resource allocation models to distribute the resources among themselves and their tenants in a way that first adheres to the SLA agreed beforehand and, second, makes effective use of the network. Resource allocation problem is seen as one of the challenges that face slicing-enabled networks according to many surveys on network slicing [5], [6].
However, not only the allocation of network resources to the different slices is considered as a crucial element towards the efficient use of the network slicing - and hence the efficient use of the network - also techno-economic challenges pop-up with the introduction of this new technique [5]. From a techno-economic

perspective, two main research questions are raised due to the use of the network slicing concept, being, first, how much cost can be saved by using network virtualization and network slicing and second, how to allocate the cost of the network to the running network slices/services?

In this paper, we propose a cost allocation model for network slicing that aims to allocate the cost of the network to the different network slices deployed on this network, as well as a cost model to derive how much can be saved by using virtualization on the core network. To this end, we present, in the next section, a literature review of the allocation models for network slicing and cost allocation models in general. In section 4.3, we describe in a detailed way the proposed model, which requires a slice dimensioning model in order to distribute the cost of the network to the running slices based on their consumption in terms of network resources. The resource allocation model or slice-dimensioning model demands the hardware requirements of each 5G network functions NFs as inputs. Thus, section 4.4 explains how we modelled the 5G network functions requirement for both the control and the data plane. Section 4.5 presents the results of the application of the developed models to a specific use case, which was selected based on availability of cost model and slice assumptions. Finally, section 4.6 concludes the paper and presents our planned next steps.

4.2 Related work

In order to allocate the cost of the network to the different slices, we have to follow an optimized resource allocation model, optimized in the sense that this algorithm assigns the needed network resources efficiently to each slice taking into consideration all Key Performance Indicators (KPIs) and dimensioning.

However, in literature, the resource allocation models are often designed in order to optimize one of the network metrics e.g. round-trip time (RTT), Signal-to-noise ratio (SNR) etc. and several of them focus only on the radio resources and spectrum sharing, not on the core network resources. For example, authors of [7] proposed a subchannel allocation algorithm that allocates radio frequencies in spectrum-sharing two-tier systems to the different tiers while considering the co-tier interference and cross-tier interference. In addition, in [8] as well, a radio resource framework is designed in order to allow tenants to dynamically and flexibly distribute base station resources among their users taking into account both the admission control and user dropping mechanisms. Authors in [9] consider that by using cloud RAN (C-RAN), network slicing at the spectrum level can be easily implemented. The C-RAN technology is based on moving the Baseband Units (BBUs) from distributed base station locations into a centralized BBU place. This centralization of BBUs aims to achieve a multiplexing gain and increase network capacity by using load balancing and cooperative processing of the traffic sent by different base stations.

On the core network side, a network slice is seen as a set of Virtual Network Functions (VNFs) which are deployed on virtual machines (VMs) or containers. The resource allocation for the core network resources is a matter of mapping VNFs to VMs based on their hardware requirements such as processing power, memory and storage. Researchers often see the allocation problem as 1) a forecasting technique that requires three steps to be implemented: a) a forecasting module to predict the traffic of a network slice based on previous traffic and user mobility, b) a network slicing admission control algorithm, c) a network slicing scheduler algorithm to fulfil SLAs and report back to the forecasting module [10] or as 2) an VNF-RA (VNF resource allocation) problem as presented in a wide survey elaborated by [6]. Such an VNF-RA problem consists of 3 stages: VNFs - chain composition, VNF-forwarding graph embedding and VNFs - scheduling. However, few researches consider the three steps of the resource allocation problem together and no research considers the VNF itself. e.g. the consumed time of the VNF installation and removal processes.

The proposed resource allocation models for the core network part somehow assume the infinite availability of network resources and consider a cloud-based network model where the allocation of resources is a matter of optimization. However, for network infrastructure providers and data centres who own the network infrastructure, allocation models are not only needed to efficiently use the resources, but also as a forecasting tool of the required investment in the future to cope with the growth of demand and as an input to the cost allocation model.

To conclude, the state of the art of the resource allocation models shows that a complete model that allocates both radio and core network resources is still missing in literature.

On the economic side, SDN, NFV and network slicing technologies add more flexibility and easy control and to the network management in addition to the efficient use of the network resources. Certainly, all these advantages have an impact in term of cost. For example, the use of NFV may reduce the Capital Expenditures (CAPEX) due to the possibility to deploy VNFs on a vendor-independent hardware. SDN can lead to a possible reduction of the Operational Expenditures (OPEX) because of the easy troubleshooting and maintenance of the network. However, this latter may also add more complexity due to the centralization of the network control, hence increasing the signalling traffic on the control plane. Therefore, the influence of these new technologies on the cost of the network must be investigated carefully. Many researches in literature investigated the impact of SDN, NFV and network slicing from a techno-economic perspective and are summarized in the following table.

Table 4-1 SOTA of cost-benefit model for network slicing

Paper Title	**Studied technology**	**Proposed model**	**Methodology and studied use case**	**Main results**
Techno-economic analysis of software defined networking as architecture for the virtualization of a mobile network [11].	SDN& virtualization	Adaptive cost model for SDN concept	Cost model for SDN networks but this model considers only the CAPEX. German reference network scenario: - Classical scenario: a distributed network architecture with distributed network control. - SDN scenario: a centralized network architecture with decoupled network control from data plane using Open-Flow as communication interface. - Sharing scenario: network virtualization and network sharing between several network operators with a FlowVisor controller.	- SDN and virtualization of the first and second aggregation stage of the network infrastructure leads to considerable CAPEX cost reductions for the mobile network operator. - A 13.81% CAPEX reduction can be achieved for the SDN scenario in comparison with the classical scenario. - A 58.04% CAPEX reduction can be achieved for the SDN based sharing scenario in comparison with the classical scenario.
Cost Modeling for SDN/NFV Based Mobile	SDN & NFV	Cost model	Comparison between Traditional network (with Virtual EPC) and cloud RAN network also with	- A 63% OPEX reduction can be achieved in comparison with the traditional scenario.

Paper Title	Studied technology	Proposed model	Methodology and studied use case	Main results
5G Networks [12].			vEPC for Sweden network. However, authors consider only for the OPEX the cost of the power consumption (no maintenance cost nor a failure cost etc..)	- The considered CAPEX could be reduced by 68% in comparison with the traditional scenario. - The considered TCO could be reduced by 69% in Comparison with the traditional scenario.
Cost Efficiency of SDN in LTE-based Mobile Networks: Case Finland [13]	SDN	Cost model	Cost model taking into account both CAPEX and OPEX applied to a Finnish reference network.	- SDN decreases the network related annual CAPEX by 7.72% and OPEX by 0.31% compared to non-SDN LTE. - The cost reduction is a small fraction of the total expenses of a Finnish MNO, but may has a important influence on the profit levels.
Life-cycle cost modelling for NFV/SDN based mobile networks [14].	NFV/SDN	Life-cycle cost (LCC) models	Cost model including both CAPEX and OPEX used to compare non-virtualized network costs with virtualized one with many options. The modelling considers the hardware requirement for each virtualized network function.	- The non-virtualized network has the highest TCO. - Among the virtualized flavours, the setup with 6WINDGate speed-up technology is cheapest
Modeling Profit of Sliced 5G Networks for Advanced	Network slicing	Revenue model	Multi-Objective Optimization Problem (MOOP) to increase operator Profit: mathematical optimization.	- Novel methodology of modeling profit generated by 5G network slices. - The expenditures and revenues of a network can be estimated according to its slice

Paper Title	**Studied technology**	**Proposed model**	**Methodology and studied use case**	**Main results**
Network Resource Management and Slice Implementation [15].				properties, such as KPI requirements and service prices. - Based on this, the slice implementations can be optimized to maximize the profit.
Network Slicing as a Service: Enabling Enterprises' Own Software-Defined Cellular Networks [16].	Network slicing	Business Model and Service Model for NSaaS	Used a detailed process of NSaaS concept by typical examples, together with the configuration process, product management of NSaaS and management APIs for customers	Detailed business and service model for the NSaaS concept that help operators to offer tailored end-to-end cellular network as a service

These researchers focused on the quantification of the cost saving resulting from the use of SDN and NFV [11]-[14], but none of them tackled the cost model for network slicing. For example, [15] presented a revenue model for network slicing based on MOOP and [16] proposed a new business and services model for network slicing as a service. However, how much exactly does an eMBB slice that requires a high network capacity costs, consume compared to an uRLLC slice that requires low latency? How can a network infrastructure provider dimension slices effectively considering the hardware requirement of each slice, in order to reduce costs and maximize revenues? Though they are very interesting research questions, to the best knowledge of the authors, there is no paper available so far that models the allocation of the network costs to the different slices.

Cost allocation models are built using two main components, being (1) the cost model of the sliced network and (2) the hardware requirements of each of the network slices. Both parts are indispensable for any operator to ensure a fast Return on Investment (ROI). To this end, we propose, in the next section, our proposal for a network slicing cost allocation model.

4.3 Network slicing cost allocation model

A slice is a logical subnetwork defined on top of the physical network to deliver a specific service such as video streaming, Internet of Things (IoT) services, etc. This logical network is customized in a way to fulfil a set of KPIs: availability, reliability, capacity, efficiency, and latency etc. To satisfy those KPIs, different types of network resources have to be reserved for this specific slice.

The network consists of three parts: core network, transport network and radio access network. An E2E slice has to be built taking into account these three different network components. On the core network side, a slice is a chain of VNFs and physical network functions. Yet, on the transport network part, the slice can be seen as a pipe or tunnel with a specific bandwidth reserved for this slice. However, on the RAN side, the slicing is made by means of a frequency subchannel reservation. Therefore, the question that raises here, as discussed previously in the introduction, is how to derive the cost of the slice given its exigence in term of these different resource types?

In addition, from a network infrastructure provider point of view, a model that allocates cost to the different slices is a useful input to their pricing models. From an MNO's point of view, such a model is also needed in order to identify what are the most consuming slices in term of hardware and radio resources.

Therefore, we propose a cost allocation model that aims, first, to fairly allocate network costs to deployed slices (services) and, second, to show that network slicing makes more efficient use of the network. The proposed model is built on the following assumptions[11]:

[11] The authors are aware that these assumptions simplify the modelling but aim to extend the model to more complex cases in a later stage (as will also be described in section 4.6).

- A network slice will support one specific service.
- A predefined throughput is reserved per service on the RAN and the transport network in a static way.
- On the core network side, a service/slice consists of a chain of (virtual) network functions, such as 5G functions besides physical network functions.
- Virtual network functions (VNFs) are running on virtual machines (VMs), which are in turn executed on servers, and one VM can only host one VNF.

The different building blocks of the suggested model are presented in Figure 4-1. The inputs of the model consist of two different type of inputs: i) available inputs which are related to the use case (e.g. slice type and the number of users) or can be derived easily from literature e.g. the chain of network functions per slice, those inputs are represented in the diagram with white boxes; ii) derived inputs from other models that we developed in order to feed this model with the required data, those are represented with the light grey boxes.

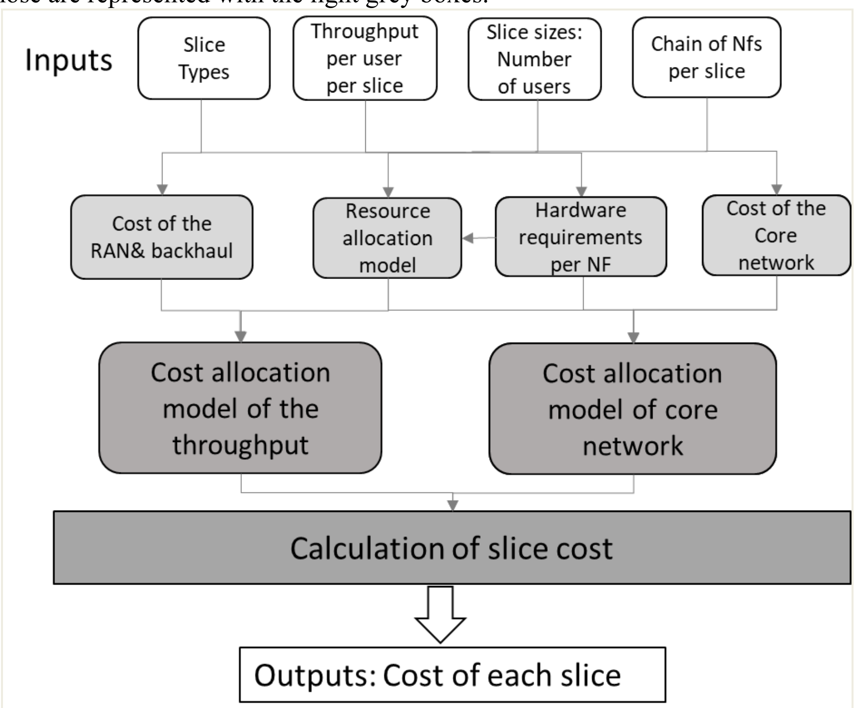


Figure 4-1: cost allocation model diagram

Both the cost model of the RAN and the backhaul network and the cost model of the core network used as cost inputs to this model are described in [17]. Since a complete model that allocates radio, transport and core network resources in the context of a sliced network is still missing in literature, and since existing models do not dive deep into the VNF characteristics-level of granularity (such as how much memory, processing power and storage each VNF needs to mitigate the

overall QoS imposed by the slice), building our cost allocation model in the light of an existing resource allocation model is not possible. Consequently, we propose an allocation model for network slicing based on the hardware requirements and the throughput required for each slice. This model aims at allocating the right amount of resources to each slice. In order to accomplish this, we have to distinguish between the three type of network resources discussed previously: RAN, transport network and core network resources. The RAN and transport network resources can be allocated by reserving a predefined throughput per slice on the base station and the backhaul network. Yet, for the core network resources this allocation is done in two steps: i) identify the hardware requirements for each 5G VNF and ii) map the 5G NFs onto VMs and then onto servers. It is noteworthy that the step of the hardware requirements identification per VNF needs a separate model that we developed and will be described in section 4.4.
Afterwards, a cost allocation model to assign the cost of each resource type to each slice is needed. The cost of the RAN and transport network can be modelled as the cost of the throughput reserved on these two network parts and is presented in section 4.3.1. For the core network part, section 4.3.2 summarizes the cost model of the core network resources. Hence the cost of a slice can be derived using Equation 4-1.

Equation 4-1:

$$C_{\text{slice}} = C_{\text{Thp}} + \sum_{i=1}^{K} C_{\text{VNF (i)}} + \sum_{j=1}^{l} C_{\text{phNF (j)}}$$

With:

C_{slice}: cost of the slice
C_{Thp} : cost of the throughput of the slice
C_{VNF} : cost of the VNF
K: number of VNFs in the slice
C_{phNF}: cost of physical NF
l: number of physical NFs in the slice

4.3.1 Cost allocation model of the throughput

For the throughput metric, a predefined throughput is reserved per slice in a static way[12]. The cost of this metric can be concluded based on consumed throughput from the overall throughput provided by the base station. The throughput required by each slice has to be reserved over the transport link as well. Hence, the cost of the throughput should incorporate a part of the cost of the transport link capacity. Moreover, each slice has its own exigence in terms of quality of service (QoS), such as priority level, packet error loss rate, packet delay etc. Thus, for each slice, we reflect the cost of the QoS requirement by the mean of a weighting coefficient.

[12] In the current version of the model, we consider that the traffic is static. Yet, in a later stage of the model we will include the dynamicity of the traffic, hence the allocation of the network resources will be dynamic as well (as specified in section 4.6).

This latter takes into consideration the required throughput of the slice as well as the desired priority level, thus the latency requirement, as shown in Equation 4-3. Therefore, the cost of the throughput for a specific slice s is calculated using Equation 4-2.

Equation 4-2:

$$C_{Thps} = (C_{BS} + C_{TrCap}) \times \frac{Qco_s}{\sum_{s=1}^{N} Qco_s}$$

Equation 4-3:

$$Qco_{Sj} = Cost_{sharC} \times \frac{Thp_{Sj}}{\sum_{i=1}^{N} Thp_{S(i)}} + (1 - Cost_{sharC}) \times (1 - \frac{L_{Sj}}{\sum_{i=1}^{N} L_{S(i)}})$$

With:

N: number of slices running

C_{BS}: cost of the base station

L_S : Latency of the slice S

Qco: weighting coefficient according to throughput and the QoS level of the slice

C_{Thps}: cost of throughput of slice s

C_{TrCap}: cost of the transport capacity link

$Cost_{sharC}$: Cost sharing coefficient

To be able to quantify the QoS weighting coefficient for each slice, we need to know the latency for the video and voice services as well as the throughput. For the former, we adopt the agreed 3GPP traffic types described in detail in [18], the latter will be provided in the scenario description used to validate the model.

4.3.2 **Cost allocation model of core network resources**

In order to fairly allocate the core network resource costs to the different slices based on the hardware requirements of each slice, we should first find a way to map those requirements to the resources used. This can be elaborated following two different approaches: an infrastructure provider or a telecom operator approach. An infrastructure provider, responsible for the operations of a data centre, tries to find a way to allocate the total cost of the data centre to the running services/slices to maximize the ROI. A telecom operator, on the other hand, leases the slices from the data centre and considers the business model of pay-per-use. A telecom operator hence only needs to know the cost of the service that he wants to offer (hence the cost of the slice that he wants to lease) beforehand to study the viability of providing such a service. Giving this, the cost allocation model differs according to whom it is designed for. In this section, we will detail the cost allocation model from a telecom infrastructure provider perspective.

In order to achieve this goal, the infrastructure provider needs to allocate resources to the different slices according to the latter's requirements in term of processing power, memory and storage capacity. This allocation consists of two steps mapping process: 1) mapping the VNFs to VMs, 2) mapping VMs to the data centre resources namely servers, as shown in Figure 4-2, and 3) the allocation of the network costs to the different slices can be achieved based on the resource consumption of each slice, which is the final step of the cost allocation model.

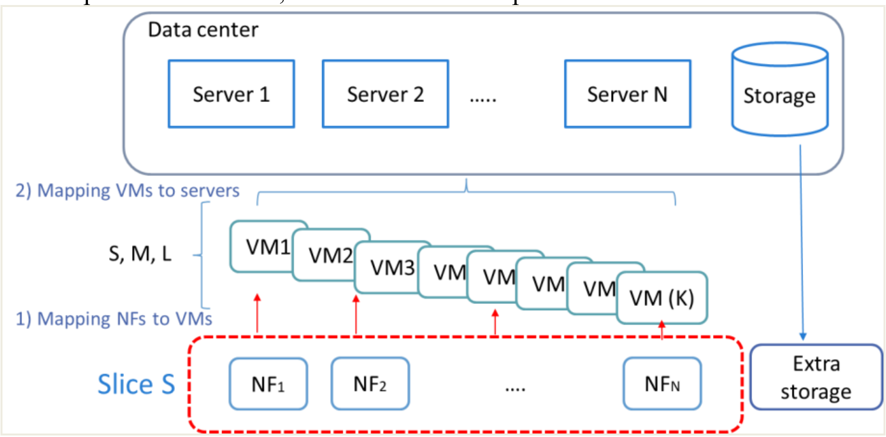


Figure 4-2: Mapping data center resources to slices

Step 1: Mapping VNF to VM

In order to accomplish the mapping of the VNFs onto VMs, we should identify first, based on the slice characteristics, the technical requirements for each VNF in term of computing, storage and networking. Giving those requirements, we can pick up the suitable VM for each specific VNF. To the best of authors knowledge, there are no references describing these requirements for the 5G networks available in the literature. Therefore, we propose, in section 4.4, our model to identify the hardware requirements of each 5G NF based on the number of users and the offered traffic of the slice.

After identifying the requirement of each NF, we should, then, be able to select the suitable VM for this NF, which is explained in the next section.

Step 2: Resource Mapping: VM to data centre resources

There are three options to map the VMs to servers and storage resources within a data centre: i) VM with fixed configuration, ii) VM with predefined categories and iii) Personalized VM according to the need of the NF. The first option was adopted in the model used in [14]; a simple mapping between the blade server resources and the resource offered to each VM is elaborated by dividing the server capacities in an equal way among VMs. However, this method is not efficient in term of resource usage; because VMs can be over- or under-dimensioned. For this, the second option with different VM classes or types was suggested. In this second option, the VMs can be classified into classes (for example: small, medium and large) such that the mapping between the VNF and VM can be more accurate.

Finally, the third option allows to create a personalized VM according the NF requirements. Though, this option assigns the required resources to each NF the most accurately, it considers only a static traffic without a margin of safety if the NF would receive more traffic than expected. This option also increases the unused resources, hence raises the question about how to allocate the unused resources cost to the different slices? Hence, for our allocation model we chose the second option of mapping VMs onto the data centre resources, being the predefined categories of VMs. Those categories are presented in Table 4-2.

Table 4-2 VM classification

	Small	Medium	Large
Core	1	4	8
RAM	8 GB	16 GB	40 GB
Temporary storage	50 GB	100 GB	1500 GB

After selecting the suitable mapping option, a server consolidation algorithm that considers the co-location of the highly communicating VMs such as [19] and [20], can be applied. It aims at placing the different VMs at the suitable server to reduce data centre network traffic and thus decrease the latency for the deployed slices. The algorithm of mapping VNFs to VMs is described in the following table:

Algorithm: NF to VM mapping algorithm

Input: Set of slices S= {S_1, $S_{2,}$ ..., S_N}; set of network functions per slice NF= {NF_1; NF_2; NF_M};each NF:{CPU(NF), RAM(NF), HDD(NF)} **set of VMs with different hardware characteristics VM={S, M, L};**

Output: allocated VMs for each slice

1: **FOREACH** slice in S **do**

2: **FOR** i=1 to M **do**

3: **FOREACH** j in **VM do**

4: vm_{ij}(CPU)=Ceil (CPU(NF_i)/ CPU_j);* nb of VM type j to satisfy the CPU required by NF_i *\

5: vm_{ij}(RAM)=Ceil (RAM(NF_i)/ RAM_j);* nb of VM type j to satisfy the RAM required by NF_i *\

6: vm_{ij}(HDD)=Ceil (HDD(NF_i)/ HDD_j); * nb of VM type j to satisfy the HDD required by NF_i *\

7: vm_j(NF_i)=MAX (vm_{ij}(CPU), vm_{ij}(RAM), vm_{ij}(HDD));

8: **ENDFOREACH**

9: vm(NF_i)=MIN (vm(NF_i)(S), vm(NF_i)(M), vm(NF_i)(L));

10: **ENDFOR**

11: **ENDFOREACH**

As inputs of the algorithm, we consider a set of slices S. Each slice S_i consists of chain of network functions NFs. In addition, each VNF has its requirements in term of CPU, RAM and HDD. Similarly, each VM from categories S, M and L has its own characteristics in terms of the same metrics i.e. CPU, RAM and HDD as well.

For each slice in the set S, we calculate the number of VMs of type “j” that satisfy the requirement of each metric (line 4 in the algorithm for CPU metric, line 5 for RAM and line 6 for the HDD). Then, the mapping of VNFs to VMS of type “j” is determined by the maximum number of VMs required based on all the metrics being CPU, RAM and HDD, and this is done for all the three VM categories (line 7).

Afterwards, we pick the category (i.e. small, medium or large VM) that gives the minimum required number of VMs (line 9).

After mapping the right amount of resources to each VNF and hence to each slice, the questions that raises is how to allocate the core network cost to the different slices based on this mapping process?

Step 3: allocate the network costs to slices

There are two options to distribute the core network costs among the slices. The first option uses the hardware requirements of each slice as a cost driver for splitting the total cost of the core network between the running slices. Here, unused resource costs will be covered by assigning them proportionally to the different services/slices based on the cost driver. The second option relies on allocating only the cost of the used resources to its slice and does not consider the cost of unused resources. Within this option, mapping VMs onto servers (step 2 of the resource allocation algorithm) is used as input to the cost allocation model, wherein only the used servers cost will be allocated to the considered slice.

For our model, we will adopt the first option. Our slices are deployed on the 5G mobile network. Hence, on the core network side, the slice is a chain of 5G NFs. For each 5G NF, we will identify a cost driver based on its key characteristic, as defined by the 3GPP spec 123.502 [21]. For example: for the User Plane Function (UPF), the cost driver is the number of packets per second, yet for the Access and Mobility Management Function (AMF), it is the number of handover requests. Secondly, we prioritize the following technical requirements (CPU, Memory, Storage, Bandwidth) according to the cost driver identified for this NF. For example, for the UPF, the cost driver of number of packets per second leads to the main technical requirements of CPU and memory.

Resulting allocated cost

Bringing the three steps together, allows to allocate the cost of the core network resources. As mentioned above, every NF is translated into hardware requirements in terms of CPU, RAM and HDD. Different virtual machines are in parallel defined in terms of the same hardware requirements, which allows mapping both. By using the hardware requirements of each NF as a cost driver, an allocated cost for the NF can be determined from the overall core network cost (as described in more detail in section 4.5.3).

4.4 Modelling the hardware requirements of the 5G NFs

In order to be able to calculate the cost of the network slices, we need to derive the cost of each 5G network function (NF) of which the network slice is composed. The cost of this latter depends on the hardware requirements of this NF (in terms of CPU, RAM and HDD). In this section, we present our model used to quantify the needed hardware requirement for both control and data plane NFs.

4.4.1 5G control plane modelling: focus on AMF and SMF

For the control plane we start by modelling the traffic handled by the AMF and Session Management Function (SMF), then derive the required processing power. The AMF and SMF form together the Mobility Management Entity (MME) in the former evolved Packet Core (EPC) of the 4G network. Therefore, in the approach presented in this section, we base ourselves on the literature that models the control traffic for virtualized MME (vMME).

Model description

Our proposed model to quantify the needed hardware requirements for the AMF and SMF is presented in Figure 4-3. It consists of two main steps. The first step is to understand the vMME modelling and use it as input to model the AMF and SMF. This step consists in identifying: i) the most frequent procedures; ii) the frequency of procedure per user and iii) the number of instructions per procedure.

The second step identifies the number of exchanged messages per procedure for both AMF and SMF and uses it as a driver to split the overall traffic of the vMME between those two functions.

It should be noted though, that when we compare the 4G and 5G service-based architecture, we conclude that there is a difference in terms of the number of messages exchanged between the MME and the rest of the 4G NFs compared to those exchanged between AMF/SMF and the rest of 5G NF. Since AMF and SMF need to communicate with much more NFs than the vMME does (as indicated with the red circles in Figure 4-4), we assume the total number of the messages for the AMF and SMF to be the same as the one of the vMME, but we include a correction factor. In [22], authors compared five different Distributed mobility management (Dmm) models of the 45/5G architecture based on the number of exchanged signalling messages for two procedures being the initial attachment and handover procedures. Results of this comparison showed that the number of signalling messages increases with an increase of the number of functional nodes in the core network. More specifically, for an SDN/NFV-based core network architecture, like the 5G core network, the number of signalling messages for the handover procedure is twice the number as for the 4G core network (12 versus 6). Based on these results, we assume a correction factor of two in our model[13].

[13] Sensitivity analysis on this factor will be performed in future work.

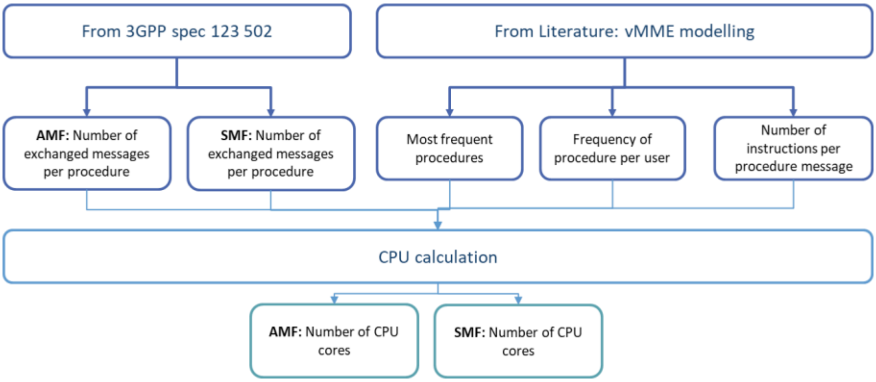


Figure 4-3 Traffic quantification model for AMF and SMF

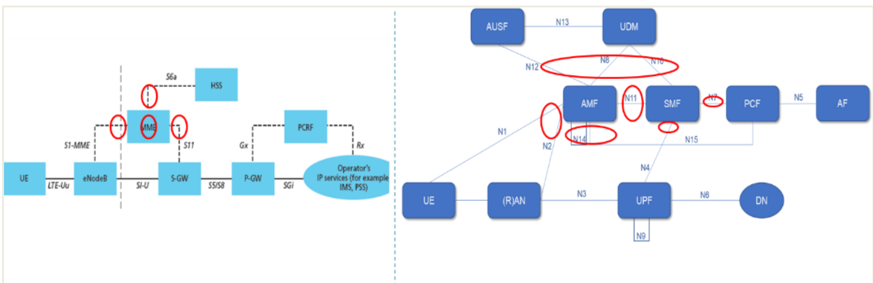


Figure 4-4 4G VS 5G service-based architecture

The AMF performs most of the functions that the MME performs in a 4G network such as terminating the RAN CP interface (N2), NAS signalling, NAS ciphering and integrity protection, Mobility Management (MM) layer NAS termination, Session Management (SM) layer NAS forwarding, authentication of UE, etc [23]. Giving the nature of the tasks performed within these functionalities we can assume that the AMF is CPU-intensive. Many researchers in literature such as [24]-[28] made this same assumption. Moreover, authors in [29] proved by calculating the CPU utilization in function of the number of users that AMF's CPU utilization is the highest among all the network functions (such as AUSF, SMF).

On the other hand, the SMF performs the session management functions that are handled by the 4G MME, SGW-C, and PGW-C: allocating IP addresses to UEs, NAS signalling for session management (SM), sending QoS and policy information to RAN via the AMF, downlink data notification, and selecting and controlling UPF for traffic routing. The UPF selection function enables Mobile Edge Computing (MEC) by selecting a UPF close to the edge of the network [23]. Based on characteristics of these functionalities and following the same reasons argued in [24]-[28], we assume that the SMF is also a CPU-intensive function and it needs a good networking interface as well.

Therefore, we will investigate for both AMF and SMF how much CPU cores are needed in order to serve a given number of users. Following the same reasons argued in [24]-[28], we assume that the most frequent procedures are the service request (SR), the service release request (SRR) and the X2-based handover (HO). Following the reasoning presented above, both AMF and SMF are CPU-intensive, which allows us to consider the CPU power as the main cost driver. Therefore, in the proposed model, only the CPU power will be calculated, from which we will derive the RAM and storage by proportionality.
As explained previously, the signalling traffic modelling for AMF and SMF requires the following inputs: the procedures' frequency, number of instructions per procedure message and the number of exchanged messages per procedure for both AMF and SMF. In the next subsections we detail each input and explain from where it was derived.

1. **Procedures' frequency**

There are two different assumptions regarding the frequency of the considered procedures. On the one hand, assuming a user inactivity of 10 s, authors in [26] calculated by the mean of a mathematical framework the signalling rates per user equipment for each signalling procedures for the vMME as presented in Table 4-3 labelled with frequency 2. On the other hand, authors in [28] derived the SR, SRR and HO frequencies from an operational network measurement as presented in Table 4-3 with frequency 1.

Table 4-3 : frequency of the considered control procedures

Procedures	**Service Request**	**Service Release**	**X2-based HO**	**Source**
Frequency 1: nb of events per user/s	0.00126	0.00126	0.00112	[28]
Frequency 2: nb of events per user/s	0.0045	0.0045	0.0012	[26]

2. **Number of instructions per procedure message**

We assume that the number of run instructions for the different control messages are the same as in [24]-[28], as presented in Table 4-4.

Table 4-4: number of instructions per user per procedure

Procedures	Service Request	Service Release	X2-based HO
Number of instructions per procedure	3,580,000	3,200,000	2,140,000

3. Number of instructions per procedure for AMF and SMF

In order to model the contribution of both AMF and SMF in the signalling traffic, we use the 3GPP specification document [21] describing the sequence diagrams for each control procedure, and from there we count the number of messages handled by AMF versus those handled by SMF for each procedure. We derive a ratio of AMF/SMF contribution in the exchanged messages for each procedure in order to be used to split the number of instructions per procedures of vMME between them. For each procedure, several assumptions were made in order to count these numbers:

Service request:

- UE Triggered Service Request is considered.
- For this procedure, the impacted SMF and UPF are all under control of the PLMN serving the UE, e.g. in the home-routed roaming case, the SMF and UPF in HPLMN are not involved.
- We assume that the Service Request was sent integrity protected.
- The UE identifies list of PDU sessions to be activated in the Service Request message
- We assume that PDU Session ID corresponds to a Local Area Data Network (LADN) and the SMF determines that the UE is within the area of availability of the LADN
- We assume that the SMF accepts the activation of uplink connection and continue using the current UPF(s);
- No N4 Session Modification is established
- No dynamic Policy and Charging control (PCC) is deployed

Service release request:

- Service release procedure corresponds to AN Release in 5G CP procedures.
- We assume that 3 PDU sessions were active at the time of initiating this procedure. This assumption is based on the fact that each network slice will be served with a separate PDU session. Thus, assuming two different slices with a PDU session for each of them and a third session for the normal calls and chat. Since we do not have a good reference that strengthens this assumption, a sensitivity analysis on this is elaborated. Results demonstrates the number of active PDU sessions does not significantly affect the required CPU cores for AMF and SMF (see Appendix for more information).
- The procedure is triggered by the User Inactivity.

X2-based handover:

- It corresponds to Xn based inter NG-RAN handover in 5G CP procedures.

The ratio of contribution of both AMF and SMF per procedure are presented in Table 4-5.

Table 4-5: number of messages for each procedure for both AMF and SMF

Procedures	**Service Request**	**Service Release**	**X2-based HO**
The ratio of AMF contribution in exchanged messages	0.533333333	0.6	0.5
The ratio of SMF contribution in exchanged messages	0.4	0.8	0.666666667

Mathematical formulation

The number of instructions per user per second for AMF and SMF for each procedure is calculated using Equation 4-4

Equation 4-4

$$\text{Nb\_Inst(NF)}_{\text{prcd /user}} = \text{Nb\_Inst}_{\text{Prcd}} \times \text{Freq(prcd)}_{\text{user /s}} \times (\text{ratio\_Nb}_{\text{Msg}}\ (\text{NF})_{\text{prcd}} \times \text{Corct}_{\text{factor}}\ )$$

The number of CPU core required to run the total instructions per procedure per network function (AMF or SMF) is calculated via applying Equation 4-5

Equation 4-5

$$\text{Nb\_CPU}_{\text{NF /prcd}} = {\text{Nb\_Inst(NF)}_{\text{prcd /user}}} \big/ {\text{CPU}_{\text{power}}}$$

The total CPU cores required for all the procedures per NF is derived using Equation 4-6.

Equation 4-6

$$\text{Nb\_CPU}_{\text{NF}} = \sum_{i=1}^{3} \text{Nb\_CPU}_{\text{NF /prcd (i)}}$$

With:

$\text{Nb\_Inst(NF)}_{\text{prcd/user}}$	number of instructions per procedures per user per second for NF	$\text{Freq(prcd)}_{\text{user/s}}$	the frequency of the procedure per user per second

Nb_Inst_{prcd}	number of instructions per procedures	$ratio\_Nb_{Msg(NF)/prcd}$	ratio of NF contribution in the exchanged messages within procedure.
$Corct_{factor}$	the correction factor of the exchanged messages between the AMF/SMF and the rest of 5G NFs.	CPU_{power}	the power of one CPU core (number of supported instructions per second)
$Nb\_CPU_{NF/prcd}$	number of CPU cores required for the procedure (prcd) within the NF.	Nb_CPU_{NF}	number of CPU cores required for NF

Results of the simulation

We simulated the proposed model to calculate the required CPU cores for both AMF and SMF and for the two frequencies while varying the number of users. We consider the same CPU power as in [26], one CPU core has the power of 2.845 x10^9 float operations per second. Results of the simulation are presented in Table 4-6.

Table 4-6: Variation of the number of CPU cores required for AMF and SMF in function of the number of users for both frequency 1 and 2

Frequency\number of users	**Network function**	**100x10^3**	**1000x10^3**	**10000x10^3**
Frequency 1	AMF	0.4	4	40
	SMF	0.5	5	50
Frequency 2	AMF	1.23	12.3	123
	SMF	1.51	15.1	151
Average of frequencies	AMF	0.82	8.2	82
	SMF	1	10	100

Results show that, for AMF, for frequency 1, half of a CPU core is needed in order to serve 100k users, yet for frequency 2, almost 1.3 CPU cores are required to satisfy the signalling traffic of the same number of users and this is. For SMF, similar results are found regarding the difference between results for the two frequencies. The difference between the two frequencies for a small number of

users is not that significant, however if we increase the number of users drastically, the gap between the results of these frequencies becomes significant.
Therefore, and since only those two frequencies are found in literature, and since we could not find a reason to prefer one of them, we decided to take the average of the two frequencies (see third row in the table above). Hence, only one core is needed to serve 100k users for both AMF and SMF and 5 cores are needed to serve 1 million users.
We can conclude from Table 4-6 that both the number of CPU cores needed for AMF and SMF scales linearly with the number of users:

$$Nb\_CPU_{AMF} = 8.2 \times 10^{-4} \times N_{users}$$
$$Nb\_CPU_{SMF} = 10 \times 10^{-4} \times N_{users}$$

With:

Nb_CPU: number CPU cores required for the NF	N_{users}: number of users

The rest of the control plane functions

The remaining control plane functions are the Policy Control Function (PCF), NF Repository function (NRF), Network Exposure function (NEF), Unified Data Management (UDM), Authentication Server Function (AUSF), Network Slice Selection Function (NSSF) and Application Function (AF) [23]. Since there is no starting point in literature to dimension those network functions, we analyse their functionalities and dimension them accordingly.

- The PCF consists of unified policy framework delivering policy rules to control plane functions and has access to subscriber information for policy decisions. Thus, it does not require significant CPU power.
- The AUSF acts as an authentication server, thus it requires a good processing power to elaborate the hashing tasks and the integrity checking.
- The UDM needs first a good memory and storage since it handles subscription management, user identification and it requires enough processing power to support the generation of Authentication and Key Agreement (AKA) credentials and access authorization.

NEF offers the exposure of capabilities and events and assures the security provisioning of information from non-3GPP application to 3GPP network, we have to think about how much external application needs to communicate with the 3GPP network but as a first order of estimate we can assume that it does not happen that often thus a modest CPU and storage resources are sufficient.

- The same holds for the NSSF, which supports the selection of the network slice instances to serve the UE and the correspondent AMF, which only happens within the service request procedure.
- For the NRF, which maintains NF profile and instances, we can assume that it requires a good processing power and memory as well.

4.4.2 5G data plane modelling

Many researchers model the data plane function (i.e. UPF) for network slicing within the LTE-A and 5G networks using virtualized middlebox network functions [30]-[36]. Middlebox network functions here refer to Network Address Translator (NAT), Firewall (FW), WAN Optimization Controller (WOC), Intrusion Detection Prevention System (IDPS), Video Optimization Controller (VOC) and Traffic Monitor (TM). Authors of these papers have identified a chain of VNFs (which are virtualized middleboxes) per service with specific requirements in terms of bandwidth, latency and hardware per VNF. The chain of network functions per service is presented in Figure 4-5. The main goals of these papers are, first, to find the best placement of the VNFs in order to optimize the network performance (e.g. minimize the latency) and, second, to investigate the number of active VNF nodes in specific scenarios in order to optimize the resource dimensioning.

Service chains considered for each service and bandwidth requirements.

Service	Chained VNFs *	Bandwidth req.
VoIP	NAT-FW-TM-FW-NAT	250 Kbps
Video Conference	NAT-FW-TM-VOC-IDPS	2 Mbps
Web Service	NAT-FW-TM-WOC-IDPS	4 Mbps

* IDPS: *Intrusion Detection Prevention*, FW: *Firewall*, NAT: *Network Address Translation*, TM: *Traffic Monitor*, VOC: *Video Optimization Controller*, WOC: *WAN Optimization Controller*

Figure 4-5: Service chains and bandwidth requirements [32]

Three main approaches are adopted in literature to dimension the UPF using virtualized middlebox functions. The first method translates the total traffic handled by VNFs to the number of concurrent operations for each VNF and deduce from that the needed hardware requirements such as in [32]. Yet, the second approach derives from the middlebox datasheets the required processing per user and given the number of users they deduce the needed CPU and others hardware requirements for each VNF [33]. However, the third approach defines a CPU-core-to-throughput relationship for each VNF and uses it to calculate the required number of CPU cores per VNF, like the method established in [35].
Since those dimensioning methods use different approaches and different assumptions, we will apply them to a specific use case (see section 4.5.2), compare their outcomes and adopt the option that gives the most realistic results that align with other findings in literature.

4.5 Application of the cost allocation model to SaT5G use case

As discussed previously, the main two goals of the developed cost allocation model are, first, to fairly allocate the network costs to the deployed slices and, second, to investigate the cost savings introduced by network slicing if any. In order to achieve these goals, two auxiliary models are developed to feed the main model with the necessary inputs. These auxiliary models are the identification of the hardware requirements of 5G NFs and the slice resources allocation. Therefore, in this section, results of these models as well as of the main model are discussed sequentially starting with the description of the specific scenario to which these models are applied. This description is elaborated in section 4.5.1. Section 4.5.2 summarizes the main outcomes of the 5G network function modelling algorithm. Section 4.5.3 presents the resource allocation algorithm's results for the considered slices. Those results feed the cost allocation model, which results are discussed in section 4.5.4. Afterwards, the cost savings on the core network side resulting from the use of NFV concept and network slicing are presented in section 4.5.5.

4.5.1 Description of the scenario

The scenario used in this paper was defined within the SaT5G project [37]. It considers providing broadband connectivity and 5G services to rural areas where currently no terrestrial network infrastructure is deployed. In order to reach those remote areas, satellite communication is used as a backhaul network that links the 5G core network to the radio access network. This latter consists of many base stations that will be installed in the considered area, which covers two villages about 5km apart connected via a rural main road. The villages are home to 350 families, with an average of 3 users per home, resulting in 1050 users. The full description of the scenario as well as the network architecture installed are detailed in [17]. The 5G services considered here are eMBB voice and eMBB video. The quality of service of these two 5G services is described in [18].

4.5.2 5G NF modelling for the specific scenario

The chain of network functions for the data plane and the required throughput per user for each service are derived based on input from literature [33] and [34], and they are presented in Table 4-7. According to the Cisco VNI Forecast Highlights Tool, Internet video traffic will be 80% of all consumer Internet traffic by 2022 compared to 73% in 2017 [38], hence we assume that the video traffic represents 80% of the overall traffic and 20% of the traffic is voice. The total generated traffic per service is calculated based on the number of served users and is presented in Table 4-7.

Table 4-7: Chain of NFs the total generated traffic per service for the considered scenario

Slice	Chain of VNFS	throughput Per user	% traffic	Total traffic (for 1050 users)
VoIP	NAT-FW-TM-FW-NAT	64kbps	20%	13.44 Mbps
Video	NAT-FW-TM-VOC-IDPS	2[14]Mbps	80%	1.68 Gbps

For the data plane network functions, we applied the three approaches described in section 4.4.2 and we compare their results to the findings in literature. Results of the model are presented in Table 4-8. The first column presents the required CPU cores for the different virtualized middleboxes based on the number of concurrent operations. The number of concurrent operations that each function needs to handle is derived by linking the original findings in [34] with the carried traffic in terms of Mbps. Results generated based on the second approach use the processing requirement per user considering the carried traffic as well. Finally, results of the third approach are driven using the CPU-core-to-throughput relationship. The comparison of our findings to those in literature [30]-[36] allows to conclude that the number of required CPU based on the throughput-to-CPU relationship (approach 3) seems to be the more realistic one.

Table 4-8: hardware requirements per VNFs per service for the three approaches

	Number of CPU cores approach 1	Number of CPU cores approach 2	Number of CPU cores from approach 3
NAT	0.35	0.966~=1	1
FW	0.7	0.945~=1	1
TM	0.35	13.965=14	2
VOC	0.7	5.67=6	1
WOC	0.35	5.67=6	1
IDPS	0.7	11.235=12	2

[14] In the original published version, there is a typo regarding the throughput per user for the video slice. It erroneously states 4 Mbps, whereas the accurate value is 2 Mbps.

Adopting the CPU-core-to-throughput approach, the hardware requirements for both eMBB voice and video slices are summarized in Table 4-9.

Table 4-9 hardware requirements of the data plane network functions for eMBB voice and video slices

VNF\Service	Voice, throughput=13.44 Mbps			Video, throughput =1.68 Gbps		
	CPU	RAM: GB	HDD: GB	CPU	RAM: GB	HDD: GB
NAT	2	2	4	1	1	2
FW	2	4	6	2	3	5
TM	1	3	2	2	6	4
VOC	-	-	-	1	1	10
WOC	-	-	-	-	-	-
IDPS	-	-	-	2	2	7

For the control plane network functions, we applied the proposed model described in section 4.4.1 assuming a user inactivity of 10s, the mobility model adopted in [26] and 1050 users. Based on the 3GPP view on network slicing, several NFs of the control plane are shared between slices. These common network functions are the AMF, NRF, NEF, UDM, AUSF and NSSF [39] and [40]. Those functions are represented with grey columns in Table 4-10. Moreover, since the video slice is carried over the satellite link, we need to deploy a satellite multicast function and a Nano-CDN node next to the edge for popular video's caching purpose.

Results of the of the model for the two slices being eMBB voice and video are recapitulated in Table 4-10.

Table 4-10: hardware requirements for the eMBB voice and video slices

	Technical requirement	UPF	AF	AMF	SMF	PCF	NRF	NEF	UDM	AUSF	NSSF	Satellite multicast function	AF (CDN)
eMBB voice	CPU	5	1	1	1	0.5	1	1	2	2	1		
	Memory	9	2	1	1	1	1	1	4	4	1		
	HDD	12	2	2	2	2	2	2	10	10	2		
eMBB video	CPU	8	2		1	0.5						1	1
	Memory	13	4		1	1						2	2
	HDD	28	4		2	2						100	100

It is clear from results presented in Table 4-10 that, similar to the RAN, the video slice requires more core network resources than the voice slice, for example, the

UPF of the video slice requires 8 CPU cores against only 5 CPU cores for the voice slice.

4.5.3 Resource allocation for video and voice slices

Given the hardware requirements of both voice and video slices presented in the previous section, we applied the proposed allocation model. The model is based on the preconfigured VMs (presented in Table 4-2 and described in section 4.3.2). The results of the required VM types for each slice as well as for the shared network functions are presented in Table 4-11.

Table 4-11: results of mapping VNFs onto VMs

VM type	**S-VM**	**M-VM**	**L-VM**
Shared Nfs	4	2	0
Video slice	4	1	1
Voice slice	3	0	1

Similar to [14] and [41], we consider a Blade server due to its capability of providing more processing power in less space which allow to simplify cabling and storage. The Blade server consists of 8CPU, 64 GB RAM, 1000 GB HDD and 4 Ethernet cards of 10 Gbps each. Moreover, for the reliability of the network, we consider a redundant VM for each NF. Hence, within the preconfigured VM option, we need 10 servers to satisfy the required hardware requirements of the two slices. Orthogonally, within the configurable VM option discussed in section 4.3.2, we need only 8 servers from the same server type.

These findings are used as an input to the cost model of the virtualized core network presented in section 4.5.5 in order to deduce the cost saving resulting from the use of network slicing.

On the other hand, we argued previously in section 4.4.1, that the SMF needs a good networking interface. Thus, to investigate if we have a limitation in term of networking interface for the Blade servers reserved based on the CPU metric, we calculate the generated traffic per user for SMF and AMF as well. We assume that the average packet size is 250 Bytes for the control messages generated by SMF and AMF (similar to the assumption for the vMME in [26]). The calculation results in 18.418 and 14.626 bits per second per user for the SMF and AMF respectively. Hence, for our scenario with 1050 users, the AMF requires 15.4 kbps to be reserved on the network interface, while the SMF requires 38.8 kbps (for both video and voice slices). Therefore, 1Gbps is allocated on the network interface card (NIC) for each VNF, it is more than enough. Hence, we do not dispose of a limitation on the networking interface resources.

4.5.4 Cost allocation for video and voice slices

The cost of the network infrastructure installed to serve the remote two villages described previously with eMBB voice and video services are taken from [17]. These cost figures and the required inputs for the cost allocation model are detailed in Table 4-12.

Table 4-12: Cost inputs for the cost allocation model

Item	Value	Reference
RAN cost (in euro)	50,067	[17]
Cost of Satellite backhaul (in euro)	160,149	[17]
Cost of 5G core network (in euro)	76,230	[17]
latency eMBB video in ms	300	latency: Packet Delay Budget from "3GPP standard QoS class identifiers" for Non-Conversational Video (Buffered Streaming) [18]
latency eMBB voice in ms	100	latency: Packet Delay Budget from "3GPP standard QoS class identifiers" for Conversational Voice [18]
Throughput eMBB video (in Mbps)	4	Assumption
Throughput eMBB voice (in Mbps)	0.64	Assumption
Cost sharing coefficient	0.7	Assumption
Qco Video	0.718	Calculated based on Equation 4-3
Qco Voice	0.282	Calculated based on Equation 4-3
CPU cost driver: video	0.59	Calculated using the model
CPU cost driver: voice	0.41	Calculated using the model

As described in section 4.3, the CPU metric will be used as a cost driver to split the total network costs between the running services. This cost driver for to the two slices being eMBB video and eMBB voice is calculated using the outcomes of the hardware requirements identification model (discussed in section 4.5.2).
The calculation of the CPU cost driver for both eMBB video and voice is presented in Table 4-12. It shows that 59% of the core network cost will be allocated to the eMBB video slice versus 41% for the voice slice. For the RAN side, the throughput and the quality of service (QoS) coefficient used to allocate the RAN cost to the two slices are also presented in Table 4-12. The QoS coefficient of the eMBB video is higher than the one of the eMBB voice (0.7 versus 0.3) since the capacity required by the video service is much more

important than the voice' capacity and since also the cost sharing coefficient (0.7) is in favour of the throughput at the expense of the latency.
Considering input values presented in Table 4-12, we run our model to allocate the cost of the network. Results of the model are represented in Figure 4-6.

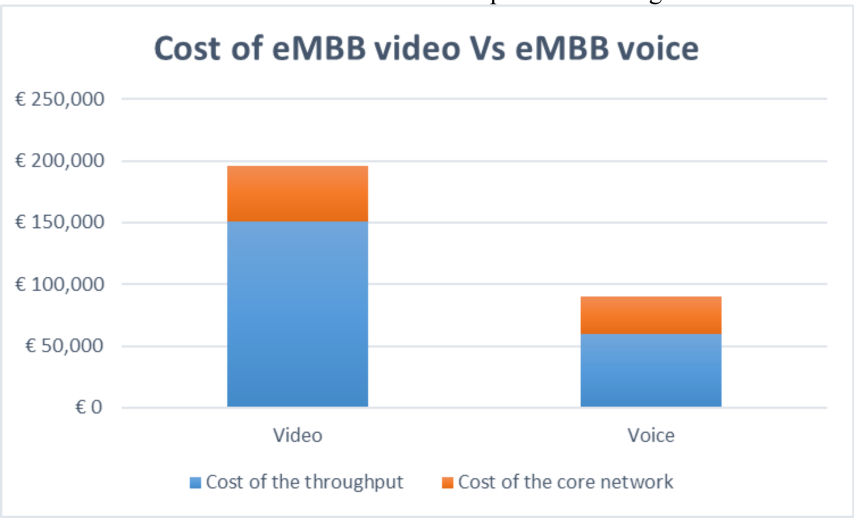


Figure 4-6 allocation of the network cost to eMBB video and voice slices

Several interpretations can be extracted from Figure 4-6. First, the eMBB video slice bears the significant amount of the network costs (68% compared to only 32% for the voice slice). This is reasonable, because it requires more throughput on the RAN and backhaul and more computing resources on the core network part as well. However, the cost of the throughput of the voice slice is still important, which can be justified with the fact that not only the required throughput is counted in the cost allocation algorithm but also the latency (Equation 4-3), for which the voice service presents a more stringent requirement.

4.5.5 Cost savings resulting from the use of NFV in the core network

NFV and network slicing paradigms promise to reduce overall network costs and allow for more cost-efficient network deployment. In this section, we investigate these promises. Given the dimensioning model for network slicing presented in section 4.5.3, we know the exact amount of resources that each slice (from the two slices eMBB voice and video) requires. Furthermore, we count redundant resources for reliability purposes. Following these inputs, the 5G core network can be dimensioned accordingly. Afterwards, calculating the cost of the core network resources and comparing it to the cost of the traditional core network (without NFV deployment) allows to calculate the cost saving of the use of virtualization

technology. Figure 4-7 presents the flow between the different sub-models that feeds the cost saving calculation.
In the appendix, we detail the cost inputs used to derive the cost of the virtualized 5G core network. We have two figures for the traditional core network cost, the first one is quoted in terms of the number of users, the second in terms of traffic, i.e. Mbps. Therefore, we calculate the cost of the core network and we quote it also per user and per Mbps.

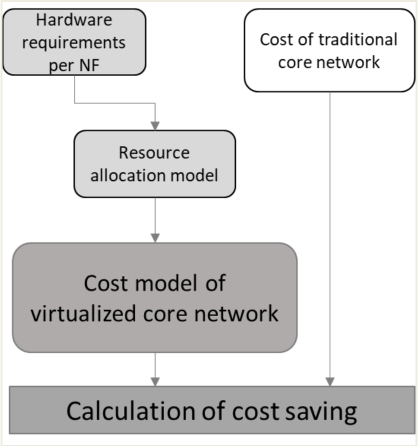


Figure 4-7 Calculation of the virtualization cost saving for the core network

The cost saving due to the use of virtualization on the core network is about 45% if the cost is quoted in Mbps compared to the cost used in [42], yet, it is only 12% if the cost is calculated per user compared to the cost found in [17]. The result of the comparison between the core network cost with NFV deployment versus without, is summarized in Table 4-13. Despite this difference, the savings can be identified as significant, and prove that the use of NFV and network slicing reduces the cost of the core network deployment.

Table 4-13 Cost reduction resulting from the use of NFV in the core network

	Cost virtualized 5G core	Cost of the traditional core network	Cost reduction
Cost 5G core per Mbps	21€	38 euro per Mbps (5 to 6 euro per Mbps for the hardware and 30 to 35 euro per Mbps for software) [42]	44.73%
Cost 5G core per user	53€	60 euro per user [17]	11.67%

4.6 Conclusions and future work

This paper proposed a novel cost allocation model for sliced networks. The model aims, first, to allocate the network cost to the deployed slices and, second, to investigate the cost reduction promises of using the NFV and network slicing technologies. Furthermore, to feed the proposed model with the required inputs and since we did not find these inputs in literature, we developed two other models. The first one is a resource allocation model that assigns the right amount of network resources to each slice aiming at fulfilling its KPI requirements, for which we needed to derive the hardware requirements of each VNF that is included in the chain of VNFs forming the slice. Therefore, we designed the second model that identified the hardware requirements of the 5G network functions, with a focus on those of the control plane. This latter is also the first of its kind in literature because, first, the control plane traffic was before often considered as a percentage of the data plane traffic and, second, because it is the first model for 5G networks. We applied the proposed model to a specific scenario considering the offer of two slices (eMBB video and voice) to a remote area. Results of the model shows that the eMBB video bears the significant part of the network cost. On the other hand, results demonstrate a cost saving of 12 to 45% (depending on the approach adopted) of the core network costs resulting from the use of NFV and network slicing. As next steps, we plan to validate our models and especially the hardware requirements of 5G NFs with, first, more scenarios, such as dense-urban scenarios, wherein higher traffic is required per slice and, second, with the testbed results within the SaT5G project. It would be a good exercise also to apply the model on two completely different types of slices (e.g. IoT versus eMBB) in order to investigate the impact of more diverse KPI requirements on the cost. Up till now, this was unfortunately not possible because of a lack of traffic model and scenario details for an IoT scenario. Finally, the traffic considered here is static, future work should extend the model to the case where the dynamicity of the traffic is considered. The same applies for the resource allocation model, as currently the assignment of the network resources is done statically based on the maximum throughput and traffic carried by the different network functions.

ACKNOWLEDGMENT

The authors would like to acknowledge all partners in the SaT5G project, funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 761413.

Appendix

1. Inputs of the cost model of the NFV-based core network:

Table 4-14: Cost inputs and calculation for the cost saving model of network virtualization

Item	Value	Reference
Blade server price	1,300	Average price in the internet
Number of servers	8	We only need 4 but 8 in total for redundancy purposes
Capex Data Centre	12,753	
Air conditioning	500	[17]
Installation	1,912.95	
Maintenance*5 year	6,376.5	
Power consumption	240	Power consumption per server per year in kWh: from the datasheet
Electricity price kwh	0.014	[17]
cost power consumption	134.4	
VNF license cost (Dollar/vCPU): 100	25,200	[43]
Nokia router 7750	1,853	[42]: 20383 euro for 11 small sites (like our scenario site)
Total Capex	14,666	
Total Opex	31,711	
Overhead costs	9,275.37	
TCO	55,652.22	
Total traffic (Mbps)	2,700	

2. Variation of the number of PDU sessions:

We assumed before in modelling the SRR procedure that 3 PDU sessions were active when this procedure is launched, but we do not have a strong assumption on this value. Hence in Figure 4-8, we vary the number of PDU sessions to visualize its effect on the required CPU cores for both AMF and SMF using the average frequency without the correction factor. From these results we can deduce, up to until 1 million users, that the number of active PDU sessions does not significantly affect the required CPU cores for AMF and SMF.

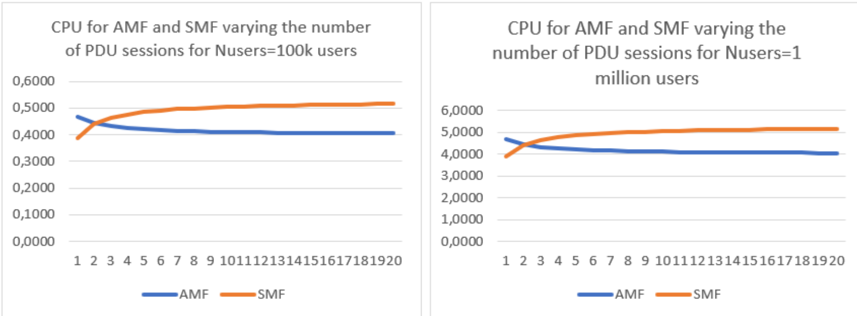


Figure 4-8 Number of CPU cores for AMF and SMF in function of number of users for different number of PDU sessions

References

[1] Wood, T., Ramakrishnan, K. K., Hwang, J., Liu, G., & Zhang, W. (2015). Toward a software-based network: integrating software defined networking and network function virtualization. IEEE Network, 29(3), 36-41.

[2] ETSI NFV, “Network functions virtualization: An introduction, benefits, enablers, challenges & call for action,” Darmstadt, Germany, SDN & OpenFlow World Congr., White Paper, Oct. 2012.

[3] Sun, S., Kadoch, M., Gong, L., & Rong, B. (2015). Integrating network function virtualization with SDR and SDN for 4G/5G networks. IEEE Network, 29(3), 54-59.

[4] Network Slicing and 3GPP Service and Systems Aspects (SA) Standard. https://sdn.ieee.org/newsletter/december-2017/network-slicing-and-3gpp-service-and-systems-aspects-sa-standard . Accessed 15 May 2019.

[5] Afolabi, I., Taleb, T., Samdanis, K., Ksentini, A., & Flinck, H. (2018). Network slicing and softwarization: A survey on principles, enabling technologies, and solutions. IEEE Communications Surveys & Tutorials, 20(3), 2429-2453.

[6] Herrera, J. G., & Botero, J. F. (2016). Resource allocation in NFV: A comprehensive survey. IEEE Transactions on Network and Service Management, 13(3), 518-532.

[7] Zhang, H., Liu, N., Chu, X., Long, K., Aghvami, A. H., & Leung, V. C. (2017). Network slicing based 5G and future mobile networks: mobility, resource management, and challenges. IEEE Communications Magazine, 55(8), 138-145.

[8] Caballero, P., Banchs, A., De Veciana, G., Costa-Pérez, X., & Azcorra, A. (2018). Network slicing for guaranteed rate services: Admission control and resource allocation games. IEEE Transactions on Wireless Communications, 17(10), 6419-6432.

[9] Checko, A., Christiansen, H. L., Yan, Y., Scolari, L., Kardaras, G., Berger, M. S., & Dittmann, L. (2014). Cloud RAN for mobile networks—A technology overview. IEEE Communications surveys & tutorials, 17(1), 405-426.

[10] Sciancalepore, V., Samdanis, K., Costa-Perez, X., Bega, D., Gramaglia, M., & Banchs, A. (2017, May). Mobile traffic forecasting for maximizing 5G network slicing resource utilization. In IEEE INFOCOM 2017-IEEE Conference on Computer Communications (pp. 1-9). IEEE.

[11] Naudts, B., Kind, M., Westphal, F. J., Verbrugge, S., Colle, D., & Pickavet, M. (2012). Techno-economic analysis of software defined networking as architecture for the virtualiazation of a mobile network. In European Workshop on Software Defined Networking (EWSDN-2012) (pp. 1-6).

[12] Bouras, C., Ntarzanos, P., & Papazois, A. (2016, October). Cost modeling for SDN/NFV based mobile 5G networks. In 2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT) (pp. 56-61). IEEE.

[13] Zhang, N., & Hämmäinen, H. (2015, March). Cost efficiency of SDN in LTE-based mobile networks: Case Finland. In 2015 International conference and workshops on networked systems (NetSys) (pp. 1-5). IEEE.

[14] Knoll, T. M. (2015, November). Life-cycle cost modelling for NFV/SDN based mobile networks. In 2015 Conference of Telecommunication, Media and Internet Techno-Economics (CTTE) (pp. 1-8). IEEE.

[15] Han, B., Tayade, S., & Schotten, H. D. (2017, July). Modeling profit of sliced 5G networks for advanced network resource management and slice implementation. In 2017 IEEE Symposium on Computers and Communications (ISCC) (pp. 576-581). IEEE.

[16] Zhou, X., Li, R., Chen, T., & Zhang, H. (2016). Network slicing as a service: enabling enterprises' own software-defined cellular networks. IEEE Communications Magazine, 54(7), 146-153.

[17] Chiha, A., Van der Wee, M., Colle, D., & Verbrugge, S. (2019). Techno-economic viability of integrating satellite communication in 4G networks to bridge the broadband digital divide. Telecommunications Policy, 101874. https://doi.org/10.1016/j.telpol.2019.101874.

[18] Policy and charging control architecture, 3GPP specification 3.203. Available on https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=810. Accessed 10 June 2018.

[19] Wu, G., Tang, M., Tian, Y. C., & Li, W. (2012, November). Energy-efficient virtual machine placement in data centers by genetic algorithm. In International conference on neural information processing (pp. 315-323). Springer, Berlin, Heidelberg.

[20] Ahmad, R. W., Gani, A., Hamid, S. H. A., Shiraz, M., Yousafzai, A., & Xia, F. (2015). A survey on virtual machine migration and server consolidation frameworks for cloud data centers. Journal of network and computer applications, 52, 11-25.

[21] 5G; Procedures for the 5G System (3GPP TS 23.502 version 15.2.0 Release 15). Available on https://www.etsi.org/deliver/etsi_ts/123500_123599/123502/15.02.00_60/ts_123502v150200p.pdf. Accessed 10 January 2019.

[22] Sun, K., & Kim, Y. (2017, July). Gap analysis for adapting the distributed mobility management model in 4G/5G mobile networks. In

2017 IEEE Conference on Network Softwarization (NetSoft) (pp. 1-5). IEEE.

[23] Grandmetric, "5G Core Network Functions," [Online]. Available: https://www.grandmetric.com/2018/03/02/5g-core-network-functions. [Accessed 10 January 2019].

[24] Prados, J., Laghrissi, A., Bagaa, M., Taleb, T., & Lopez-Soler, J. M. (2019). A Complete LTE Mathematical Framework for the Network Slice Planning of the EPC. IEEE Transactions on Mobile Computing.

[25] Prados-Garzon, J., Ramos-Munoz, J. J., Ameigeiras, P., Andres-Maldonado, P., & Lopez-Soler, J. M. (2016, March). Latency evaluation of a virtualized MME. In 2016 Wireless Days (WD) (pp. 1-3). IEEE.

[26] Prados-Garzon, J., Ameigeiras, P., Ramos-Munoz, J. J., Andres-Maldonado, P., & Lopez-Soler, J. M. (2017, May). Analytical modeling for virtualized network functions. In 2017 IEEE International Conference on Communications Workshops (ICC Workshops) (pp. 979-985). IEEE.

[27] Prados-Garzon, J., Ramos-Munoz, J. J., Ameigeiras, P., Andres-Maldonado, P., & Lopez-Soler, J. M. (2017). Modeling and dimensioning of a virtualized MME for 5G mobile networks. IEEE Transactions on Vehicular Technology, 66(5), 4383-4395.

[28] Hirschman, B., Mehta, P., Ramia, K. B., Rajan, A. S., Dylag, E., Singh, A., & McDonald, M. (2015). High-performance evolved packet core signaling and bearer processing on general-purpose processors. IEEE Network, 29(3), 6-14.

[29] Buyakar, T. V. K., Agarwal, H., & Tamma, B. R. (2019, June). Prototyping and Load Balancing the Service Based Architecture of 5G Core Using NFV. In 2019 IEEE Conference on Network Softwarization (NetSoft) (pp. 228-232). IEEE.

[30] Savi, M., Hmaity, A., Verticale, G., Höst, S., & Tornatore, M. (2016, July). To distribute or not to distribute? Impact of latency on virtual network function distribution at the edge of FMC networks. In 2016 18th International Conference on Transparent Optical Networks (ICTON) (pp. 1-4). IEEE.

[31] L. Ruiz, R. J. Durán, I. d. Miguel, P. S. Khodashenas, J.-J. Pedreno-Manresa, N. Merayo, J. C. Aguado, P. Pavon-Marino, S. Siddiqui, J. Mata, P. Fernández, R. M. Lorenzo and E. J. Abril., "Genetic Algorithm for Effective Service Mapping in the Optical Backhaul of 5G Networks," in 20th International Conference on Transparent Optical Networks (ICTON), 2018.

[32] Pedreno-Manresa, J. J., Khodashenas, P. S., Siddiqui, M. S., & Pavon-Marino, P. (2017, July). Dynamic QoS/QoE assurance in realistic NFV-enabled 5G access networks. In 2017 19th International Conference on Transparent Optical Networks (ICTON) (pp. 1-4). IEEE.

[33] Savi, M., Tornatore, M., & Verticale, G. (2019). Impact of processing-resource sharing on the placement of chained virtual network functions. IEEE Transactions on Cloud Computing.

[34] Savi, M., Tornatore, M., & Verticale, G. (2015, November). Impact of processing costs on service chain placement in network functions virtualization. In 2015 IEEE Conference on Network Function Virtualization and Software Defined Network (NFV-SDN) (pp. 191-197). IEEE.

[35] Gupta, A., Habib, M. F., Mandal, U., Chowdhury, P., Tornatore, M., & Mukherjee, B. (2018). On service-chaining strategies using virtual network functions in operator networks. Computer Networks, 133, 1-16.

[36] Gupta, A., Jaumard, B., Tornatore, M., & Mukherjee, B. (2018). A scalable approach for service Chain mapping with multiple SC instances in a wide-area network. IEEE journal on Selected Areas in Communications, 36(3), 529-541.

[37] Liolis, K., Geurtz, A., Sperber, R., Schulz, D., Watts, S., Poziopoulou, G., ... & Fitch, M. (2019). Use cases and scenarios of 5G integrated satellite-terrestrial networks for enhanced mobile broadband: The SaT5G approach. International Journal of Satellite Communications and Networking, 37(2), 91-112.

[38] CISCO VNI, “VNI Forecast Highlights Tool,” 8 3 2019. [Online]. Available: https://www.cisco.com/c/m/en_us/solutions/service-provider/vni-forecast-highlights.html#.

[39] IETF, “Network Slicing Use Cases: Network Customization and Differentiated Services.,” [Online]. Available: https://tools.ietf.org/id/draft-netslices-usecases-01.html..

[40] IETF, “Network Slicing–3GPP Use Case,” IETF, [Online]. Available: https://tools.ietf.org/id/draft-defoy-netslices-3gpp-network-slicing-02.html. [Accessed 5 January 2019].

[41] Economou, D., Rivoire, S., Kozyrakis, C., & Ranganathan, P. (2006). Full-system power analysis and modeling for server environments. International Symposium on Computer Architecture-IEEE.

[42] SAT5G, “Internal document,” 2017.

[43] Dieye, M., Ahvar, S., Sahoo, J., Ahvar, E., Glitho, R., Elbiaze, H., & Crespi, N. (2018). CPVNF: Cost-efficient proactive VNF placement and chaining for value-added services in content delivery networks. IEEE Transactions on Network and Service Management, 15(2), 774-786.

5

Cooperative, Connected and Automated Mobility (CCAM) services provisioning in cross-border settings: techno-economic analysis in the light of technical challenges

Asma Chiha, Frederic Vannieuwenborg, Benoît Denis[15], Sofie Verbrugge, and Didier Colle

Published in the Transport Policy Journal

Chapter 5 shifts the focus from eMBB services, studied in the previous chapters, to URLLC services, specifically cooperative, connected, and automated mobility (CCAM) services. To provide reliable and low-latency network requirements for CCAM services, 5G technology is identified as the optimal provider. However, the provision of CCAM services involves multiple stakeholders from different fields, such as MNOS, OEMs, and regulators. This chapter studies potential business and collaboration models that benefit from emerging 5G technologies such as network slicing and MEC. To assess the viability of the identified business models, a cost-benefit analysis is conducted. Furthermore, a cost allocation model is developed to share 5G deployment costs among the various business actors involved.

***

[15] CEA-Leti, Université Grenoble Alpes,Grenoble, France

Abstract—Cooperative, connected and automated mobility (CCAM) has a great potential to change drastically the mobility landscape and improve safety on the roads by assisting drivers to take the best decisions in given circumstances or by supporting autonomous driving features aboard the car, among other expected benefits. Yet, providing CCAM services poses not only technical but also business challenges, especially in cross-border environments. In this paper, we present a general techno-economic methodology intended first, to identify the cross-border ecosystem and the potential business models to enable CCAM services provision and second, to derive the Total Cost of Ownership (TCO) of deploying the required Vehicle to Infrastructure (V2I) and Vehicle to Network (V2N) infrastructure. As an illustration, we then apply this methodology to a canonical Cooperative Lane Merging (CLM) scenario in a representative border environment (i.e., the Brenner pass between Italy and Austria). Finally, a cost allocation model is developed to link these deployment costs to the key stakeholders within the identified business models. Results aim to give insights into the different deployment strategies in typical cross-border environments and settings.

Keywords: CCAM services, 5G, SDA, Network slicing, MEC, Business model, Cost model, Techno-economics, Cross-border environments.

5.1 Introduction

In general, borders can have a variety of effects on the economy, with negative effects such as barriers to trade and investment being the most recognized. Borders can create barriers to trade, such as tariffs, which can increase the cost of goods and reduce trade. They can also create barriers to investment, such as regulations and political instability, which can discourage investment and reduce economic activity which can lead to fragmentation of production and supply chains, reducing efficiency and productivity [1] and [2].

Therefore, the European Union (EU) has adopted in 1993 the concept of the European Single Market, also known as the Internal Market. The Internal Market is a single market created by the EU to allow the free movement of goods, services, people, and capital within the EU. It is one of the EU's main achievements and is considered as a key driver of economic growth and job creation [3]. To reach the objectives of the single European market, the European mobility corridors have received a lot of attention from the European Commission (EC) due to their crucial role in connecting regions and cities, and in supporting economic growth and integration. Additionally, mobility corridors can foster the development of new technologies, such as autonomous vehicles and smart logistics, which can further improve the efficiency and sustainability of the transportation network [4].

On the other hand, the 5G paradigm consists of a set of promising technologies such as network virtualization, network slicing, edge computing etc. These technologies help not only Information Technology (IT) companies but also other sectors to expand their markets, create new business opportunities and reduce the cost of launching new services, among other advantages [5]. These technologies can typically fulfil the requirements of various demanding classes of services, such as high-capacity communications (i.e. enhanced Mobile BroadBand (eMBB)), low latency and high reliability communications (ultra Reliable Low Latency Communications (uRLLC)) or the ability to support huge numbers of connected devices, typically up to 1 million connected device/km^2 (i.e. massive Machine Type Communication (mMTC)) [6]. Enabling the related use cases is expected to generate new revenues in different economic sectors. According to several recent surveys and research [7] and [8], most of the generated 5G revenues are expected in smart cities, connected factories and automotive domains. For the last past years, even more particularly, the latter domain has received enormous attention from industry players such as car manufacturers and Original Equipment Manufacturers (OEMs), but also from research communities, regulators, and other policy makers.

The European Commission (EC), for example, has adopted several initiatives (C-ITS, CCAM [9] and GEAR 2030 [10]), agendas (5G strategic Deployment agenda for CCAM [11]) and programs (CEF [12] and CEF2 digital [13]) in order to accelerate the adoption of Cooperative Connected and Automated Mobility (CCAM) in Europe by 2025 to 2030. This specific interest in CCAM deployment from the EC is raised by major expected benefits of 5G-enabled CCAM along the mobility corridors, such as a significant improvement of road safety, traffic efficiency and comfort of driving, which can also boost further the single digital market agenda of Europe besides the anticipated considerable decrease in CO2 emission on the roads (e.g. by adopting CCAM green use cases) [9].

However, these promising initiatives and agendas to cover all Europe' corridors with 5G networks for the support of CCAM services are facing a major challenge, being the handling of cross-border environments. The latter indeed pose not only technical issues related to 5G connectivity, but also unprecedented business-related questions.

In fact, among the numerous possibilities offered by 5G, CCAM services rely mainly on ultra reliable and low latency communication capabilities. Yet, crossing countries' borders induces a disruption of the connectivity and hence, of their supported services as well, which is mostly due to roaming. Besides, possible gaps in mobile networks coverage in cross-border areas also represent an issue to be resolved in this context. At the business level, CCAM and V2X services in general have a quite complex ecosystem involving multiple stakeholders from different backgrounds, such as service providers, Mobile Network Operators (MNOs), road operators, car manufacturers, regulators etc., whereas innovative and sustainable business models that could take into account technical and legislative aspects are obviously still missing [14]. Moreover, this CCAM ecosystem becomes even more complex in cross-border contexts. Indeed, (i) some of the involved stakeholders are needed on the two sides of the border, (ii) different legislations with respect to

5G networks deployment may apply (e.g., different auction timelines and network sharing policies) and (ii) road regulations would also need to be harmonized across the involved countries (i.e., Europe in our case).

Therefore, to ensure seamless connectivity and services across the borders, all these stakeholders need to work together, what poses several questions such as "who will pay for the required network infrastructure?", "how will the revenues be shared among these stakeholders?", etc. It is then crucial to define new sustainable business models that can also integrate the technical dimension in such cross-border environments. To solve out some of the cross-border challenges mentioned above (in terms of both technical solutions and business models), the EC has funded several research projects (i.e., 5G-CARMEN, 5G-Mobix and 5G-Croco) under the umbrella of H2020 program [15].

In this paper, we thus present a techno-economic framework specifically developed in the scope of 5G-CARMEN to answer the following research questions:

- Who are the involved actors in the CCAM ecosystem?
- What are the potential business models?
- What is the Total Cost of Ownership (TCO) of the different/possible deployment scenarios depending on the used technology (5G, C-V2X)?
- What is the impact of new technologies (i.e., MEC and Network Slicing) on the business case?
- How to allocate the deployment cost to the involved stakeholders?

Accordingly, we first identify potential business models for the provision of cross-border CCAM services in the light of their related technical challenges. Then we derive the Total Cost of Ownership (TCO) of deploying the V2I (Vehicle to Infrastructure) and/or V2N (Vehicle to Network) communication infrastructure required to fill network coverage gaps close to the border, given a canonical use case example. Finally, we allocate the network deployment cost to the involved key stakeholders, by using a cost allocation model. Note that in this paper, we only focus on research questions related to connectivity and multi-stakeholder issues, thus legislative and regulatory aspects will not be studied.

The remaining of the paper is structured as follows. In Section 5.2, CCAM business models proposed in the literature will be analysed in the light of the cross-border challenges. Then, Section 5.3 introduces the suggested techno-economic framework to identify new business models and assess them qualitatively and quantitatively and presents the selected use case. Section 5.4 discusses the results of our analysis in term of identified business models and cost model results for both V2I and V2N deployment scenarios, as well as the cost allocation to key stakeholders.

5.2 Related work

To support the deployment of effective and secure CCAM services, it is important to focus not only on the technical requirements and challenges, but also on how these services will disrupt the current way of working, on their impact on the market, as well as on the different market stakeholders. Hence, there is an obvious need for establishing new business models that could better capture and predict the great potential of these foreseen CCAM services.

It has been noticed in the literature that only few works have really investigated CCAM services (and beyond, V2X-based services in general) from a pure business model perspective [14]. In [14] , the authors classified relevant papers and contributions according to various aspects, such as communication, security, legislation and business. Among eight references, only one work partially focuses on the business aspects of V2X services [16]. Therefore, the authors in [14] tried to tackle these aspects by proposing a new business model that accounts for some advantages of 5G technologies, namely network virtualization and technology neutral solutions, by inventing the new role of "Joint Venture" that assumes a tight coordination between all the involved stakeholders.

Due to the limited number of publications focusing on identifying V2X business models, we shortly review best practices from other 5G V2X-related projects hereafter. The 5G-NORMA project [17] for instance identifies the business model for vehicular connectivity with key service providers, MNOs, Car manufacturers and consumers. They suggest that a network slice could be provided to each vehicle manufacturer on behalf of its customers; however, vehicles from different manufacturers may also want to communicate with each other. Service providers will hence need to determine the most efficient way of addressing this issue. Communication infrastructure and V2X systems may necessitate to be shared between government agencies and operators as well. Therefore, we conclude that a lot of issues are still open under this model. Concerning the revenue models, they proposed that network operators may agree on a charging model with automotive companies. Typically, a charge per active embedded device would cover the connectivity needed for a core set of services, particularly autonomous driving and safety measures. This could thus be used as a rental charge to recoup the extra cost of the coverage network (including road-side sensors and units) necessary for these services to work. But roadside units are not deployed or operated by MNOs, hence the proposed revenue model lacks the description of the revenue sharing between MNOs and road operators/authorities.

In the 5GCAR project [18], the objective of the analysis of the V2X business models is to identify how 5G could enable new business models in the V2X ecosystem. The study considers three different categories of V2X services, namely 1) existing services, i.e. services already available on the market for the connected car (emergency call, remote diagnostic, etc.), 2) autonomous driving features and 3) convenience services intended to enhance the user experience. Each service is described with several business model parties involved in the value chain and necessary for service provision. However, due to the highly dynamic nature of both telecom and automotive industries, the study does not describe the whole

business model, rather it highlights what are the important elements for creating innovative business models for each service.
In the METIS II project [19], the business model of reference is the wholesale-based model on a Business to Business to Consumer (B2B2C) basis. Accordingly, the user pays the manufacturer for the service rendered and the car manufacturer shares revenues with the service provider. In an alternative model, the end user pays the MNO for the connection and buys the car from the automaker.
In addition to the previous EC-funded projects, the 5GPPP Automotive working group has released several white papers studying the 5G V2X deployment from both technical and business feasibility standpoints, like [20] and [21]. For instance, the latter white paper [21] studies the business feasibility of such a deployment, and discusses five different network infrastructure sharing models, being:

1. Passive infrastructure sharing: only passive infrastructure elements such as base station location, masts, power generators, and air conditioning equipment can be shared between operators. These elements could be deployed by a joint MNOs group, road Infrastructure Operators or third-party infrastructure providers. Savings are estimated by [22] to be 16% to 35% CAPEX and 16% to 35% OPEX.
2. Active infrastructure sharing excluding spectrum: in this model, active elements of the mobile network can be shared like antennas, BBU, RRU etc. However, each operator uses her/his own spectrum. Savings are assumed to be 33% to 35% CAPEX and 25% to 33% OPEX according to [22].
3. Active infrastructure sharing including spectrum sharing, in addition to active equipment sharing as discussed in the previous model; radio frequencies can be shared between the different operators as well. One operator can deploy the Radio Access Network (RAN) and open it up for the use of the other operators, including the spectrum. This RAN is connected to the different operators' core networks. This scenario is described in more detail in [23] "Scenario1: Multiple core networks sharing common radio access network in R99". Savings are estimated by BEREC to be 33% to 45% CAPEX and 30% to 33% OPEX [22].
4. Core network sharing: 5G core network elements such as computing, and storage resources can be shared on top of active and passive infrastructure. This model is also described in [23] as a "Multiple radio access networks sharing common core network". Sharing only core network elements shows a marginal cost saving [22], yet if also RAN elements are shared the cost saving will be similar to that of the previous sharing model.
5. National Roaming: It can be also seen as an active network sharing wherein one single operator deploys her/his network in the service area and serves the other operators' customers based on an agreement made between them. This

model is referred to in [23] as a "Geographically split network using national roaming between operators".

Although relevant business models have been discussed in the literature, in 5G V2X projects and in 5GAA white papers, they mainly focus on the deployment of CCAM services within a unique country and discuss ways to enhance the 5G network deployment in order to support such CCAM services. In a cross-border scenario, the service area is composed of several sub-areas belonging to different countries and operators. This makes some of the business and sharing models mentioned previously infeasible due to spectrum license issues, country-specific regulations and laws, or they simply cannot be adopted because the underlying assumptions would not ensure the continuity of CCAM services. For example, the active sharing model excluding the spectrum cannot be adopted in a cross-border situation as an operator cannot use the active cellular antennas of another operator in a different country without owning the spectrum license. Even the active sharing model including the spectrum is hardly applicable in this scenario for CCAM services without assuming additional tighter collaboration between the involved mobile network operators. This is because if different RANs belonging to different operators are used, there would be an obvious service discontinuity, hence requiring a roaming procedure when crossing the border.

5.3 Techno-economic analysis methodology and use case

Techno-economic analysis is a method that combines technical and economic factors to evaluate the feasibility of a project. By considering both technical and economic factors, the techno-economic analysis provides a comprehensive evaluation of the feasibility and potential profitability of providing CCAM services using 5G networks.

5.3.1 Techno-economic methodology

With a clear focus on business models and their viability for providing CCAM in a cross-border environment, we developed a generic methodology that consists of four main consecutive and iterative steps (illustrated in Figure 5-1):

1. Value network and business models identification: this can be achieved by firstly defining the different business roles and stakeholders involved, and secondly the interactions between them. Different ways of interactions result in different value network configurations, and accordingly in different individual business models for the stakeholders.
2. Business case viability study: A Total Cost of Ownership (TCO) model as well as revenue assumptions are used to calculate an Average Revenue Per User (ARPU), which can later be benchmarked to similar CCAM services prices.

3. Impact of new technologies on business case: this step consists in investigating the impact of the innovative technology on the business case. Here, specifically the impact of network slicing as well as MEC placement will be studied.
4. Sensitivity analysis is elaborated to assess the degree of uncertainty that links the model outputs to the inputs.

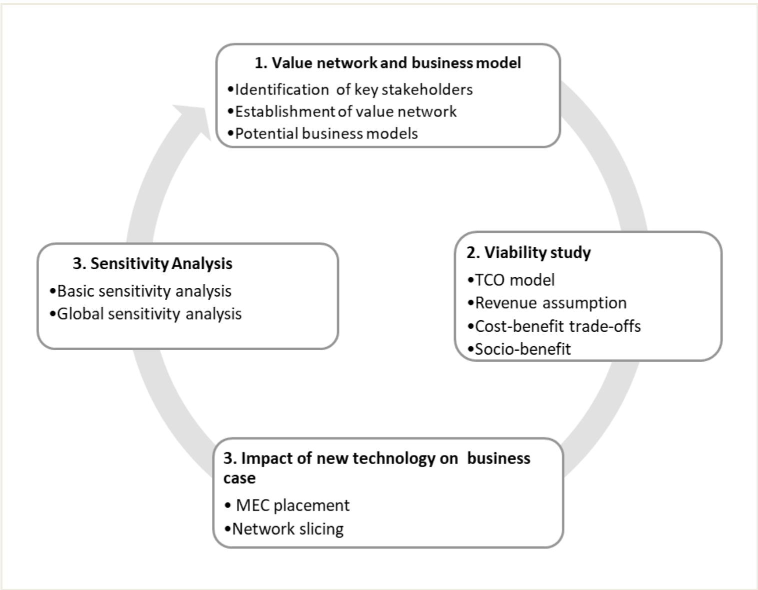


Figure 5-1 Techno-economic analysis methodology for CCAM in cross-border scenario

In this paper we mainly focus on 1) the business model identification (first step) and 2) the TCO of the different network deployment scenarios (part of the Viability study step) and 3) the impact of MEC placement on the cost of the different deployment scenarios which will be integrated in the cost modelling of the different V2N deployment scenarios. In addition, a cost allocation model is developed to allocate the cost of the different network deployment scenarios to the key stakeholders identified within the different business models.

5.3.2 Use case and cross-border corridor under study

Four families of use cases were initially considered in 5G-CARMEN namely cooperative manoeuvring (CM), situation awareness, video streaming, and green driving [24]. For our study, we focus on the cooperative manoeuvring, which has been selected more recently as a flagship use case in alignment with the choice made in the simulation and field-trials [25]. The CM use case addresses the coordination of trajectories of a group of vehicles in close proximity by first, sharing in-vehicle produced data, e.g., from on-board cameras, lidar, radar and the different on-board sensors and second, gathering other vehicles' information with combination of a precise positioning and traffic information. The goal is to help the driver (or the autonomous driving system) to have a clear and better view of the surrounding environment [same previous reference].

More specifically, the Cooperative Lane Merging (CLM) sub-case is considered as a reference scenario because it relies on 5G-aided cooperative position awareness and raises a variety of technical challenges in view of ensuring service continuity in a cross-border context. The CLM scenario considers a set of vehicles in which one of them wants to change its lane, due to any condition preventing the continuation of the driving on that lane (e.g., a line end) to an adjacent lane in which other vehicles are already driving. In a highway section, this situation may occur in different locations simultaneously. One could then set the start of the lane merge at different points with regard to the start of the cross-border area (i.e. and accordingly, assuming that the CLM procedure is initiated and completed before, while or after crossing the border), as represented in Figure 5-2 [26]. Two deployment options have been identified for the CLM system architecture. The first option is the centralized mode where either V2N connectivity (via LTE or 5G NR) using eNBs (eNodeB) /gNBs (gNodeB), or V2I connectivity via PC5-Mode4 using Roadside Units (RSUs) is used to connected to a centralized resource hosting the manoeuvre management application (ideally a MEC). The second option consists of a decentralized CLM mode where only V2V connectivity via PC5-Mode4 is used whereas the manoeuvre management application is hosted on-board in each vehicle [27]. In this paper we consider only the first option being the Centralized CLM, as illustrated in Figure 5-3.

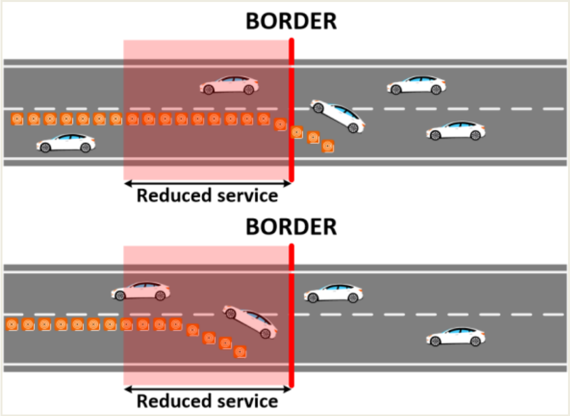


Figure 5-2: Illustration of Cooperative Lane Merging scenario. It is represented an imaginary border line followed by a special area where the communication service may be interrupted or altered. In the upper part of the figure, the lane merge occurs before the crossing of the border area but one of the vehicles involved is after the border line. In the lower part, the merging occurs within the border area, although one vehicle involved has not crossed the border line yet [26].

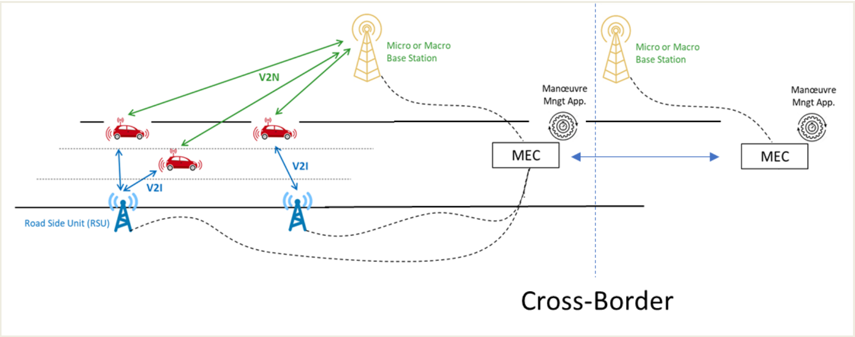


Figure 5-3: The two centralized CLM deployment options;1) the green arrows present the V2N-based connectivity where a base station is involved in the communication between the vehicles to execute the centralized CLM; 2) the blue arrows present the V2I-based connectivity using roadside units for the communication between the vehicles to execute the centralized CLM [26].

The considered highway in this study is the motorway linking Bologna to Munich over 600 km and crossing three countries namely Italy, Austria, and Germany. More specifically, we started our analysis with the Brenner pass (border between Italy and Austria), which is a 4.5 km junction of the considered highway.

5.4 Techno-economic analysis results

By using the techno-economic methodology outlined, we initiate our analysis by identifying the potential business models that can support the delivery of CCAM services. This is followed by a comprehensive evaluation of the costs associated with deploying the network that will provide these services, taking into account various deployment scenarios and the technologies adopted. The final step of this analysis requires a cost allocation exercise that links the costs of different deployment scenarios with the stakeholders involved in the identified business models.

5.4.1 Business model identification

In order to identify potential business models for providing CCAM services in cross-border settings, we use the Value Network analysis of Verna Allee [28]. Based on this framework, the first step in of the analysis is to identify the ecosystem. This can be achieved by following several steps:

- First, the *business roles,* indicating "what is being done", should be identified.
- The second step consists of the establishment of the value network, which depicts the interactions between these roles.
- The next step is the identification of *business actors*, showing "by whom" the roles are executed. In this step, actors are identified based on real world examples. These business actors are mapped onto the value network, resulting in a value network configuration.

After generating several network configurations, from a specific value network configuration, a *business model per actor* can be deduced, representing the incoming and outgoing value streams of this actor.

Key business roles and Value network

As a result of several rounds of discussion and feedback with the 5G-CARMEN consortium partners who represent the CCAM-related industries such as mobile network operators, car manufacturers, road operators etc. the business roles of CCAM ecosystem in cross-border settings have been identified. These business roles are classified into two categories being:

1) key business roles which represent those who are directly impacted by CCAM provision in the cross-border (summarized in the left part of Table 5-1) and

2) important business roles which are required to support (e.g. by providing adaptable regulations or required equipment etc..) the key business roles in their operations. These business roles are: *Regulations/ Policies* (role-8) both a national and international level and are developed to codify CCAM ecosystem; *Telecommunication standardization* (role-9): follows the increasing technological needs, aiming at having common rules and definitions of each technology among

countries. Implementing these standards across the telecommunication industry facilitates the inter-operability between different networks and systems; *Third-party infrastructure provision* (role-11): provides network infrastructure i.e., the tower company, which in many scenarios acts as a Neutral Host operator of infrastructure and as such (i) enables the passive/active infrastructure sharing by MNOs and other operators, (ii) hosts the edge computing nodes at the towers located close to the highway; Network Orchestration provision (role-12): designs and develops new solutions to solve orchestrations and management issues in the specific cross-border setting.

Table 5-1 Summary of the key business roles presented in the left-hand side of the table and the business actors presented in the right-hand side of the table (discussed in the next section) involved in the CCAM ecosystem.

Business roles	**Description**		**Business actors**	**Description**
End user the CCAM service consumption (role-1)	Consumes CCAM services in the cross-border area		End user	car driver
Service provision (role-2)	Provisions CCAM services e.g., CLM		CCAM service provider	this role can be taken by different entities depending on the provided service (e.g. cooperative merging; video streaming etc.) and will be detailed in the value network configuration subsection.
Long-range Connectivity provision Home country (role-3)	Provides long-range connectivity (4G/5G) in the home country of the end user.		road operator	Deploy and maintain road infrastructure and provide short-range connectivity like the Autostrada del Brenner who is operating the A22
Long-range Connectivity provision	Provides long-range connectivity (4G/5G) in the visited country.		Mobile network operator (MNO)	Deploys and maintains mobile networks and provides the long-range connectivity

Business roles	Description		Business actors	Description
Visited country (role-4)				and can be MNO-home country or MNO-visited country like Deutsche Telekom (Germany), Telecom Italia (Italy) and Magenta (Austria). MNOs can also provide short-range communication in its LTE/5G version C-V2X.
Short-range Connectivity provision (role-5)	In this work, focus is on 5G-networks, thus for short-range communication, C-V2X (more specifically in the long term, 5G NR-based C-V2X) variant is adopted. Since the short-range communication (5.9 GHz) is an unlicensed band, we do not make the distinction between home country and visited country because no spectrum license (which is national) is required to operate on this frequency band.		Car manufacturer	Manufactures connected cars like BMW, Fiat etc.
Mobile network operation Home country (role-6-1)	deployment, operation and maintenance of mobile networks (4G/5G) in the home country of the end user.		Regulators/ Policy makers	such as the European commission, BEREC, governments
Mobile	deployment,		Standards	like 3GPP, 5GPPP,

Business roles	Description		Business actors	Description
network operation Visited country (role-6-2)	operation and maintenance of mobile networks (4G/5G) in the visited country.		organizations	IETF etc.
Road infrastructure deployment and maintenance (role-7)	deploys safety instructions and roadside units and maintains road infrastructure		5G-CARMEN platform provider	the platform is developed by the 5G-CARMEN project and owned by the European commission. This later can provide the designed 5G-CARMEN platform to European MNOs in order to encourage them to open their networks and collaborate in align with the Digital Single Market agenda or can be adopted by a standardization body such as 3GPP and then MNOs will follow the standards in this regard (CCAM services in 5G networks).
Car manufacturing (role-10)	designs and produces cars that support CCAM services.			

Since the CCAM environment is very complex, some assumptions have been made to simplify the value network and to depict the most important actors, as follows:

- We assume that cars are equipped with on-board units that contain all the needed sensors and modems to be able to serve the selected use cases and that the car manufacturer is the responsible of that, hence, we do not include equipment vendors (e.g., OEMs).

- Car insurance companies are not included: as the focus of the work is on providing CCAM services in the cross-border setting, the role of this stakeholder is addressed in many other H2020 projects such as [29].
- The relation between the third-party infrastructure provision (role-11) and Mobile network operation (role-6) is not depicted in this value network same as the relation between the equipment vendors and Car manufacturing to simplify the value network.

The generic value network that depicts the relation between the different business roles involved in the CCAM ecosystem is presented in Figure 5-4.

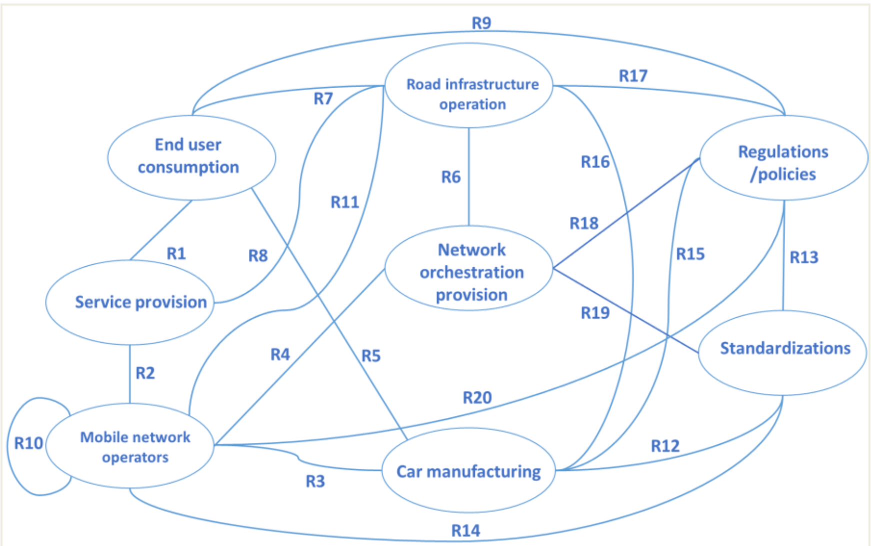


Figure 5-4: CCAM generic value network which depicts the relation between the involved business roles in the CCAM ecosystem; Rx indicates that there is a relation that links two business roles, yet the exact description of this relation will become clear in the value network configuration step discussed in the next section.

Business actors, value network configuration and business models

The key business actors identified by the 5G-CARMEN consortium have been mapped onto specific business roles, resulting in a value network configuration. This configuration illustrates the interactions and value streams between stakeholders. However, the focus of this study is on four key stakeholders: mobile network operators, road operators, car manufacturers, and service providers. These actors must collaborate in order to provide seamless connectivity across the corridor. Additionally, as this study considers a centralized system architecture that relies on either V2N or V2I connectivity, the value network configuration and business model of the mobile network operator and road operator will be specifically addressed.

Value network configuration 1: Mobile Network Operator (MNO) as CCAM service provider

In this value network configuration, the Mobile Network Operator (MNO) takes on the roles of Service Provision (role-2), Long-range Connectivity Provision in the Home Country (role-3), and Mobile Network Operation in the Home Country (role-6-1) and provides CCAM services to end users. For services that require short-range connectivity in addition to long-range connectivity, there are two options. The first is that the MNO leases this connectivity from road operators, while the second option is that the MNO leases only the road infrastructure to install its own equipment to deploy the short-range communication in its C-V2X variant. This latter option corresponds to passive network infrastructure sharing. The MNO agrees with car manufacturers on the required equipment to install in cars, such as the On-Board Unit (OBU) and 5G modem, before providing the service. This agreement can be made directly, via a consortium where several MNOs and car manufacturers are members (such as 5GAA), or via a mediator entity. Additionally, on the other side of the border, the MNO providing the service (MNO-H, role-3) needs to use the 5G connectivity of the visited country's MNO (MNO-V, role-4) through a roaming agreement or other forms of collaboration models, as discussed in the next section. These interactions are illustrated in Figure 5-5.

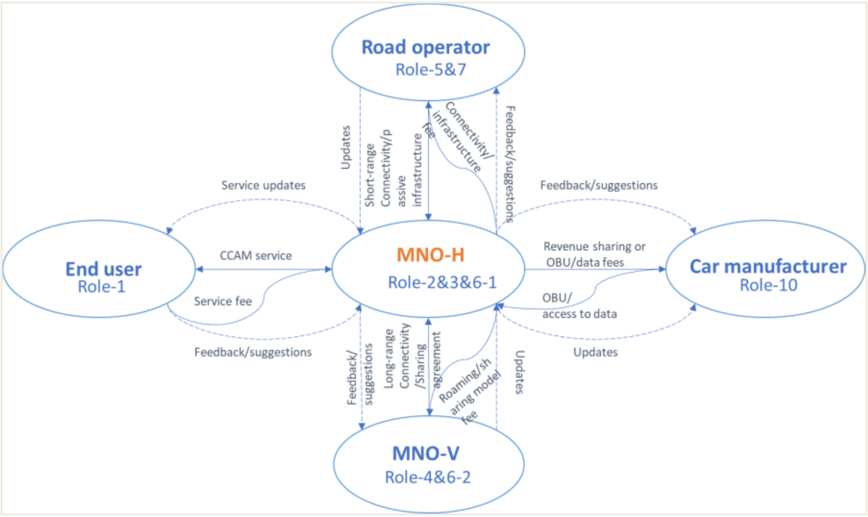


Figure 5-5: Value network configuration of MNO as CCAM service provider with role-1: End user the CCAM service consumption; role-2: Service provision; role-3: Long-range Connectivity provision Home country; role-4: Long-range Connectivity provision Visited country; role-5: Short-range Connectivity provision; role-6-1/2: Mobile network operation Home/Visited country; role-7: Road infrastructure deployment and maintenance; role-10: Car manufacturing.

Value network configuration 2: Road Operator (RO) as a CCAM service provider

In this case, the road operator (RO) plays the role of Service Provider (role-2) and also handles the deployment and maintenance of the road infrastructure (role-7) to provide CCAM services to end users. There are two categories of services the RO can provide: those that only require short-range communications (V2I-based services) and those that require long-range communications in addition to short-range connectivity (V2N and/or V2I-based services).

V2I-based services: In this case, the RO deploys Roadside Units (RSUs) along the corridor to ensure full coverage of short-range connectivity (C-V2X, ITS-G5 or a combination of both in a hybrid mode). Since many ROs only operate within one country, it is unlikely that the RO from the home country (RO-H) has the right to operate in the visited country. In such cases, an agreement must be made between the RO-H and the RO of the visited country (RO-V) for the deployment of RSUs. This agreement can be based on a passive or active sharing model. In the passive sharing model, the RO-V hosts the active equipment of the RO-H in its road infrastructure, whereas in the active sharing model, the RO-V provides connectivity or access to its active equipment to the RO-H. This option is presented as "**option1**" in Figure 5-6.

V2N and/or V2I-based services: In this case, the RO requires long-range connectivity (i.e. 5G connectivity) to offer CCAM services. As a result, the RO can lease this connectivity as a slice from MNO-H and MNO-V. The business model for leasing slices from MNOs in a cross-border environment is discussed in more detail in the next section. This second option of the value network configuration is presented as “**Option 2”** in Figure 5-6.

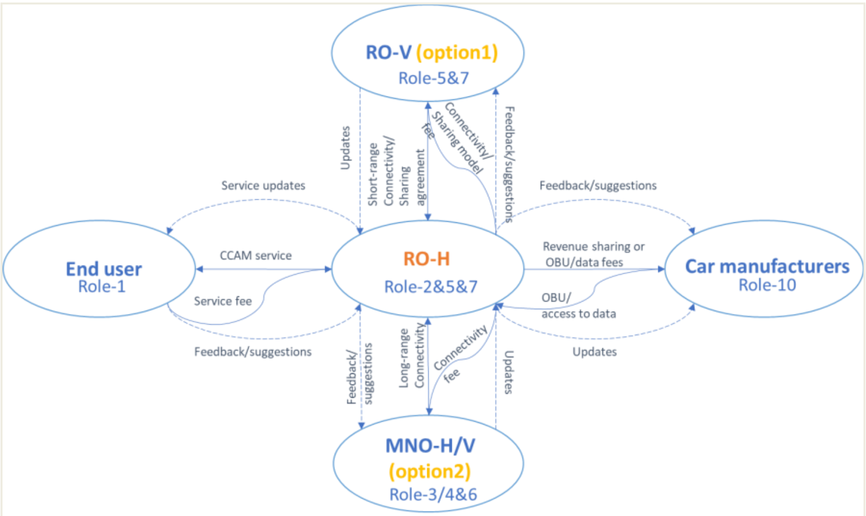


Figure 5-6: Value network configuration of Road Operator (RO) as a CCAM service provider with: role-1: End user the CCAM service consumption; role-2: Service provision; role-3: Long-range Connectivity provision Home country; role-4: Long-range Connectivity provision Visited country; role-5: Short-range Connectivity provision; role-6-1/2: Mobile network operation Home/Visited country; role-7: Road infrastructure deployment and maintenance; role-10: Car manufacturing.

Collaboration and 5G connectivity provision business models

When crossing the border, a handover has to be made between the network cell (base station) to which the car is connected to, and which belongs to the home MNO, to a new cell in the visited country that belongs to another MNO. This switch of connectivity is known as roaming. In 5G networks, we have two types of roaming, roaming with Local Breakout and Home Routed roaming. These two roaming types are discussed in detail in the section “Gap analysis with state-of-the-art systems” of [30]. The latter discussed why these two roaming types (in their current implementation) cannot fulfil the key KPIs of the CCAM use cases such as ultra-low latency and service continuity, hence the solution would be to rely on the new technologies coming with 5G networks such as MEC and network slicing. Therefore, new business and collaboration models using these new technologies should be investigated to resolve this issue from also a business model perspective.

However, the availability of these technologies namely MEC and network slicing depends on the 5G networks deployment. It becomes somehow a consensus, that the 5G networks will be rolled out following two phases namely short and long-term deployment following the standardization and the market readiness [31]. The first phase is the Non-Standalone (NSA) 5G, which is like a transition phase between the 4G and 5G networks, yet the second phase is a Standalone version (SA) of the 5G networks. Therefore, in our study, for each phase different business/collaboration models have been identified.

Short-term deployment

In the short-term deployment, the non-standalone option is prevailing, where a 5G- New Radio (5G-NR) in addition to 4G RAN are connected to the EPC (4G core network), thus an end-to-end network slicing is not possible. The technical architecture of 5G-CARMEN described in [32] proposes that both operators at the two border sides activate the inter-MEC interface to coordinate the handover prior to its execution time. This proposal assumes a strong coordination between the two operators; yet it is not always feasible given the different operators' priorities and visions. Therefore, another solution would be that a Mobile Virtual Network Operator (MVNO) leases infrastructure and spectrum from both countries and provides connectivity in the border to CCAM service providers or provides CCAM services themselves to end users. However, adding an MVNOs to the picture might lead to more complexity since the actual agreement for roaming should be executed by MNOs that serve this MVNO in addition to several regulation restrictions regarding leasing and re-selling radio spectrum. It might be a good alternative that in first two to three years of kicking-off CCAM services using V2I and V2V technologies (which will lead to offer of a limited set of CCAM services, those that do not require a high-capacity connectivity). Therefore, road operators will take a leading role in such a business model.

Long-term deployment

Based on the previous discussion, new business models must be identified to ensure the continuity of services for the CCAM users. We assume that we have 5G networks deployed next to the border and MNOs' networks support the end-to-end slicing and MEC is deployed near the border in the three countries that the highway is crossing.

Here we propose three new business models from an MNO perspective building on top of the *Value Network configuration: Mobile Network Operator (MNO) as CCAM service provider* described in section 5.4.1, these business models are based on the idea of using network slicing to support CCAM services, this concept is already used in [33], [34] and [35]:

1. **Collaborative business model**: In this case, the involved network operators establish an agreement between them stating that they collaborate to design an end-to-end slice that covers the entire corridor. This slice is (or multiple slices are) designed and customized in such a way that they meet the requirements of the offered services (typically an uRLCC slice for safety-related services and an eMBB slice for infotainment and video streaming services). Within the 5G network architecture, there are two types of handovers that are slice-aware, being the Xn-based handover and the NG/5GC-handover [36], which allow the switch between base stations while preserving the connection to the same slice. Challenges resulting from this business model are i) the definition of cost-revenue

sharing models that satisfy all the collaborating parties, and ii) the identification of responsibilities for maintaining and troubleshooting the joint network slices. The sharing model here is an active sharing model where slices of the network are shared. A slice consists of computing resources and storage on the core network side, a pipe on the transport network part and a pipe on the radio resource side as well (e.g., reserved piece of the operator's spectrum). The corresponding collaborative business model is depicted in Figure 5-7.

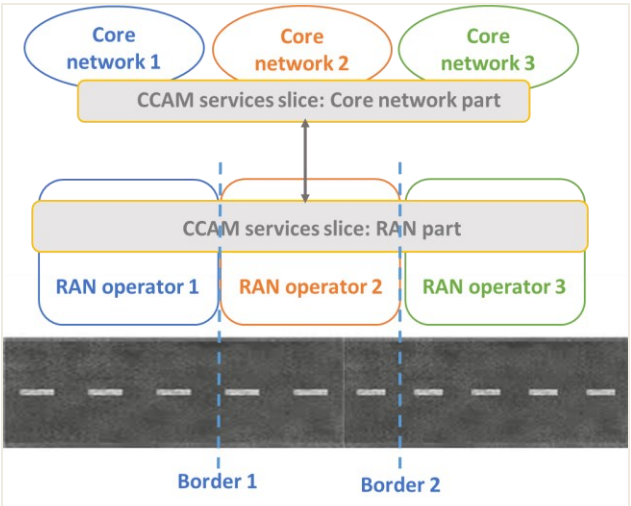


Figure 5-7: Technical architecture of the collaborative business model, here the figure depicts the example of three MNOs in 3 neighboring countries designing a common network slice to support the provision of CCAM services in the cross-border areas.

2. **Slice-as-a-service (SaaS) business model**: operators that intend to cover the corridor have to lease a network slice from the operators that cover different parts of the corridor where they do not have network infrastructure deployed. One of the technical challenges that the operators (who wants to provide services to the end user) face is how to establish the continuity between different slices coming from different operators. Imagine that we have the corridor that traverses Germany, Austria and Italy and an Austrian network operator wants to cover this entire corridor. In this case, this MNO needs to lease a slice from a German MNO and another one from an Italian MNO, these two slices should be integrated into its own slice to ensure the continuity of the provided services. Usually, the agreement that should be made between the involved operators should clearly state the type of the slice, the required KPIs, the Quality of Service (QoS) class, and the type of the resource allocation model for this slice (static or dynamic allocation). All these

requirements can be described in a Service Level Agreement (SLA). The Slice as a service business model is illustrated in Figure 5-8.

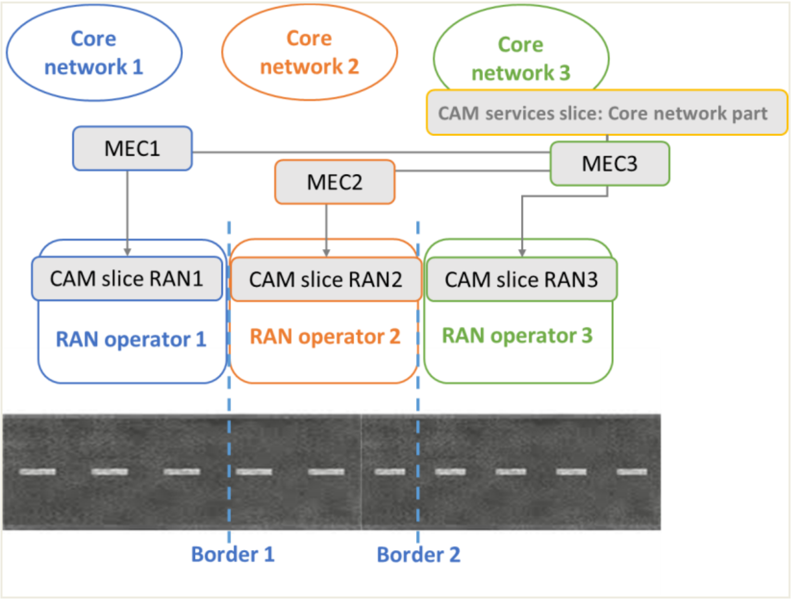


Figure 5-8: Technical architecture of the SaaS business model, here the figure depicts the example of three MNOs in 3 neighboring countries and leasing a network slice on their networks to MNO 3.

To understand the technical challenges of integrating the 3 slices (or even more), it is needed to know that individual slice types are recognized by the so-called ‘Single Network Slice Selection Assistance Information’ (S-NSSAI), which consists of two parts, namely the Slice Service Type (SST) and the Slice differentiator (SD). The SST defines the specific slice type with its supported features and services, and the SD, on the other hand, is used to distinguish between multiple slice instances with the same SST-identifiers. Network operators can create and use their own SSTs, but they might also choose to support the standardized SSTs shown in Table 5-2.

Table 5-2: Standardized Slice Service Types (SSTs) [37]

Slice/Service type	SST value	Characteristics.
eMBB	1	Slice suitable for the handling of 5G enhanced Mobile Broadband.
URLLC	2	Slice suitable for the handling of ultra- reliable low latency communications.
MIoT	3	Slice suitable for the handling of massive IoT.

Using one of the standardized SSTs can help to simplify roaming scenarios, where users are connected to a visited PLMN (Public Land Mobile Network) that could not support or know the S-NSSAIs from the user's home PLMN. In this case, the visited PLMN needs to translate the S-NSSAI of the user's home PLMN to a recognized one and connects the user to it.

In our case, the three slices coming from the three operators, in 99% of usual cases, have different S-NSSAIs, which will lead us again to the case of roaming and the discontinuity of the service. That is why the third operator (in our example) has to figure out a solution to integrate the three slices to be seen as one single slice with a single identifier (S-NSSAI). One solution could be, that telecommunication regulators (e.g., BEREC) fixe a range of SDs that are reserved only for cross-border usage, to simplify the integration of the different sub-slices.

This business model (Slice-as-a-Service) is gaining more and more interest from several verticals such as transportation and logistics companies, healthcare providers, factories and energy, automotive and utilities providers [38]. According to the same source, the global market size of network slicing is forecasted to grow from USD 161 million in 2020 to USD 1,284 million by 2025.

3. **Non-collaborative model**: in this case, every network operator deploys its own network to cover the corridor independently from the others. However, the network deployment could be non-cost effective due to the huge investment that should be made to cover the entire corridor and the cost of the spectrum licenses that the network operator must buy or lease in all the countries that the corridor traverses. Additionally, the operator must comply with immense legal obligations and costs of operating a mobile network in an additional country, unless a harmonized European solution will be established. Besides, the roaming problem is just transferred to another place in the visiting country because sooner or later a switch between networks has to be executed.

In general, some challenges faced by the three discussed business models to be realized in the real world are i) the lack of network coverage near country borders and/or network interferences with foreign MNOs; ii) the adoption of different deployment strategies of 5G networks (in both non-standalone and standalone deployments) in the involved countries and the different timelines of these deployments; iii) the necessary support of end-to-end slicing by the involved MNOs networks, as well as MEC and finally iv) regulations that put restrictions on buying and re-selling portions of the spectrum.

Based on confidential interviews made with telecom operators, the preferable business model among the three discussed above is the collaborative business model but with a detailed and defined cost-revenue sharing model. The second business model seems a bit difficult to realize since the integration of the three slices is technically very challenging, while somehow shifting the continuity issue from the border to the edges of each slice. Finally, the third business model is not realistic, even if it is still relevant to mention it for the sake of completeness in our analysis.

5.4.2 Cost modelling of the studied use cases

To evaluate the identified business models discussed in the previous section. We developed a generic cost model and we then apply it to two sets of deployment scenarios to support the centralized canonical use case, namely CLM, using V2I and/or V2N connectivity respectively. These two sets of deployment scenarios are studied in the Brenner pass for a highway portion of 4.5 km.

Generic cost model and scenario analysis

To evaluate the economic feasibility of the selected scenarios and identified business models, a generic Total Cost of Ownership (TCO) model is constructed. The model considers both Capital Expenditure (CAPEX) and Operating Expenditure (OPEX), as well as overhead costs such as marketing and helpdesk. The TCO model is composed of two main sub-models: the dimensioning process module and the cost model. The dimensioning process module uses inputs such as the network design principles, existing infrastructure on the corridor, traffic densities, base station coverage, RSU coverage, and the project horizon to determine the necessary network infrastructure to cover the entire corridor and meet the defined Key Performance Indicators (KPIs). It is noteworthy that the dimensioning process will be elaborated using system-level simulation tools (as described in [27]). The cost model, which takes the bill of materials as well as CAPEX and OPEX elements as inputs, calculates the TCO of the network deployment needed to achieve full connectivity coverage of the corridor and support CCAM services. The model is generic and is presented in Figure 5-9. This cost model aims to calculate the cost of the network deployment needed to have a full connectivity coverage of the considered corridor to support the provision of CCAM services. A cost allocation model will be used when evaluating the business case of each key stakeholder. In addition, considering the existing infrastructure of the short-range communication (V2I infrastructure) and long-range communication (V2N mobile networks) as a baseline and varying the number of new RSUs and base stations to be installed beyond, as well as the timeline of this deployment, will help defining different deployment scenarios.

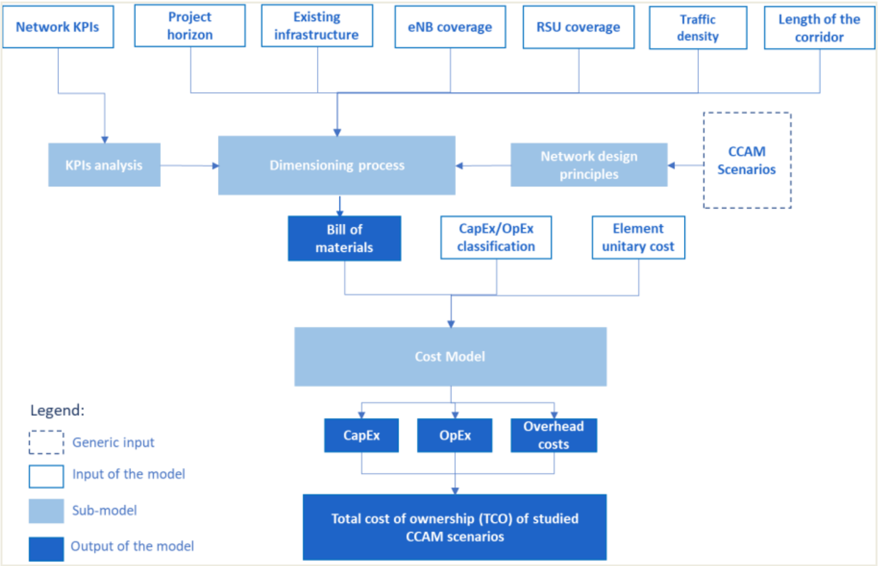


Figure 5-9: Generic cost model structure

For the viability study, the TCO model will be applied to a different set of scenarios which is known as a scenario analysis exercise. The scenario analysis technique has been chosen because of its best fit in strategic planning to evaluate the potential outcomes of different input parameters on a quantitative model, providing valuable insights for decision making, particularly when accurate predictions are difficult to make. In our case being deploying 5G network to support CCAM provision, scenario analysis can be helpful in assessing the demand for networks and making deployment decisions [39]. The specific use of scenario analysis in this study involves defining various deployment scenarios (V2I and V2N) to consider different fleet penetration assumptions.

Cost modelling results of V2I deployment scenarios

For the V2I links with respect to PC5 Mode4 – enabled RSUs, three deployment scenarios have been defined in alignment with the system-level simulations [27], where service availability has been assessed as a function of different traffic density figures. These three deployment scenarios are as follows:

- The baseline V2I scenario corresponds to the physical deployment currently available in the studied corridor: 2 RSUs in total in the test portion of highway of length ~ 4.5 km (i.e., ~ 0.5 RSU/km).
- The dense V2I deployment scenario: this scenario corresponds to a doubled baseline density: 4 RSUs in total in the same environment (i.e., ~ 1 RSU/km).

- The ultra-dense V2I deployment scenario: this scenario corresponds to a quadrupled baseline density: 8 RSUs in total in the same environment (i.e., ~ 2 RSUs/km).

Note that the names attributed to these different scenarios are defined based on the densification of RSUs comparing to the baseline V2I scenario.

Model inputs and assumptions

Inputs used to run the cost model are classified as public values that can be shared (derived from literature) and confidential (internal project data). For the latter however, we provide practical intervals: 2 to 6 k€ for the cost per RSU, 20 to 50 W for the average power consumption per RSU and 10 to 15 k€ for the linear cost of 1 km of fiber. In addition, the cost model is based on several assumptions that should be made to get results as realistic as possible, as follows:

- A lifetime of 10 years is considered for RSUs [40];
- Project horizon is 10 years (hence no hardware renewal is anticipated)[16];
- Hardware installation cost is 15% of the hardware costs [41];
- Maintenance cost for V2I infrastructure is assumed to be 2% to 5% of the original hardware and labor costs (CAPEX) according to [40];
- The overhead cost is around 22% on top of the sum of the CAPEX and OPEX costs [70].
- Yearly costs are also corrected for inflation (time value of money) with 5% as growth rate.
- Assuming a kWh price of 10.85 €-cents based on the average prices of electricity for the industry in Belgium from 2008 to 2019 [43].
- The cost of electrification and cabling for the road infrastructure is quoted per kilometer. This cost is there to support not only the V2I infrastructure but also several other services such as public lightening etc. Hence, a cost allocation metric is needed to allocate a part of that cost to the V2I infrastructure. In this first version of the cost model, we start with 15% as a reasonable allocation metric. This will be validated and updated afterwards, if necessary.
- The cost of fiber installation is quoted also per Km, assuming that fiber cable will be installed along the entire corridor portion under study with the first deployment scenario (to serve the first RSUs to be deployed). Hence, we assume that there is no need for additional fiber installation when deploying new RSUs.
- The TCO is the total cost of ownership for the project lifetime (hence, costs are not discounted).

Model simulation

We simulated the cost model described in Section 5.4.2 for the three deployment scenarios described previously namely the baseline, the second and the third

[16] The lifetime of most equipment is almost 10 years thus the lifetime of the project is also assumed as 10 years without renewal of hardware.

scenario. CAPEX, OPEX, Overhead and the TCO have been studied accordingly. The TCO at year N is the cumulative cost of N years including capital and yearly operational expenditures as well as overhead costs.

1. Results of the baseline V2I scenario

This scenario corresponds to the existing deployment at Brenner Pass (i.e., at the border between Italy and Austria). Although, this V2I infrastructure is already deployed, results of the cost modelling of this scenario can give insights into the deployment costs of similar infrastructure in other locations (i.e., the rest of the corridor where no V2I infrastructure is already present). In addition, the operational costs of the existing infrastructure have to be included in the total cost of ownership of the final deployment (existing deployment plus additional deployment required to support CCAM services at the cross border). Results of the baseline scenario are depicted in Figure 5-10. From these results we can conclude that the CAPEX is a significant cost to deploy the V2I infrastructure. To understand the dominating component of the CAPEX, a cost breakdown is made and illustrated in Figure 5-10. The figure shows that the two main cost components are the fiber and electrification costs, yet the cost for RSUs is a minor cost. Recurring costs are the OPEX and overhead costs. The TCO of deploying and keeping up and running two RSUs in the considered part of the corridor for 10 years is 286K euro, thus the cost per km for the same period is 63.5K euro.

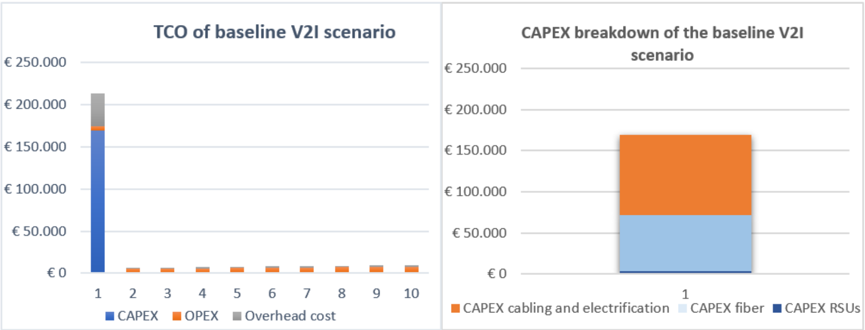


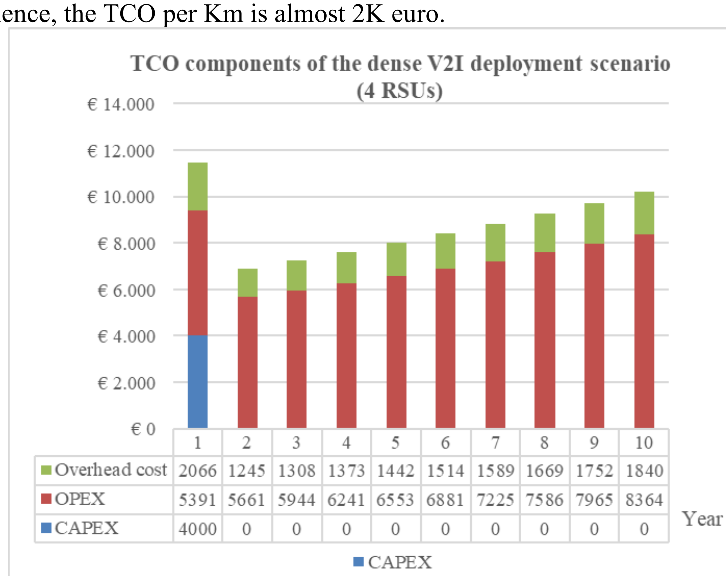
Figure 5-10: Different TCO cost components of the baseline scenario (in left-hand side) and CAPEX breakdown of the baseline scenario (in the right-hand side).

It is noteworthy that the cost of cabling and electrification of the considered corridor portion is a significant cost and has a major impact on the TCO of the baseline scenario. Moreover, in the absence of a strong assumption about the allocation metric to be used to allocate this cost to the V2I infrastructure, results of this analysis are considered as preliminary and would require further validation.

2. Results of the dense V2I deployment scenario

The dense V2I deployment scenario corresponds to 1 RSUs/km (i.e., 4 RSUs in the test sub-area of 4.5 km), thus adding two new RSUs comparing to the baseline scenario. We simulated accordingly the cost model to derive two types of results. First, we consider 1) the cost to deploy 4 RSUs at the Brenner Pass (4.5 Km of the corridor) and keep them up and running for 10 years. This cost includes the CAPEX of the two new RSUs (i.e., assuming we have already 2 RSUs installed), the OPEX of the total deployed RSUs (i.e. 4RSUs). Second, we consider 2) the incremental deployment cost (i.e., incremental comparing to the baseline scenario), which means the CAPEX, OPEX and overhead costs of the two new RSUs only.

The TCO and its cost components of deploying 2 additional RSUs and keep the total 4 RSUs up and running for 10 years, two of them are already installed are summarized in Figure 5-11. Results show that the CAPEX is not a dominant cost component as is the case in the baseline scenario and this is because the considered portion of the corridor (4.5 km on the Brenner Pass) is already equipped with fiber and electricity. On the other hand, the OPEX of the second scenario is slightly higher than the OPEX of the baseline scenario since on top of the latter the power consumption and maintenance of the two new RSUs are added. The total cost of ownership (TCO) and the TCO per Km for deploying 4 RSUs (two of them are already installed) and keep them up and running for 10 years are 87.6K euro and 19.5K euro respectively. Yet, the incremental TCO for 10 years of adding two new RSUs to the existing deployment is thus around 8K euro, hence, the TCO per Km is almost 2K euro.



	1	2	3	4	5	6	7	8	9	10
Overhead cost	2066	1245	1308	1373	1442	1514	1589	1669	1752	1840
OPEX	5391	5661	5944	6241	6553	6881	7225	7586	7965	8364
CAPEX	4000	0	0	0	0	0	0	0	0	0

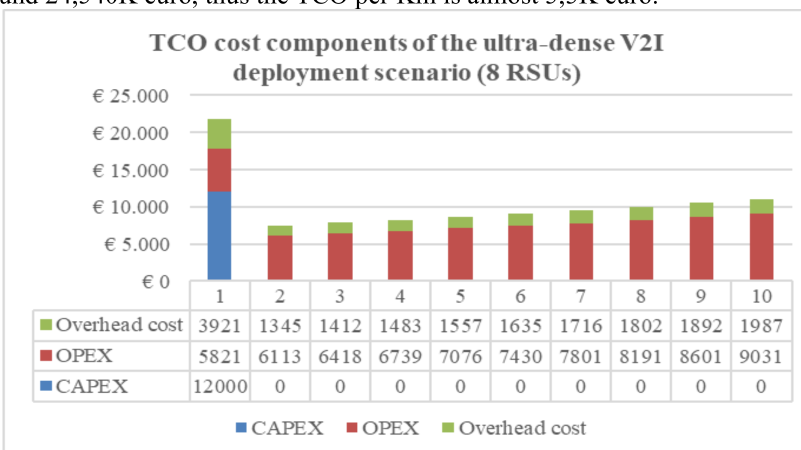
Figure 5-11: TCO cost components of the dense V2I deployment scenario: CAPEX of the two new and the OPEX of the total deployed RSUs i.e., 4 RSUs.

3. *Results of the ultra-dense V2I deployment scenario*

The ultra-dense V2I deployment scenario corresponds to a quadrupled baseline density which means 2 RSUs/km (i.e., 8 RSUs in the tested sub-area), hence adding 6 new RSUs comparing to the baseline scenario. Similar to the dense V2I scenario, we have simulated the cost model to derive two types of results:

1) The cost of deploying 8 RSUs in the Brenner Pass (4.5 Km of the corridor) and keep them up and running for 10 years. This cost includes the CAPEX of the 6 new RSUs (we have already 2 RSUs installed), the OPEX of the total deployed RSUs i.e., 8 RSUs.
2) The incremental deployment cost (incremental comparing to the baseline scenario), which means the CAPEX, OPEX and overhead costs of the 6 new RSUs.

The TCO and its cost components of deploying 6 additional RSUs and keep the total 8 RSUs up and running for 10 years, two of them are already installed are summarized in Figure 5-12. The Cumulative TCO of this scenario is 103,970K euro which results in 23,104K euro per km. Like the dense V2I deployment scenario, results show that the CAPEX is not a dominant cost component as in the baseline scenario. On the other hand, the OPEX of this scenario is higher than the OPEX of the baseline scenario since on top of the latter, the power consumption and maintenance of the 6 new RSUs are counted. Yet, the incremental TCO for 10 years of adding 6 new RSUs to the existing deployment in the Brenner Pass is around 24,540K euro, thus the TCO per Km is almost 5,5K euro.



	1	2	3	4	5	6	7	8	9	10
Overhead cost	3921	1345	1412	1483	1557	1635	1716	1802	1892	1987
OPEX	5821	6113	6418	6739	7076	7430	7801	8191	8601	9031
CAPEX	12000	0	0	0	0	0	0	0	0	0

Figure 5-12: TCO cost components of the ultra-dense V2I deployment scenario: CAPEX of the 6 new and the OPEX of the total deployed RSUs i.e., 8 RSUs.

As a conclusion of the cost modelling of the three studied scenarios, installing the first RSUs (i.e., baseline scenario) has a significant cost since we need to install fiber and electricity cables which are considered as the dominant investment cost component based on Figure 5-10. Beyond, adding new RSUs will not have an important capital expenditure since the required cabling to connect the additional RSUs is already there. This conclusion can clearly be observed in Figure 5-13.

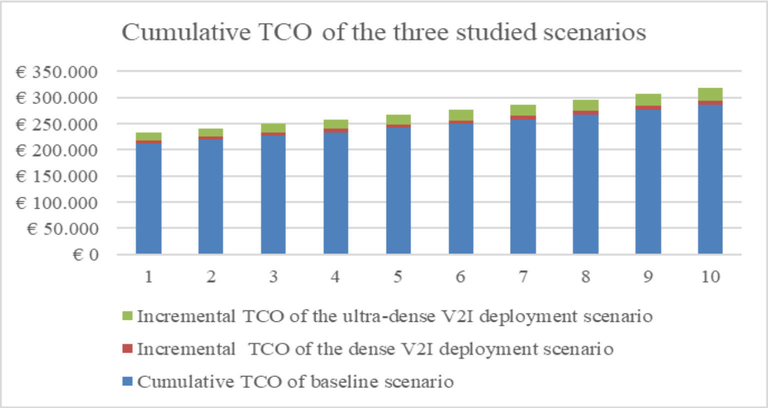


Figure 5-13: TCO and Incremental TCO of the three considered scenarios: figures shows that the baseline scenario bears the higher TCO among the two other scenarios since installing fiber and electricity cables to serve RSUs is the dominant cost component.

The three studied deployment scenarios were simulated to occur in the same year. However, it is an interesting exercise to vary the timeline of the deployment according to the adoption rate of CCAM services and to study its impact on the TCO. This will be elaborated based on the inputs from the market analysis and will be subject of future works.

The impact of the variation of the infrastructure deployment in terms of performance has been studied also from a system-level simulation angle. Results show that the service availability, defined as the rate of completed CLM negotiations when all the required intermediary messages have been correctly received, can be satisfied (around 99.9%) with high road traffic densities (i.e., 83 vehicle/km) if we consider the ultra-dense V2I deployment scenario (8 RSUs), whereas low road traffic densities (i.e., 28 vehicle/km) can be supported with the dense V2I deployment scenario (4 RSUs) around 98.4% as service availability. Yet, the baseline scenario (2 RSUs) results in a modest service availability 95.8% and 71.9% for the road traffic densities 28 vehicle/km and 83 vehicle/km respectively [44].

Cost modelling results of V2N deployment scenarios

V2N deployment scenarios will also support the provision of the CLM centralized option (i.e., using mobile networks based on 4G/5G connectivity) in the studied corridor, the Brenner pass (border between Austria and Italy). Two initial scenarios have been identified. As MEC technology is crucial in fulfilling the latency KPIs discussed previously, two variants per scenario have been considered, where MEC is installed either in a distant Data Center (DC), next to the core network, or alternatively, locally next to the tower (e/gNB). The difference between these two variants is depicted in Figure 5-14.

- Baseline scenario V2N: corresponds to the physical deployment currently available in the studied corridor portion, which means 2 installed eNBs, working on 800 and 1800 MHz
    - Baseline V2N scenario-a: MEC are installed next to the towers; thus, in total we have 2 MEC nodes
    - Baseline V2N scenario-b: MEC nodes are hosted in Data Center; hence, one in Austria and one in Italy; in total we have 2 MECs.
- Dense V2N scenario: in addition to the current deployment, an additional eNB will be installed (working on 800/1800 MHz).
    - Dense V2N scenario-a: MEC are in the Tower; thus, in total we have 3 MEC nodes
    - Dense V2N scenario-b: MEC nodes are in the DC; one in Austria and one in Italy; in total we have 2 MECs.

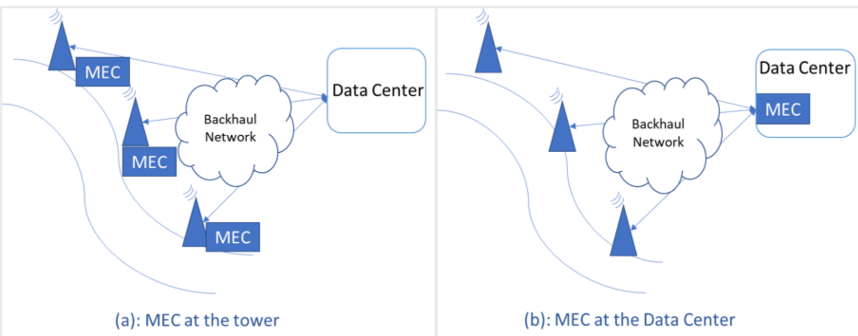


Figure 5-14 Placement of the MEC (a) at the tower and (b) at the Data Center

Model inputs and assumptions

Besides the cost inputs and assumptions mentioned in the previous section, additional data is used to run the cost model for these V2N scenarios. Some of these data are confidential cost figures (i.e. project internal data) but fall in the following intervals: Cost of building a new macro site (only the tower infrastructure) is 30 to 50 k€; Cost of a 2G/3G/4G BS is 25 to 45 k€ (antennas excluded); Cost for upgrading to 5G a 2G/3G/4G BS is 40 to 55 k€ including active antennas; maintenance of mobile network equipment is assumed to be 10% of the original hardware and labor costs (CAPEX) [41]. The cost of a MEC edge server of type "gigabyte-h242-z10" is 10 k€ [45]; cooling system and cabinet costs for MEC deployment are around 2.2 k€ and 2.5 k€ respectively [17].

Model simulation

We have simulated the cost model described in Section 5.3.2 for the two scenarios, described previously, Baseline V2N scenario-a&b and Dense V2N scenario-a&b. CAPEX, OPEX, Overhead and the TCO have thus been studied for each of the considered scenarios. The TCO at year N is the cumulative cost of N years including capital and yearly operational expenditures as well as overhead costs.

Results of the first scenarios being the Baseline V2N scenarios show that the CAPEX of Baseline V2N scenario-a is higher than in Baseline V2N scenario-b, since installing MEC next to the towers requires installing cabinets to host both the MEC equipment and its cooling system, which are not required if we deploy it in the data center. The different TCO cost components are illustrated in Figure 5-15. Results of the cost model for the Dense V2N scenario-(a & b) when an additional eNB is installed are presented in Figure 5-16. Compared to the Baseline V2N scenario, adding a new eNB adds around 500 k€ as an incremental cumulative TCO for 10 years with 77 k€ as incremental CAPEX, 400 k€ as incremental cumulative OPEX and 100 k€ as incremental cumulative Overhead cost. The OPEX is a dominant cost in this scenario due to the site rental cost required to host the new macro cell (eNB). On the other hand, deploying MEC next to the towers means installing 3 MEC nodes in Dense V2N scenario-a compared to only 2 MEC nodes in Dense V2N scenario-b, thus the CAPEX is slightly higher as shown in Figure 5-16.

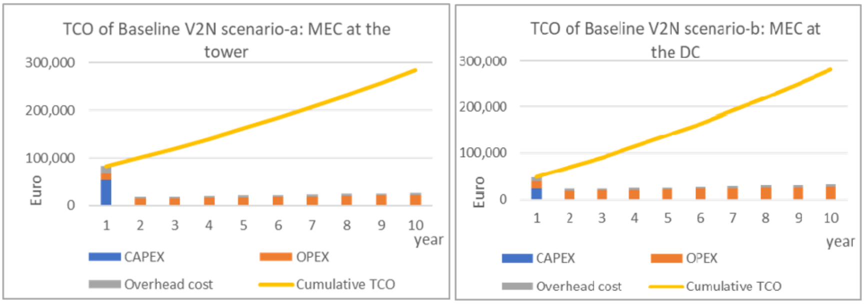


Figure 5-15 TCO cost components of Baseline V2N scenario-a: MEC at the tower (at the left-hand side) and Baseline V2N scenario-b: MEC at the DC (at right-hand side)[17], Comparing the two TCOs we can conclude that installing MEC at towers is more expensive than installing it at the DC and the dominant cost component here is the CAPEX of the MEC nodes (Blue color).

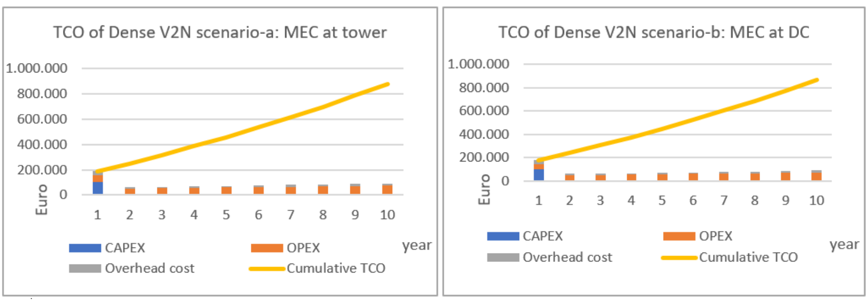


Figure 5-16 TCO cost components of Dense V2N scenario-a: MEC at the tower (at the left-hand side) and Dense V2N scenario-b: MEC at the DC (at right-hand side), Comparing the CAPEX of the two scenarios we can conclude that installing MEC at towers is more expensive than installing it at the DC but this small difference is cancelled out over 10 years since the dominant cost component here is the OPEX of the eNBs (Orange color)which is equal in the two scenarios.

One could think about installing MEC nodes at different positions of the network topology, that is, not only next to towers or in data centers. Therefore, MEC clustering models is planned to be studied as a future work.

5.4.3 **Cost allocation model**

To evaluate the proposed business models for the key stakeholders namely the MNO and the RO, a cost allocation model is required to link the cost of the different V2I and V2N deployment scenarios to these stakeholders.

[17] It is worth mentioning that these results are under validation by the concerned MNOs

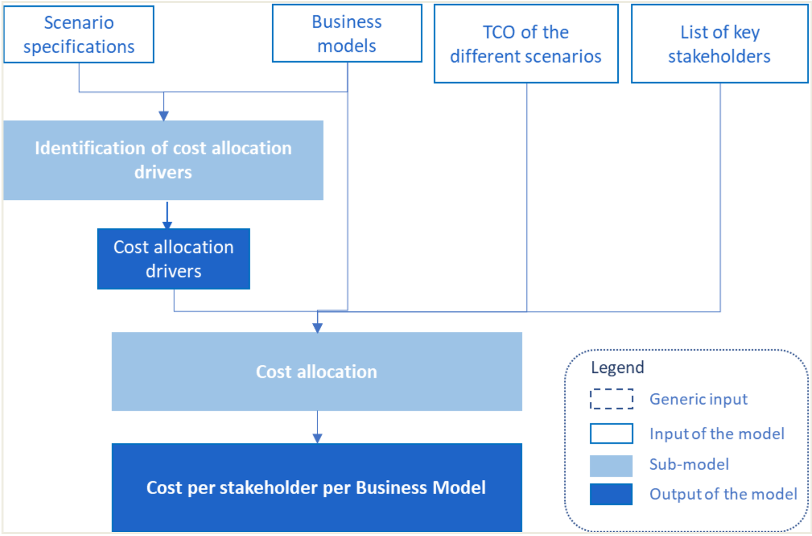


Figure 5-17 Building blocks of the cost allocation model

The allocation model suggested here consists of two steps as presented in Figure 5-17 with the light blue (sub-models). First, the identification of the cost allocation drivers which requires as inputs the scenario specifications (e.g., type of connectivity used) and the business models (e.g., in collaborative business models, the degree of collaboration helps in defining the cost allocation driver) as well. In the second step, the actual allocation process is executed where the TCO of the different scenarios is divided among the key stakeholders based on the cost allocation drivers identified in the previous step. The outcome of this step is the cost per stakeholder per business model. The building blocks of the cost allocation model are depicted in Figure 5-17.

We applied this model to our case starting first by the identification of the cost allocation drivers where the scenario specifications and the identified business models are used as key inputs, results are summarized in Table 5-3.

Table 5-3 Identification of the cost allocation drivers

Input	Description	Cost allocation drivers	Explanation
Scenario specifications	V2I and V2N deployment scenarios	Type of connectivity	Allocate the cost of the deployment to the operator that usually provides the type of connectivity studied in that specific deployment scenario (e.g., 5G connectivity to MNO)
	Geographical location of the deployment	Country of the deployment	Under national regulations and in cross-border environment, operators

			should be operating within their home country.
Business models	Non-collaborative model	100% allocation to one stakeholder.	In the non-collaborative business model, each stakeholder bears the entire cost of the deployment scenario considered.
	Collaborative model	Degree of collaboration and network sharing (**passive** versus **active** sharing)	Using network sharing models results in different cost savings [22]. Passive infrastructure sharing could save from 16% to 35% CAPEX and 16% to 35% OPEX. Active infrastructure sharing excluding spectrum could result in savings from 33% to 35% CAPEX and 25% to 33% OPEX. Yet savings in the active sharing including spectrum sharing model would range from 33% to 45% CAPEX and 30% to 33% OPEX.

For the second step of the cost allocation model, the list of the key stakeholders and the TCO of the different deployment scenarios are needed. The key stakeholders (previously identified in section 0) are the Mobile Network Operator (MNO) and the Road Operator (RO). As the corridor segment under study is the Brenner pass between Austria and Italy, we assume that MNO-A-1, MNO-A-2, RO-A-1 and RO-A-2 are operating in the Austrian side of the Brenner pass whereas, MNO-I-1, MNO-I-2, RO-I-1 and RO-I-2 are operating in the Italian side. The TCO and the different cost components of studied deployment scenarios are summarized in Table 5-4.

Table 5-4 TCO and cost components of the different considered deployment scenarios

Scenario	Cumulative TCO (€)	CAPEX (€)	Cumulative OPEX (€)	Incremental Cumulative TCO (€)	Incremental CAPEX (€)	Incremental Cumulative OPEX (€)
Baselin	286,068	169,37	65,107	/	/	/

Scenario	Cumulative TCO (€)	CAPEX (€)	Cumulative OPEX (€)	Incremental Cumulative TCO (€)	Incremental CAPEX (€)	Incremental Cumulative OPEX (€)
e V2I scenario		5				
Dense V2I scenario	87,610	4,000	67,812	8,180	4,000	2,750
Ultra-dense V2I scenario	103,970	12,000	73,221	24,540	12,000	8,114
Baseline V2N scenario-a	284,717	53,400	179,975	/	/	/
Baseline V2N scenario-b	282,003	22,000	209,150	/	/	/
Dense V2N scenario-a	920,961	119,100	635,786	636,244	65,700	455,811
Dense V2N scenario-b	865,699	99,400	610,189	583,695	77,400	401,038

As clearly shown in Table 5-4, we studied seven scenarios, three of them are V2I connectivity-based scenarios and the four remaining scenarios are V2N connectivity-based. Hence, given the cost allocation driver “type of connectivity” (discussed in Table 5-3), the three V2I scenarios are allocated to ROs and the four V2N ones are assigned to MNOs.

As for the V2I connectivity-based scenarios, three scenarios have been studied in Section 5.4.2, namely the baseline V2I scenario (2RSUs), which corresponds to the existing infrastructure on the Italian side of the border (deployed by RO-I-1); the dense V2I scenario (4 RSUs) and the ultra-dense V2I scenario (8 RSUs). If the Austrian RO (for example RO-A-1) would like to cover the Austrian side with the same deployment as RO-I-1, within the non-collaborative business model, the cost of the V2I deployment would be the TCO of the baseline V2I scenario. Whereas,

if RO-A-1 and RO-A-2 agree to share the investment cost different investment cost savings could result from the different sharing models (passive versus active sharing). For the dense V2I scenario, two additional RSUs must be installed. This can be done by any ROs on the two sides, if the non-collaborative model is adopted, the RO deploying these additional RSUs will bear the incremental TCO, yet if one of the sharing models is adopted, savings on the incremental cost can be gained. For the ultra-dense V2I scenario, where 6 additional RSUs will be installed comparing to the baseline scenario, similar analysis is found. Allocated costs as well as achieved cost savings are summarized in
Table 5-5. MONs on both sides could also realize cost savings as ROs if they decide to also provide the short-range (V2I) connectivity and adopt network sharing with ROs operating in the considered corridor.
For the V2N deployment, MNOs are responsible for the network deployment, yet if ROs want to provide the CLM service using 5G connectivity, a slice can be leased from MNOs based on either the collaborative (common slice designed by MNOs on the two border sides) or the SaaS business models described previously in section 5.4.1, but the network coverage needs to be complete (only in Dense V2N scenario) and the price of leasing a slice is still unknown for the moment. This is considered as a revenue for MNOs, which can help them to reach the ROI (Return On Investment) faster. As for V2N connectivity, according to the Baseline V2N scenario, we consider two existing eNBs, one of them is in the Austrian side and the other one in the Italian side, with MEC at the tower (a) and MEC at DC (b). Hence, the TCO can be divided, 50% being allocated to MNO-A-1 and 50% to MNO-I-1. In the dense V2N scenario, we consider an additional eNB is installed on the Italian side since there is a network coverage gap there, thus Dense V2N scenario-a with 3 eNBs and MEC at the tower and Dense V2N scenario-b with 3 eNBs and MEC at the DC. In the non-collaborative business model one of the MNO-I will bear the incremental cost comparing to the first scenario. Sharing models can be adopted by the Italian MNOs to save on the investment cost. Cost allocation figures for the different scenarios are summarized in Table 5-5.

Table 5-5 Allocation cost to MNOs and ROs based on the adopted business model

Allocation model	**Austria**		**Italy**	
	MNO	**RO**	**MNO**	**RO**
Baseline V2I Scenario	/	RO-A-1: 286 k€ as cumulative TCO for 10 years (**non-collaborative model**) RO-A-1&2 can save from 27.1 to 59.3 k€ CAPEX	/	/

		and 10.4 to 22.8k€ OPEX (**passive network sharing**), Thus, RO-A-1 and RO-A-2 need to invest from 102k€ to 124.3 k€ each as TCO for 10 years. RO-A-1&2 can save from 55.9 to 60 k€ CAPEX and 16.3 to 22.8 k€ OPEX (**active sharing**). Hence, RO-A-1 and 2 need to invest from 102k€ to 107 k€ each as TCO for 10 years.		
Dense V2I scenario	/	Deploying two new RSUs on top of the baseline scenario results in 8 k€ as TCO for RO (**non-collaborative model**). RO-A-1&2 can save from 0.64 to 1.4 k€ CAPEX and 432€ to 950 € OPEX (**passive network sharing**). Thus, RO-A-1 and RO-A-2 need to invest from 3k€ to 3.6 k€ each as incremental TCO. RO-A-1&2 can save from 1.32 to 1.4 k€ CAPEX and 680€ to 950€ OPEX (**active sharing**). Hence, RO-A-1 and RO-A-2 need to invest from 3k€ to 3.1 k€ each as incremental TCO.	/	Deploying two new RSUs on top of the baseline scenario results in 8 k€ as TCO for RO (**non-collaborative model**). RO-I-1&2 can save from 0.64 to 1.4 k€ CAPEX and 432€ to 950€ OPEX (**passive network sharing**). Thus, RO-I-1 and RO-I-2 need to invest from 3k€ to 3.6 k€ each as incremental TCO. RO-I-1&2 can save from 1.32 to 1.4 k€ CAPEX and 680€ to 950€ OPEX (**active sharing**). Hence, RO-I-1 and RO-I-2 need to invest from 3k€ to 3.1 k€ each as incremental TCO.

Ultra-dense V2I scenario	/	Deploying two new RSUs on top of the baseline scenario results in 24.5 k€ as TCO for RO (**non-collaborative model**). RO-A-1&2 can save from 2 to 4.2 k€ CAPEX and 1.3 to 2.8k€ OPEX (**passive network sharing**). Thus, RO-A-1 and RO-A-2 need to invest from 8.7k€ to 10.7 k€ each as incremental TCO. RO-A-1&2 can save from 4 to 4.2 k€ CAPEX and 2 to 2.8k€ OPEX (**active sharing**). Hence, RO-A-1 and RO-A-2 need to invest from 8.8k€ to 9.3 k€ each as incremental TCO.	/	Deploying two new RSUs on top of the baseline scenario results in 24.5 k€ as TCO for RO (**non-collaborative model**). RO-I-1&2 can save from 2 to 4.2 k€ CAPEX and 1.3 to 2.8k€ OPEX (**passive network sharing**). Thus, RO-I-1 and RO-I-2 need to invest from 8.7k€ to 10.7 k€ each as incremental TCO. RO-I-1&2 can save from 4 to 4.2 k€ CAPEX and 2 to 2.8k€ OPEX (**active sharing**). Hence, RO-I-1 and RO-I-2 need to invest from 8.8k€ to 9.3 k€ each as incremental TCO.
	MNO	**RO**	**MNO**	**RO**
Baseline V2N scenario-a	MNO-A-1: 142.4 k€ (**non-collaborative model**).	/	MNO-I-1: 142.4 k€ (**non-collaborative model**).	/
Baseline V2N scenario-b	MNO-A-1: 141 k€ (**non-collaborative model**).	/	MNO-I-1: 141 k€ (**non-collaborative model**).	/
Dense V2N scenario-a	MNO-A-1: 142.4 k€ (**non-collaborative model**).	/	MNO-I-1: 778.6k€ (**non-collaborative model**). MNO-I-1&2 can save up to 195 k€ cumulative	/

			TCO (**passive network sharing**) And save up to 272.5 k€ cumulative TCO (**active network sharing**)	
Dense V2N scenario-b	MNO-A-1: 141 k€ (**non-collaborative model**).	/	MNO-I-1: 724.7 k€ (**non-collaborative model**). MNO-I-1&2 can save up to 181.2 k€ cumulative TCO (**passive network sharing**) And save up to 253.6 k€ cumulative TCO (**active network sharing**)	/

After conducting a thorough analysis of the deployment cost for each of the seven studied scenarios, it has become evident that the collaborative business model is the most economically promising option. This model is particularly attractive when considering the implementation of passive and active network sharing strategies. In this model, operators pool their resources and work together, leveraging each other's strengths to provide 5G-based CCAM services. The cost of deployment can be significantly reduced, as the expenses are shared among multiple parties. Moreover, by adopting passive/active network sharing, companies can reduce the cost of equipment, maintenance, and operations, further increasing the profitability of this business model.

5.5 Conclusion and future steps

In this paper, a full techno-economic analysis of providing CCAM services in cross-border environments was elaborated, based on CLM service as a canonical case study. This analysis shows how challenging the cross-border environment is, due to first, the technical considerations such as network coverage issues and the need for new network technologies and implementations to fulfill the CCAM KPIs and second, to business issues such as feasible and innovative business models that enable cross-border collaborations to support the provision of the studied services. The proposed business models rely on the upcoming technologies within 5G networks namely network slicing and MEC. These business models were identified considering the two-phase roll out of 5G networks being the non-standalone and standalone deployment. The collaborative business model seems to be the most promising model based on MNO’s feedbacks but requires a review of the current regulations such as the limitation of the transmit power of base stations close to the border. The NSaaS (Network Slicing as a Service) business model needs a strong intervention from the telecommunication regulators to be realized in cross-border settings. Cost modelling results provide insights into the TCO of the different deployment scenarios for both V2I and V2N types of connectivity, which can be used later as an order of magnitude to derive the required funding by the EC to incentivize MNOs to densify their 5G deployment in the border regions. In addition, a cost allocation model was developed to link the different scenarios costs to the key stakeholders i.e., MNOs and ROs on the two sides of the considered border (the Brenner pass in our case) within the identified business models. Results of the allocation model show that adopting collaborative business model (i.e., with infrastructure sharing) allows the involved operators to save on the investment cost of the considered deployment scenario. The simulation of the deployment scenarios to be rollout at the same year can be considered as a limitation of this study. However, it could be worthwhile to investigate the effect of varying the deployment timeline based on CCAM adoption and examine its impact on TCO. In the future, a focus on revenue models will be necessary to complete the business case evaluation. Additionally, the study could be broadened to encompass CCAM in smart cities.

ACKNOWLEDGMENT

This work has been carried out in the frame of the 5GCARMEN project, which is partly funded by the European Commission (H2020 GA No.825012). The authors would like to thank all consortium partners for providing data used in this study.

References

[1] A. Caragliu, "Better together: Untapped potentials in Central Europe," Papers in Regional Science, vol. 101, no. 5, pp. 1051-1085, 2022.

[2] J. McCallum, "National borders matter: Canada-U.S. regional trade patterns," American Economic Review, vol. 85, pp. 615-623, 1995.

[3] Bottasso, A., & Sembenelli, A. (2001). Market power, productivity and the EU Single Market Program: Evidence from a panel of Italian firms. European Economic Review, 45(1), 167-186.

[4] J. M. Pereira, "5G for Connected and Automated Mobility (CAM) in Europe: Targeting Cross-Border Corridors," IEEE Network, vol. 35, no. 3, pp. 6-9, 2021.

[5] Chiha, A., Van der Wee, M., Colle, D., & Verbrugge, S. (2020). Network slicing cost allocation model. Journal of Network and Systems Management, 28, 627-659..

[6] 5GPPP report (2017), "5G innovations for new business opportunities,".

[7] Belgian Institute for Postal Services and Telecommunications (BIPT). (2020). "Evolution of mobile data in the Belgian mobile licensed spectrum and the impact on media presence".

[8] Statista, (2020). "5G services with the greatest commercial potential 2018,".

[9] European Commission (2016), "Cooperative, connected and automated mobility (CCAM)," [Online]. Available: https://ec.europa.eu/transport/themes/its/c-its_en. [Accessed 03 2021].

[10] European Commission, "Commission launches GEAR 2030 to boost competitiveness and growth in the automotive sector," 2016. [Online]. Available: https://ec.europa.eu/growth/content/commission-launches-gear-2030-boost-competitiveness-and-growth-automotive-sector-0_en. [Accessed 03 2021].

[11] European Commission, "Europe sets 5G investment agenda for Connected and Automated Mobility," 2020. [Online]. Available: https://digital-strategy.ec.europa.eu/en/news/europe-sets-5g-investment-agenda-connected-and-automated-mobility. [Accessed 03 2021].

[12] European Commission, "Connecting Europe Facility," [Online]. Available: https://ec.europa.eu/inea/connecting-europe-facility. [Accessed 03 2021].

[13] European Commission, "Connecting Europe Facility (CEF2) Digital," [Online]. Available: https://ec.europa.eu/digital-single-market/en/connecting-europe-facility-cef2-digital. [Accessed 03 2021].

[14] Usman, M., Qaraqe, M., Asghar, M. R., Gebremariam, A. A., Ansari, I. S., Granelli, F., & Abbasi, Q. H. (2020). A business and legislative perspective of V2X and mobility applications in 5G networks. IEEE Access, 8, 67426-67435.

[15] European Commission, "5G cross-border corridors," 2021. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/cross-border-corridors. [Accessed 03 2021].

[16] Campolo, C., Molinaro, A., Iera, A., Fontes, R. R., & Rothenberg, C. E. (2018, June). Towards 5G network slicing for the V2X ecosystem. In 2018 4th IEEE conference on network softwarization and workshops (NetSoft) (pp. 400-405). IEEE..

[17] 5GNORMA, "Deliverable D2.3: Evaluation architecture design and socio-economic analysis - final report," 2017.

[18] 5GCAR, "Deliverable D2.3 Automotive use cases and connectivity challenges, business models and Spectrum related aspects," 2019.

[19] Metsi. II, "Deliverable D1.1: Refined scenarios and requirements, consolidated use cases, and qualitative techno-economic feasibility assessment," 2016.

[20] 5GPP Automative working group, "A study on 5G V2X Deployment," Version 1.0, February 2018.

[21] 5GPP Automative working group, "Business Feasibility Study for 5G V2X Deployment," 2019.

[22] BEREC, "BEREC Report on Infrastructure Sharing," June 2018.

[23] 3GPP, "TR 22.951 version 10.0.0 Release 10".

[24] 5GPP H2020 Projects, "5G Trials for Cooperative, Connected and Automated Mobility along European 5G Cross-Border Corridors Challenges and Opportunities," 5G PPP, 2021.

[25] 5G-CARMEN Deliverable, "5G-CARMEN Preliminary Pilot Report: Deliverable D5.2 of 5G-CARMEN project," 5G-CARMEN, 2020.

[26] 5G-CARMEN, "Preliminary market analysis and business models : Deliverable D6.2 of 5G-CARMEN project," 5G-CARMEN project, 2020.

[27] Poli, F., Denis, B., Mannoni, V., Berg, V., Martín-Sacristán, D., García-Roger, D., & Monserrat, J. F. (2021, April). Evaluation of C-V2X sidelink for cooperative lane merging in a cross-border highway scenario. In 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring) (pp. 1-5). IEEE..

[28] V. Allee, "Value Network Analysis and value conversion of tangible and intangible assets," Journal of Intellectual Capital, vol. 9, no. 1, pp. 5-24, 2008.

[29] Concorda, "M31 Studies – Preliminary results available – A5.1 Business Modelling," [Online]. [Accessed 4 2020].

[30] 5G-CARMEN, "Deliverable D2.2 5G-CARMEN Preliminary System Architecture and Interfaces Specifications," 2019.

[31] Liu, G., Huang, Y., Chen, Z., Liu, L., Wang, Q., & Li, N. (2020). 5G deployment: Standalone vs. non-standalone from the operator

perspective. IEEE Communications Magazine, 58(11), 83-89..

[32] Malinverno, M., Avino, G., Casetti, C., Chiasserini, C. F., Malandrino, F., & Scarpina, S. (2019). Edge-based collision avoidance for vehicles and vulnerable users: An architecture based on MEC. IEEE vehicular technology magazine, 15(1), 27-35.

[33] Campolo, C., Molinaro, A., Iera, A., & Menichella, F. (2017). 5G network slicing for vehicle-to-everything services. IEEE Wireless Communications, 24(6), 38-45..

[34] Khan, H., Luoto, P., Bennis, M., & Latva-aho, M. (2018, May). On the application of network slicing for 5G-V2X. In European Wireless 2018; 24th European Wireless Conference (pp. 1-6). VDE.

[35] Zhang, H., Liu, N., Chu, X., Long, K., Aghvami, A. H., & Leung, V. C. (2017). Network slicing based 5G and future mobile networks: mobility, resource management, and challenges. IEEE communications magazine, 55(8), 138-145.

[36] ETSI, "5G; NR; Overall description; Stage-2 (3GPP TS 38.300 version 15.3.1 Release 15)," ETSI, 2018-10.

[37] 3GPP, "3GPP TS 23.501 V15.0.0 System Architecture for the 5G System (Stage 2)," (2017-12).

[38] marketsandmarkets, "network slicing market," [Online]. Available: https://www.marketsandmarkets.com/Market-Reports/network-slicing-market-120515704.html. [Accessed February 2021].

[39] Oughton, E. J., & Lehr, W. (2022). Surveying 5G techno-economic research to inform the evaluation of 6G wireless technologies. IEEE Access, 10, 25237-25257..

[40] itscosts, "Intelligent Transportation Systems Joint Program Office," [Online]. Available: https://www.itscosts.its.dot.gov/ITS/benecost.nsf/ID/CA235D82593B9B288525819F0050C9D1?OpenDocument&Query=CApp. [Accessed 13 7 2020].

[41] K. Casier, Techno-economics of FTTH deployment in the presence of competition (Doctoral dissertation), 2010.

[42] "Atkearney," [Online]. Available: https://www.atkearney.com/documents/20152/434204/Telco_Fitness.pdf/427dcc76-15f9-b9a3-0fd5-5c2e157918d6.

[43] statista, "Prices of electricity for the industry in Belgium from 2008 to 2018," [Online]. Available: https://www.statista.com/statistics/595775/electricity-industry-price-belgium/.

[44] Chiha Ep Harbi, A., Poli, F., Denis, B., Vannieuwenborg, F., Garcia-Roger, D., Monserrat, J. F., & Aigner, W. (2021). Techno-economic and simulation study of a V2I-based cooperative manoeuvring case in a cross-border scenario. In 94th IEEE Vehicular Technology Conference (VTC-Fall). IEEE.

[45] thinkmate, "Gigabyte h242-z10," [Online]. Available:

https://www.thinkmate.com/system/gigabyte-h242-z10. [Accessed 01 2021].

[46] ETSI, "Multi-access Edge Computing (MEC)," ETSI, [Online]. Available: https://www.etsi.org/technologies/multi-access-edge-computing. [Accessed 03 2021].

[47] 5GAA, "MEC for Automotive in Multi-Operator Scenarios," 5GAA Automotive Association, Munich , 2021.

6

Business Case Evaluation of Cooperative, Connected and Automated Mobility Service Provision in cross-border settings [INVITED PAPER – Project 5G-CARMEN]

Asma Chiha, Thibault Degrande, Sofie Verbrugge, Didier Colle , George Avdikos[18], Walter Aigner[19], Benoît Denis[20], and David Garcia-Roger[21]

Published in the proceedings of 2022 IEEE Future Networks World Forum.

In the previous chapter, Chapter 5, due to the lack of data regarding penetration rates of connected cars, we simulated different 5G deployment scenarios all starting in the same year as a reference year. However, predicting the penetration rate of connected cars is uncertain due to outdated studies and the impact of the Covid crisis on the automotive industry. To overcome this issue, we proposed a scenario-based approach with three different scenarios: low, medium, and high penetration rates. This approach provides more specific predictions and the techno-economic analysis using this approach is presented in this chapter.

***

[18] Eight Bells, Athens, Greece
[19] HiTec, Vienna, Austria
[20] CEA-Leti, Université Grenoble Alpes,Grenoble, France
[21] Universitat Politècnica de València, Valencia, Spain

Abstract— In this paper, we present a techno-economic analysis of providing a Cooperative, Connected and Automated Mobility (CCAM) use case, namely Cooperative Lane Merging (CLM), in a cross-border environment. Multiple network deployment scenarios are proposed to provide Vehicle to Infrastructure (V2I) connectivity with respect to PC5 Mode 4 – enabled RSUs. Total cost of Ownership (TCO) model together with four revenue models are developed to assess the viability of providing CCAM services in the studied settings. Results show that the higher the number of simultaneous connected cars, the higher the TCO of the required deployment needs to be to meet the defined KPIs and especially for the green field deployment with no existing fibre cable or electricity facilities. Another important insight from this analysis is that only with a high fleet penetration rate of connected vehicles, a viable business case can be achieved.

Keywords—TCO model, techno-economic analysis, 5G networks, C-V2X, CCAM services, revenue models, cross-border.

6.1 Introduction

Cooperative connected and automated mobility (CCAM) is one of the next big trends and necessities in the transportation industry. This became clear with the different European incentives and agendas with regards to connected mobility corridors in Europe [1], [2]. However, providing CCAM services at continental scale raises many challenges, amongst which the seamless cross-border continuity of CCAM services is the important one. The latter is not only challenging from a technical perspective but also from a business model one, due to the multi-stakeholders' nature of the CCAM ecosystem in the cross-border settings (including car and equipment manufacturers, road operators, mobile network operators, service providers, institutional entities...). To this end, the European Commission has funded several projects to research these challenges and the 5G-CARMEN project in which this study was carried out is one of them [3]. On the other hand, to realize the societal benefits CCAM promise, penetration of connectivity in passenger cars is essential. However, as for now, the automotive industry remains divided on the technology aspect, hampering adoption of CCAM in passenger cars [16]: "the stakeholders of the automotive industry (e.g., car manufacturers and road operators) are still skeptical about the capability of the telecom industry to take the lead in a market that has been dominated by dedicated intelligent transport systems (ITS) deployments.". Therefore, current adoption in new vehicles is significantly lower than initially predicted, and it is highly uncertain for upcoming years. Currently, it is increasingly considered as a societal issue to provide assistive systems immediately for all and not only for car owners of new cars who can afford increasing cost due to assistive services and advanced levels of automated driving. Yet, several telecom network operators have openly communicated, that investment into deploying the fifth generation of mobile networks (5G) infrastructure in Europe will be targeted towards industrial uses of

Internet of Things (IoT) and toward agglomerations with high population density. European rural areas and highways without big agglomerations or significant industrial IoT usage are openly rated as third level priorities in deploying 5G – even significantly later than 2030 or 2035. This low prioritization of connectivity deployment in highways especially next to country borders is justified by the high investment cost required especially for low latency-sensitive services such as CCAM services and for the unclarity around sustainable CCAM business models.

Therefore, the need for clear CCAM revenue models, that capture the revenue streams and show how key CCAM stakeholders can pay off their initial investment, is becoming more prominent. Hence, there is a need to shed more light on identifying revenue models for CCAM service provision since in the literature there are few papers, only discussing the investment costs and not the revenue part as such [4], [5],[6].

In this paper we perform a techno-economic analysis to evaluate the provision of a CCAM use case namely Cooperative Lane Merging (CLM) in the Brenner pass corridor considering the use of Vehicle-to-Infrastructure (V2I) connectivity with respect to Road Side Units (RSUs) based on the Cellular-V2X (C-V2X) sidelink (i.e., PC5-Mode4). For this purpose, we first identified three deployment scenarios with regards to different fleet penetration rates and developed a Total Cost of Ownership (TCO) model to study the investment cost for these three deployment scenarios (discussed in section 6.2). In section 6.3, we proposed four different revenue models to evaluate the viability of the business case for the different penetration rates.

6.2 CCAM deployment scenarios and Total cost of ownership

We study the CLM use case detailed in [7], where one vehicle intends to move onto a side lane containing already some vehicles. Hence, these side vehicles should create a gap as such the initiator vehicle can insert between them. This manoeuvre is centralized because a centralized manoeuvring application running in the Multi-Access Edge Computing (MEC) or a local server (i.e., on local RSU). The execution of the centralized CLM is based on the collection of the positional information of the different vehicles. In our specific operating context, the CLM procedure can be initiated and completed before, while or after crossing the border. The highway segment considered in this study is the Brenner pass between Austria and Italy of approximately 4.5 km of length [5].

6.2.1 **Penetration rate assumptions**

It would be unrealistic to assume a 100% of penetration rate of connected cars from the starting year of the deployment. However, there is no common agreement within the research and concerned communities about convergent forecasts of the connected car penetration rate in the upcoming years. There are outdated studies with findings ranging from pessimistic [8] to optimistic [9] penetration rates which cannot be used as such since the automotive industry was hardly hit by the

Covid crisis [10]. One way out is to follow two different approaches in overcoming this issue:

1. Generic and abstract approach based on a variation of penetration rates:

In this approach, we work with a variation of penetration rates which are not linked to specific starting year, i.e., working with abstracted years (year 1 to 10) and hence give generic results that road operators or CCAM providers can use and extrapolate when a clearer idea about the penetration rates and their evolution over time can be conducted. The results of the techno-economic analysis using this approach have been published in [5].

2. Assuming three penetration rate scenarios:

In this approach we define three different scenarios: Low, Medium and High scenarios in line with the Strategic Deployment Agenda (SDA) study [15]. The two first forecasts namely the low and medium scenarios are taken from the CEDR's MANTRA project [8], [15]. For these two scenarios they defined "Low scenario" as the "business as usual" scenario, where the automated driving use cases are taken into use as in usual market economy, utilizing solutions based on the utility or economic value to the customer or user. In this low scenario, the fleet penetration rate varies from 1.19% (2021) to 1.30% (2030). Our Medium scenario is the "High scenario" in the MANTRA study where it is assumed an acceleration of automated driving use cases via financial incentives such as reduced taxation or via regulatory actions, for instance by mandating automated driving in specific conditions. In the MANTRA study, they started the forecasts from year 2030, and in this study, we consider 2021 as the starting year of the deployment hence we consider a 1% yearly increase in sales for vehicles equipped with V2X technologies to predict the missing figures. The penetration rate hence varies from 6.03% (2021) to 6.60% (2030). The High scenario is the most optimistic scenario and is taken from IHS Mark-It study [11], where fleet penetration rate is rising from 6.82% (2021) to 39.12% (2030).

6.2.2 Network deployment scenarios and resulting TCO

For the V2I links with respect to PC5 Mode4 – enabled RSUs, three deployment scenarios have been defined. These three deployment scenarios are as follows:

- The baseline V2I scenario corresponds to the physical deployment currently available in the studied corridor: 2 RSUs in total in the test portion of highway of length ~ 4.5 km (i.e., ~ 0.5 RSU/km).
- The dense V2I deployment scenario: this scenario corresponds to a doubled baseline density: 4 RSUs in total in the same environment (i.e., ~ 1 RSU/km).
- The ultra-dense V2I deployment scenario: this scenario corresponds to a quadrupled baseline density: 8 RSUs in total in the same environment (i.e., ~ 2 RSUs/km).

Note that the names attributed to these different scenarios are defined based on the densification of RSUs comparing to the baseline V2I scenario.
The TCO model as well as input data used to run the cost model are described in detail in [5].
Based on the number of simultaneous connected users/cars the required deployment for each fleet penetration rate scenarios are derived to fulfil the service availability KPI depicted in Table 6-1. For the low scenario a low deployment scenario is defined where 2 RUSs are installed in the first two years (for the green field deployment/ and maintain the existing 2 RSUs for the Brenner pass corridor i.e., case of pre-equipped field) and by the third year an additional 2 RSUs are installed to ensure a low rejected CLC request rate. Yet, for the medium scenario a medium deployment scenario is considered where from year one we deploy 4 RSUs (4 new RSUs in the green field scenarios, but in the pre-equipped where we have already 2 existing RSUs we add 2 additional ones). To mitigate the high traffic generated by the high number of connected cars in the high fleet penetration rate scenario 4 RSUs are needed in the first year but by the second year 4 new RSUs are required to have 8 RSUs in total (considering the existing infrastructure, 2 RSUs in the pre-equipped field scenario).

Table 6-1 The rate of rejected CLM requests based on the variation of the connected cars for the three infrastructure density scenarios derived by means of system-level simulations [5]

Fleet penetration rate	**RSU deployment**		
	2 RSUs	**4 RSUs**	**8 RSUs**
5%	14.9%	0%	0%
10%	17.1%	3.4%	0.1%
50%	21.3%	5.2%	0.5%
100%	28%	6.4%	0.7%

The cumulative TCO per km (CUM_TCO_perKM) for the three deployment scenarios low, medium and high for two different type of fields i.e. green field and pre-equipped field are depicted in Figure 6-1 and Figure 6-2 respectively.

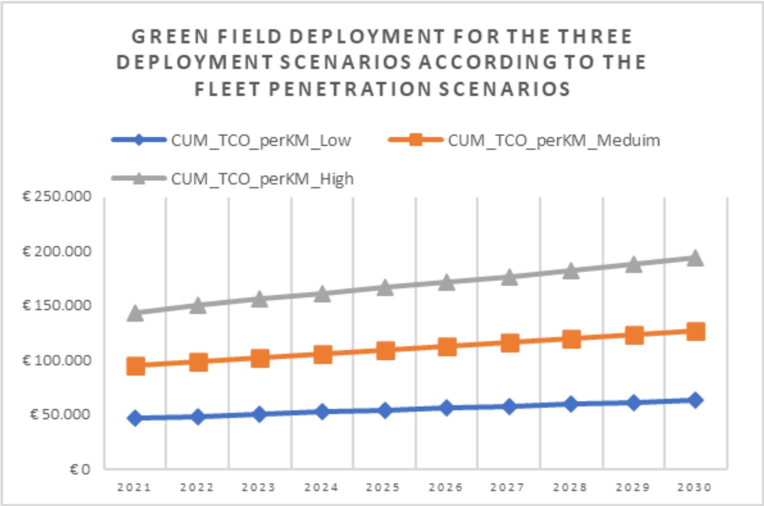


Figure 6-1: Cumulative TCO per km for the low, medium, and high deployment scenarios for a green field corridor segment

As expected, the higher the number of simultaneous connected cars, the higher the TCO of the required deployment needs to be to meet the defined KPIs. Results show how expensive it is to deploy in a green field of the highway where no existing fibre cable or electricity facilities. The green field deployment is three times more costly than the pre-equipped field deployment.

6.3 Proposed Revenue models

To evaluate the economic viability of the proposed CCAM services in the cross-border environment, revenue assumptions have to be taken. We hereafter work with four revenue models reflecting four different assumptions, aiming at giving as much as possible of useful insights into the viability of the use case using different revenue models for CCAM connectivity providers in the cross-border area.

1. **Revenue model 1:** this model uses the assumption taken from the 5GPPP automotive Working Group white paper [12], where they assume a revenue of 0.5 euro per user for each use of the highway segment (which has a length of 100 km). This connectivity fee can be collected via the toll fees on the highway or via another approach depending on the operator providing the connectivity. For our case we divide this 0.5 euro by 100 to derive the revenue per user per use per km.
2. **Revenue model 2:** this model assumes that drivers of cars supporting V2X connectivity would pay a yearly subscription of 250 euro (for Europe-wide V2X services). This payment is done when buying new car such there is a

lumpsum payment (for the car) and a yearly fee to benefit from the V2X services. To have a unified metric to compare between the different revenue models, we derive here as well a revenue per km using the average driven distance in Europe per year, being 12,000 km [13]. Connected and assisted driving functions significantly alter how driving on highways is perceived. It is not lost travel time but for a significant part gained convenience with the safe and legal capability to communicate. Assuming that out of average driven 12000 kilometres in the first years only a rather small fraction of 30 per cent would be driven under full V2X connectivity supporting cars on highways; we conservatively calculated that in distressing lower speed conditions the support functions would be on for most of the time (some 80 per cent) and this would rapidly amount to some 50 hours significantly improved convenience for the driver. So, if we anticipate Euro 5 per hour for less distressing stop and go traffic or rather boring highway travel this would lead to an estimated Euro 250 per year. Populations who drive significantly higher distances in new upper-class cars would even have more attractive fictitious average rate per convenient hour.

3. **Revenue model 3:** since we are studying the cross-border segment (the Brenner pass) where an extra-investment is needed to ensure the CCAM service continuity, it is reasonable to assume that in order to pay-off this extra-investment we need to charge more for this specific segment in the toll fees. Therefore, in this model we assume that the revenue of 0.5 euro per user for each use of the corridor segment (of revenue model 1) is for the 4.5 km portion of the cross-border segment.
4. **Revenue model 4:** in this model a cost-based pricing is considered. Starting from the cost of the deployment needed and an Average Cost Per User (ACPU) per km is calculated, thus adding a profit margin on top results in an Average Revenue Per User (ARPU) per km. Usually the margin profit of Telco's ranges from 10% to 30% [14]. Using this revenue model, we guarantee to break-even before the end of the project period (10 years in this study).

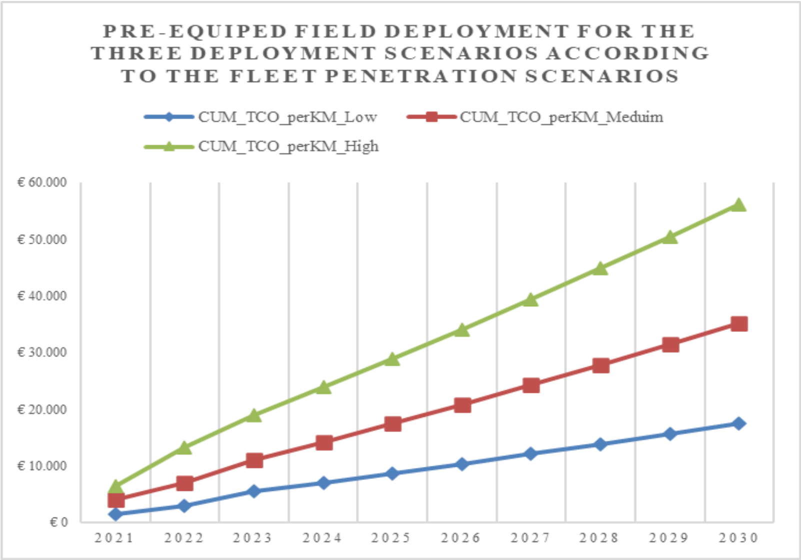


Figure 6-2: Cumulative TCO per km for the low, medium, and high deployment scenarios for a pre-equipped field corridor segment

We run the simulation for these 4 revenue models considering all the different parameters and taking into account the three fleet penetration rate scenarios presented in section 6.2, in order to derive the cumulative TCO per deployment scenario (low fleet penetration scenario results in a low deployment scenario and so on..) and also to calculate the number of users using the service to derive the cumulative revenue per model and per fleet penetration scenario.

For the deployment scenarios we have, as usual, the two alternatives: deploying on a green field segment of the corridor or on a pre-equipped one.

For all the scenarios we assume that:

- There is a way to collect the revenues from only the connected users (those using the V2X connectivity in the studied segment) not all the users on the road.
- For revenue model 4 we assume a profit margin of 10%.
- Since in the revenue model 4 the ACPU is used to derive the cumulative revenue, we can only compare the cumulative revenue for the green field deployment to the cumulative TCO of the green field deployment and the same applies for the pre-equipped field. Yet, for the other revenue models they can be compared to both deployment TCO.

Starting with the low fleet penetration scenario where we assume that the cumulative TCO per km used is the one of the low deployment scenarios for the green and pre-equipped field as presented in section 6.2.2.

Figure 6-3 presents the cumulative revenue per km of all the revenue models as well as the cumulative TCO per km of the two low deployment scenarios (green/pr-equipped field). Results show that using revenue models 1, 2 and 3 would not allow to break-even neither during the project lifetime nor even in the near future if the fleet penetration rate remains low. However, using revenue model 4, we break-even around the fourth year for the pre-equipped field deployment and around year 8 for the green-field deployment.

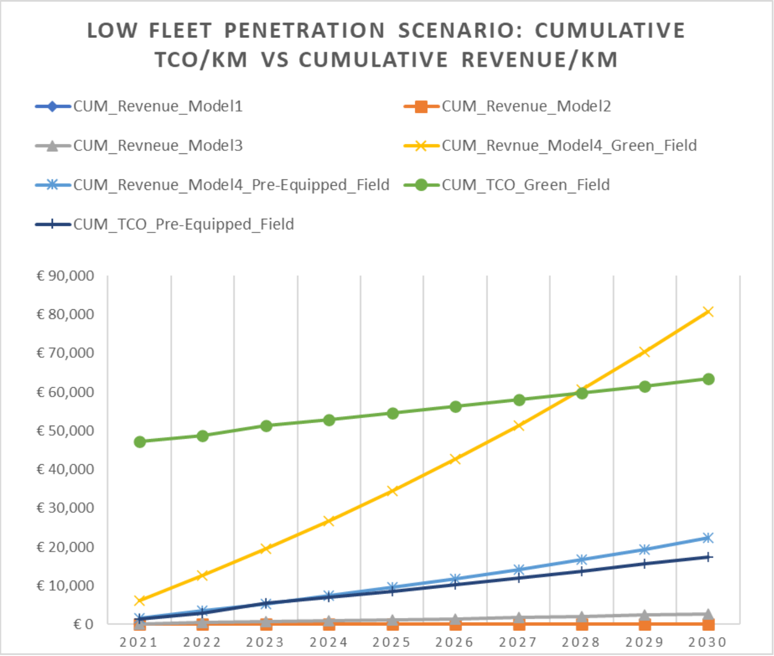


Figure 6-3: Low fleet penetration scenario: Cumulative TCO/km VS Cumulative Revenue/km for the four revenue models

For the medium fleet penetration scenario, we assume that the cumulative TCO per km used is the one of the medium deployment scenarios for the green and pre-equipped field as presented in section 6.2.2.

Similar to the low scenario, results presented in Figure 6-4, which presents the cumulative revenue per km of all the revenue models as well as the cumulative TCO per km of the two medium deployment scenarios (green/pr-equipped field), show that using revenue model 1, 2 and 3 would not allow to break-even during the project lifetime and even in the near future if the fleet penetration rate remains medium, except for revenue model 3 which could break even around year 13 to 14 (2033 to 2034). However, using revenue model 4, we break-even around the fifth year for the pre-equipped field deployment which takes longer than the low fleet scenario and this is explained by the fact that the yearly increase of fleet penetration (which drives the revenue) is very low to cover for the additional required deployment which is double the required deployment for the low fleet penetration scenario. For the green-filed deployment scenario the break-even for the medium scenario is reached around year 8 for.

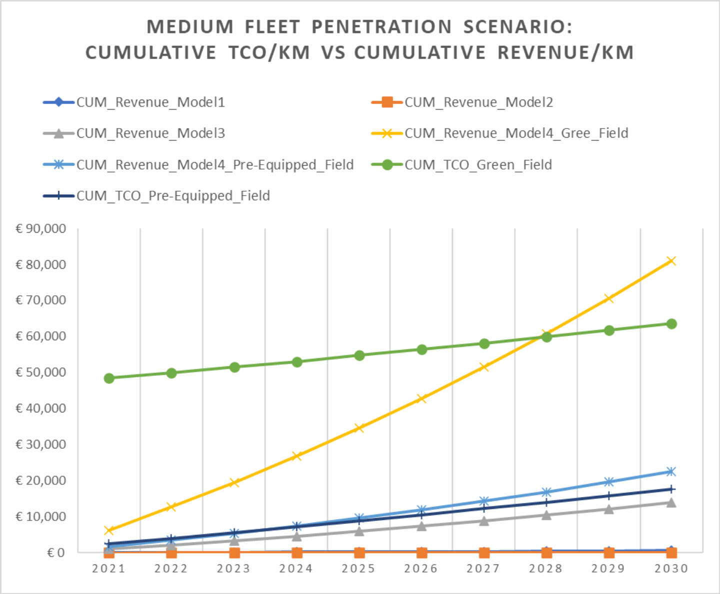


Figure 6-4: Medium fleet penetration scenario: Cumulative TCO/km VS Cumulative Revenue/km for the four revenue models

For the high fleet penetration scenario, we assume that the cumulative TCO per Km used is the one of the high deployment scenarios for the green and pre-equipped field as presented in section 6.2.2.

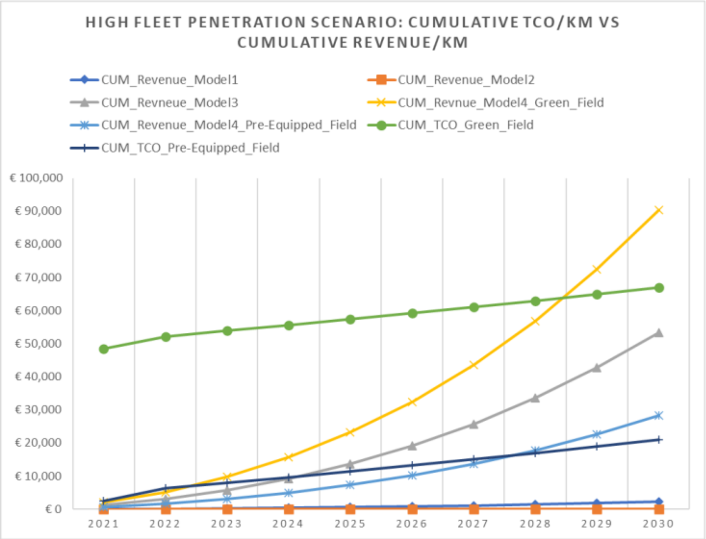


Figure 6-5: High fleet penetration scenario: Cumulative TCO/km VS Cumulative Revenue/km for the four revenue models

Figure 6-5 shows that even with high penetration rate (which means high number of users) revenue models 1 and 2 cannot pay-off the investment cost. But, for the revenue model 3 and unlike the low and medium scenarios, the break-even point can be reached around the fourth year (2024) for the pre-equipped field deployment and around year 2031/2032 for the green-field deployment.

However, for the revenue model 4, we see that the break-even point is reached later compared to the low and medium scenario, around year 8 for the pre-equipped field deployment (comparing to year 4 in the previous two scenarios) and between year 8 and 9 for the green field deployment (comparing to only at year 8 in the previous scenarios). This can be explained by the fact that the revenue model 4 is based on the average cost per user which is very high in the high deployment scenario comparing to the low and medium ones with a low profit margin of only 10%. With higher profit margin, higher revenues can be reached as shown in Figure 6-6.

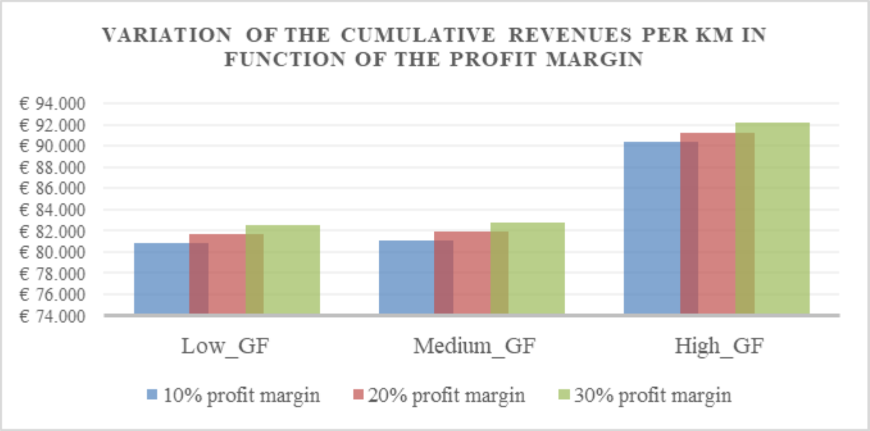


Figure 6-6: Variation of the Cumulative revenues per km in function of the profit margin for the three fleet penetration scenarios with a green-field deployment

If we assume that the revenue per user is charged as part of the toll fees, it is difficult to differentiate between road users who are using the V2X connectivity from those who are not. Therefore, we used our model to derive the ARPU for the three penetration scenarios, low, medium, and High scenarios (which implies low, medium, and high infrastructure deployment) while counting only the connected users and then while counting all the road users (users of the corridor segment under study). The results of this comparison are presented in Figure 6-7 and Figure 6-8.

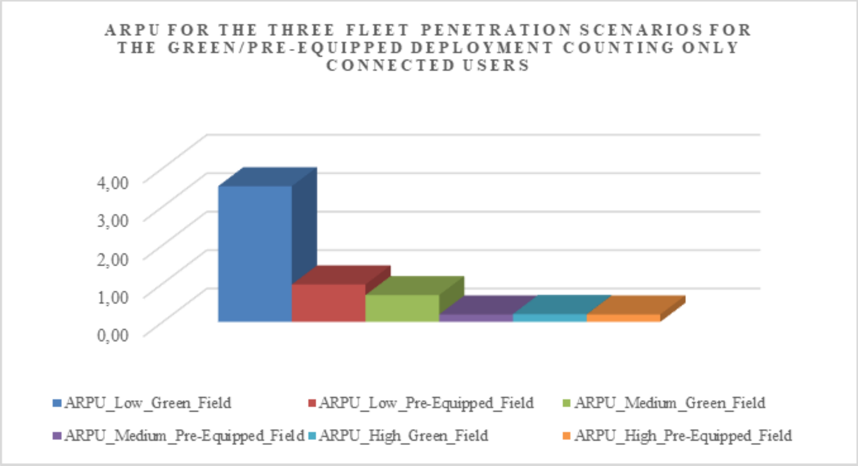


Figure 6-7 : ARPU for the three fleet penetration scenarios for the green/pre-equipped deployment counting only connected users

It is important to mention that the two figures Figure 6-7 and Figure 6-8 have different Y axis since the ARPU in the case where all the corridor segment users are taken into account are way too small (in the order of 4 Euro-cents for the green filed deployment) comparing to the case where only the connected users are counted (3.5 Euro for low penetration rate in a green field deployment scenario).

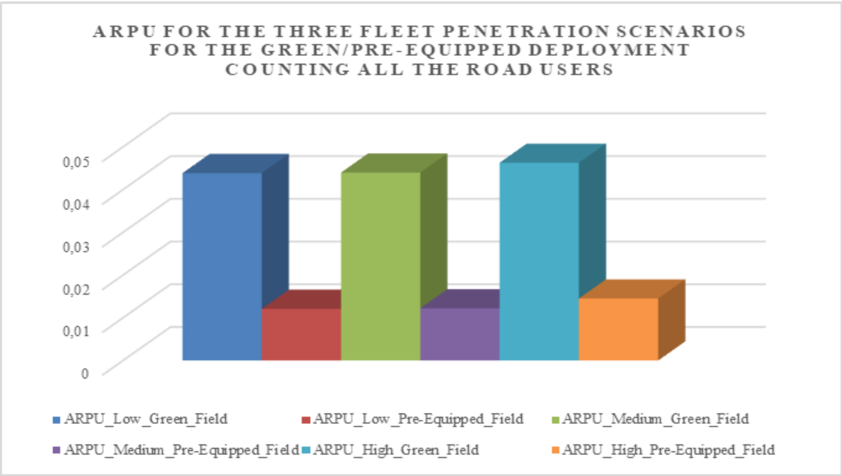


Figure 6-8 : ARPU for the three fleet penetration scenarios for the green/pre-equipped deployment counting all the road users

Results also show a comparison between the ARPU/km for a green field versus pre-equipped field deployment, which demonstrate how low is the ARPU (which is based on the ACPU) that needs to be collected to pay-off the deployment cost in a pre-equipped (brown) field comparing to a green field one.
In order to have a clear idea how counting all corridor segment users can reduce the connectivity fees to be charged to the road users, we take the example of the low penetration scenario: for the green field with counting only connected users the fee requested per user for the 4.5 km of the Brenner pass is 4.5*3.5=15.75 Euro versus only 4.5*0.0438= 0.197 Euro (19 Euro-cents) if we consider all the corridor segment users. For the high fleet penetration scenario for pre-equipped field deployment, we need to request 4.5*0.19=0.855 Euro (85 Eurocents) when only connected user are considered versus only 4.5*0.0145=0.0652 Euro (6.5 Eurocents) when all users are counted.
Hence it seems reasonable to incorporate the connectivity fee in the toll fees while distributing it among all users rather than charge it directly to the connected users.
The main takeaways from this exercise are:

- For the low and medium fleet penetration scenarios, only revenue model 4 (the one based on the ACPU) guarantee to break-even before the end of the project lifetime.
- For the high fleet penetration scenario and using the revenue model 3 the break-even point can be reached around the fourth year (2024) for the

pre-equipped field deployment and around year 2031/2032 for the green-field deployment (after the project lifetime), but models 1 and 2 are always non-viable even with high penetration rate. Yet using model 4 guarantee to break-even before the end of the project for the two types of deployment.

- Using a high profit margin in revenue model 4 help collecting more revenues but on the expenses of the end user since a higher ARPU will be charged.
- It seems realistic/feasible to incorporate the connectivity fee in the toll fees while distributing it among all users rather than charge it directly to the connected users.

6.4 Conclusion

In this paper we performed a techno-economic analysis of providing CLM use case in a cross-border environment. The network deployment cost of three deployment scenarios have been discussed with respect to the service availability KPI. Results showed that the higher the number of simultaneous connected cars present, the higher the TCO of the required deployment is, and discussed how expensive it is to deploy in a green field of the highway where no existing fibre cable or electricity facilities. The green field deployment is three times more costly than the pre-equipped field deployment. We proposed four revenue models reflecting four different assumptions, aiming at giving as much as possible of useful insights into the viability of the use case using different revenue models for CCAM connectivity providers in the cross-border area. Findings suggested the use of a cost-based revenue model to ensure reaching a fast break-even point. A future research topic might be to study the incorporation of the connectivity fee in the toll fees to distribute it among all users rather than charge it directly to the connected users. Finally, our future work will focus on the study of the use of Vehicle to Network (V2N) connectivity link instead of V2I one.

ACKNOWLEDGMENT
This work has been carried out in the frame of the 5G-CARMEN project [3], which is partly funded by the European Commission (H2020 GA No.825012).

References

[1] "Digital Single Market program of the European Commission [Online]," Available: https://ec.europa.eu/digital-single-market/en/, access: 2022-07-19.

[2] "Connecting Europe Facility" [Online], https://ec.europa.eu/inea/en/connecting-europe-facility, access: 2022-07-19.

[3] "5G-CARMEN [Online]," Available: https://5gcarmen.eu/, access: 2022-07-01.

[4] G., L., F., J. L., Hicks, D., Mercer, M., & Thompson, K. (2018). Intelligent transportation systems benefits, costs, and lessons learned: 2018 update report (No. FHWA-JPO-18-641). United States. Dept. of Transportation. ITS Joint Program Office.

[5] A. C. Ep Harbi et al., "Techno-economic and Simulation Study of a V2I-based Cooperative Manoeuvring Case in a Cross-border Scenario," 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall), 2021, pp. 1-7, doi: 10.1109/VTC2021-Fall52928.2021.9625407.

[6] D., T, et al. "C-ITS road-side unit deployment on highways with ITS road-side systems: A techno-economic approach." IET Intelligent Transport Systems 15.7 (2021): 863-874.

[7] F. Visintainer et al., "5G-CARMEN Use Cases and Requirements," Deliverable D2.1 of 5G-CARMEN project, May 2019.

[8] W. Aigner et al., MANTRA project deliverable: "Vehicle fleet penetrations and ODD coverage of NRA relevant automation functions up to 2040" [Online]," Available: https://www.mantra-research.eu/wp-content/uploads/2020/03/MANTRA_Deliverable_D2.1_1.0.pdf", access: 2022-07-19

[9] 5G-NORMA, "Deliverable D2.3 Evaluation architecture design and socio-economic analysis - final report", Dec. 2017, access: 2022-07-19

[10] COVID-19: "Automotive industry signs joint Code of Business Conduct to support re-start of production [Online]," Available: https://www.acea.auto/press-release/covid-19-automotive-industry-signs-joint-code-of-business-conduct-to-support-re-start-of-productionC

[11] EETAsia, "BEVs, Connected Cars & Software?", February 2021. URL: https://tinyurl.com/3ptknrrt/, access: 2022-07-19

[12] A. Laya, K. Manolakis, G. Vélez, M. Fallgren, S. He, J. Favaro, P. Syros, M. Paolino, B. Sayadi, B. Altman, M. Gharba, M. Dillinger, L. Gomes Baltar, J. F. Monserrat, "Business Feasibility Study for 5G V2X Deployment". 5G PPP Automotive Working Group, Version 1, February 2019

[13] "CHANGE IN DISTANCE TRAVELLED BY CAR" [Online],Available: https://www.odyssee-mure.eu/publications/efficiency-by-sector/transport/distance-travelled-by-car.html, access: 2022-07-19

[14] Investopedia (2018). average-profit-margin-company-telecommunications. Retrieved from https://www.investopedia.com/ask/answers/060215/what-averageprofit-margin-company-telecommunications-sector.asp, access: 2022-07-19

[15] G. Avdikos et al., "5G-CARMEN deliverable D6.4: Connected, Cooperative, and Automated Mobility: Market analysis, go-to-market strategy, and business models" [Online], Available:" https://5gcarmen.eu/, access: 2022-07-01.

[16] F. Bouali et al., "5G for Vehicular Use Cases: Analysis of Technical Requirements, Value Propositions and Outlook," in IEEE Open Journal of Intelligent Transportation Systems, vol. 2, pp. 73-96, 2021, doi: 10.1109/OJITS.2021.3072220.

7

Techno-Economic Analysis of MEC Clustering Models for Seamless CCAM Service Provision

Asma Chiha, Benoît Denis[22], Sofie Verbrugge, and Didier Colle

Published in IEEE Communications Magazine.

The MEC paradigm is crucial in the 5G network architecture for minimizing E2E latency when providing CCAM services. Nonetheless, MNOs need to assess various deployment models carefully to achieve optimal MEC deployment in terms of both E2E latency and cost. In this chapter, we propose several MEC deployment models and evaluate them from both E2E latency and cost perspectives. Our research outcomes can guide MNOs in making informed decisions about MEC technology deployment strategies.

***

[22] CEA-Leti, Université Grenoble Alpes,Grenoble, France

Abstract—The latency requirements of delay-sensitive applications such as Cooperative, Connected and Automated Mobility (CCAM) services challenge the capabilities of traditional vehicular radio access technologies, i.e., IEEE 802.11p and cellular networks. To this end, the fifth-generation (5G) cellular network is adopting the Multi-access Edge Computing (MEC) paradigm. Yet, the use of this technology comes with several challenges. In this paper, MEC placement challenges and their impact on network deployment costs are studied. We propose three MEC clustering models and compare them from a cost perspective. Results show that all the MEC clustering models outperform the non-clustering approach. In addition, the conditions in which specific clustering models yield the most optimal results are analyzed. Results aim at providing insights into cost-effective MEC deployment models.

***Index Terms*—5G mobile networks, CCAM services, clustering models, TCO cost model, MEC, MEC placement, techno-economics.**

7.1 Introduction

Multi-access Edge Computing (MEC), Software-Defined Networking (SDN), and Network Functions Virtualization (NFV) are seen as the key enabler technologies for meeting the fifth-generation (5G) promised Key Performance Indicators (KPIs) [1]. High capacity, low latency, and high reliability are among the requirements needed to provide the three famous 5G use cases: ultra-Reliable Low Latency Communications (uRLLC), enhanced Mobile Broadband (eMBB), and massive Machine Type Communication (mMTC). One sector that has strict requirements concerning the abovementioned KPIs is the automotive sector, and even more so specifically when it comes to Cooperative, Connected and Automated Mobility (CCAM) use cases. In particular, CCAM safety-related services are delay-sensitive applications [2]. MEC technology, with its capacity to reduce latency by hosting services close to users and data generators, is considered essential for providing CCAM services [2].

The European Telecommunications Standards Institute (ETSI) stated that “MEC offers application developers and content providers cloud-computing capabilities and an IT service environment at the edge of the network. This environment is characterized by ultra-low latency and high bandwidth as well as real-time access to radio network information that can be leveraged by applications” [3].

Furthermore, according to the 5G Automotive Association (5GAA), “MEC technology is substantial for the realization of the V2X use cases that require low latency on the one hand and are expected to increasingly generate data that can be collected and processed closer to the user on the other hand” [4].

Within the scope of the 5G-CARMEN project, several CCAM use cases require low latency and high reliability, such as cooperative and automated lane-change maneuvers and cooperative and automated in-lane maneuvers. The latter have been investigated in a more specific cross-border context, where roaming will introduce a significant additional latency [5]. Therefore, the developed architecture relies on the use of MEC technology to meet these KPIs [6].

Although the MEC paradigm puts computation and storage capabilities close to the edge or a Radio Access Network (RAN), suitable MEC placement within the network topology raises significant challenges for Mobile Network Operators (MNOs). From a cost perspective, a transition from centralized hosting models to distributed/decentralized models results in significant deployment costs. From a technical perspective, a highly distributed MEC deployment (i.e., next to each cell tower) introduces security and technical challenges, such as the need for dynamic service placement optimization and control of the impact on the latency KPI due to the required inter-MEC switches. Therefore, a trade-off between latency and cost needs to be reached to benefit from the potential of this new technology while keeping the cost of network investment as low as possible.

Numerous papers in the literature have studied the MEC paradigm adoption to support delay-sensitive services from different perspectives, such as challenges and opportunities in [1] and [2], and dynamic service placement and migration between MEC servers for all types of services in [7] and CCAM services in [8]. However, there is a need to shed more light on the MEC placement issue from a cost perspective while fulfilling required CCAM service KPIs.

In this paper, we develop a Total Cost of Ownership (TCO) model to assess the impact of MEC placement at different yet fixed levels within the network topology on the network investment cost to support 5G-CARMEN use cases.

Furthermore, to find a trade-off between CCAM-required latency and the cost of MEC deployment, given certain parameters like the existence of certain network infrastructure or the possibility of active/passive network sharing, we propose three MEC clustering models. Concretely, we propose two variants of the star topology model and the collapsed star topology model. We use the developed TCO model to compare these three proposed MEC clustering methods. In addition, since the star topology model requires links between the MEC node and the next-generation Node-B (gNB) under the same cluster, an optimization algorithm is developed to find the optimal MEC node position that minimizes the total required fiber length.

7.2 Total Cost of Ownership model

7.2.1 Structure of the generic Total Cost of Ownership model

The generic TCO model structure for MEC deployment is presented in Figure 7-1. The proposed model considers both the Capital Expenditures (CAPEX) and Operational Expenditures (OPEX) as well as overhead costs such as marketing and helpdesk costs. The main inputs of the model are related to the costs of the hardware equipment. More specifically, these costs concern the MEC servers, the cabinets to host the MEC hardware, the cooling system, the fiber backhaul, and the cost of renting sites to host MEC nodes. These inputs feed into the CAPEX calculation where costs of hardware as well as installation costs are considered. The OPEX, which represent the costs to keep the network operational, consist of site rental, power consumption, and hardware maintenance. The cost of software licensing is excluded from the calculation due to a lack of inputs. Next, overhead costs, which cover marketing, helpdesk services, human resources, finance, etc., are approximated, based on the work of Casier [9], by adding 20 percent to the sum of the CAPEX and OPEX. Finally, summing up CAPEX, OPEX, and overhead costs results in the TCO of the MEC deployment. It is worth mentioning that the application of this model is case-specific. In some cases, we do not need to deploy new fiber to connect the MEC node to the rest of the network since a fiber backhaul is already in place. In other cases, there is enough space next to the tower (gNB) to host the cabinet, so there is no need to rent new sites to deploy the MEC nodes. In addition, based on the placement of the MEC nodes in the network, the need for a cabinet and a new cooling system is to be decided.

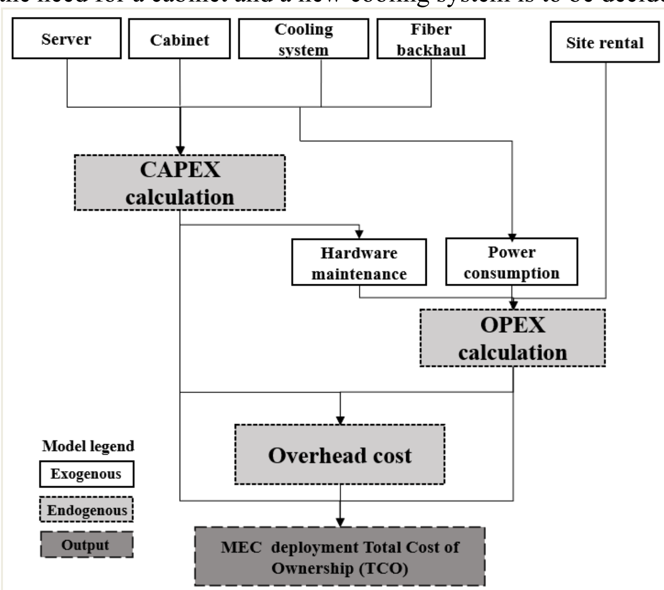


Figure 7-1 MEC cost model structure

7.2.2 Application of the TCO model to different MEC placements within the network architecture

Given a generic network architecture, as presented in Figure 7-2, we studied the cost of placing MEC nodes at the different levels of the network: at the MNO data centers (DCs), at the core transmission sites, at the radio aggregation sites, or at the RANs. We applied the cost model to the entire Bologna-Munich corridor case, a corridor under study in the 5G-CARMEN project that covers 600 kilometers (km) and spans three countries. Assuming an average inter-gNB distance of 3 km, we have around 200 gNBs in total installed next to the corridor. Since three

countries are involved, three DC sites operated by three different MNOs are considered. In the reference mobile network architecture shown in Figure 7-2, which has been put forward in collaboration with two MNOs active in the 5G-CARMEN project, there are 100 aggregation sites, which corresponds to almost one for every second radio site. For every ten aggregation sites, there is one core transmission site, which yields ten for the network in total. We considered the use of one server per MEC node when the node is installed at a radio or aggregation site. In case the MEC node is installed at the core transmission site or the DC, we consider the use of two servers per MEC node, assuming that spare DC processing resources can be used to support the processing of peaks in traffic.
We used the following costs to run the proposed TCO model: the cost of a MEC edge server of type "gigabyte-h242-z10" is €10,000 [10]; cooling system and cabinet costs for the MEC deployment are around €2,200 and €2,500, respectively [11]; maintenance of mobile network equipment is assumed to be 10 percent of the original CAPEX [8], and the price of electricity is €0.1085 per kWh [12].

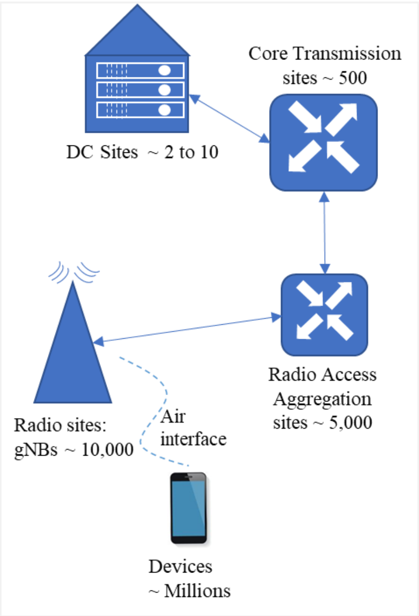


Figure 7-2 Reference mobile network architecture

Table 7-1 shows a breakdown of the cost per cost type related to a MEC deployment over a period of 10 years for each MEC location option. Note that the cost results presented in Table I do not include the cost of fiber cables that link the RANs to the aggregation sites. Even though this cost has been excluded, which represents a significant cost due to the cable installation cost, deploying MEC at the RANs is the costliest option. We refer to the latter deployment option in the rest of the paper as the non-clustering MEC deployment. Yet, moving the MEC placements up in the hierarchy of the network topology, i.e., closer to the DC, significantly decreases the cost of the MEC deployment due to the lower number of required MEC nodes. The latter impacts the end-to-end (E2E) latency for the CCAM services, however. By placing the MEC nodes higher up in the network topology, multiple hops are introduced between the connected vehicle and the CCAM application instance, which is hosted at the MEC server. Therefore, the round-trip time is affected, hence the need to find a relevant balance between targeted latency and network deployment costs.

Nonetheless, rather than choosing one location type, one solution could consist of performing MEC clustering, where one MEC node serves multiple gNBs at the same time within the so-called cluster. Different clustering models are identified and detailed in the next section.

Table 7-1 Different MEC placement costs

Item	Unit	Radio sites	Radio access aggregation sites	Core transmission sites	Core data center sites
Number of sites	#	200	100	10	3
MEC CAPEX	€k	2,940	1,220	244	74
Cumulative MEC OPEX over 10 years	€k	3,070	1,285	295	89
Cumulative overhead costs over 10 years	€k	1,202	501	108	33
MEC TCO (10 years)	€k	7,212	3,006	647	195

7.3 Proposed MEC clustering models

In addition to representing a costly solution, installing MEC nodes at each tower has other disadvantages. First, it introduces security challenges, among which are the need for physically securing each MEC node on the corridor and securing it against intrusions at the software level. The exact required security measures in a distributed deployment model are still unknown since MEC is currently being deployed at MNO DCs. Accordingly, the related additional costs are still unknown. Second, installing MEC nodes at each tower has an impact on the E2E latency because, if we increase the number of MEC nodes with a short inter-distance, latency increases since switching from one MEC to another requires the transfer of service-related data and sessions, which induces new latency of roughly 1–2 milliseconds (ms) per MEC switch. Therefore, MEC clustering could be viewed as a solution that considers cost, security, and latency requirements. But the question of how to define a MEC cluster still holds.

7.3.1 Star topology model

In this model, the MEC node is situated at the same distance from all the gNBs within the MEC cluster and connected directly point to point with the gNBs with two possible options:

- Option 1: The number of gNBs within the same MEC cluster is different from the number of gNBs that are served by the radio access aggregation site, so hosting MEC at the aggregation site will not make sense. In this case, we need to lease a new site between the aggregation site and the radio access sites to host the MEC node, as illustrated in the top left-hand side of Figure 7-3. In addition, installing fiber cable is needed to connect each tower to the MEC node.
- Option 2: The number of gNBs within the same MEC cluster is the same as the number of gNBs that are served by the radio access aggregation site, so it can be placed at the radio access aggregation site. This option is depicted in the top right-hand side of Figure 7-3.

In both cases, we have to consider the cost of linking the MEC nodes to each other with fiber cable to ensure inter-MEC communication.

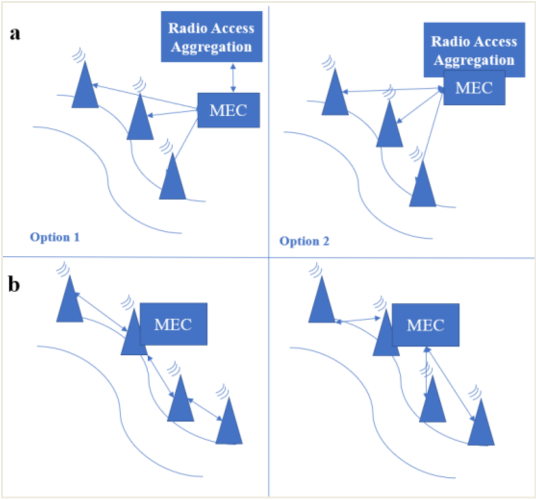


Figure 7-3 a) star topology with option 1 on the left-hand side and option 2 on the right-hand side; b) bus topology on the left-hand side and collapsed star topology on the right-hand side.

7.3.2 **Collapsed star topology model**

An alternative option to deploy MEC clustering is to co-allocate the MEC node with one of the gNBs in its cluster instead of leasing a new site. The right-hand side of Figure 7-3 presents this model. Connecting the MEC node to the different gNBs can be done in two different ways:

- Option 1: As the bus topology, represented in the bottom left-hand side of Figure 7-3, introduces more hops in the link between gNBs and the MEC, we have decided to exclude this option for the sake of preserving latency.
- Option 2: Point-to-point connections between each gNB within the cluster and the MEC node could be installed, as illustrated in the bottom right-hand of Figure 7-3. Hence, fiber cables are needed for the different links. An optimization algorithm is developed to determine the best position to host the MEC node while aiming at minimizing the total fiber cable needed given the certain number of gNBs per MEC cluster.

Another option could consist of leasing fiber cables from road operators if they have already been installed along the corridor. These two different options will be compared from a cost perspective to give more insights to MNOs for the MEC deployment options in section 7.5.

7.4 End to End latency for different MEC deployment options

MNOs and MEC service providers have been looking for deployment models that help reach trade-offs between the required service KPIs and deployment costs. In order to illustrate the impact of MEC placement on E2E latency, system-level simulations were carried out by Poli et al. [13] within a centralized Cooperative Lane Change (CLC) use case involving three vehicles. The targeted round-trip latency is typically in the order of tens of ms, while the exchange delay requirement should be kept lower than 100 ms from the generation time to the reception time [15]. In these simulation-based evaluations, different road traffic conditions were considered as well as a varying proportion of connected road users generating their own CCAM data traffic (i.e., with up to 25 percent of the simulated vehicles active in transmission, besides the three CLC vehicles under test). This comparison encompassed decentralized, clustered, and centralized MEC deployment options (respectively with one MEC per gNB, one MEC per radio aggregation site, or one MEC per MNO DC). RAN latency was simulated through ns-3 up to the application layer, including jitter at the radio level depending on link quality, resource scheduling, and network load effects. Conversely, higher-level latency components beyond radio access were set a priori as abstract bounds, for example, by integrating typical extra delays related to backhaul, core network, and possibly fast inter-MEC connections in the overall latency budget. For instance, for backhaul latency and core network latency terms, a distance-driven penalty of 2 ms was assumed for each component. These assumptions are typically based on a coarse estimate of 1 ms per 100 km for fiber transport. In addition, ultra-fast inter-MEC connections of typically 1–2 ms per MEC switch for transferring service-related data/sessions were also considered, whenever needed.

As expected, the overall E2E latency was hence shown to increase as a function of network load (e.g., 6–12 ms for uplink messages, depending on the tested configurations and radio settings). Moreover, the latency magnitude specifically imputable to radio access seemed to dominate by far all other higher-level latency components present in the overall budget (1–6 ms, depending on the MEC placement assumption). This simplified illustration suggests that a fully decentralized MEC option would be only marginally advantageous while requiring more costly orchestration efforts and associated latency (not accounted for in the previous study though), and hence, even larger latency. Therefore, a clustered MEC deployment is likely to offer a relevant balance between performance and cost.

7.5 Cost comparison between the proposed MEC clustering models

In order to find the best MEC clustering model in specific circumstances, we run the MEC TCO model, as described in the second section, for each of the three models described in the previous section, namely:

- **M11**: star topology model option 1, MEC node hosted in between the radio sites and the aggregation sites;
- **M12**: star topology option 2, MEC hosted in the aggregation site;
- **M2**: collapsed star topology, MEC hosted next to one of the towers within the MEC cluster.

Multiple initial condition scenarios per MEC clustering model can be identified and are parametrized in the models. They range from having no infrastructure in place at all to already having part of the required infrastructure installed. For models M11 and M2, two parameters have been used to describe four scenarios per model. Specifically, one parameter captures the existence of fiber cables that link the gNBs to the MEC node, and the other captures the need to lease a site to host the MEC node. For model M12, however, the MEC node will be hosted in the aggregation site, so there is no need to lease an additional site to host it. The M12 clustering model instead requires fiber linking the different aggregation sites to ensure a direct inter-MEC connection. The prior existence of this link is captured by an additional parameter, yielding four scenarios for M12 as well. In this paper, we only compare the worst and the best scenarios of the three models. The worst scenarios represent the greenfield deployments, which means no prior availability of the infrastructure. Yet, the best scenarios consider the availability of some of the required infrastructure. The worst scenarios of the three models combining the abovementioned parameters are as follows:

- **M11**: no prior availability of fiber from the gNBs to the MEC node and no prior availability of hosting site;
- **M2**: no prior availability of fiber from the gNBs to the MEC node and no prior availability of hosting site;
- **M12**: no prior availability of fiber from the gNBs to the MEC node (hosted in the aggregation site) and no prior availability of fiber between aggregation sites.

The best scenarios of the three MEC clustering models combining the aforementioned parameters are:

- **M11**: prior existence of fiber from the gNBs to the MEC node and prior availability of hosting site;
- **M2**: prior existence of fiber from the gNBs to the MEC node and prior availability of hosting site;
- **M12:** prior existence of fiber from the gNBs to the MEC node (hosted in the aggregation site) and prior existence of fiber between aggregation sites.

The corridor under consideration has 200 gNBs spanning 600 km of highway from Munich to Bologna. Since the required number of gNBs that can be clustered per

MEC node is not known, the total required number of MEC nodes is unclear. Thus, a range of values is analyzed between four gNBs per MEC node and 24 gNBs per MEC node. Twenty-four gNBs are chosen as a higher bound because of the hardware performance of the assumed MEC server, which can support traffic of up to 24 gNBs [10].

The results of the cost models are visualized in Figure 7-4 and Figure 7-5.

Figure 7-4 shows the cost results of the greenfield deployment, which means no prior availability of the infrastructure. As shown in Figure 7-4, M12 is the costliest clustering model. It is 32–44 percent and 19–89 percent more expensive than M11 and M2, respectively, depending on the number of gNBs per cluster. This is due to the total fiber length needed to connect the 200 gNBs to the aggregation sites and the aggregation sites to each other. Contrast this to clustering model M11, in which the fiber links to the MEC node are aggregated, and only one cable is required to link the MEC node to the aggregation site, reducing the total length of expensive fiber connections required. Another important observation is that the collapsed star topology (M2) is the most cost-effective model as long as fewer than 17 gNBs are grouped per MEC cluster. If more gNBs are clustered per MEC node, the total length of required fiber increases significantly, which makes the model more costly than M11. This can be explained by the fact that if we increase the number of gNBs per MEC cluster, the average distance between gNBs and the MEC node in a collapsed star topology becomes bigger than in the one using the star topology model.

When a significant number of gNBs per MEC cluster is considered, an inconsistent relationship between the number of MEC clusters and the total cost is noticed in M2. In fact, increasing the number of MEC clusters by one can be cheaper than having more gNBs per MEC cluster and their related fiber connections. This phenomenon can be observed starting from point (17, 11) on the M2 curve in Figure 7-4.

Especially for the star topology model (M11), it seems rather unrealistic that fiber cables would already be in place. In contrast, in the collapsed star model (M2), where the MEC node is hosted next to one of the gNBs within the MEC cluster, it is more likely that fiber cables are already present. These connections would generally run along the roadside and might already be deployed by road operators or MNOs. Therefore, we simulated a variant of the M2 model with different prior infrastructure availability assumptions. Concretely, it is assumed that the MNO will save on fiber cables by sharing the infrastructure in a passive or active way with a different operator, which could be a road operator or an MNO. Using network-sharing models results in different cost savings. According to a report by BEREC [14], passive infrastructure sharing could save 16–35 percent of CAPEX and 16–35 percent of OPEX, while active infrastructure sharing could result in savings of 33–35 percent of CAPEX and 25–33 percent of OPEX. Taking these assumptions into account, we compared the collapsed star clustering model (M2)

without prior infrastructure with active and passive fiber-sharing options to the worst scenarios in Figure 7-4.
The results show that by sharing the fiber installation, the TCO can be reduced by up to 31 percent (Figure 7-4). Active sharing yields more savings than passive sharing: 32 vs. 25 percent. Importantly, even with a high number of gNBs per MEC cluster, the collapsed star topology model (M2) with active sharing is the least costly of all clustering models.

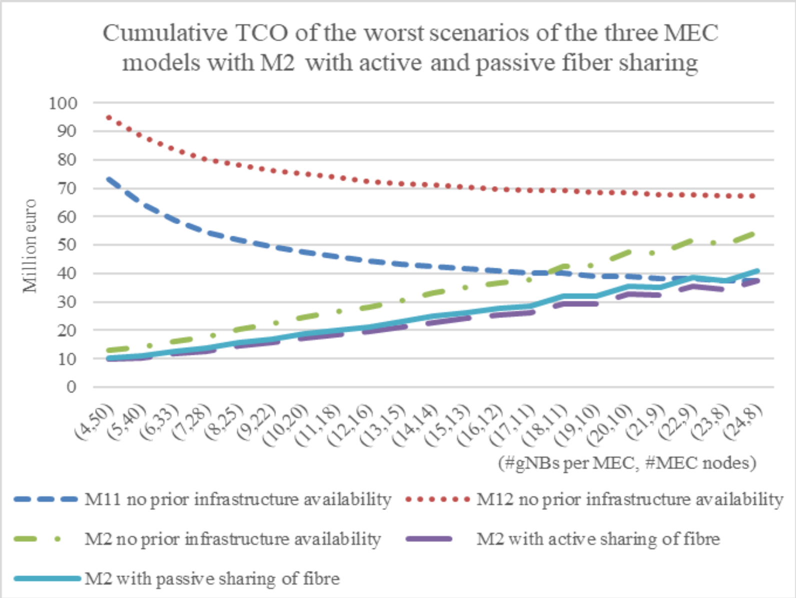


Figure 7-4 The Cumulative TCO plotted against the number of MEC nodes for the three MEC models, given no prior infrastructure availability, as well as for two variations on M2 with active and passive sharing of fiber cable.

What if some of the required infrastructures for the MEC deployment are already in place? Figure 7-5 depicts the best scenarios with the prior infrastructure available for all models. It is found that the clustering model M12, in which MEC nodes are hosted at the aggregation sites, is the most cost-effective clustering option given prior infrastructure availability by 17 percent as a cost reduction compared to the two other models. The source of the cost advantage of M12 over M11 and M2 is the cost of the cabinet required to host the MEC hardware, since it is assumed that there is enough space to host the required hardware at the aggregation site. Additionally, models M11 and M2 are identical from a cost perspective when prior infrastructure is maximally available. This makes sense because exactly the same hardware needs to be deployed, since fiber connections are already present and MEC hosting sites are available. Importantly, even the most expensive clustering models M11 and M2 are less costly compared to the non-clustering option described earlier. They have a TCO of €2 million, given the highest number of MEC nodes required, compared to €7 million for the non-clustering model presented in Table 7-1, which resulted in 71 percent cost saving on the TCO.

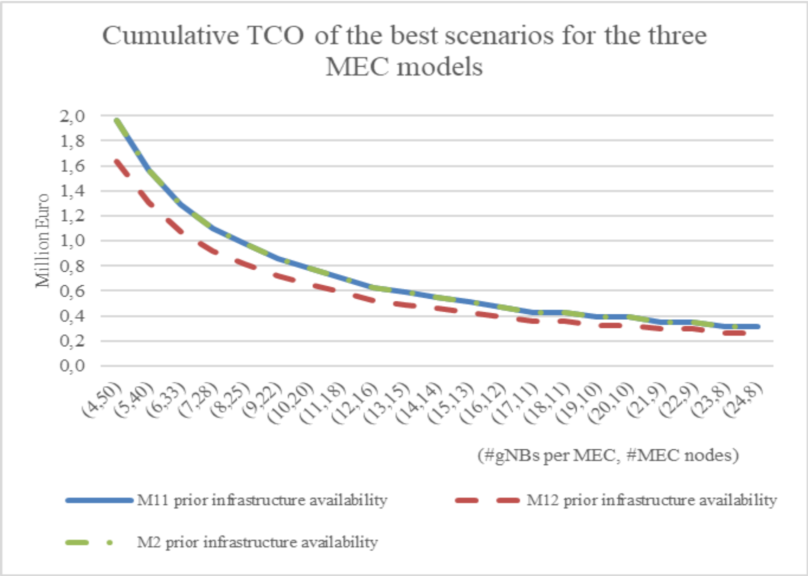


Figure 7-5 The cumulative TCO as a function of the number of MEC nodes for best scenarios of the three MEC models.

The main takeaways from this analysis are as follows. First, it is found that all the MEC clustering models are less costly than the non-clustering approach, where MEC nodes are hosted next to each gNB. Second, when no prior infrastructure is available, the star topology model with MEC nodes hosted at the aggregation site (M12) is the costliest model, while the collapsed star topology model (M2) is the most cost-effective model as long as 17 or fewer gNBs are grouped per MEC cluster. When more than 17 gNBs are grouped per MEC cluster, the star topology model with MEC nodes hosted in between aggregation sites (M11) is the least costly. Finally, when prior infrastructure is maximally available, the star topology model with MEC hosted at the aggregation site (M12) is consistently the most cost-effective clustering model.

In practice, no MEC clustering model fits all networks. The derived cost results can help network operators decide which option to adopt given the status of their network in the considered deployment region. In areas where a good fiber infrastructure is already present, the star topology clustering model is likely to be the optimal one. Yet, for remote areas where no fiber is installed, the collapsed star topology model is likely the optimal clustering method, especially if passive or active network-sharing strategies can be adopted.

7.6 Conclusion

In this paper, three different MEC clustering models addressing the MEC placement challenge are analyzed from a cost perspective, given latency requirements. The analysis showed that all MEC clustering models outperform the non-clustering deployment in both cost-effectiveness and efficiency: MEC clustering models allow a saving of up to 71 percent on the TCO and a reduction of E2E traffic latency of 1–6 ms. Given the prior availability of network infrastructure along the considered highway, certain models are cheaper than others. For instance, the collapsed star topology (M2) is 19–89 percent less expensive than the other models when no prior infrastructure is available. Yet, the star topology with hosting MEC at the aggregation site (M12) is 17 percent cheaper when some of the infrastructure is available. Network sharing in both its active and passive flavors significantly decreases the cost of the MEC deployments by 31 and 25 percent, respectively. Since a comparison of latency was only performed between centralized, clustered, and decentralized models, future work could consist of performing system-level simulations to compare the E2E latency between the three clustering models. In addition, future work could address the impact of the variation of bandwidth requirements and the network traffic loads over time on the dimensioning of the number of servers needed per MEC node.

ACKNOWLEDGMENT

This work has been carried out in the frame of the 5G-CARMEN project, which is partly funded by the European Commission (H2020 GA No.825012).

References

[1] B. Blanco, J., Fajardo, I. Giannoulakis, E. Kafetzakis, S. Peng, J. Pérez-Romero, I. Trajkovska, P. S. Khodashenas, L. Goratti, M. Paolino, E. Sfakianakis, F. Liberal, G. Xilouris, "Technology pillars in the architecture of future 5G mobile networks: NFV, MEC and SDN", Computer Standards & Interfaces, Volume 54, Part 4, 2017, Pages 216-228, ISSN 0920-5489; https://doi.org/10.1016/j.csi.2016.12.007

[2] R. Soua, I. Turcanu, F. Adamsky, D. Führer and T. Engel, "Multi-Access Edge Computing for Vehicular Networks: A Position Paper," 2018 IEEE Globecom Workshops (GC Wkshps), 2018, pp. 1-6, doi: 10.1109/GLOCOMW.2018.8644392.

[3] ETSI, "Multi-access Edge Computing (MEC)," ETSI, [Online]. Available:https://www.etsi.org/technologies/multi-access-edge-computing. [Accessed 03 2021].

[4] 5GAA, "MEC for Automotive in Multi-Operator Scenarios," 5GAA Automotive Association, Munich, 2021.

[5] Pereira, Jorge M. "5G for Connected and Automated Mobility (CAM) in Europe: Targeting Cross-Border Corridors." IEEE Network 35.3 (2021): 6-9.

[6] Malinverno, M., Avino, G., Casetti, C., Chiasserini, C. F., Malandrino, F., & Scarpina, S. (2019). Edge-based collision avoidance for vehicles and vulnerable users: an architecture based on MEC. IEEE vehicular technology magazine, 15(1), 27-35.

[7] S. Lee, S. Lee and M. -K. Shin, "Low Cost MEC Server Placement and Association in 5G Networks," 2019 International Conference on Information and Communication Technology Convergence (ICTC), 2019, pp. 879-882, doi: 10.1109/ICTC46691.2019.8939566.

[8] Q. Yuan, J. Li, H. Zhou, T. Lin, G. Luo, and X. Shen, "A joint service migration and mobility optimization approach for vehicular edge computing," IEEE Trans. on Veh. Tech., Jun. 2020.

[9] Casier, K. "Techno-economics of FTTH deployment in the presence of competition", (Doctoral dissertation), (2010).

[10] Thinkmate, "Gigabyte h242-z10," [Online]. Available: https://www.thinkmate.com/system/gigabyte-h242-z10. [Accessed 01 2021].

[11] 5G-NORMA, "Deliverable D2.3 Evaluation architecture design and socio-economic analysis - final report", Dec. 2017.

[12] Statista, "Prices of electricity for industry in Belgium from 2008 to 2021", [Online]," https://www.statista.com/statistics/595775, [Accessed: 01 2021].

[13] F. Poli, N.-L. Dinh, B. Denis, V. Mannoni, "Evaluation of 5G-NR V2N Connectivity in a Cooperative Lane Change Scenario", IEEE Vehicular Technology Conference 2022 – Spring (IEEE VTC-Spring'22), Helsinki, June 2022.

[14] BEREC, Report on Infrastructure Sharing, [Online]," berec.europa.eu/eng/document_register/subject_matter/berec/reports/8164-berec-report-on-infrastructure-sharing, [Accessed: 01 2021].

[15] D. Garcia-Roger, E. E. González and J. F. Monserrat, "Regional Multi-RAT Dual Connectivity Management for Reliable 5G V2X Communications," 2022 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit), 2022, pp. 320-325, doi:0.1109/EuCNC/6GSummit54941.2022.9815600.

8

Conclusions and future work

"The only way to be truly satisfied is to do what you believe is great work, and the only way to do great work is to love what you do. If you haven't found it yet, keep looking, and don't settle. As with all matters of the heart, you'll know when you find it." - Steve Jobs

Fast and reliable mobile broadband connectivity is crucial for end-users who demand high-speed connections, anytime and anywhere, to access anything they need. 5G networks offer the potential to meet these expectations, with new promising technologies like Massive MIMO, NFV, SDN, network slicing, and MEC. However, to ensure the success of 5G networks, not only must these technologies be developed and optimized, but also the cost-effective rollout of 5G on a large scale while adopting these technologies is crucial. To make this possible, significant attention must be given to the economic implications of 5G network deployment, including identifying the factors driving high costs and selecting cost-saving strategies. By prioritizing cost-effectiveness alongside technical advancements, 5G networks can be deployed at scale, meeting the high expectations of end-users and unlocking new opportunities for businesses and society as a whole.

Moreover, the rapid advancement of 5G rollout in prominent cities and transportation routes is showing promising economic prospects, as elaborated in section 1.3. However, hard-to-reach, rural and cross-border areas, present hurdles in achieving swift returns on investment. This is due to the substantial costs involved and the sparse population in these regions. In light of these challenges, we have chosen these areas as our study's focal points to examine the economic feasibility of 5G deployment. This approach aims to provide valuable insights for MNOs and policymakers in making informed decisions.

Therefore, in this dissertation, we conducted a comprehensive techno-economic analysis of the 5G rollout in the selected use cases, with a focus on addressing five key research challenges. The challenges we tackled in this study include:

1. Assessing the feasibility of 5G deployment for different types of regions and services.
2. Quantifying the cost savings associated with various 5G technologies.
3. Developing new cost allocation models that are better suited to the unique characteristics of 5G technologies to share costs among services and stakeholders.
4. Identifying potential business and collaboration models that can enable successful 5G deployment.
5. Identifying technology choices and timing decisions for optimized 5G deployment Strategies.

Throughout the different chapters of this dissertation, we addressed one or more of these research challenges. In the following section, we provide a summary of the main conclusions drawn from each chapter with respect to these key research challenges.

8.1 Summary of the main conclusions

1. Assessing the feasibility of 5G deployment for different types of regions and services

The deployment of 5G networks is dependent on the region's characteristics, such as rural or urban, and the specific service KPIs, such as low latency and high capacity. However, many techno-economic analyses in the literature do not adequately consider these two factors together. Therefore, our research addressed this gap by focusing on both factors to evaluate the feasibility of 5G deployment.

In **Chapter 2**, we assessed the viability of providing **eMBB** for **rural areas** in Europe to bridge the digital divide. We proposed a new broadband solution using satellite communication for the backhaul network part and 4G networks for the access and core network parts. A TCO model was developed to study the feasibility of integrating satellite communication within the 4G network. **The results of the model showed that the proposed solution is economically viable for the universal broadband service (2Mbps) based on the comparison of the ACPU to the WTP of the end user, but is still expensive for high-speed broadband (30Mbps) even with caching**. Sensitivity analysis showed that population density has a significant impact on the resulting ACPU. In fact, **similar to other broadband solutions, the cost per user of deploying the satellite-4G solution in urban areas is much cheaper than in rural areas.**

While Chapter 2 focused on rural and remote areas, **Chapter 3** addressed inflight connectivity (**eMBB**) or the so-called "**internet in the sky**". We studied the feasibility of a similar broadband solution but this time based on 5G networks. The 5G core network on the ground is connected to the 5G access network installed onboard via a satellite link**. The TCO model's key findings indicate that integrating satellite communications into the 5G network enables the provision of inflight broadband service with a good throughput per user**

(ranging from 2 to 5 Mbps) at reasonable prices, thus demonstrating its feasibility.

Previous chapters focused on the same service (i.e., **eMBB**) but for **different types of regions**. In contrast, **Chapter 5** investigated the provision of **URLLC** services. We conducted a techno-economic analysis of providing CCAM services in **cross-border environments** based on CLM service as a canonical case study. The results of cost modelling offer valuable insights into the TCO associated with various deployment scenarios for both vehicle-to-infrastructure (V2I) and vehicle-to-network (V2N) connectivity. These insights can be utilized as a rough estimate to determine the funding required by the EC to incentivize MNOs to increase their 5G deployment in border regions. To evaluate the feasibility of various deployment scenarios, we conducted an additional study that examined several revenue models. The results, which are presented in **Chapter 6**, compare the costs of deployment with the revenue generated. To provide a comprehensive understanding of the viability of the use case, we proposed four revenue models based on different assumptions. Our findings indicate that a **cost-based revenue model stands out as the most effective approach for achieving a rapid break-even point**. In contrast, **subscription-based revenue models did not demonstrate the potential to ensure profitability over the investment's considered lifetime**, spanning 10 years. Consequently, this situation leads to an unprofitable deployment of 5G within the scope of the case study.

The findings reveal that the deployment of a 5G network aimed at delivering high QoS, encompassing robust throughput for eMBB services or minimal latency for CCAM services, comes at a significant cost in terms of both the ACPU and the TCO per kilometre.

2. Quantifying the cost savings associated with various 5G technologies

The literature lacks comprehensive quantitative assessments of the potential cost savings of 5G technologies such as network virtualization, content caching, and network slicing. Many papers fail to provide sufficient analysis to support their claims about cost savings. We contributed to addressing this research gap by quantifying the cost savings and identifying the optimal technology for specific conditions.

In **Chapter 2** where an integrated satellite-4G networks solution has been evaluated, we compared two scenarios to assess **the potential cost savings from adopting a strategy of caching popular content on the MEC**, one without and the other one with the use of MEC. Results show that caching popular content on the edge **can significantly save on OPEX costs (60%) and thus decrease the required Average Cost Per User (ACPU).** Sensitivity analysis showed that the caching rate has a significant impact on the resulting ACPU. Indeed**, the CAPEX increases with the use of caching due to the cost of the MEC deployment**. However, **there is a significant decrease in OPEX as the caching rate increases**, due to a reduction in traffic sent via satellite link and a subsequent decrease in satellite capacity cost. It may be beneficial to use a high rate of caching to decrease operational costs, but a detailed analysis of consumer traffic and optimization algorithms for popular content is required to determine the optimal caching rate.

Furthermore, we have examined the impact of using caching popular content on the business case of proving inflight connectivity in **Chapter 3**. Findings have demonstrated that **MEC-based caching techniques can significantly decrease OPEX and consequently, the ACPU by 25 to 32%, depending on the caching rate employed**. Nonetheless, there is a trade-off to be made between the level of content personalization available on the aircraft and the caching rate, as higher caching rates can result in reduced content personalization.

In addition to the potential cost savings of MEC technology, **the impact of** network function virtualization **(NFV)** on deployment cost has been assessed in **Chapter 4**. Results demonstrated a **cost saving of 12 to 45% (depending on the approach adopted) of the core network costs resulting from the use of NFV and network slicing.**

In general, the incorporation of cost-saving technologies like network virtualization and Multi-Access Edge Computing (MEC) yields substantial reductions in the overall expenses associated with deploying 5G infrastructure. However, the timing and methodology of integrating these technologies markedly influence the extent of cost savings achievable. For instance, in scenarios where backhaul capacity costs are high—such as in the case of satellite capacity—maximizing popular content caching proves advantageous. Nonetheless, the precise selection of content and caching quantities necessitates the implementation of a robust forecasting algorithm for accurate determination.

3. Developing new cost allocation models that are better suited to the unique characteristics of 5G technologies to share costs among services and stakeholders

To ensure the feasibility of deploying 5G networks, new cost allocation models are needed that take into account the unique and complex nature of 5G technologies. However, the existing literature does not adequately address this issue. To fill this gap, we have developed new cost allocation models that allocate deployment costs to services and stakeholders involved.

In **Chapter 4**, we proposed **a novel cost allocation model** for **5G-sliced networks**. It allocates the network costs to the deployed slices. To feed the proposed model with the necessary inputs, we developed two other models. The first one is a resource allocation model that assigns the appropriate amount of network resources to each slice to fulfil its KPI requirements. For this, we needed to derive the hardware requirements of each VNF included in the chain of VNFs forming the slice. Therefore, we designed the second model that identified the hardware requirements of the 5G network functions, with a particular focus on those of the control plane. **This latter model is the first of its kind in literature and is essential for allocating network resources efficiently**.

We applied the proposed model to a specific scenario where **two slices**, **eMBB video, and voice**, were offered to a remote area. The results of the model demonstrate that eMBB video bears a significant part of the network cost.

As previously discussed, in **Chapter 5**, we elaborated a techno-economic analysis to evaluate the provision of CCAM services in cross-border environments. The analysis included the development of **a cost allocation model** that shares the costs of different scenarios **among the key stakeholders**, namely (MNOs) and Road Operators (ROs) on both sides of the border. The study focused on **identifying various key cost allocation drivers** taking into account both the identified business models and the type of connectivity used. **The results of the allocation model indicated that adopting a collaborative business model, which involves infrastructure sharing, can significantly reduce the investment cost of the considered deployment scenario for the operators involved**.

In general, emerging technologies that facilitate cost savings and innovative business models, such as network slicing and MEC, necessitate the development of novel cost allocation models. These models are essential for evaluating the feasibility of offering specific services or adopting particular business strategies. Consequently, the development of these new cost allocation models holds paramount significance in ensuring a comprehensive evaluation of the integration of these technologies.

4. Identifying potential business and collaboration models that can enable successful 5G deployment

The emergence of 5G networks offers promising opportunities for businesses to innovate, but the complexity of the ecosystem presents a challenge in establishing sustainable and viable business models. Therefore, we investigated potential business and collaboration models that assist stakeholders in leveraging the introduction of 5G technologies.

In **Chapter 3, we explored various business models** for providing broadband service to onboard users through integrated satellite-5G networks. We identified different business models depending on who will provide the service to the end user, namely **wholesale versus retail models,** with the choice based on the level of control and risk the airline company is willing to take. The retail model is more beneficial to customers as it offers a shorter value chain and higher quality assurance. For low-cost structures, **the sponsorship model** is suggested. **The freemium wholesale and retail models** offer branding opportunities and free services to both MNOs/airline companies and passengers but require a high enough conversion rate from free to premium to work effectively. Premium airline companies are better suited to **the complementary service model**, where customers pay for the full premium service that includes connectivity. Furthermore, **we proposed various pricing schemes** for each business model to ensure maximum profitability.

On the other hand, **Chapter 5** highlighted the complexities of the cross-border environment for providing mobility services, which stem from technical and business considerations. Technical challenges include network coverage issues and the need for new technologies to meet the CCAM KPIs. Business challenges include developing feasible and innovative models that promote cross-border collaboration to deliver services. To address these challenges, **we proposed three business models that rely on emerging 5G technologies, specifically network slicing and MEC**, **and account for the phased rollout of 5G networks (non-standalone and standalone deployments**). **Among the various models identified**, **the collaborative business model shows the most promise** based on MNO feedback, but it requires regulatory review, including a review of limitations on base station transmission power near the border. Additionally**, the NSaaS (Network Slicing as a Service) business model** requires strong regulatory intervention to be realized in cross-border settings.

In essence, identifying potential business models represents only the first step in the process of finding a viable business case for the rollout of 5G. Equally crucial are endeavours to delve into the value generated for all stakeholders involved, coupled with an exploration of the evolving obstacles within the ecosystem that necessitate resolution, as explained in sections 8.2.1 and 8.2.2.

5. Identifying technology choices and timing decisions for optimized 5G deployment strategies

This research question investigated the most efficient deployment strategies for 5G networks, taking into account not only the impact of various regions and service KPIs (as addressed in the first research question) but also deployment timing as well as the choice of 5G technologies.

In **Chapter 5**, where a techno-economic analysis of providing CCAM services in a cross-border environment was elaborated, **two different types of connectivity** were evaluated namely **V2I and V2N** with different densification of the RAN part. **For the V2I scenario**, we assessed **three different densifications** of RSUs per kilometre. We concluded that installing the first RSUs incurs a significant cost because we need to install fibre and electricity cables, which are the dominant investment cost component. However, adding new RSUs will not have a significant capital expenditure since the required cabling to connect the additional RSUs is already in place (as discussed in Section 5.4.2).

For the V2N scenario, we evaluated **two different densifications** of macrocells in the studied area for two scenarios where MEC is to be installed next to the towers or in the data centre. Our results showed that adding additional macrocells resulted in a high cumulative cost, where OPEX is the dominant cost due mainly to the site rental cost required to host the new macrocell. When deploying MEC next to the towers, the CAPEX is slightly higher compared to the case where MEC is hosted in the data centre (as discussed in Section 5.4.2).

Although the three deployment scenarios were simulated to take place in the same year, exploring **the impact of varying the deployment timeline based on the adoption rate of CCAM services on the TCO** can be an interesting exercise, which was not addressed in Chapter 5 due to the lack of data regarding penetration rates of connected cars. To solve the issue of predicting the penetration rate of connected cars, which is uncertain due to outdated studies and the impact of Covid on the automotive industry. **We suggested two approaches**: **a generic and abstract approach based on a variation of penetration rates**, and **a scenario-based approach with three different scenarios: low, medium, and high penetration rates**. The techno-economic analysis using the generic approach has been published in [5], and the scenario-based approach which provides more specific predictions is presented in **Chapter 6**. Results of the latter approach showed that as expected, **the higher the number of simultaneously connected cars (high penetration rates), the higher the TCO of the required deployment needs to be to meet the defined KPIs**. Results demonstrated also how expensive it is to deploy in a green field of the highway where no existing fibre cables or electricity facilities. **The green field deployment is three times more costly than the pre-equipped field deployment**.

The impact of installing MEC next to the tower (next to the RAN) or in the data centre on the TCO has been addressed in **Chapter 5**. Yet, it was outlined there that more deployment options should be studied to achieve a trade-off between the desired latency for CCAM services and the deployment cost. **We tackled this research limitation in Chapter 6** by proposing three MEC clustering models and evaluating them from both cost and E2E latency perspectives. **The analysis revealed that deploying MEC clustering models outperforms non-clustering deployment in terms of both cost-effectiveness and efficiency**. With **MEC clustering models, you can save up to 71% on the TCO and reduce E2E traffic latency by 1-6 ms**. However, certain models are more cost-effective than others depending on the availability of prior network infrastructure. For example, the collapsed star topology is 19-89% less expensive than other models when no prior infrastructure is available, while the star topology with hosting MEC at the aggregation site is 17% cheaper when some infrastructure is already available. **Network sharing in both its active and passive forms can also significantly reduce the cost of MEC deployments by 31% and 25%, respectively**.

Overall, the rollout strategy of the 5G network and its enabling technologies can have a substantial influence on the eventual deployment costs. Various parameters can be fine-tuned to achieve a cost-effective implementation, including considerations such as MEC clustering models and the strategic timing for deployment.

8.2 Insights and recommendations for the key stakeholders in the 5G ecosystem

In the preceding subsections, we highlighted the main conclusions regarding the five studied research questions. However, in this section, we will delve into the major findings from the conducted research work. Our focus here is deliberately set on pinpointing the primary beneficiaries of these insights and expounding upon the recommendations we advocate in light of these momentous findings.

8.2.1 Insights and recommendations for policy and decision-makers

Our Findings proved that providing **rural areas** with broadband connectivity (**eMBB**) at a minimum bitrate of 30Mbps, as per the European digital agenda's target, **remains a significant challenge**, even with the use of 5G technologies. To address this, **relevant authorities could explore several paths including:**

- **Adopt funding mechanisms aimed at subsidizing broadband connectivity services in rural regions** like done in the United States. Presently, funding frameworks are implemented at the Member State level, exemplified by the case of Spain [1]. Nonetheless, **the adoption of a more unified and harmonized framework could foster the 5G rollout in these underserved areas**.
- Promote the establishment of **5G private network initiatives** to accelerate the adoption of 5G for business applications and extend coverage to regions with a lower commercial appeal [2], like the one adopted in the 5G+ Strategy to Realize Innovative Growth of South Korea [3]. **One approach could involve allocating spectrum bands for 5G private network deployment at reduced fees exclusively for rural areas and cross-border corridors.**

On the other hand, to establish **sustainable business models for connected mobility services**, not only the economic feasibility of 5G network deployment must be carefully assessed, but also several challenges resulting from the novelty of these services and the supporting technologies.

Our findings have shown that 5G deployment costs to support CCAM services in cross-border areas are significant, especially when an important number of simultaneously connected cars is considered, which translates into a higher network infrastructure needed to meet the required service KPIs. Yet, this also means a faster reach of the breakeven point, indicating that the investment in 5G network deployment for connected mobility services can be profitable in the long run if **an appropriate pricing model is adopted.** However, the number of connected cars is following a low penetration curve in the European market, this **will demotivate MNOs to invest in the deployment of 5G networks along the roads and highways to support connected mobility services.** Moreover, the complexity of the 5G ecosystem for CCAM services acts as a deterrent to involved stakeholders, hindering the swift initiation of these services. This hesitation arises from various uncertainties, including the allocation of liability in the event of an accident and the parties' accountability, as well as the lack of clarity surrounding data sharing and storage protocols.

Relevant authorities could explore several paths including:

- **Devoting financial resources towards bolstering CCAM services, particularly within safety-critical ones, in border regions**. Leveraging our research outcomes, **these findings can be instrumental in approximating the requisite funding necessary** to incentivize MNOs for the enhanced deployment of 5G infrastructure in these specific zones. The present regulatory landscape mandates certain responsibilities onto MNOs, such as achieving comprehensive national coverage or a substantial portion thereof. Consequently, MNOs often extend their services to encompass less commercially alluring zones, albeit at the expense of compromised Quality of Service (QoS), failing for example to meet the QoS required by CCAM services. However, the direct allocation of funds designated for these regions, coupled with an unwavering focus on achieving prescribed QoS levels, holds the potential to propel MNOs toward expeditious and impactful endeavours.

- **The promotion and facilitation of infrastructure sharing,** especially within the passive layer of access networks (like towers and small cells), are essential to reduce deployment expenses for operators. Moreover, granting access to public spaces for the installation of 5G infrastructure, like street furniture will eliminate a significant obstacle to cell deployment [2] and [4]. This is also valid for other use cases e.g. urban deployment, not only cross-border environment. The example of network sharing in China had shown promising cost savings for operators, as discussed in section 1.3.1.a.
- **Establish frameworks designed to foster dialogue among pertinent stakeholders, facilitating an in-depth exploration of the obstacles surrounding the initiation of CCAM services**. Within these discussions, stakeholders can collaboratively identify and present strategic roadmaps aimed at effectively addressing challenges like liability attribution and the intricacies of data sharing.

Additionally, the journey of this research **has revealed a notable insight**: when delving into the discourse on **Service Level Agreements (SLAs) and their impact on potential business models and deployment costs**, MNOs often perceive it as a complex research subject beyond the scope of the current study. In light of this, we contend that a two-fold approach should be adopted, starting with policy considerations followed by regulatory actions:

- Initiate an R&D project focusing on the identification of innovative SLA formats tailored for 5G applications targeting both B2C and B2B markets. These SLA formats should also seamlessly incorporate the innovative technologies introduced by 5G, such as network slicing and edge computing, ensuring they are effectively accommodated within the framework of the new agreements.
- The harmonization of regulatory endeavours at the European Union (EU) level holds paramount significance in ensuring harmonious cross-border alignment and establishing the foundation for the Digital Single Market. An exemplary instance is the cohesive implementation of Service Level Agreements (SLAs) designed for slice-aware and capable roaming. This component is pivotal for safeguarding seamless service continuity across borders, particularly for applications with safety-critical implications. Present roaming agreements solely support the "Best-effort" service

class. Ensuring service continuity across borders mandates that the SLA established in the home country is equally upheld in the visited one.

An additional valuable insight emanating from this research underscores its potential relevance to policymakers. Notably, the absence of full 5G coverage in border regions emerges sometimes from legislative constraints governing permissible signal strength within these areas. In response, **we recommend a comprehensive reassessment of such regulations, accompanied by the formulation of alternative frameworks. These revised frameworks should ideally accomplish dual objectives: ensuring seamless coverage across border regions while simultaneously safeguarding against any potential interference among emitting cells**.

8.2.2 Insights and recommendations for MNOs

As discussed in the section summary of the main conclusions, section 8.1, several findings can help MNOs take informed decisions in their journey with the 5G rollout, but in what follows, we list several highlights and recommendations:

- Adopt cost-saving strategies such as network sharing, collaborative models, MEC clustering models, and network virtualization, as previously discussed in section 8.1.
- Our results furnish a cost estimation for 5G deployment across various use cases. Moreover, the formulation of viable business models depends on the meticulous evaluation of suited pricing structures, as expounded upon in Chapter 6. These business models must further encompass factors like liability and the implications of Service Level Agreements (SLAs) in instances where network shortcomings or unattainable Quality of Service (QoS) levels arise. This particular consideration has also been examined in-depth within the preceding subsection.
- Adopt 5G technologies that enables new business models such as the use of MEC and network slicing as highlighted in section 8.1. For instance, the adoption of network slicing can help to overcome some of the complexity of the 5G network ecosystem, particularly for connected mobility services. This technology enables the creation of independent virtual networks tailored to meet specific requirements, including low latency, high reliability, and security, among others. As a result, new business models, such as **the collaborative business model** that involves creating a joint slice for connected mobility **among different operators,** or **the network slice as a service (NSaaS)** business model, can be realized through network slicing.
- However, the adoption of network slicing comes with various challenges that need to be addressed to ensure the sustainability of business models. **One of the challenges that we addressed in this dissertation is related to the cost allocation of 5G deployment costs to the offered slices**. We **proposed a novel cost allocation model that allocates the cost to**

slices based on their network requirements. This model helps **MNOs** and **service providers** in pricing the different end-to-end slices and hence services.

Throughout the research, we conducted a comparison between different 5G deployment scenarios based on various adopted technologies, which provide **MNOs** with valuable insights into the optimal deployment strategies. For instance, we proposed three MEC clustering models and demonstrated that clustering models consistently outperform non-clustering MEC deployments, resulting in significant cost savings. Furthermore, we conducted a thorough investigation into the circumstances under which certain models are more cost-effective than others.

8.2.3 Insights and recommendations for Small and Medium-sized Enterprises (SMEs)

Throughout this study, valuable insights have been gathered, poised to assist SMEs in harnessing the benefits of 5G deployment. Below are a few key highlights:

- Recognize gaps in value creation within the 5G ecosystem, leading to the formulation of novel business models that capitalize on the integration of emerging technologies. The 5G new technologies such as MEC, network slicing, the network virtualization amongst others open up opportunities for a wide range of business models. For example, our findings proved that integrating satellites into 5G networks could be a viable solution for in-flight connectivity, where the onboard capacity-demanding services are typically video streaming services. Caching some of the popular content on the edge, using MEC, can save on the required bitrate per user and hence make the solution more affordable, which enables new business models for **airliners**, **sponsors**, and other stakeholders in the ecosystem, as explored in Chapter 3.
- **Adopt initiatives like OpenRAN**[23] which are rapidly gaining momentum and are poised to have a profound, long-term impact on the industry. This initiative is guided by the foundational principles of open radio access networks (RANs) which centre on open interfaces, pivotal for vendors and operators to provide their services and tailor networks, thereby fostering multivendor deployment and a dynamic supplier ecosystem. Within this framework, smaller entities, **including SMEs, are presented with opportunities to innovate on certain network functions and processes, enabling them to compete effectively with established players in the technology vendors' marke**t [2].

[23] OpenRAN is a project aiming at creating an open ecosystem of vendor-neutral RAN solutions. See: https://telecominfraproject.com/openran/

8.2.4 Insights and recommendations for researchers

Initially appearing as distinct entities, the developed models, in reality, interconnect to establish a cohesive foundation, collectively constituting a comprehensive framework for conducting a holistic techno-economic analysis of 5G deployment. Subsequently, we will elucidate the practical utilization of the diverse models, outline the necessary data prerequisites, and underscore key insights gleaned from data acquisition and model application. This exposition serves as guidance for fellow researchers, urging them to consider these insights when engaging with the developed models or embarking on the creation of analogous ones.

The developed cost model in Chapters 2 and 3: within the context of network dimensioning in the two cost models, the focal point revolves around the average bitrate per user. However, if your network dimensioning pertains to services demanding a guaranteed Quality of Service (QoS), a shift to utilizing the peak bitrate is imperative. This transition applies to Equation 2-1 in Chapter 2 and Equation 3-3 in Chapter 3.

The developed cost model in Chapter 5: the formulation of the cost model incorporates the car traffic density, based on varying levels of low, medium, and maximum traffic density observed on the examined road (based on data collected from the road operator). Additionally, the model integrates the inherent physical constraints of the road, encompassing the lane count in each direction and the stipulated safety distance between vehicles. Consequently, should you attempt to replicate the model without adhering to these essential conditions, the resultant outcomes are unlikely to yield meaningful insights.

Data collection: to generate insightful results using these models and eventually any other model, accurate data should be used. In this dissertation, we developed a two-step strategy to derive accurate data, see section 1.5 for more information. Remarkably, an intriguing revelation emerged during the course of data acquisition that the time invested in data collection should not be underestimated. While constructing the model itself or generating results may appear time-intensive, it is actually the process of collecting data from engaged collaborators within the research project that proves to be the most time-consuming aspect. This phenomenon becomes even more pronounced when delving into the exploration of nascent technologies like 5G. The realm of uncertainty enveloping diverse factors, such as the deployment strategies adopted by various actors like MNOs and the intricate cost considerations encompassing several equipment and installation methods, accentuates this reality. Hence, **our recommendation** is to secure data from a minimum of two partners, as they often receive distinct quotations from equipment vendors. Given the confidential nature of this data, partners may provide a range of costs rather than precise figures. Consequently, it is advisable to further enhance your approach by juxtaposing the gathered data against existing information in the relevant literature. This comparative analysis will enable the derivation of an averaged dataset, solidifying the foundation of your study, as explained in our data strategy in section 1.5.

Adoption of the TCO approach: within this research endeavour, our methodology was based on the utilization of the TCO model. Given the multitude of advantages associated with the TCO model, as expounded in section 1.5, we propose a pivotal principle: techno-economic analysis should not be solely based on CAPEX or OPEX. Relying solely on either of these elements can result in overlooking substantial costs. Therefore, we strongly advocate the adoption of the TCO framework. However, our recommendations extend further. To enhance the analytical rigour, we propose two key augmentations to the TCO model:

1. **Scenario Analysis:** By integrating scenario analysis, the analysis broadens beyond confined parameters, fostering the generation of multifaceted insights.
2. **Sensitivity Analysis:** Furthermore, supplementing the TCO model with sensitivity analysis is advised. This meticulous step serves to evaluate the uncertainties inherent in the assumptions underpinning the TCO model.

These two practices we diligently incorporated into our adopted methodology, as described in section 1.5.

Nonetheless, it is imperative not to overlook a significant aspect: the TCO approach's inherent limitation in accounting for indirect benefits. Researchers need to augment the TCO model with an additional framework that quantifies these indirect benefits and seamlessly integrates them into the analysis of the overall benefits.

8.3 Future work

In this section, we will shortly discuss several future trajectories for the completed research.

Throughout this dissertation, a comprehensive exploration of the MNO perspective has been diligently undertaken in the majority of cases. However, it is imperative to recognize that a multi-actor perspective is equally indispensable. This inclusive approach is necessary to comprehensively examine the impediments obstructing the realization of the broader societal and economic advantages associated with 5G deployment. Several of these challenges, as previously delineated, encompass concerns like liability, intricate data sharing protocols, and the formulation of concrete SLA formats.

Extension to other 5G use cases: in this dissertation, we focused on two of the three 5G use cases, namely eMBB and URLLC, due to the availability of input data. However, we recognize the importance of evaluating the third use case, mMTC, as it can provide valuable insights into the impact of use case requirements on the cost of 5G network deployment. The mMTC use case has distinct network requirements that may necessitate different network dimensioning and lead to different cost drivers of the deployment, making it an essential aspect to consider for a comprehensive cost-effective deployment of 5G networks. Therefore, extending this research to include the mMTC use case can offer valuable insights and further enhance the practical implications of our findings. The study of mMTC can also help validate the network slicing cost allocation model due to its different KPIs. In addition, the latter model (i.e., network slicing cost allocation model) was simulated only with static traffic per slice, future work should **extend the model to the case where the dynamicity of the traffic is considered**. The type of traffic is crucial in determining the necessary network resources needed for different services. However, relying solely on static traffic analysis may result in network under-dimensioning if average distributed traffic is used or over-dimensioning if peak traffic hours are the reference points. Therefore, it is essential to consider the dynamic nature of traffic over time to accurately dimension the network and allocate resources effectively. The same applies to the resource allocation model, as currently the assignment of the network resources is done statically based on the maximum throughput and traffic carried by the different network functions (see Chapter 4).

Furthermore, **the extension of our analysis to other case studies** would be interesting. For example, evaluating the provision of 5G connectivity to other moving platforms such as trains and ships would be an interesting piece of follow-up work (related to Chapter 3). Additionally, **extending the analysis to other types of regions** by assessing for example the deployment of 5G networks in urban areas to encompass CCAM in the context of smart cities for example (related to Chapter 5), as now it was analysed in the cross-border settings.

In this dissertation, we conducted a techno-economic analysis of 5G networks and their enabler technologies. However, we recognize the importance of also evaluating the socio-economic benefits of the 5G rollout and its potential impact on achieving the Sustainable Development Goals (SDGs).

The SDGs stand for the Sustainable Development Goals. They are a set of 17 global goals adopted by the United Nations in 2015 as part of the 2030 Agenda for Sustainable Development. The SDGs aim to guide the world toward sustainable development by balancing social, economic, and environmental factors. The goals address a range of issues including poverty, hunger, education, gender equality, affordable and clean energy, decent work and economic growth, industry, innovation and infrastructure, reduced inequalities, and sustainable cities and communities among other goals. The SDGs represent a shared vision for the future of our planet and are designed to be achievable by all countries and all stakeholders, including governments, the private sector, civil society, and individuals. The goals are interconnected and require collaboration and innovation to achieve [5].

Therefore, an important **future research trajectory** is to **conduct a socio-economic analysis** of the 5G rollout and its potential impact on achieving the **SDGs**. By considering the broader societal and environmental implications of 5G technology, it would be possible to provide a more holistic understanding of its potential impact and contribute to the ongoing discourse around sustainable development.

References

[1] European Commission. (2023, June 15). State aid: Commission approves €680 million Spanish scheme under the Recovery and Resilience Facility to support roll out of 5G mobile networks in rural areas [Press release]. Brussels.
https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3044

[2] European Investment Bank. (2021). Accelerating the 5G transition in Europe: How to boost investments in transformative 5G solutions [Report].

[3] Sun, P. (2023). 5G Development Around the Globe. In A Guidebook for 5GtoB and 6G Vision for Deep Convergence (Chapter 2). Posts & Telecom Press. https://doi.org/10.1007/978-981-99-4024-0_2

[4] BEREC, "BEREC Report on Infrastructure Sharing," June 2018.

[5] United Nations: Sustainable Development Goals . Retrieved from https://www.un.org/sustainabledevelopment/sustainable-development-goals/