CELTIC-NEXT
Next Generation Telecommunications

Project Report

AIMM

CELTIC-NEXT AIMM Project

WP2: Use Cases and Business Requirements

D2.3

Authors: Adrian Sharples (BT), Rui Inacio (Vilicom), Arman Shojaeifard (InterDigital), Cliff Ellement (ThinkRF)

Project Acronym:	**AIMM**
Project Full Title:	**AI-enabled Massive MIMO**
Project Coordinator:	**Arman Shojaeifard (InterDigital)**
Project Duration:	**24 months (Oct. 2020 – Sep. 2022)**
Submission Date:	**30st September 2022**
Dissemination Level:	**External**

Abstract

This document is the final report of Workpackage 2 (WP2, use cases and business requirements) of the AIMM project. The aim of WP2 is to study the business background to the technologies developed in the entirety of the project. This includes business requirements, commercial viability, ethics, trust and security, and system management methods. The overall objective is to ensure the relevance to business of all other aspects of the project. An initial assumption of the AIMM project is that the target deployment architectures will be open, with the emerging ORAN standards used as a reference.

The linkage between the technical workpackages of AIMM and WP2, involved the definition of a number of use cases. These use cases allow WP2 to quantify the benefit of the output of this technical work, predominantly in terms of financial value. The position of the benefit in the mobile network value chain has been considered by defining the flow of value between functions in the mobile ecosystem. Preliminary models have been put in place to quantify these value flows in terms of cost reduction and revenue enhancement and the necessary input data for each of the use cases have been defined. It is the nature of such models that they can be updated and enhanced in future but the work of WP2 has established a baseline against which further work can provide enhancements.

In addition to financial benefit, this document includes consideration of the safeguards that should be put in place when considering the use of AI in technical solutions and the security considerations of a move to more open access architectures. The activities of most of the technical workpackages and results of AIMM do not involve access to personal data of the end customers but the collection of data to understand and explain unintentional bias has been identified as a requirement for future applications of AI techniques.

Executive Summary

This report describes the work carried out within AIMM Workpackage 2 (WP2) to enable the societal and business impacts of the AIMM project to be described and quantified. The main approach to achieving this is the definition of a number of use cases which the technical work of the project is addressing. In parallel, WP2 will continue through the lifetime of the project to assess the benefits that arise from solutions to the use cases.

The work of WP2 therefore falls into four principal tasks:

- The definition of use cases against which the value of the benefits can be assessed.
- Benchmarking of existing Radio Access Network (RAN) architectures and solutions to provide a baseline against which the output of AIMM can be measured.
- Quantification of the benefits that AIMM brings and where these benefits contribute.
- Consideration of the ethical and security issues arising from the use of AI techniques using data derived from the network and their deployment on more open RAN architectures.

One of the early decisions of WP2 was an assumption that the target RAN architecture should be compliant with the emerging standards of the O-RAN Alliance. Some aspects of AIMM extend beyond the current O-RAN plans and these should be considered as evolutions of the O-RAN architecture.

A set of use cases was defined in the early phase of the project. These use cases have been mapped to the most appropriate AIMM workpackage and formed the basis of a mechanism to relate the developed technical improvements to the resulting commercial and societal benefit.

In terms of network deployment, WP2 considered three main areas: wide area public mobile networks with primarily outdoor installations; private networks targeting industry verticals with a combination of indoor and local outdoor deployments; indoor deployments that could involve neutral host operation of public access networks, private network provision or a combination of both. The differences between these network deployments from both a technical and commercial perspective are described in this document. In addition the availability of different data sources to enable a consideration of benefits is also presented.

A framework for supply and value chains in the mobile network ecosystem has been devised. This framework is used to identify the members of the ecosystem most likely to gain benefit from the output of AIMM and the flow of value through the ecosystem, from the customer, through network operators and service providers, to component suppliers. Different supply chain mapping has been produced for each of the network deployments and the location within the supply chain where the AIMM use cases will contribute have been identified. Examples of this ecosystem and value flows mapping are included in the document.

Consideration of ethical aspects of the use of AI in the RAN has been carried out. Through the project and the potential use of AI that have been considered in other workpackages, it has become apparent that the personal data of end users is not involved and therefore many of the ethical considerations normally associated with AI implementations do not apply. However the possibility of introducing unintentional bias has been identified as a potential issue resulting in the recommendation that sufficient data from training AI algorithms is retained to enable decisions to be audited if bias is suspected.

To create quantifiable values for the benefits arising from AIMM, a number of models have been created. Particularly in the area of revenue generation where a number of different factors result in a customer selecting a particular product or service, a significant number of assumptions are required to enable any form of quantification to occur. This report details the rationale and values for a number of these assumptions, mapping them to the different use cases. In the case of AIMM solutions that result in cost reduction, quantification is easier and this has been based on public domain models of network deployment, most notably that produced by the UK regulator, Ofcom. These public domain models have been extended and example results for different AIMM improvements are included in this document

It is inevitable for a project like AIMM that the outcomes that were anticipated in the early phases of the work have evolved through the project. Some use cases have progressed in more detail than others and an approach to realising solutions has developed in new directions. An example of this is the activity around federated AI in WP5. For this reason the report concludes with recommendations for future areas of study beyond AIMM, identifying areas that, in the opinion of the WP2 contributors, are of significant importance to the realisation of the future network optimisation and enhancement.

Table of Contents

Abbreviations

Abbreviation	Definition
3G	Third generation cellular
3GPP	Third Generation Project Partnership
4G LTE/LTE-A	Fourth generation cellular Long Term Evolution/Long Term Evolution Advanced
5G NR	Fifth generation cellular New Radio
A1	O-RAN interface between Non-RT RIC and Near-RT RIC
AAS	Active antenna system
AI/ML	Artificial Intelligence/Machine Learning
ARPU	Average Revenue Per User
BS	Base station
CAPEX	Capital expenditure
CoMP	Coordinated multipoint
CPRI	Common Public Radio Interface
CQI	Channel Quality Indicator
CSI	Channel State Information
CS-RS	Cell-Specific Reference Signal
CU	Centralised unit
DCI	Downlink Control Indicator
DM-RS	Demodulation Reference Signal
DoS	Denial of Service
DPB	Dynamic Point Blanking
DPC	Dirty-paper-coding
DPS	Dynamic Point Selection
DU	Distributed unit
E2	O-RAN interface between Near-RT RIC and CUs/DUs
eCPRI	Enhanced Common Public Radio Interface
EM	Electromagnetic
eNB	eNodeB (4G LTE/LTE-A base station)
EVM	Error vector magnitude
F1	3GPP interface between CU and DU
FD MIMO	3D/full-dimension MIMO
FR1	Frequency range 1
FR2	Frequency range 2
gNB	gNodeB (5G NR base station)
HLS	Higher-layer-split
IPR	Intellectual property rights
IRS	Intelligent Reflecting Surface
KPI	Key-performance-indicator
L#	Layer number # on the protocol stack
LLS	Lower-layer-split
LOS	Line-of-sight
MAC	Medium Access Control
MDT	Minimisation of drive test
MIMO	Multiple-input multiple-output
MRT	Maximum-ratio-transmission
M-TRP	Multi transmission/reception points
Near-RT	Near-real-time
Non-RT	Non-real-time
OPEX	Operational expenditure
O-RAN	O-RAN Alliance
Open RAN	Ecosystem for open standardised interfaces implementation
PA	Power amplifier
PBCH	Physical Broadcast Channel
PDCP	Packet Data Convergence Protocol
PDSCH	Physical Downlink Shared Channel
PHY	Physical Layer
PRACH	Physical Random Access Procedure

PSS	Primary Synchronisation Signal
PUSCH	Physical Uplink Shared Channel
QoE	Quality-of-experience
QoS	Quality-of-service
RAN	Radio access network
rApp	An application designed to run on the Non-RT RIC
RIC	O-RAN RAN Intelligent Controller
RIT	Radio Interface Technology
RLC	Radio Link Control
RRC	Radio Resource Control
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RT	Real-time
RU	Radio unit
SINR	Signal-to-interference-plus-noise ratio
SISO	Single-input single-output
SLA	Service Level Agreement
SON	Self-organising-network
SRIT	Set of Radio Interface Technologies
SSB	System synchronisation block
SSS	Secondary Synchronisation Signal
TXRU	Transceiver chain
UE	User equipment
vRAN	Virtualised RAN
X2	3GPP interface between eNBs
xApp	An application designed to run on the Near-RT RIC
Xn	3GPP interface between gNBs
ZF	Zero-forcing

1 Introduction

This report describes the work of AIMM Workpackage 2 (WP2) to enable the societal and business impacts of the AIMM project to be described and quantified. The main approach to achieving this has been the definition of a number of use cases which the technical work of the project are expected to address.

Additional work for WP2 includes the potential impact of AIMM on aspects of society. The use of Artificial Intelligence (AI) within Radio Access Networks (RAN) is one of the fundamental activities of AIMM and public perception of this technology is receiving increasing attention in the public domain. To address this issue, a review of the societal requirements for the use of AI within the RAN has been included in WP2.

One of the early decisions of WP2 was to determine the target RAN architecture that would be assumed by the project. This is necessary to provide guidance to the technical workpackages and also for WP2 to identify the position within the mobile network customer, operator and supplier ecosystem where the benefits of AIMM will apply. Since much of the work of AIMM is based on applying flexible and targeted intelligence within the RAN, the decision has been made to adopt the emerging architecture of the O-RAN Alliance as the assumed architecture. Within O-RAN, the introduction of the RAN Intelligent Controller (RIC), provides a function where data derived from the RAN can be collected. New RAN management algorithms provided by the equipment vendor, the network operator or potentially third parties can also be introduced and quickly modified via the RIC. The activities of WP5 in particular make use of the availability of a RIC and the network optimisation capabilities that arise from this.

A set of use cases was defined in the early phase of the project. These use cases have been mapped to the most appropriate AIMM workpackage and formed the basis of a mechanism to relate the developed technical improvements to the resulting commercial and societal benefit.

To enable consideration of the different uses of RAN in different market segments, WP2 has defined three types of network deployment for further consideration considered. These have been defined as:

- Wide area public access networks, initially delivered by tower or rooftop macrocells but increasingly moving to small cell deployments both at street level and indoor.
- Private networks for indoor and outdoor deployments on campuses or large industrial sites.
- Private networks for indoor deployments in office environments, including neutral host solutions to support customers accessing multiple core networks (public and private).

The adoption of O-RAN as a target architecture does lead to an increased number of interfaces within the RAN where new hardware and software components from a diverse range of suppliers can be connected. For this reason, and due to the increase reliance of mobile radio networks for a variety of new services including machine to machine, security of the RAN is considered to be an area for new solutions to be identified and proposed. WP2 has considered where these security concerns will arise and considered mechanisms to mitigate their impact.

The work of WP2 therefore falls into four principal tasks:

- The definition of use cases against which the value of the benefits can be assessed.
- Benchmarking of existing Radio Access Network (RAN) architectures and solutions to provide a baseline against which the output of AIMM can be measured.
- Quantification of the benefits that AIMM brings and where these benefits contribute.
- Consideration of the ethical and security issues arising from the use of AI techniques using data derived from the network and their deployment on more open RAN architectures.

This report describes the work carried out to consider each of these different aspects within WP2. It is inevitable for a project like that the outcomes that were anticipated in the early phases of the work have evolved through the project. Some use cases have progressed in more detail than others and an approach to realising solutions has developed in new directions. An example of this is the activity around federated AI in WP5. For this reason the report concludes with recommendations for future areas of study beyond AIMM, identifying areas that, in the opinion of the WP2 contributors, are of significant importance to the realisation of the future network optimisation and enhancement.

The structure of this document follows is as follows:

Section **Error! Reference source not found.** provides a list of the use cases, how these are mapped to the different workpackages within AIMM and proposes the deployment scenarios and KPIs which WP2 could use as a benchmark for value assessment. In some cases the technical workpackages have chosen to use alternative KPIs and these have been accommodated in the later methods of quantification defined by AIMM.

Section 3 describes the different network deployment scenarios considered in WP2, highlighting differences from both a technical and commercial perspective and how these will be affected by the output of the technical AIMM workpackages.

Section 4 considers the relationships between different stakeholders in the mobile ecosystem, identifying the flow of value across the ecosystem and mapping the AIMM use cases to this ecosystem.

Section 5 considers the data requirements needed to quantify specific benefits for each of the defined use cases in turn. This section also considers the type of modelling required to define benefits and outlines some of the assumptions that feed into this modelling approach.

Section 6 includes a description of the models that have been used within WP2.

Section 7 includes example results as the output of the models and data assumptions presented in the previous two sections.

Section 8 details the ethical considerations that have been reviewed to assess aspects of the societal impact of the use of AI within AIMM.

Section 9 provides information relating to the security aspects of changes to the RAN and customer requirements in the context of AIMM.

Section 10 includes proposals for additional areas of study beyond the conclusion of AIMM that, in the opinion of the WP2 contributors, are of significant importance to the realisation of the future network optimisation and enhancement.

2 Use cases

Use cases are fundamental to a project such as AIMM, representing the link between the technical detail considered by other workpackages and the business and mobile network ecosystem requirements defined by WP2. Early in the project a set of use cases were defined and reviewed.

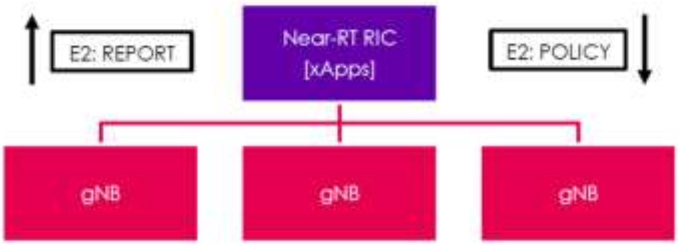
Details of the use cases are included in a project working document (1). Brief descriptions of these use cases are included in the following subsections.

2.1 List of use cases

The identified use cases are listed below.

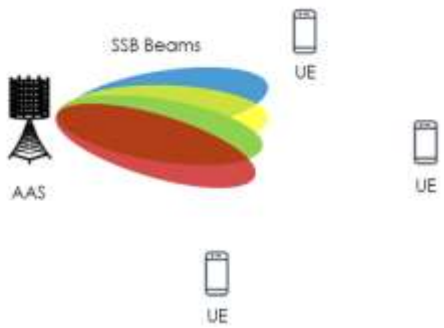
(i) Smart Interference Management for QoS optimisation

This use case is defined around research and development of advanced interference and handover techniques which can be deployed in the form of xApps on the standardised O-RAN near real-time (Near-RT) RAN Intelligent Controller (RIC) and E2 interface.



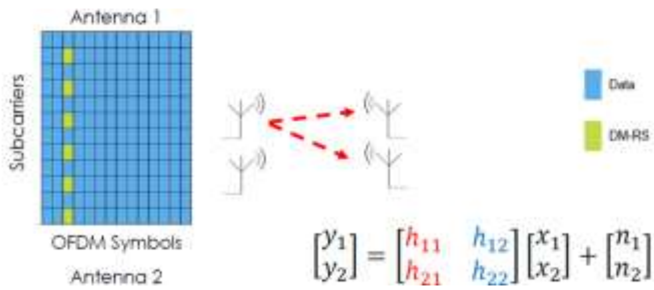
(ii) Broadcast Beam Optimisation for Coverage Enhancements

This use case looks at the optimisation of the Synchronisation Signal Block (SSB) beam patterns using AI techniques trained on real network data.



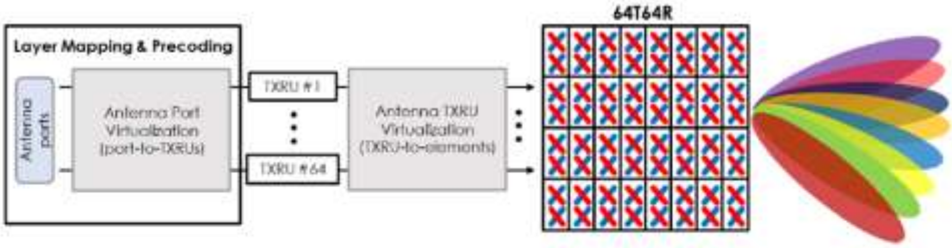
(iii) AI-based Channel Estimation & Detection

This use case focuses on the applications of AI/ML for channel estimation and detection in Massive MIMO systems.



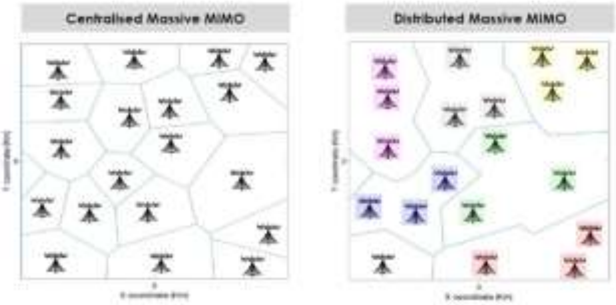
(iv) AI based Massive MIMO Precoding and Scheduling

The main KPI for this use case is spectral efficiency improvement through the use of improved precoding techniques for massive MIMO systems.



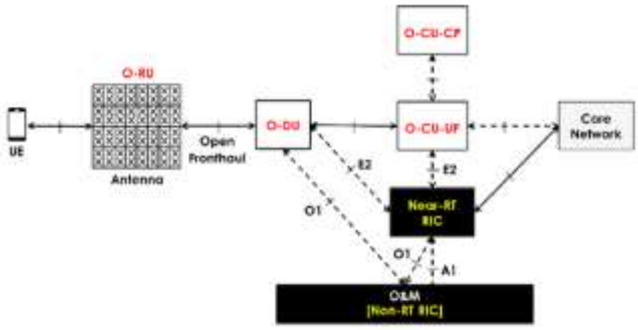
(v) Distributed and Cell-less Massive MIMO

This use case looks at coherent joint transmission solutions to improve spectral efficiency and reliability, potentially moving away from the cell based approach of mobile networks to a more user-centric deployment of RAN.



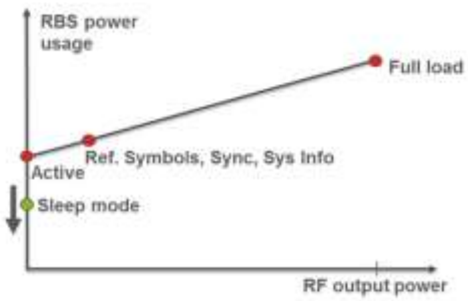
(vi) Disaggregated and Open Massive MIMO

This use case looks at the performance of Massive MIMO in the context of the O-RAN Lower Layer Split (LLS). The impact of radio performance will be considered in the context of required changes to the RAN architecture and deployment techniques.



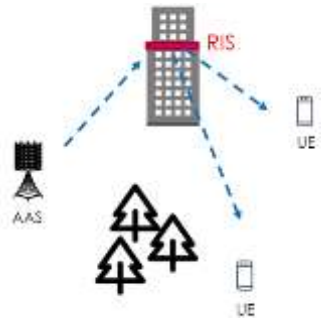
(vii) AI for RAN Energy Efficiency

The main KPI for this use case is RAN energy efficiency. Data derived from live networks will be used to identify the circumstances in which appropriate management of the RAN will enable a reduction in energy consumption.



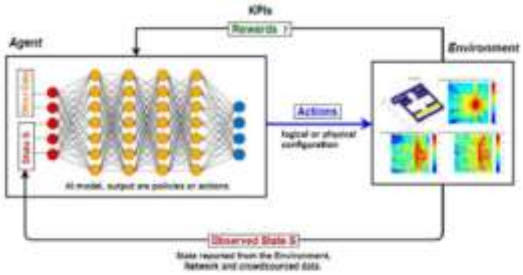
(viii) Reconfigurable Intelligent Surfaces

This use case looks at the interplay between Reconfigurable Intelligent Surfaces (RIS) and conventional massive MIMO antennas. The expectation is that the use of such surfaces can provide coverage enhancement and spectral efficiency improvement.



(ix) Self-organising In-Building Small Cells

AI based, data driven coverage improvement, **using** machine learning techniques to calibrate RF prediction models and identify the best physical system design.



(x) Massive MIMO Combined Pre-distortion Architectures for Power Amplifiers

This use case aligns to Use Case (vii) since it will result in improved energy efficiency

(xi) AI for RAN Security

This use case will consider the use of AI to detect and locate adversarial devices and attacks. The use case will also consider wider security issues relating to open distributed networks.

2.2 Mapping use cases to workpackages

As described above, the use cases are seen as the primary mechanism for relating the work of the technical workpackages to that of WP2. The relationship between the workpackages is shown in **Error! Reference source not found.** below.

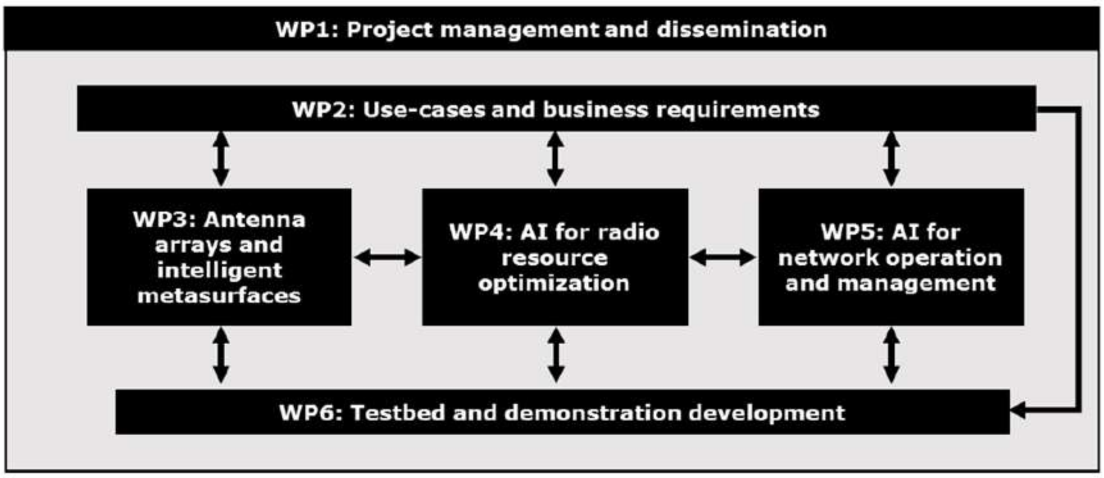


Figure 1. Relationship between AIMM Workpackages

The mapping of the use cases to the workpackage that was identified as most relevant to consider the subject from a technical perspective is shown in Figure 2. In this diagram a distinction is made between primary and supporting workpackages. In general, WP3-5 made the technical contributions to AIMM. Quantification of the benefits arising from the output of WP3-5 is a principal role of WP2 and described in the subsequent sections of this document. WP6 is responsible for testbed implementations to prove the technical solutions. Both WP2 and WP6 are therefore shown as being supporting contributors to all use cases.

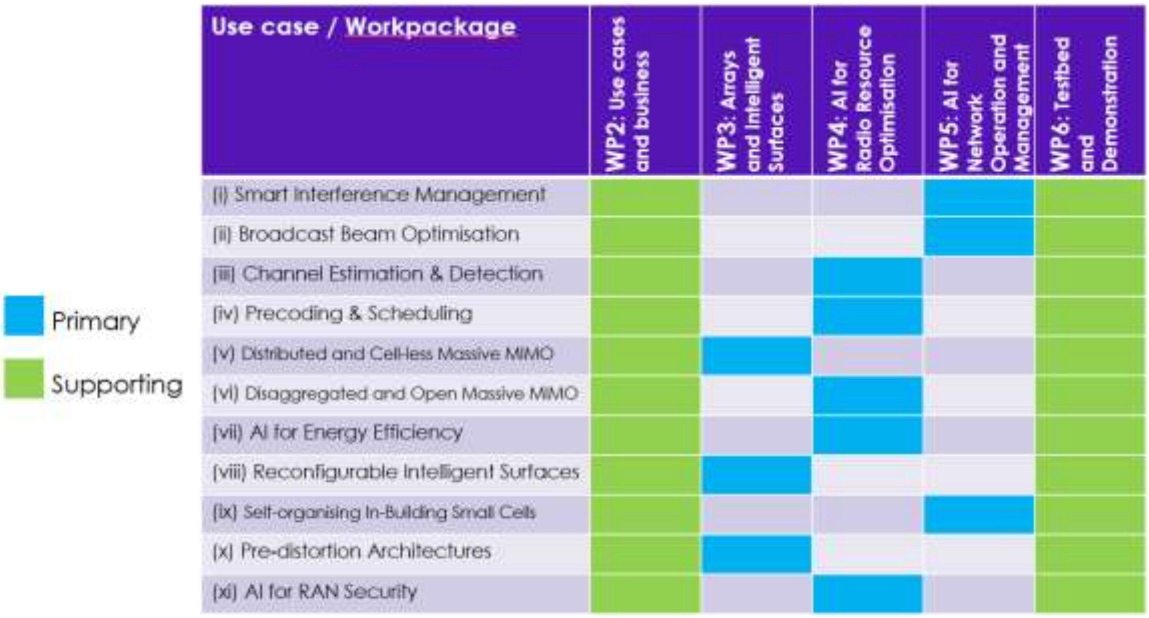


Figure 2. Mapping of use cases to Workpackages

2.3 KPIs for value assessment

To assist in the work of WP2 to assess the value of the AIMM project, a set of objective simulation scenarios have been defined against which the workpackages can be tested. The use of these scenarios was not a mandatory requirement on the other workpackages where several contributing members will already have established techniques. In many cases these existing techniques have been based on proposals by standards bodies, including 3GPP and ITU. To avoid duplication of activities, the simulation scenarios defined by WP2 were based on existing definitions, predominantly 3GPP TR 38.901.

In addition to the simulation scenarios, KPI definitions were defined to provide a common definition across the different workpackage activities. The primary aim of this action was to ensure a commonality of definition for aspects such as spectral efficiency, which is a defining metric for several of the use cases. Here again, WP2 has chosen to align with definitions that have already been agreed with the ITU.

Details of the simulation scenarios and KPIs are included in a working project document (2). The five different deployment scenarios are:

- Urban macrocell (Uma) with antennas placed on rooftops and high masts with a mix of indoor and outdoor users at different levels of mobility.
- Urban microcell (UMi) with a layer that lies beneath the macrocells, with radio units typically located on items of street furniture (e.g. lamp posts), below the height of the surrounding buildings.
- Rural macrocell (RMa) with antennas on high towers in areas with significantly lower traffic density and higher mobility.
- Indoor office primarily aimed at larger office scenarios with a number of radio access points located at ceiling level. The definition of whether the area is open plan office or mixed office space is included in the statistical assumptions of the path loss models.

- Indoor factory where a large open area contains a variety of industrial machines and radio access points are considered to be fitted at different heights within the area.

Some of these simulation definitions have been integrated into the simulator developed for use by AIMM project collaborators as part of WP5.

A summary of the KPIs is presented in Table 1.

Table 1. KPI definitions for value assessment

Criteria	**Assessment Method (ITU-R M.2412)**	**Comment**
Peak data rate	Analytical	Peak data rate is the maximum achievable data rate under ideal conditions (in bit/s), which is the received data bits assuming error-free conditions assignable to a single mobile station, when all assignable radio resources for the corresponding link direction are utilized (i.e. excluding radio resources that are used for physical layer synchronization, reference signals or pilots, guard bands and guard times).
Peak spectral efficiency	Analytical	Peak spectral efficiency is the maximum data rate under ideal conditions normalised by channel bandwidth (in bit/s/Hz), where the maximum data rate is the received data bits assuming error-free conditions assignable to a single mobile station, when all assignable radio resources for the corresponding link direction are utilized.
User experienced data rate	Analytical for simple scenarios, simulation for complex scenarios	User experienced data rate is the 5% point of the cumulative distribution function (CDF) of the user throughput. User throughput (during active time) is defined as the number of correctly received bits, i.e. the number of bits contained in the service data units (SDUs) delivered to Layer 3, over a certain period of time.
5^{th} percentile user spectral efficiency	Simulation	The 5^{th} percentile user spectral efficiency is the 5% point of the CDF of the normalized user throughput. The normalized user throughput is defined as the number of correctly received bits, i.e. the number of bits contained in the SDUs delivered to Layer 3, over a certain period of time, divided by the channel bandwidth and is measured in bit/s/Hz.
Average spectral efficiency	Simulation	Average spectral efficiency is the aggregate throughput of all users (the number of correctly received bits, i.e. the number of bits contained in the SDUs delivered to Layer 3, over a certain period of time) divided by the channel bandwidth of a specific band divided by the number of TRxPs and is measured in bit/s/Hz/TRxP.
Area traffic capacity	Analytical	Area traffic capacity is the total traffic throughput served per geographic area (in Mbit/s/m^2). The throughput is the number of correctly received bits, i.e. the number of bits contained in the SDUs delivered to Layer 3, over a certain period of time.
User plane latency	Analytical	User plane latency is the contribution of the radio network to the time from when the source sends a packet to when the destination receives it (in ms). It is defined as the one-way time it takes to successfully deliver an application layer packet/message from the radio protocol layer 2/3 SDU ingress point to the radio protocol layer 2/3 SDU egress point of the radio interface in either uplink or downlink in the network for a given service in unloaded conditions, assuming the mobile station is in the active state.
Control plane latency	Analytical	Control plane latency refers to the transition time from a most “battery efficient” state (e.g. Idle state) to the start of continuous data transfer (e.g. Active state).
Connection density	Simulation	Connection density is the total number of devices fulfilling a specific quality of service (QoS) per unit area (per km^2).
Energy efficiency	Inspection	Network energy efficiency is the capability of a radio interface technology to minimize the radio access network energy consumption in relation to the traffic capacity provided. Device energy efficiency is the capability of the

		RIT/SRIT to minimize the power consumed by the device modem in relation to the traffic characteristics.
Reliability	Simulation	Reliability is the success probability of transmitting a layer 2/3 packet within a required maximum time, which is the time it takes to deliver a small data packet from the radio protocol layer 2/3 SDU ingress point to the radio protocol layer 2/3 SDU egress point of the radio interface at a certain channel quality.
Mobility	Simulation	Mobility is the maximum mobile station speed at which a defined QoS can be achieved (in km/h).
Mobility interruption time	Analytical	Mobility interruption time is the shortest time duration supported by the system during which a user terminal cannot exchange user plane packets with any base station during transitions.
Bandwidth	Inspection	Bandwidth is the maximum aggregated system bandwidth. The bandwidth may be supported by single or multiple radio frequency (RF) carriers. The bandwidth capability of the RIT/SRIT is defined for the purpose of IMT-2020 evaluation.

3 Network Deployments and Architectures

3.1 Network build

In the introduction to this document it was noted that different types of network deployment would be considered including:

- Wide area public access networks, initially delivered by tower or rooftop macrocells but increasingly moving to small cell deployments both at street level and indoor.
- Private networks for indoor and outdoor deployments on campuses or large industrial sites.
- Private networks for indoor deployments in office environments, including neutral host solutions to support customers accessing multiple core networks (public and private).

Each of these network types have different characteristics in terms of the rate of network build and the requirement to upgrade over time. As a result the benefits of the AIMM project are will apply differently to the network types. The remainder of this sub-section will consider the factors that affect the network evolution.

The approach to building a wide area public RAN is to assume that base station sites are constructed for one of two purposes. In the early stages of network rollout, sites are quickly installed to provide coverage over the whole of the service area. With this coverage comes an inherent level of capacity that, in the early days of deployment, is far greater than the traffic that could be offered by the customer base since these customers will not buy new devices if the coverage is not available to justify their purchase.

As the number of users coming onto the system grows, and the individual usage increases due to the new applications that are available on the network and devices, the capacity of the coverage layer becomes exhausted. The fundamental capacity of the layer is due to the technical capability of the network and the volume of spectrum. Spectrum is considered in more detail below.

Once the initial coverage layer is exhausted, the second reason for base station site construction, namely the provision of additional capacity is necessary. In the absence of alternative techniques to increase capacity, a growth in traffic necessitates the construction of additional sites. This growth in site numbers can be delayed by the release of additional spectrum or the improvement of the spectral efficiency of the existing system.

For most public access mobile networks in developed countries, the initial build for coverage was completed by the end of the 20th century. Exceptions to this do occur when a public policy initiative from government aims to extend the coverage to remote areas that are not commercially viable. For private networks, the design lifetime, the user base and the application requirements are more easily specified at the time of network build. This enables the initial coverage phase to have the necessary capacity designed into the initial system.

Calculating the number of sites required for network build is also affected by other practical considerations. The first of these is that the range of a base station will vary depending on its environment and the coverage requirements. Coverage provision in urban areas is affected by the shadowing of locations by the density of buildings in the area. Similarly the buildings in urban areas are larger, with more signal loss experienced to achieve coverage to all locations within a building from external base stations. Both of these factors reduce range. In more rural areas there is little blockage by buildings and the margin required to compensate for signal loss when penetrating a building is lower, resulting in larger base station ranges. However in all areas, local building regulations can impact the height of the base station installation.

For public network coverage build it is normal to provide the sites with the largest coverage area, resulting in macrocells on masts or high buildings. The subsequent build of sites for capacity, has also traditionally involved additional macrocells since this site type provides the largest coverage area and so covers sufficient traffic to justify the cost of build. Ultimately, in city centres, the point is reached where there are no suitable rooftops remaining or the density of macrocells is such that interference between them reduces their traffic carrying capacity. The established route to achieve greater capacity is therefore to build additional base stations at street level. These small cells, or microcells, have their coverage constrained by surrounding buildings and so have a smaller area over which to capture traffic. Similarly the cost of a small cell is sensitive to the cost of acquiring a suitable site and providing power and backhaul connectivity back to the core network. For this reason the deployment of small cells varies significantly between different countries where regulatory and commercial pressures differ, altering the cost of deployment.

For private network build, for example on a campus or industrial estate, where the deployment of base station sites and provision of connectivity is controlled by a single organisation, small cells could be deployed without an initial macrocell layer if this is justified by the density of traffic.

Finally, in-building systems provide a high level of capacity with multiple indoor small cells, or picocells. These also improve coverage within the building and generally involve working with a single organisation to gain access to a site. If the in-building system is to provide service for public access networks then these are generally shared, or neutral host, systems with the cost of installation being shared between the public network operators or provided by a neutral host network provider through an agreement with the building owner. The provision of such neutral host systems is one of the target deployment models for AIMM.

Within AIMM, several of the activities of WP3, 4 and 5 consider improvements to spectral efficiency, which will improve the user data carrying capacity of radio spectrum and so improve the utilisation of the base station and spectrum assets of the network operator. WP3 also considers changes to the network architecture, introducing new network nodes including reconfigurable intelligent surfaces, which will reduce the need for new base stations by improving coverage and capacity of an existing base station network.

3.2 Network architecture

A traditional RAN architecture is one where the RAN assets are provided by a single company in a geographic area and, in the case of 4G, these assets are located at the base station. As mentioned in the introduction, the assumed target architecture is the disaggregated, virtualised OpenRAN approach currently being specified by O-RAN. The difference between the two approaches is illustrated in Figure 3 where, in the upper diagram, the RAN comprises a 4G eNodeB at a macrocellular site with 3GPP defined backhaul connecting back to the core network. The near real time control functions of the RAN are embedded in the eNode B and fall in the domain of the main RAN equipment supplier.

In OpenRAN, the RAN is disaggregated and virtualised. The functions of the RAN are split into Radio Unit (RU), Distributed Unit (DU) and Centralised Unit (CU). The O-RAN defined interfaces between these components, fronthaul, midhaul and backhaul, lead to the possibility that the hardware could be supplied by different vendors, as indicated by the different colours shown for these components. Virtualisation also enables the DU and CU to be partitioned between hardware and software, with software components from different suppliers being deployed on generic hardware or cloud platforms. Given the uncertainty around the potential cost of these different sub-components, for the purposes of AIMM it will be assumed that RU, DU and CU are individually supplied as complete units combining hardware and software.

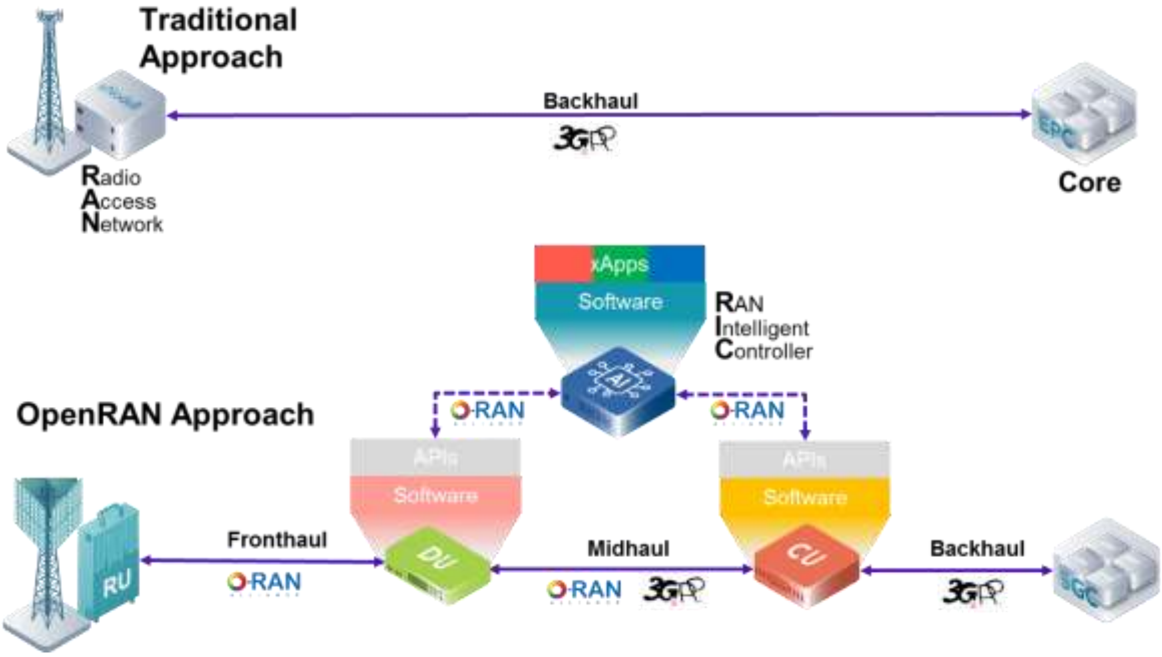


Figure 3. The difference between traditional and OpenRAN architectural approaches

Also shown in Figure 3 is the RAN Intelligent Controller (RIC). This O-RAN defined function will gather network measurement data from other components that can be used to train AI algorithms. These algorithms are themselves installed on the RIC in the form of xApps, which can be supplied by the equipment vendor, the network operator or a third party optimisation specialist. This is illustrated in Figure 3 as different colours representing the different sources of xApps. Many of the functions developed by WP5 would make use of the data collected by a RIC and then be implemented as xApps.

With the disaggregation possible for the RAN, the positioning of the different components brings in the possibility for centralisation and sharing of resources between nodes further to the edge of the RAN. This is illustrated in Figure 4, where multiple RUs are shown connected to a single DU. In turn multiple DUs are connected to a single CU.

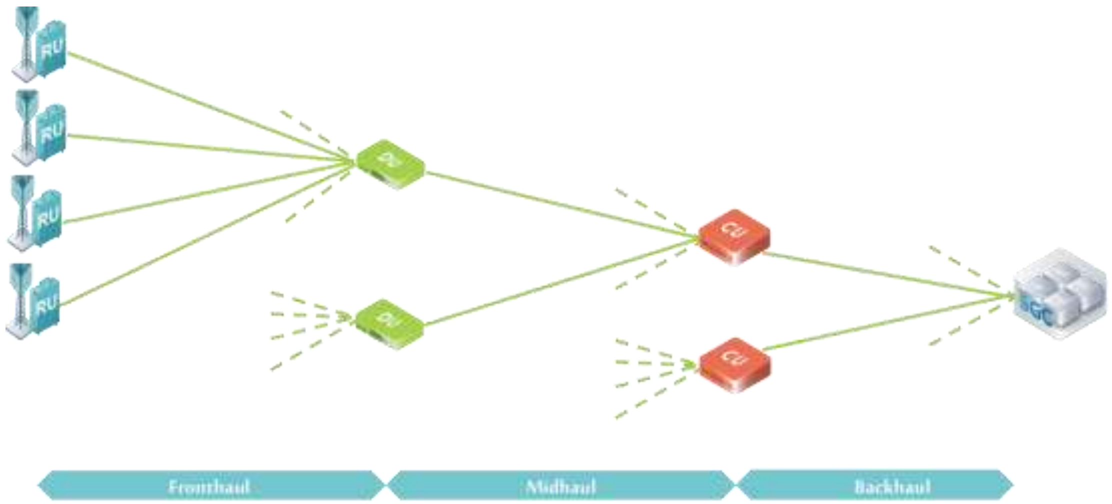


Figure 4. Node hierarchy in OpenRAN

In reality it is assumed that, to achieve collocation benefits, disaggregated CU units will be located with the core network nodes. This mimics the approach followed by 3G where, in many countries, the Radio Network Controller (RNC), logically part of the RAN and controlling multiple base station installations, was often physically located with the core network nodes.

The disaggregated approach of OpenRAN also provides a mechanism for providing effective in-building coverage, with multiple RUs covering the floor of a larger building and routing back to a shared DU, which itself is connected back to a CU. In many instances it is expected that some core network functionality will be located with the CU to enable traffic to leave the network and access local applications. This concept is illustrated in Figure 5.

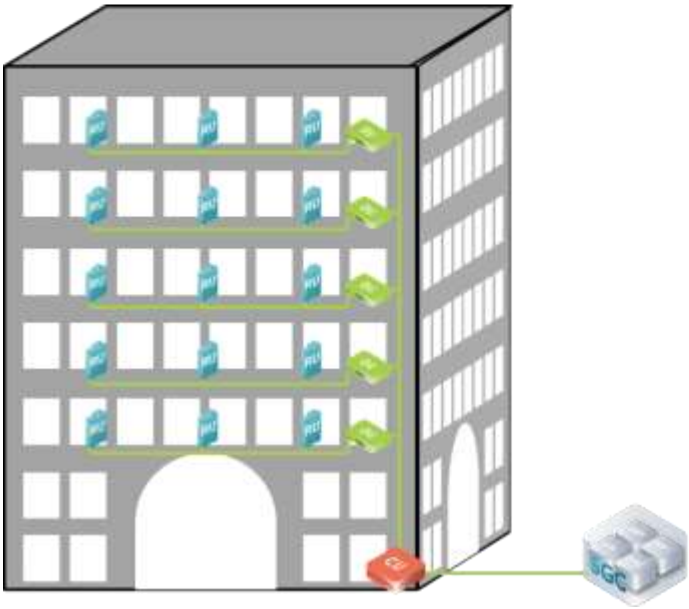


Figure 5. RAN disaggregation for in-building systems

In particular with such in-building systems, the O-RAN approach with virtualisation of the network functions is suitable for neutral host implementations. In this case the same hardware implementation can support multiple software instances of the DU, CU and RIC, with each operator having full control of the services and settings on their segment of the shared physical infrastructure. This reduces one of the drawbacks perceived with neutral hosts systems which is a reduction in service differentiation between operators.

Sharing of network assets, mainly physical, also takes place on existing public wide area networks. The benefits model therefore allows for sharing of network assets on different technologies between different operators to be accounted for in different geographies and at different levels in the RAN network e.g. sharing macrocells but not microcells.

3.3 Subscriber and traffic growth

Appropriate treatment of subscribers, which in turn feeds into the overall growth and distribution of traffic across a network is primarily an issue for public access networks. In private networks there are no flows of subscribers between competing operators, fewer legacy networks and end user devices that are often supplied directly as part of the service.

Accounting for the movement of subscribers between different levels in the network is performed using four principal metrics:

i) The total subscriber base in terms of the total number of active SIMs in the market, expressed in terms of market penetration of the whole population. This will vary with time as new devices are brought to market.

ii) The market share of different operators and how this varies in time. In many cases this is difficult to determine since it is dependent on many factors, including the success of marketing campaigns. For the purposes of AIMM it is assumed that the total market share is split evenly between competing operators at all times.

iii) The churn rate of the market, expressed as the number of subscribers who make a new device purchase decision in every year. A consideration of how minimising churn can alter the reveue generating capability of a network is included in Section 7.

iv) The proportion of subscribers making a purchase decision in any year who would choose a specific device with different network capabilities.

As an example for point (iv), in the early phases of a 5G rollout, 5G coverage will be limited and the devices will typically be charged at a premium price. As a result, very few customers making a purchase decision will choose to purchase a 5G handset, preferring a 4G device where there is more choice and competition has forced prices lower. As the 5G coverage progresses, the number of customers choosing a 5G device when their contract comes to an end will increase. Even when 5G is available there will remain some customers who will choose a very low cost 2G handset for voice if this continues to be supported on the network.

The combination of the four metrics above will dictate the speed of movement of the customer base between technologies and so the timescales in which new techniques will realise benefits that rely on device availability. Historically in the UK, it has taken 5-7 years for half of the customer base to adopt a new 3GPP generation.

In addition to the speed of adoption of a new technology by the customer base, there is usually a financial requirement for a public operator to depreciate its assets over a period of time to reflect the value of the network in its accounts. To replace a piece of equipment before its value has been depreciated in the accounts would require the operator to realise a financial loss. If a new technique or capability requires a hardware change in the network then, unless the benefits are sufficient to offset the financial loss from early replacement, the new capability will need to wait until the existing equipment has been fully depreciated. The timescales over which a piece of equipment are depreciated does vary by the nature of the equipment and the rate at which technology is advancing and therefore making the equipment obsolete. For RAN network equipment, the depreciation timescale is typically in the region of eight years.

The speed of migration of the customer base and the requirement to depreciate existing network assets will affect the rate at which the new developments from AIMM will be deployed and provide benefit into a public RAN.

To calculate the total traffic load on the network from the subscriber base, an estimate of the average traffic generated for each device type, e.g. 4G, 5G handsets and machine to machine devices, is required. For data services these are normally expressed in gigabytes per month since this is often the way in which the network contracts are marketed with specified data bundles. Several public domain sources exist providing estimates of the traffic per subscriber into the future, e.g. CISCO VNI, and these have been used in AIMM to consider the timing of benefits arising from new capabilities.

3.4 Spectrum and capacity

As mentioned in Section 3.1, the available spectrum will have an effect on the base station range that can be achieved but primarily will have an impact on the capacity can be achieved on a base station. Spectrum availability will also depend on how many legacy networks, each with their own spectrum requirements, are operating at any moment in time and how much traffic they are required to carry. As customers purchase devices that can access newer generations of equipment, the spectrum supporting the legacy generations will become vacant and so can be repurposed, or “refarmed”, from one technology generation to another.

The previous section described how the model will account for the traffic that must be carried on the different technology generations at each point in time. The model can also enable traffic to be refarmed from one generation to another, in addition to any assumptions that can be made about new spectrum being made available, typically through a government auction.

The refarming and acquisition of new spectrum is often the lowest cost method by which an operator can increase the overall capacity of the RAN. By transferring spectrum from less efficient legacy networks, such as 3G, to more efficient 4G or 5G, the operator can achieve an instant uplift in the ability to carry traffic since most devices 4G and 5G devices in the customer base will already be able to operate in these legacy frequency bands. Several of the AIMM use cases target increased capacity, through improved spectral efficiency. The timing of the ability of these techniques, compared to the timing of the capacity requirement that cannot be met through existing approaches, will determine the point at which the new benefits will be achieved.

Spectrum in the low and mid-range frequency bands is increasingly scarce and any new spectrum is expected to be made available at higher frequencies. These frequencies have inherently shorter range and so the standard technique of deploying new spectrum on existing macrocell sites is less appropriate. For this reason high frequency, including mmWave, installation is more likely to be in the microcell and picocell environment. For the moment these high frequencies will not be considered in the AIMM benefits model.

Many of the aspects of AIMM concern the improvement of spectral efficiency, using advanced techniques to increase the volume of user data that can be transferred through a given amount of spectrum. Within the model there is an opportunity to define the average data throughput that can be transferred through a unit of spectrum. This figure can be different in terms of technology, network layer (macrocell, microcell and picocell), geotype and also vary over time as new features and capabilities evolve over the lifetime of a technology generation. By applying this spectral efficiency figure and taking account of the volume of spectrum at different base stations following new spectrum acquisition and refarming, the ability of the network to absorb the predicted traffic growth can be implemented. This aspect will be of greatest applicability to the public access networks where traffic growth and penetration of device types into the customer base is outside the direct control of the operator. The same features will also be used to consider the upgrade path of private network deployments.

4 Benefits and flows of value

It is the responsibility of WP2 to quantify the benefits of the AIMM results to the telecommunications industry and society as a whole. Within WP2, a template has been created to consider how value flows through the mobile network ecosystem. Depending on the precise nature of a deployment, the RAN architecture used and the contractual obligations between the different ecosystem members there can be a wide range of different configurations. For this reason, different versions of the template have been produced for different deployment scenarios.

In the following subsections, examples are presented of the form of the ecosystem template to illustrate how these have been applied to the use cases of AIMM. The purpose of WP2 is to quantify benefits of the work of AIMM and this is achieved by modelling the interactions between the entities in the ecosystem template to determine where new value arises.

4.1 Ecosystem template - Public access network

Figure 6 below shows a version of the ecosystem template. In this diagram a number of ecosystem classifications are included:

- Customer – the end user who provides the revenue into the ecosystem and gains personal and societal benefits
- Service providers – the organisation that contracts with the end customer
- Technical services – services provided to the services provider from a variety of sources, including in-house. These are both network build and operate services.
- Systems – part of the network comprising a set of components
- Components – the fundamental components contributing to a system
- Suppliers – providers of the technical services or components

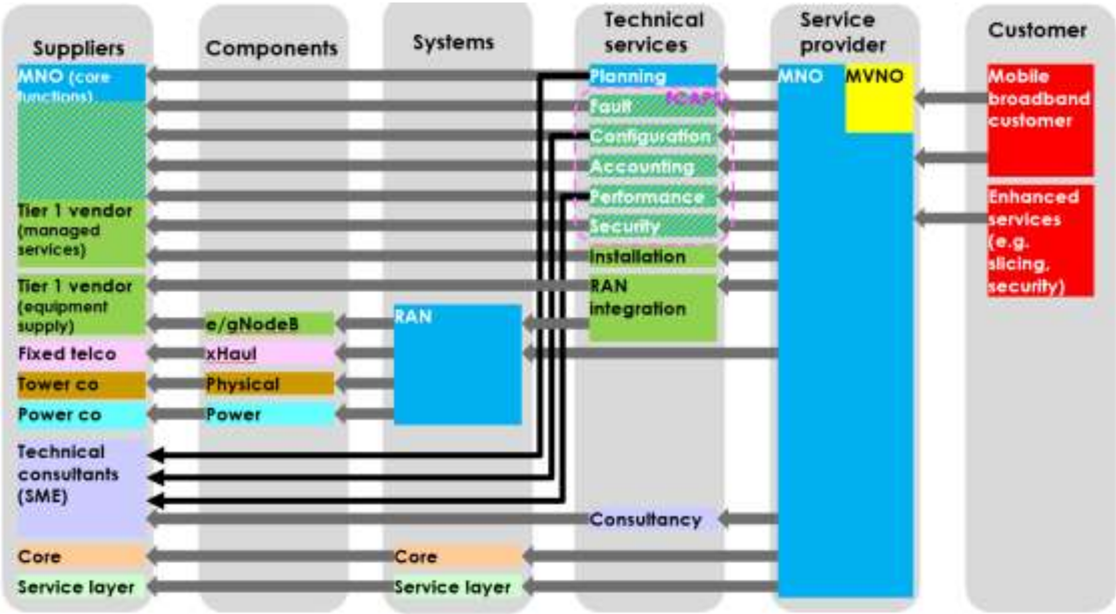


Figure 6. Ecosystem template for a benchmark public access network

In this model, value flows from the customer at the right through the various ecosystem classifications. Value is retained in some of the classifications since all members of the ecosystem must remain financially viable. The colouring of the blocks in the chart represents elements that are the responsibility of the same organisation. Where an element is shaded by multiple colours this shows that the result could vary depending on contractual negotiation. For example, in Figure 6, the Mobile Network operator (MNO), maintains responsibility for the RAN and strategic planning activities. Other operational and installation activities could be the responsibility of the MNO or could, as is often the case in deployed networks, be subcontracted to the main equipment vendor. Since the focus of AIMM is the RAN, most elements relate specifically to radio aspects. Core network and service network layer facilities are included to indicate that network costs other than the RAN are required to be paid by the network operator.

With this ecosystem template, the activities of AIMM can influence the value that flows along the arrows in the diagram . To illustrate this, the use cases presented in Section 2.1 are also mapped to the value flows in Figure 7 where the different use case labels, UC(i) – (xi) are associated with the different value flows. This mapping can be converted to consideration of the activities of the AIMM workpackages by referring to Section 2.2.

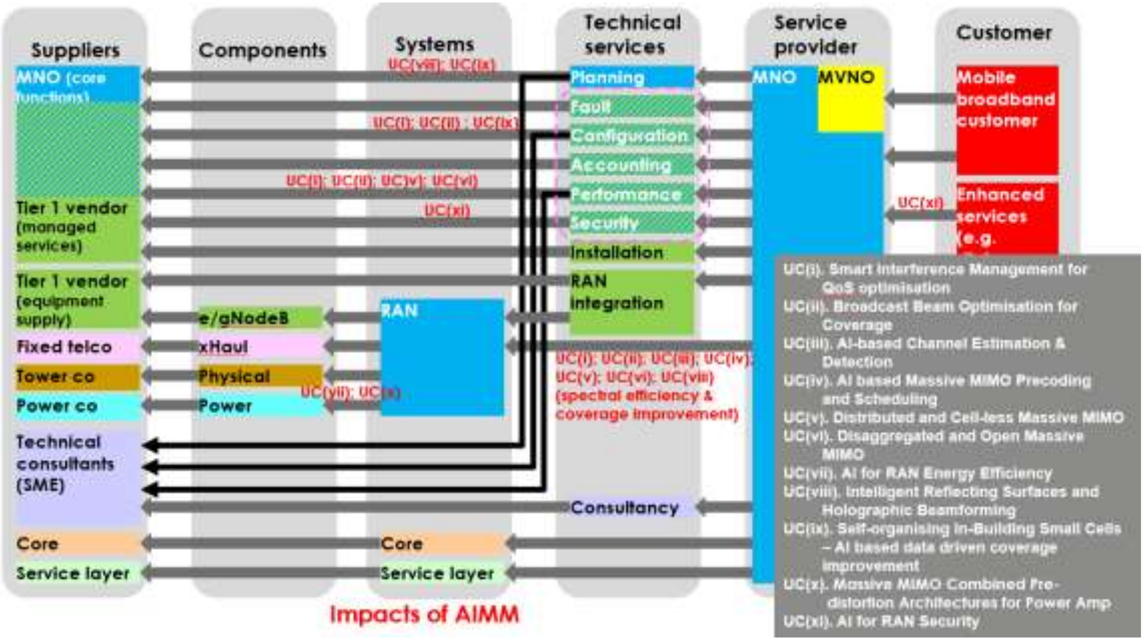


Figure 7. Mapping of the use cases to value flows in the ecosystem template

As a value flow enters a category from the right, it represents revenue to that category. Flows that leave from the left of a category represent a cost. In some cases the cost will be internal. For example, in Figure 7, a flow exists that passes from the Mobile Network Operator (MNO) to the Radio Access Network (RAN) indicating that the RAN is a cost to the MNO and is a revenue to the RAN provider. In this example the RAN as a system is the responsibility of the MNO and is coloured blue, reflecting that the RAN is an internal cost to the MNO.

The mobile network ecosystem can be further complicated depending on the contractual relationships between the different companies involved. Again in Figure 7, there is a flow from the MNO to a RAN integration function and from there to the RAN. In this case the RAN integration is subcontracted to the main RAN supplier and so is coloured green to reflect that the ultimate supplier is the Tier 1 equipment vendor. However RAN integration requires RAN components to integrate and so there is a flow from the RAN integration category to the RAN system itself. This does lead to a circular responsibility that shows that the diagram sometimes requires flexible interpretation. Within WP2 it is considered that representing the ecosystem in this way, although imperfect, does provide a framework for assessment that is valuable to the project as a whole.

Many of the use cases within AIMM relate to an improvement in efficiency of operation, either through spectral efficiency or through improved automation of the management functions. For this reason, many of the use cases are mapped to the value flow between the MNO and the RAN. By improving efficiency, the cost of the RAN reduces for a given level of capability. This capability can be expressed in terms of more capacity from the deployed base station and spectrum assets. These improvements in efficiency can be quantified using the KPI definitions of Section 2.3.

The reduction of the cost of the RAN to the MNO could be retained by the MNO to increase profitability or, potentially, could be passed to the customer by having a reduction in the flow between the customer and MNO. If the saving is passed on this represents a reduced cost to the customer and therefore reduced revenue to the MNO.

Where AIMM provides for new services and therefore new revenue, the flow of value into the MNO will increase. The cost of providing the new service will appear in one of the additional flows, most likely in the form of increased RAN equipment and core network costs. It should be noted that increased revenues from the customer can result from direct payment for the additional service or alternatively from the availability of a new service proving more attractive than the offering of competitor networks and leading to an increase in market share. It is this indirect growth in customer revenue that is more likely to arise from improvements to quality of experience. This approach is used to consider the value of network improvements on revenue in Section 5.2.1.

The value flow template was used extensively to identify the different flows in value affected by the use cases and applied to different workpackages, as described in Section 5.

In the following sub-sections, different models will be considered for the deployment of the RAN infrastructure. This will be followed in Section **Error! Reference source not found.** by a consideration of the values that lie in the different value flows.

4.2 Ecosystem template – O-RAN public access

As described in the introduction, one of the early decisions of WP2 was that the target RAN architecture to be considered in AIMM is that arising from the O-RAN standardisation activities or potential future evolutions of that standard. Within the context of AIMM, one of the primary reasons for that choice is the ability to use AI and ML techniques to optimise specific networks through the capabilities introduced with the RAN Intelligent Controller (RIC). This is particularly the case for the work of WP5. Claims for cost reduction in RAN deployment by increasing the supplier base through more open architectures are primarily an effect of market economics that lie outside the scope of AIMM.

The introduction of O-RAN compliant components change the ecosystem template for a public access network to that shown in Figure 8. The principal difference between this and the benchmark template of Section 4.1 is that the eNodeB / gNodeB have been replaced by the Centralised Unit (CU), Distributed Unit (DU) and Remote Unit (RU) along with the introduction of the RIC. Although these functions are shown as discrete components with single suppliers, in reality these functions could be a mix of hardware and software components. Particular with the CU, this hardware can be shared with other functions of the architecture that are part of the core network, to enable the break out of user traffic close to the edge of the network. In addition the RIC also provides for the support of management applications, xApps for near real-time applications and rApps for non-real time applications. These applications could be provided by the RIC supplier, by the MNO themselves or by a third party. Several of the AIMM use cases are expected to be realised by the availability of these xApps and rApps.

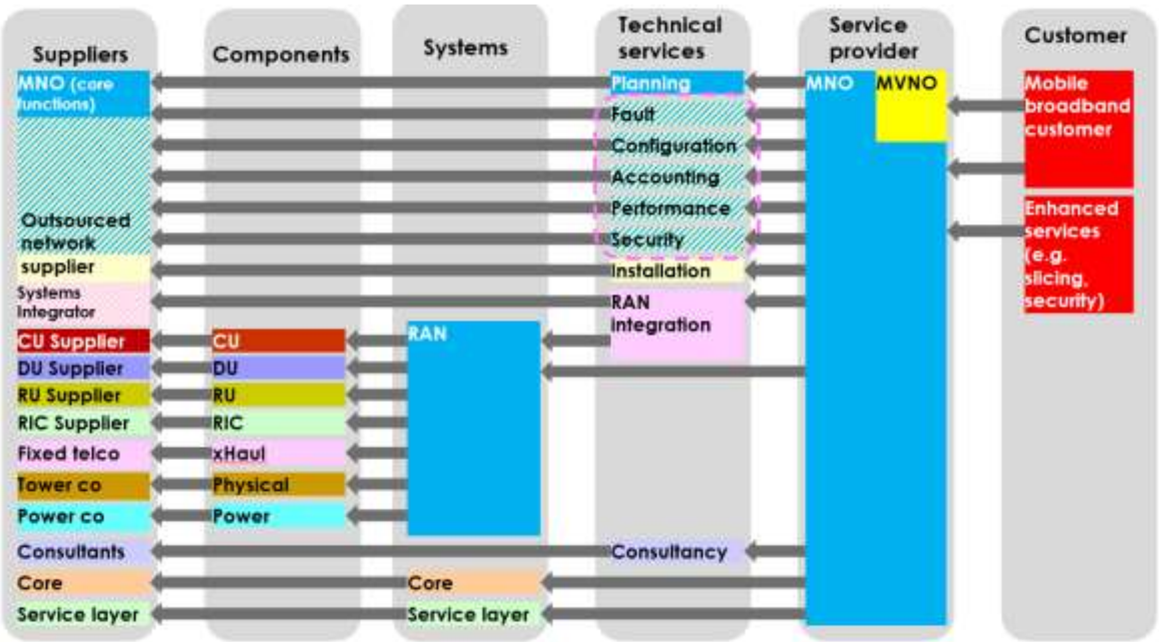


Figure 8. Ecosystem template for an O-RAN compliant public network

Although the xHaul remains shown as a single component supplied by a fixed line telco, depending on the physical location of the CU, DU and RU functions, aspects of this xHaul could be supplied by the MNO within MNO buildings.

An additional change compared to the benchmark ecosystem template of Section 4.1 is the alteration of some other supplier functions. One of the most important of these is the RAN integration. In traditional deployments this is usually carried out by the main equipment vendor in a particular geography, who selects components from their own portfolio of products. With the O-RAN architecture the possibility exists for different RAN components to be selected from different suppliers to meet individual deployment requirements. Integration of these components to deliver a solution is now ultimately the responsibility of the MNO themselves. Although it is not yet clear exactly which model of operation will be dominant in public access networks using the O-RAN architecture, within Figure 8 it is assumed this is a separate system integrator, with the system integrator functions coloured in pink. It is also assumed that the operation of the network is carried out by an outsourced network operating company. In this illustration the MNO itself has reduced to an organisation that outsources most of its network functions, with only strategic planning activities remaining as core functions. A more detailed discussion of the potential business models for future public access networks is outside the scope of AIMM.

4.3 Ecosystem template – O-RAN private networks

A final example showing how the ecosystem template could be used in different deployment models is presented in Figure 9. Here, an O-RAN compliant private network is supplied to an end customer.

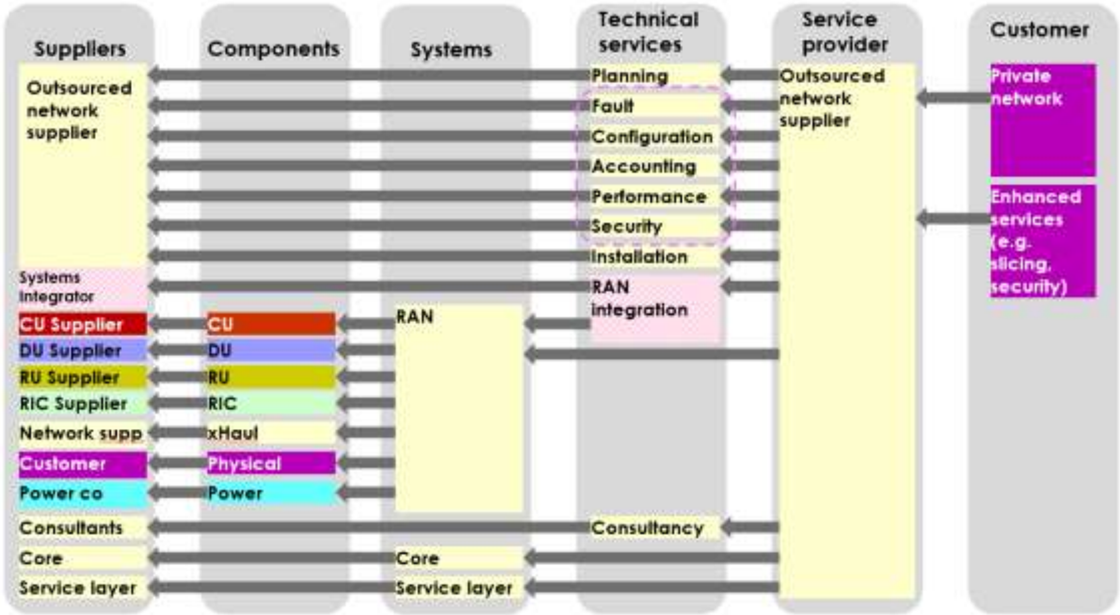


Figure 9. Draft ecosystem template for a private network deployment with an O-RAN compliant architecture

In this case, the private network supplier could provide most of the elements of the deployment themselves, as shown by the extensive expanse of yellow boxes. The private network supplier could also perform the systems integration although, in the example in Figure 9, the possibility has been included that this would be carried out by a separate systems integration organisation.

In the private network, the customer is a single organisation and so the provision of new services will provide new revenue directly. There is no scope for increasing market share by enhancing the network once it has been deployed. The ability to quickly modify and enhance the RAN through its lifetime, primarily by altering techniques with the RIC, will however increase the attractiveness of the private network at the time of initial purchase.

To quantify the value of the different AIMM use cases requires a view of the point in the value chain where these will have an effect. This is the approach that has been adopted in Section 5. As mentioned in an earlier section, changes to network efficiency will reduce the costs in the RAN, a benefit that might be passed in part to the end customer.

5 Input Data and Assumptions

5.1 New revenue opportunities

The research of new revenue opportunities focused on identifying the potential for monetisation and to capture value across the ecosystem and supply-chain continuum. Figure 10 depicts the opportunities identified in WP2 research for AIMM. This diagram shows, following a demand and supply logic, how market opportunities (including both market push and market pull generated opportunities) is linked with possible new business models, new technical services, new systems and new supplier opportunities.

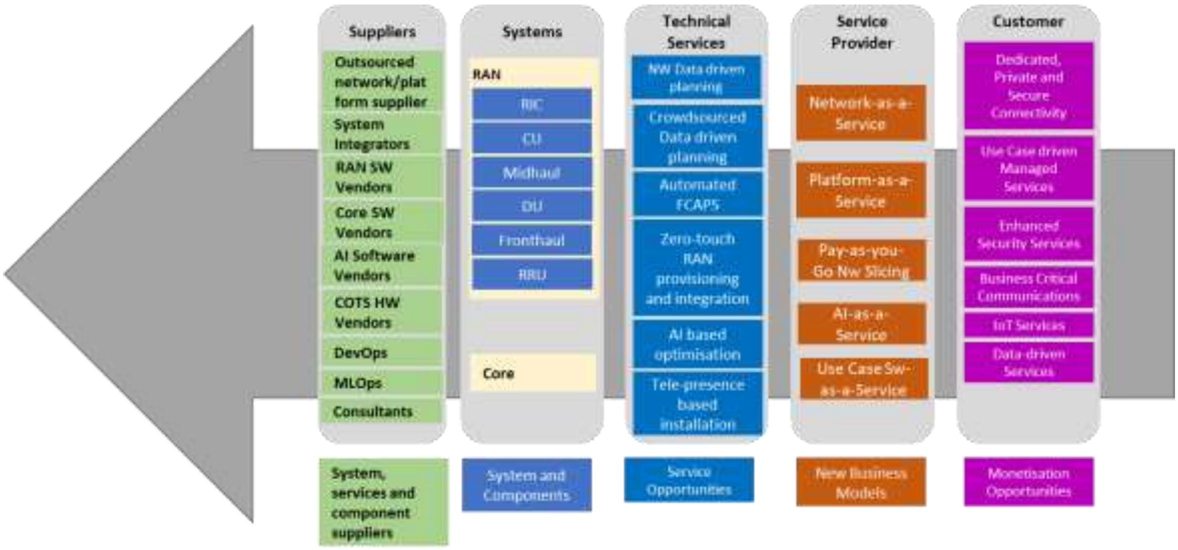


Figure 10 - Monetisation opportunities across the ecosystem and supply-chain

Customer (Market Pull and Market Push)

- Customers going through the process of digital transformation and automation will pull for new types of services that combines the power of wireless communication to underpin digitisation use cases and requirements. In other cases, new innovations coming from the CSP side will generate products/services that create extra value to the end-customer, thus enabling the creation of market differentiation for the CSP.
- Dedicated, Private and Secure Connectivity
  - Customers looking for premium Wireless Connectivity services that offer them assured coverage, system capacity and capability dimensioning tailored as per their requirements and dedicated to its users, ensuring mobile communications that are private and secure through the use of a system that relies on well accepted and tested standards and offers future and feature proof roadmaps.
- Use Case driven Managed Services
  - The customer may be looking to buy a service that addresses a business, industrial or organisational digital use case, that might require to allows be coupled with high capability, premium quality and assured experience mobile connectivity. This allows for the Communication Service Provider to offer the market differentiated services/products, to the end customer, that combine network services, with data collection and Software-as-a-Service helping organisations on their digital transformation journeys.
- Enhanced Security Services
  - With the advent of Digital Transformation of organisation's administrative, operational, and productive processes, it is fundamental for the customer of a mobile communication service to rely of security and privacy of communications. There will

be customers that due to the nature of their business, industry and related activities will require extra levels of assured security this including at Cyber and Radiofrequency layers.

- Business Critical Communications
    - Business Critical communications demand from the mobile service extra levels of reliability for the most relevant KPIs from a Use Case point-of-view. This requires a great level of flexibility, from the network, regarding the functions' configurability, programmability, and operability. The requirements for this type of communication call for a self-organised, cognitive, and self-healing network.
- IoT and Data-driven Services
    - Digital transformation and automation of business processes, operations and services rely heavily on the use of sensors, data collected and machine-type communications. The demand for these services/systems raise opportunities for the development of new partnerships between CSPs and IoT and Software providers to create new products and services that blend all these components to the offered value-proposition.

Service Providers:

- Existing service providers are encouraged to use new technologies to address new market pull opportunities or to push to the market new products, services and business models that will generate extra value for customers. The reluctance of some of the more established CSP players in adopting such technologies and market positioning will create opportunities for new entrants.
- New business models that integrate the different technology components, capabilities and competences will be created to deliver products and services that address the market opportunities with the right functionality, cost structure and value proposition.

Technical Services:

- These new business models generate demand for: new types of technical services; software products/features; new hardware technologies that enable innovations in terms of network/service design and deployment and for the exponential growth of AI-based systems and services.

Systems and Suppliers:

- New technologies enable the emergence of new network systems and service components that integrated enable the delivery of the business models and related technical /technology capabilities and competences, which creates the opportunity of suppliers pushing innovations across the supply chain. On the other hand, the CSPs' new business models will demand from the supply chain a special focus on new components such as O&M and Self-healing capabilities, Network Management automations and Configuration/Functionality Auto-tuning, which increases the demand for the use of AI and network programmability.

This initial qualification of the new revenue opportunities across the supply-chain allows AIMM partners to better understand the business opportunities that this project may enable to them in the future.

In collaboration with WP5, WP2 has considered the following questions:

- The size of the market and revenue opportunities to those developing RIC software;
- The value creation and capture by those developing and selling rApps and xApps i.e. what are the viable value-propositions and business models for value creation and capture? What is the potential size of the market for these models?
- The value creation and value capture by those CSPs employing the whole O-RAN architecture, including RIC, and rApps/xApps to automate Network Management through network programmability to underpin the deliver the products and services that customers require?
- What is the value proposition for the Open-Radio, Software-Defined-Radio (SDR) and (Reconfigurable Intelligent Surface) RIS technologies and the size of these markets?

- What is the value-proposition and viable delivery models of enhanced Network and Service security? What is the size of the market across the supply-chain components?
  - Security for Critical-Communications over 5G NR
    - RF Security, threat detection, network reconfiguration to adapt against an RF attack
    - Cyber-security, NFV secure deployment, implementation of security across Disaggregated Architectures

5.2 Individual Use Case Assessments

For each of the use cases specified in Section 2.1, a table was produced and these are presented in the following sub-sections. Separate sections in the table are included for cost savings and new revenue opportunities. In each of these sections there is a set of numbered descriptions relating to the relevant opportunity. This numbering scheme is then carried through to identify the Key Performance Indicators (KPIs) that are expected to determine the value of the opportunity. Finally the necessary inputs and expected outputs for the use case are identified.

In many cases a benchmark value for various costs or new revenues is required to be able to quantify the benefit of the AIMM solution. In the sub-sections below the use case tables are followed by a description of the data that has been identified as suitable to fulfil the data requirements. Each table also includes an indication of the technical workpackage that is most closely associated with the use case.

5.2.1 Smart Interference Management for QoS optimisation

UC1	Smart Interference Management for QoS optimisation	WP5
Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Improvement in radio capacity through improved interference management and handover. 2. Automation of interference management and handover leading to a reduction in operating costs. 3. In building and private network installations, that are purpose-built to meet specific B2B customer (e.g., enterprises, industrial venues, etc.) requirements, the reduction of interference may translate into a reduction of number of radio points, thus reducing the overall cost of ownership.	1. Reduction in capacity upgrade costs from increase in spectral efficiency. 2. Reduction in optimization and operation costs expressed as a reduction in required manpower to achieve improvement. 3. Reduction in the total cost of ownership of B2B purpose-built networks.	**Input requirements** 1. (a)Baseline capacity upgrade costs and capacity upgrade volume estimates. (b) Improvement in spectral efficiency through new techniques. 2. Estimate of people time reduced by the use of automated optimization. **Output** 1. Reduction in capacity upgrade costs on public & private networks.
Revenue		
Opportunity	**KPI**	**Input/Output**

1. Market share improvement through the provision of improved capacity per unit cost and optimisation. 2. Value assigned to algorithm developer (which could be the network provider). 3. In B2B purpose-built networks, the improvement of spectral efficiency characteristics of a network-service offering, may reduce the total cost of ownership of such services, thus contributing to the growth of market share and even enabling the creation of new markets across the Industry 4.0 and Digitisation opportunities.	1. Market share improvement through "best network" claims leading to increased revenues. 2. New revenue considered as a proportion of the cost savings made available to the operator. 3. Market share improvement through reduction of total cost of ownership to the customer and thus able to massively expand the market.	**Input requirements** Percentage of cost saving that could be claimed by the algorithm developer. **Output** Estimate of value that could be gained by algorithm developers for public and private network deployments.

i) Baseline capacity upgrade costs and capacity upgrade volume estimates – Public Networks

Per site baseline capacity costs estimates are derived from (3) and a representative set are shown in Figure 11.

Strategy	LTE availability	Cost type	Capex (GBP)	Capex time trend	Opex (GBP)	Opex time trend	Source
Integrating spectrum into the macrocelluar network	Site with 4G LTE	Additional carrier on current BS	15,000	-3%	1,800	0	MTC 2015
	Site with no 4G LTE	Deploying a multicarrier BS	40,900	-3%	3,898	-5%	MTC 2015
		Site lease	-	0	5,000	3%	MTC 2015
		Civil works	18,000	0	-	3%	5G NORMA (2016)
		Fibre backhaul Urban: 1 km	20,000 per km	0	-	0	Provisional assumption
		Fibre backhaul Suburban 1: 2km, Suburban 2: 4km	20,000 per km	0	-	0	Provisional assumption
		Fibre backhaul Rural 1: 8km Rural 2: 10km, Rural 3: 20km, Rural 4: 30km	20,000 per km	0	-	0	Provisional assumption
Network densification through small cells		Small cell equipment	2500	-3%	350	-5%	5G NORMA (2016)
		Small cell civil works	13300	0	0	0	5G NORMA (2016)
		Small cell site rental	-	0	5,000	0	Provisional assumption
		Small cell backhaul	-	0	1,000	3%	5G NORMA (2016)
Core upgrade cost on all strategies			10% mark-up on RAN deployment cost	0	-	0	Provisional assumption

Figure 11. Baseline radio infrastructure cost elements

All analysis regarding coverage and capacity growth in AIMM WP2 has been derived for the UK. The processes deployed will be similar for other countries.

Public network coverage extension in the UK is currently limited to growth in rural areas as part of an arrangement with the UK government extending coverage to remote areas (the Shared Rural Network, SRN). Recent reports indicate that three UK operators are collaborating to install 220 additional sites to meet this requirement (4). Although this is a significant number of sites, it is the growth in capacity and increase in quality of experience that is expected to drive the increase in site numbers on wide area public mobile networks. For this reason, rural coverage extension is not included in the assessment of WP2.

The calculation of the benchmark for public network capacity sites is based on traffic growth forecasts for the UK. These are available from a number of sources but for the purposes of WP2, the forecasts will be derived from extrapolations of results from Analysys Mason (5). These forecasts also give a view of the time taken for the customer devices to migrate from one technology to another and therefore give the opportunity to define the point at which AIMM solutions will create a benefit if these require new features in devices. An overview of the handset penetration and traffic growth that will be assumed is presented in Figure 12 and Figure 13 below. The dramatic growth in 5G traffic is a function of both early adoption of 5G handsets by those who require good data performance and the effect of the higher data rates and increased capacity that 5G provides, encouraging increased usage per user. The growth is also affected by the speed of rollout of 5G networks.

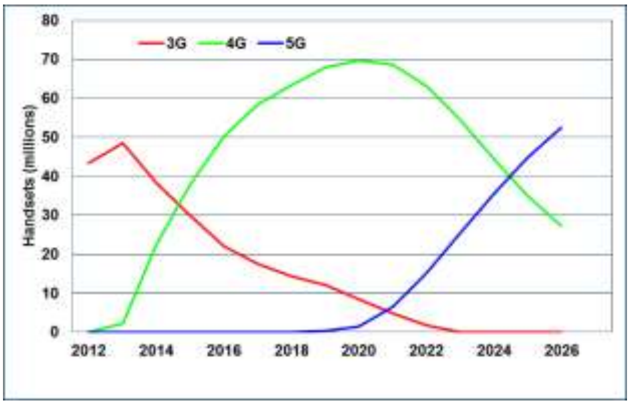


Figure 12. Assumed distribution of devices between generations over time – UK. Note that 4G was deployed relatively late in the UK resulting in faster adoption by the customer base compared to 5G. UK operators are indicating that 3G service will be discontinued from 2023.

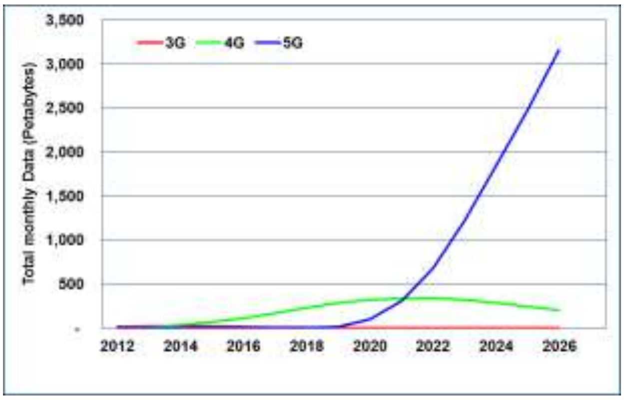


Figure 13. Total monthly traffic growth by technology - UK.

To carry the additional traffic growth, operators will migrate spectrum from previous generations to 4G and 5G, although the slow relative decline in 4G traffic in Figure 13 indicates that the growth in 5G traffic will occur more quickly than 4G traffic declines, resulting in a delay in spectrum migration, as discussed in Section 3.4.

Conversion of the traffic growth profile for public networks to an equivalent growth in base station sites will be achieved using a version of the model developed by Ofcom in the UK (6) which has been extended in WP2 to include 5G deployment and network rollout. The capex and associated opex for the resulting national public network is generated from the size of the network and unit costs equipment costs.

ii) Baseline capacity upgrade costs and capacity upgrade volume estimates – Private and Indoor Networks

Private networks are one of the main areas of benefit of 5G. Support for specific services with lower latency, high bandwidth requirements in a local area have been designed into the early stages of both the radio and core network architectures. It is the nature of private networks that they are designed to meet a particular customer's requirements and therefore the ability to produce a benchmark for a standard installation and potential growth in a diverse market is difficult to create.

For the purposes of WP2, the growth in the private networks market is considered in aggregate. Forecasts do exist of the growth in the market over time, in both the operator and equipment provider market and in the number of networks. This is illustrated in Figure 14, which shows the number of dedicated private networks globally of different sizes over time. In this chart a small network comprises installations that occupy an area of less than 500m square. This category is expected to encompass most single building industrial installations. The medium installations involve larger sites comprising multiple industrial, educational or healthcare buildings on a small campus. Large sites are generally in remote locations and include large installations associated with mining, oil and gas processing sites and larger transport hubs, including ports. Extra large sites include extensive mining and utilities sites.

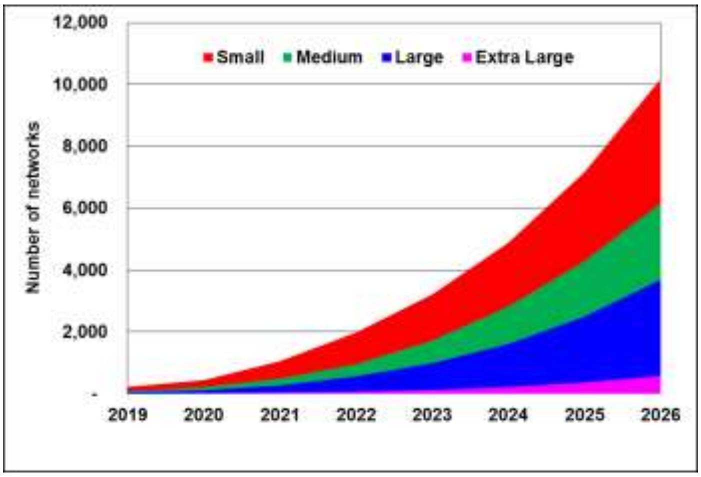


Figure 14. Forecast of the global number of private networks over time

Using this information and assuming an assignment of capex and opex by area between the size of the different private networks shown in Figure 14, the proportion of cost relating to the radio network elements can be derived from publications such as that shown in Figure 15 (7).

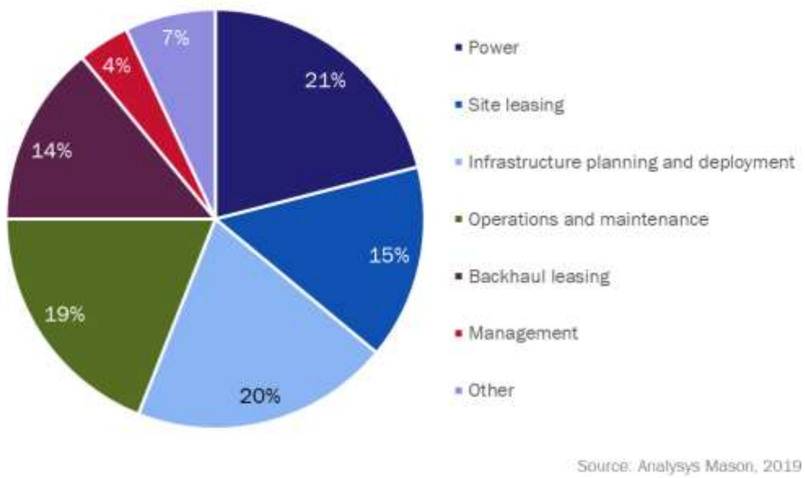


Figure 15. The composition of radio access network opex cost *(7)*

For the purposes of AIMM WP2, it is assumed that only dedicated private network sites are considered. A number of different private network implementations involve passive network sharing (e.g. towers, backhaul connectivity) with public networks or by supporting a network slice on the public network. These approaches, illustrated in Figure *16* from ACIA (8), are in the early stages of definition of both technical and commercial implementation and so are excluded from the current benefits assessment in AIMM. The different options are shown here since they are considered in the discussion of potential future work in Section 10.

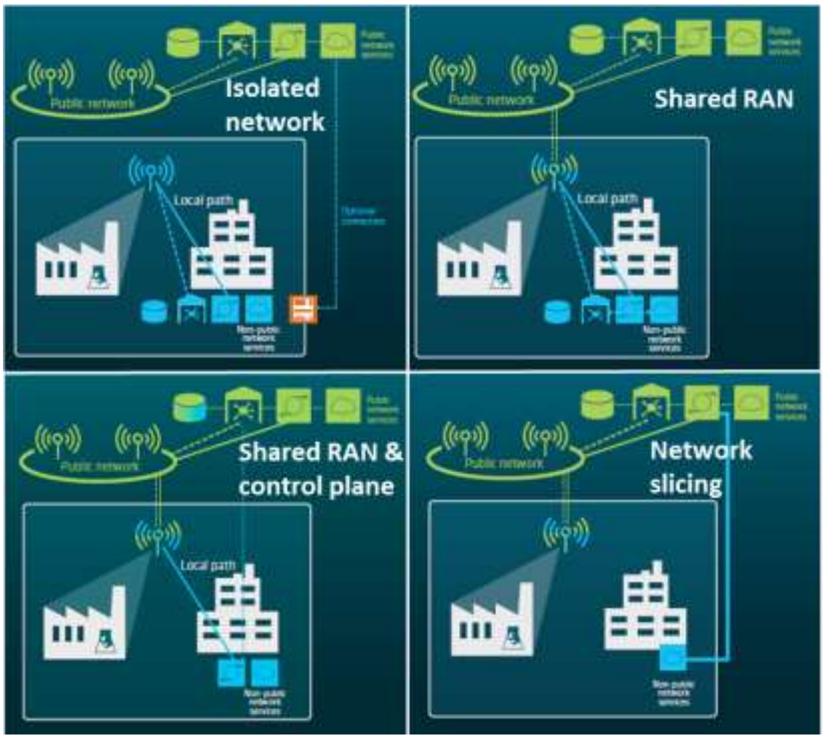


Figure 16. Different configurations of Public (green) and Private (blue) networks *(8)*. AIMM WP2 analysis only considers the isolated network approach.

iii) Spectral efficiency improvement through interference optimisation

In terms of benefit to the public network operator, the greatest cost benefit is expected to be in the reduction of capex by improving the spectrum efficiency of the RAN. Improving spectral efficiency will expand the data transport capability of existing physical network and spectrum assets and so reduce the upgrades for capacity that are required to accommodate the growth in traffic. These cost savings will be quantified by amending the spectrum efficiency assumptions included in the Ofcom model described above. Estimates of the spectral efficiency improvements will be provided through the outputs of WP3, 4 & 5.

Traffic growth on private networks is controlled through the contracts between network supplier and the end customer. As such this cost reduction will not be quantified in the analysis.

iv) Market share improvement through “best network” claims

This category of data is more difficult to quantify since the movement of market share over time varies due to a number of factors, including service pricing, marketing as well as the customer perception of the quality of the network. However, growth in market share does generate new revenue compared to the earlier discussions where the benefit is realised in terms of cost reduction.

To realise a definite financial value for public networks, the increase in market share will be coupled to the churn rate, which represents the proportion of the customer base that will leave a network in any year. The reciprocal of the churn rate therefore gives the average lifetime of a customer on the network. This, coupled with the Average Revenue Per User (ARPU), gives the average value over time of each of the customers on the network. To maintain compatibility with the analysis on cost, the figures for public network value will be realised from figures for the UK. Typical ARPU for UK mobile networks is currently £12.29 per month (9), excluding revenue associated with handset sales. This figure covers both contract and Pay As You Go (PAYG). If PYG customers are excluded, the figure is £14.75 and has stayed relatively constant years following the introduction of 5G.

The churn rate varies between operators showing the effect that operator choices relating to pricing, marketing and network quality have on customer behaviour.

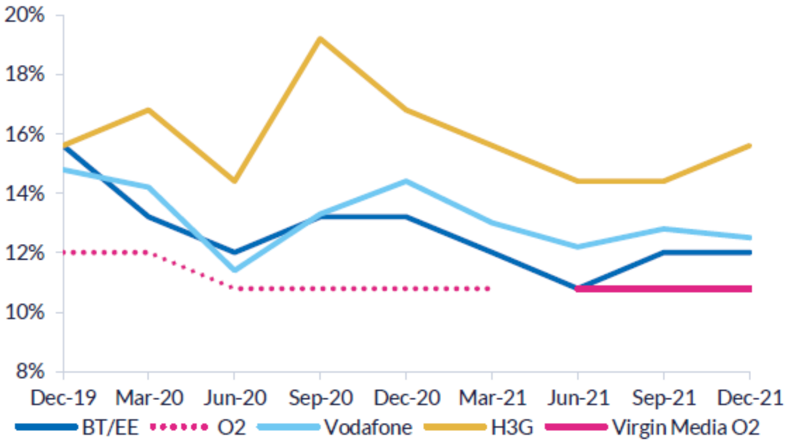


Figure 17. Annual churn rate of contract customers by UK operator (company reports and Enders Analysis)

Figure 17 shows the annual churn rate across UK operators, indicating a variation from approximately 12% (indicating customers stay with the company for 8 years on average) to approximately 16% (indicating customers stay for 6 years on average). Assuming that the pricing in the market is similar across operators, this indicates a lifetime value of a contract customer ranging from £1400 to £1050 base on the ARPU. If the market share between operators remains relatively constant, the operator with the higher churn rate would need to spend more in marketing and potential discounts to acquire customers to offset those that are leaving.

It is difficult to determine the level of churn that is due to the quality of the network. Analysis by others suggests that the correlation between a customer's likelihood to recommend a service following an improvement with different aspects of service offering varies by country (10).

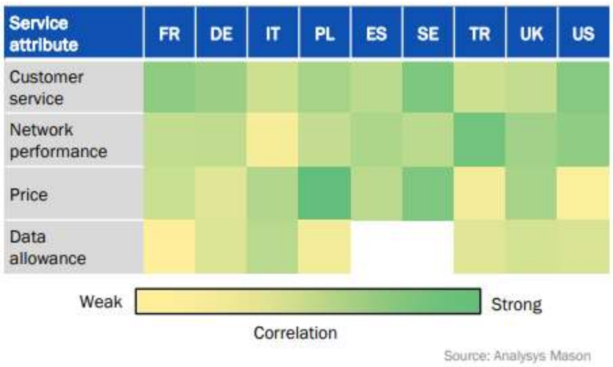


Figure 18. Correlation of different services features to the customer willingness to recommend across different countries *(10)*

Figure 18 shows the correlation of different factors to a customer recommending a service. For the UK this indicates a reasonable level of correlation with network performance although how that performance is perceived by the customer is again, difficult to quantify. For a customer, network performance could encompass data speed, coverage and the availability of different services. Also, the customer's view of the value of churning would also be dependent on the relative position of the competition. Even if the network performance of their existing supplier was considered to be poor, if all other operators provided a similar level of service that would not necessarily induce a customer to churn.

For the purposes of AIMM, only an indicative estimate of the value of increased revenue through churn reduction can be provided. Taking the assumption that there is a background level of churn for all operators, set at approximately 12% for the best operator in Figure 17, then any differential improvement in network quality will only erode the 4% difference between the best and worst operator. With the reasonable level of correlation between network performance and likelihood to recommend an operator, we assume that the improved network will contribute to 1.2% of churn reduction and only a proportion of that will be associated with improvements in the RAN, which will be assumed at 0.25%.

.

The value of a network improvement to reduce churn cannot be considered to occur every year for the same enhancement. Over time, competitors will also improve their networks if there is the perception that this is an area of differentiation between operators. It is assumed that the benefits of any change will be time limited to one or two years at most.

v) Percentage of cost reduction that can be claimed by an algorithm developer

The proportion of any benefit value that could be claimed by a supplier of a feature will be dependent on the relative commercial position of the network operator and the supplier. If a similar feature is available from multiple suppliers then the operator can achieve a lower price through the competition. Given the uncertainty of any particular outcome, for the purposes of AIMM it is assumed that the value of any benefit arising from a new feature is shared equally between the operator and the supplier.

5.2.2 Broadcast Beam Optimisation for Coverage

UC2	**Broadcast Beam Optimisation for Coverage Enhancements**	**WP5**
Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Avoidance of costs associated with improving coverage with additional network installations. 2. Increased capacity through interference reduction at cell edge if inter-base station co-operation is involved. 3. QoS improvement through clearer coverage boundaries in inter-cell base station co-operation is involved.	1. Reduction in cost to augment coverage by additional site build enabled by beam optimisation. 2. Reduction in capacity upgrade costs from increase in spectral efficiency. 3. Reduction in optimization costs from clearer cell boundaries.	**Input requirements** 1. (a) Anticipated increase in range or level of building penetration (dB) for macrocell and small cell deployments. (b) Value of coverage compared to the equivalent additional network build. 2. (a) Spectral efficiency improvement (b) Capacity upgrade costs 3. Reduction in manpower for optimization. **Output** Coverage, capacity and optimisation cost reductions.
Revenue		
Opportunity	**KPI**	**Input/Output**
1. Market share improvement through the provision of improved coverage and potentially network slicing. 2. Value assigned to algorithm developer (which could be the network provider).	1. Market share improvement through the provision of improved coverage and potentially network slicing. 2. Value assigned to algorithm developer (which could be the network provider).	**Input requirements** Percentage of cost saving that could be claimed by the algorithm developer (**BT, Vilicom**) **Output** Estimate of value that could be gained by algorithm developers for public and private network deployments.

i) Anticipated increase in range or level of building penetration (dB)

This information has been considered by WP5 and represents the benefit of increased gain from optimising the Synchronisation Signalling Block (SSB) beams to meet particular user distributions. The penetration loss values for different buildings and environments are taken from 3GPP document TR 38.901 (section 7.4.3) (11).

ii) Value of coverage compared to the equivalent additional network build

The increase in building penetration will be used to quantify the benefit of a reduced number of base station sites for coverage using the models defined to achieve the benchmark network figures described for use case 1.

There will be additional benefit potentially arising from increased market share due to improvement in coverage. This improvement will be included with the overall calculation of network improvement benefit described in use case 1.

iii) Spectral efficiency improvement

Any increase in coverage will also potentially provide an increase in spectral efficiency as SSB beams could be used to target specific areas of traffic demand at the boundary between cells and hence reduce interference. An estimate of spectral efficiency improvement will be requested from WP5.

iv) Capacity upgrade costs

Savings in capacity upgrade through spectral efficiency improvements will be calculated in the method described in use case 1.

v) Reduction in manpower for optimisation

Reduction in operational costs through optimisation using AI to enhance coverage is covered in the detail provided in us case 6.

vi) Percentage of cost saving that could be claimed by the algorithm developer

As with use case 1, we will assume that benefits are shared equally between the operator and the supplier of a new feature.

5.2.3 AI-based Channel Estimation & Detection

UC3	AI-based Channel Estimation & Detection	WP4
Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Radio capacity increase through improved assessment of the radio channel.	1. Reduction capacity upgrade costs achieved through improved spectral efficiency.	**Input requirements** 1. a) Baseline capacity upgrade costs and capacity upgrade volume estimates. (b) Improvement in spectral efficiency through new techniques. **Output** Capacity cost reductions.
Revenue		
Opportunity	**KPI**	**Input/Output**
1. Licensing value from technology generation 2. Improved market share for equipment vendors	1. Direct revenue from licensing techniques.	**Input requirements** 1. Estimate of value of capacity improving IPR.

(*assumes that this is not a part of the standard*)?	2. Increased market share for equipment vendor resulting in higher value.	2. Estimate of value of improved vendor market share arising from improved algorithms. **Output** 1. Licensing value. 2. Value of equipment vendor market share increase

i) Baseline capacity upgrade costs and capacity upgrade volume estimates

See use case 1.

ii) Improvement in spectral efficiency through new techniques

This information is being considered by WP4. The improved spectral efficiency will feed into the models as described in use case 1.

iii) Value of capacity improving IPR

Benefits through the acquisition of IPR by a network algorithm supplier will be included in the 50% share of any financial benefit that is assumed to be claimed by supplier. The remaining 50% of benefit is assumed to be retained by the operator.

iv) Improved market share through improved algorithms

This aspect of the potential revenue for an operator will form part of the assessment of churn reduction considered in use case 1. Any benefit arising from this improvement will be considered to be split equally between the operator and supplier as described in use case 1.

5.2.4 AI based Massive MIMO Precoding and Scheduling

UC4	**AI based Massive MIMO Precoding and Scheduling**	**WP4**
Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Radio capacity increase through improved assessment of the radio channel.	1. Reduction in capacity upgrade costs achieved through improved spectral efficiency.	**Input requirements** 1. (a)Baseline capacity upgrade costs and capacity upgrade volume estimates. (b) Improvement in spectral efficiency through new techniques. **Output** Capacity cost reductions.
Revenue		
Opportunity	**KPI**	**Input/Output**
1. Improved operator revenue / market share through the ability to provide improved RAN service slicing.	1. Value of improved provision of RAN slicing in terms of service revenues and market share.	**Input requirements** **1.** High level estimate of correlation between market

2. Licensing value for algorithm developers (*assumes that this cannot be achieved through xApp*) 3. Improved market share for equipment vendors (*assumes that this is not a part of the standard*)	2. Direct revenue from licensing techniques. 3. The level of increased market share for the equipment vendor resulting in higher value.	share and network performance. **2.** Estimate of value of IPR relating to capacity improvement IPR. **3.** Estimate of value of improved vendor market share arising from improved algorithms. **Output** 1. Value of operator market share increase. 2. Licensing value. 3. Value of equipment vendor market share increase.

i) Baseline capacity upgrade costs and capacity upgrade volume estimates

See use case 1.

ii) Improvement in spectral efficiency through new techniques

This information will be the output of WP4. The improved spectral efficiency will feed into the models as described in use case 1.

iii) High level estimate of correlation between market share and network performance

See the churn analysis in use case 1.

iv) Estimate of value of IPR

Benefits through the acquisition of IPR by a network algorithm supplier will be included in the 50% share of any financial benefit that is assumed to be claimed by supplier. The remaining 50% of benefit is assumed to be retained by the operator.

v) Improved market share through improved algorithms

This aspect of the potential revenue for an operator will form part of the assessment of churn reduction considered in use case 1. Any benefit arising from this improvement will be considered to be split equally between the operator and supplier as described in use case 1.

5.2.5 Distributed and Cell-less Massive MIMO

UC5	**Distributed and Cell-less Massive MIMO**	**WP3**
Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Radio capacity increase through improved radio path diversity between network and device. 2. Coverage improvement from multi-connectivity reducing the number of sites required for coverage.	1. Reduced number of base station installations to achieve growing capacity requirements (cost reduction potentially offset by increased fronthaul requirements). 2. Reduced number of base stations for coverage at a specified level of quality for	**Input requirements** 1. (a)Baseline capacity and coverage costs and capacity upgrade volume estimates. (b) Improvement in spectral efficiency through new techniques. (c) Fronthaul bandwidth and latency definition.

	both public and private networks.	**2.** Increase in base station range through multi-connectivity. **Output** Assessment of reduction in capacity and coverage costs.
Revenue		
Opportunity	**KPI**	**Input/Output**
1. Increased availability through multi-connectivity of devices enabling new services. Potential deployments are indoor Industrial IoT and campus private networks. 2. Revenue for fronthaul connectivity suppliers through the requirement for more capable connections.	1. Increased revenue arising from new, high availability services (public networks and private networks). 2. Potential upgrade revenue for high capability (high bandwidth, low latency) fronthaul services.	**Input requirements** 1. (a)Estimate of the value of new services made available through increased availability. (b) Estimate of the size of the market for new high availability services. 2. Literature survey of the cost of fronthaul with different capability in different countries. **Output** 1. Value of new high availability services enabled by distributed and cell-less massive MIMO. **2.** Value of increased fronthaul capability.

i) Baseline capacity upgrade costs and capacity upgrade volume estimates

See use case 1.

ii) Improvement in spectral efficiency through new techniques

This information is considered by WP3. The improved spectral efficiency will feed into the models as described in use case 1.

iii) Fronthaul bandwidth and latency definition

In many markets direct (dark) fibre is available and therefore there is no direct impact of cost on the nature of what is carried. In some markets, including the UK, fibre is available from the local operator and purchased by bandwidth, for example with Openreach, the largest provider of fibre connectivity in the UK (12). However, the price of bandwidth purchased in this way has declined over time, as shown in Figure 19, so for the purposes of WP2 it is assumed that a direct fibre product is available.

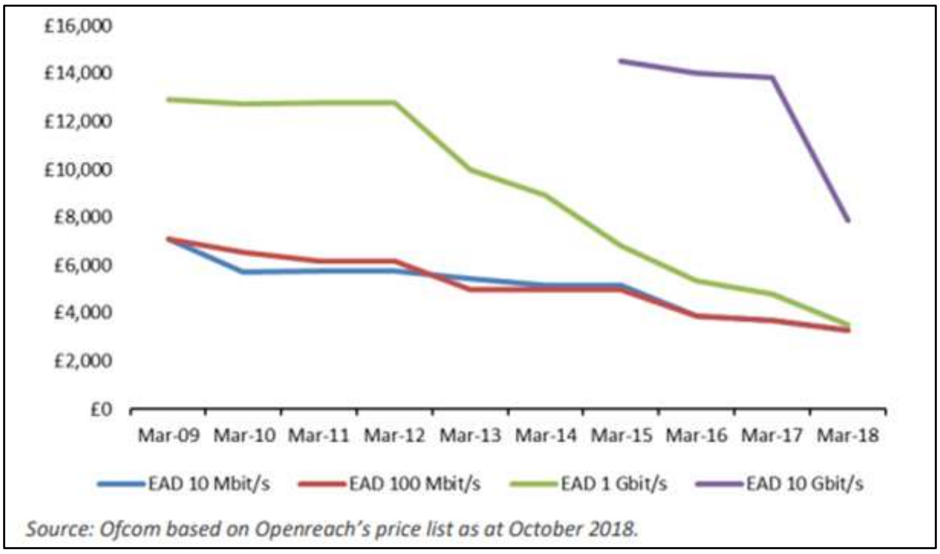


Figure 19. Annualised Total Cost of Ownership (TCO) for Openreach Ethernet Access Direct at different bandwidths

iv) Increase in base station range by multi-connectivity

The effect of the increase in range will be included in the model discussed for use case 1 in terms of a reduced requirement for radio access point installations.

v) Estimate of the value of new services made available

The assessment of the value of new services is difficult to assess without knowledge of the nature of the services. For this reason the value of new services will be considered as part of the assessment of churn reduction associated with the perception of a good network, as discussed in Section 5.2.1.

5.2.6 Disaggregated and Open Massive MIMO

UC6	**Disaggregated and Open Massive MIMO**	**WP4**
Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Cost reduction through wider vendor supply market. 2. Reduced network planning, optimization and operational costs arising from O-RAN RIC. 3. Reduced requirement for additional base stations driven increasing traffic demand. 4. Reduced cost in site acquisition due to reduced installation footprint.	1. Estimate of potential for cost reduction in RAN supply market generally. 2. Estimate of potential percentage cost reduction through improved automation. 3. Improvement in spectral efficiency arising from a move to disaggregated MIMO. 4. Relative physical size of installation footprint for traditional and new RAN architectures.	**Input requirements** 1. Literature review of cost reduction estimates through wider vendor supply. 2. Literature review of cost reduction through increased automation, using output of other use cases as proof points. 3. Baseline infrastructure requirement and reduction due to disaggregated RAN. 4. Installation and cost sensitivity to installation footprint. **Output** 1. Qualitative discussion of impact on RAN equipment

		market of O-RAN architectures. 2. Estimate of cost reduction through increased automation.
Revenue		
Opportunity	**KPI**	**Input/Output**
1. Value assigned to optimisation algorithm developer (which could be the network provider). 2. Revenue to new vendors entering an open architecture market. 3. Planning services from organisations that determine the locations for disaggregated & open MIMO deployment (public, private and in-building). 4. Installation service revenue for organisations that acquire sites, install and initially disaggregated MIMO	1. New revenue considered as a proportion of the cost savings made available to the operator through the reduction in manpower requirements. 2. Cost and number of sites that are deployed new or swapped from existing architectures. 3. The number of sites that are deployed new or swapped from existing architectures. 4. The number of sites that are deployed new or swapped from existing architectures.	**Input requirements** 1. Percentage of optimization cost saving that could be claimed by the algorithm developer. 2. Equipment cost for disaggregated & open MIMO installations. 3. Cost per site for planning services. **4.** Cost per site for installation and configuration services. **Output** 1. Estimate of value that could be gained by algorithm developers based on costs savings. 2. Per site revenues for planning, installation and optimization services.

i) Review of cost estimates through wider vendor supply

There have been a number of organisations producing cost estimates through the use of ORAN solutions. These suggest up to 40% capex and 35% opex reduction when compared to the equivalent cost of traditional RAN (13), leading to a total cost of ownership reduction of approximately 30% (14). The overall saving varies with the type of disaggregated architecture assumed and is illustrated in Figure 20.

Additional information at a more granular level, comparing the costs of OpenRAN with that of a traditional build, have been included in recommendations of the FCC in the US for the purposes of establishing the cost of replacing equipment in a network (15). For AIMM WP2 we assume that the stated percentage reductions in capex and opex are valid.

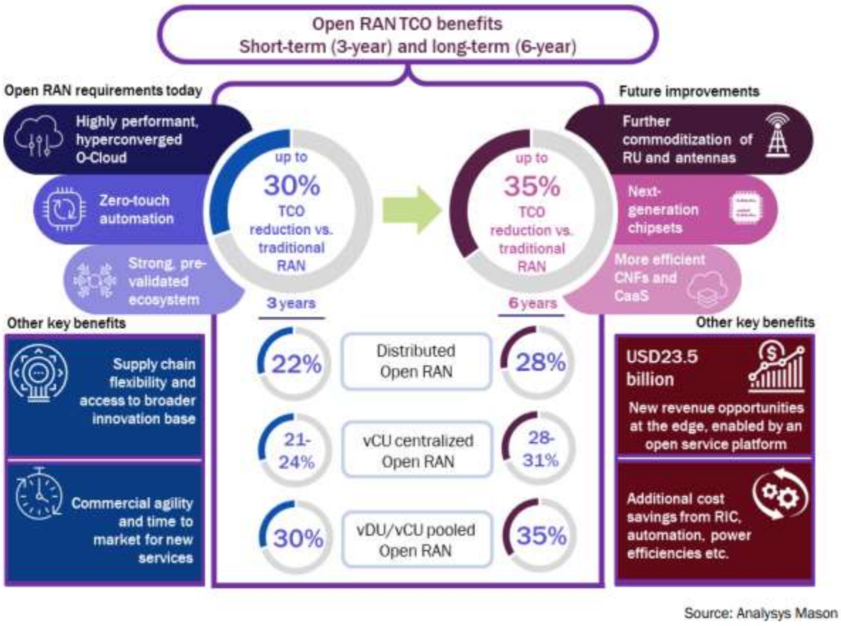


Figure 20. Sources of TCO cost reduction *(14)*

ii) Review of cost estimates through increased automation

This information is included in the high level estimates provided in the documents referred to in the previous item.

iii) Baseline infrastructure requirement and reduction due to disaggregated RAN

The baseline network costs will follow the approach in use case 1. For the purposes of AIMM it is assumed that disaggregated RAN makes the benefits of use cases 1, 2 and 5 available.

iv) Installation and cost sensitivity to installation footprint

The assumption relating to this item is that disaggregated RAN hardware will result in radio units that are physically smaller and that a centralised architecture will lead to more efficient use of baseband components.

v) Percentage of optimization cost saving that could be claimed by the algorithm developer

In line with the earlier use cases, it will be assumed in AIMM WP2 that the value of benefits is split equally between the beneficiary (often the operator) and the developer.

vi) Equipment cost for disaggregated & open MIMO installations

This information will be taken from the FCC publication (15).

vii) Cost per site for planning services

Data for this aspect has been derived from WP2 contributors.

viii) Cost per site for installation and configuration services

Data has already been collected that shows a high degree of variability for installation costs at different sites. WP2 estimates have been based on WP2 contributor submissions.

5.2.7 AI for RAN Energy Efficiency

UC7	AI for RAN Energy Efficiency	WP4 /5
Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Reduction in the opex (estimated at 20% of total network opex) that is associated with power, through the use of improved network and radio resource management. 2. Reduction in power opex through changes in network hardware through the use of O-RAN architectures and advanced algorithms. 3. Reduced energy cost of private purpose-built networks, that typically is covered by the end-customer.	1. Percentage reduction in network power requirement arising from the use of AIMM algorithms. 2. Percentage reduction in power requirement associated with O-RAN compliant hardware architectures on which new algorithms can be deployed. 3. Percentage reduction in network power requirement arising from the use of AIMM algorithms.	**Input requirements** 1. (a) Figures for power use and cost in public / private networks at the lowest level of granularity available (e.g. per nodeB, RRU/BBU). (b) Power reduction percentage based on the use of AIMM techniques. 2. Hardware percentage power reduction associated with a move to O-RAN compliant architectures. **Output** Value of power cost reduction in public and private networks.
Revenue		
Opportunity	**KPI**	**Input/Output**
1. Value assigned to optimisation algorithm developer (which could be the network provider). 2. Potential benefit to end customers based on an increase in device battery life, including remote IoT devices, leading to improved market share. 3. The reduction of energy costs in private purpose-built networks, might improve the operators' market share and product/service proposition.	1. New revenue considered as a proportion of the cost savings made available to the operator. 2. Potential market share improvement through improved battery life (unlikely to be quantifiable). 3. Potential market share improvement through reduction of energy bill of the customer of private purpose-built networks.	**Input requirements** 1. Percentage of cost saving that could be claimed by the algorithm developer 2. Qualitative assessment of the value of improved battery life. **Output** 1. Estimate of value that could be gained by algorithm developers based on costs savings and market share improvement. **2.** Estimate of impact of device battery life improvement on the customer perception of a network operator.

i) Figures for power use and cost in networks at the lowest level of granularity available

Recent work reported by NGMN (16), provides a view of the distribution of power consumption between the elements of an end to end mobile network. The output of this work is predominantly based on submissions from public network operators but for the purposes of AIMM WP2, the same distribution is used to characterise private networks. NGMN also references work from GSMA (17), an example of which is included in Figure 21 below

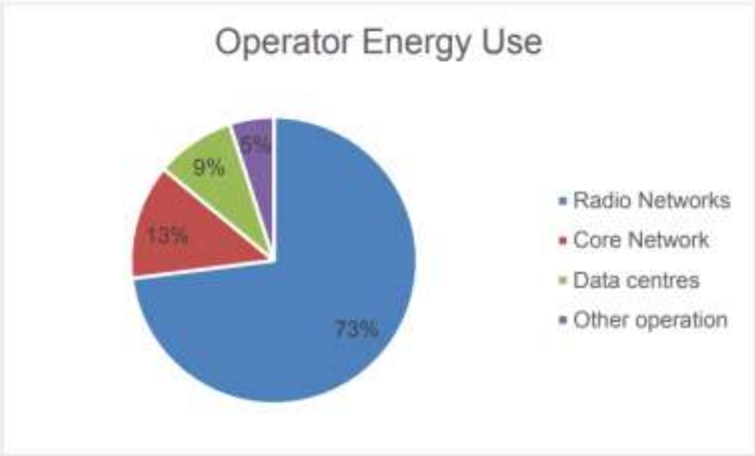


Figure 21. Distribution of power consumption across the elements of a mobile network operator *(16)*

The total power consumption is reported as representing approximately 20% of the total network opex of the organisation, as shown in Figure 22 from GSMA (18). Opex in this definition is associated with network equipment. Site rentals are included in the Selling, General & Administrative (SGA) and cost of goods associated with sales, often including handsets for a public MNO.

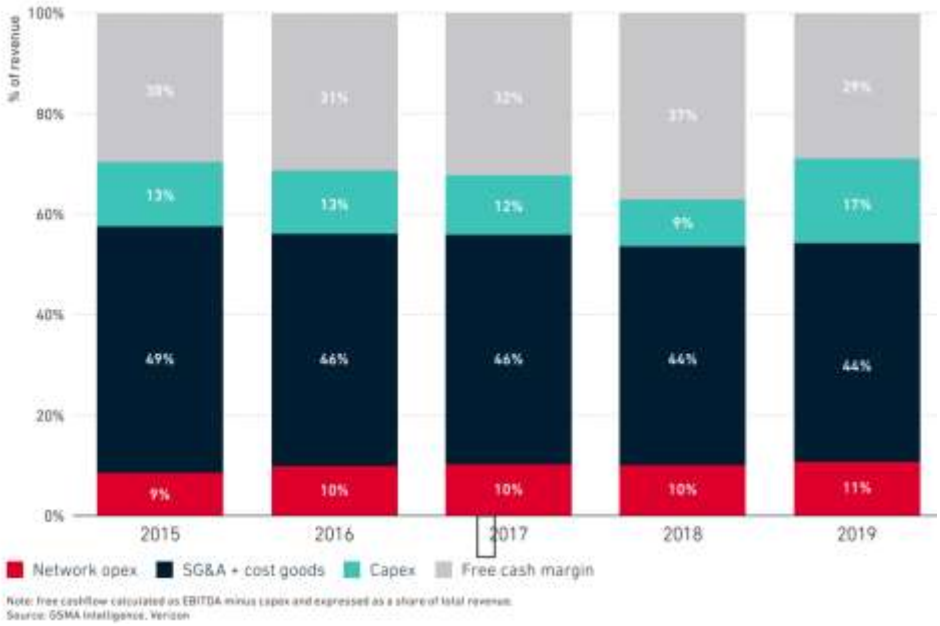


Figure 22. Breakdown of network operator costs relative to revenues *(18)*

Within this definition, the energy cost represents approximately 90% of network opex, as shown in Figure 23.

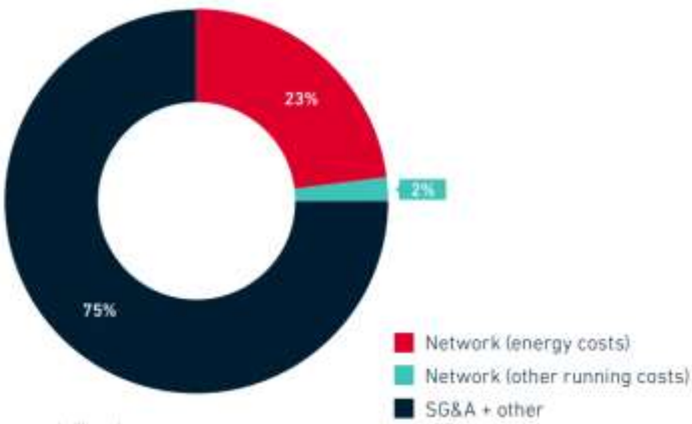


Figure 23. Further split of network opex costs between power and other network running costs *(18)*

The NGMN report also further subdivides the energy consumption of the radio network into baseband and rf, as shown in Figure 24. This information shows a high proportion of the power being involved with air conditioning, which itself is required to dissipate the heat that is being generated by the RAN equipment. The NGMN figures are derived from a number of networks with different climatic conditions. This, coupled with the fact that the method of deployment, e.g. mast mounted, cabinet or indoor installed equipment, has an impact on the required cooling. However the values in Figure 24 are considered to be a reasonable representation of the RAN as a whole.

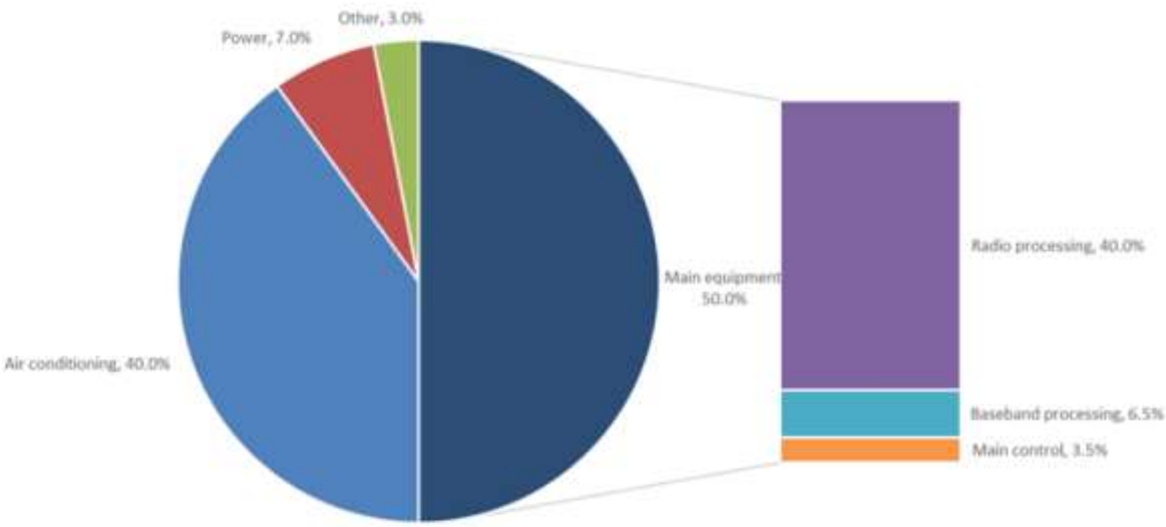


Figure 24. Distribution of power consumption requirements in the RAN *(16)*

Within the base station equipment itself, the majority of power is related to the layer 1 processing and amplification for the radio layer. This is illustrated in Figure 25 which differentiates between the power dissipated by the power supply, the power amplifiers and the small signal and digital intermediate frequency hardware.

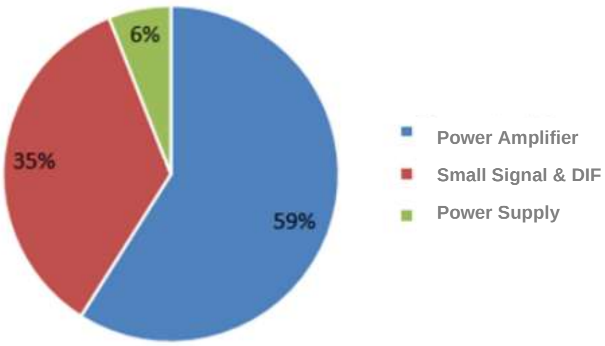


Figure 25. The split of power in the radio system of a modern 5G system *(16)*

Using the above information it is possible to create an estimate of the overall cost of different components relating to the consumption of power in the RAN for different network operators. Energy efficiency improvements in the solutions created through AIMM enable a value to be placed on the resulting benefits..

ii) Power reduction percentage based on the use of AIMM techniques

Indicative values for the reduction in power due to improved system level management of the users within the RAN has been considered by WP5.

iii) Hardware percentage power reduction associated with a move to O-RAN compliant architectures

Analysis of early O-RAN implementations show that a move to an O-RAN compliant solution will increase the power requirement for a single operator by 10-15% (14). This is primarily based on the assumption that the O-RAN components are based on Commercial Off-the-shelf (COTS) processors which are less energy efficient that existing RAN products based on ASICs. Projects are underway with a number of suppliers to address this issue by considering alternative compute platforms on which to operate the O-RAN system. For the purposes of AIMM WP2, it is assumed that there are no short term energy efficiency benefits to be derived from O-RAN but that in the medium term the hardware will approach parity with traditional RAN equipment. The inherent support in O-RAN for centralised architectures implies a more efficient use of signal processing compared to a traditional distributed approach. For this reason, at the moment AIMM WP2 will assume that hardware energy efficiency benefits from a move to O-RAN are neutral.

Where O-RAN will provide benefits is in the introduction of the ability to have greater control of the RAN through the RIC. These benefits will be included.

iv) Percentage of cost saving that could be claimed by the algorithm developer

In line with the assumptions made in previous use cases, it will be assumed that the operator and any supplier will each take a 50% share of the benefits arising from an algorithm resulting in energy efficiency.

v) Qualitative assessment of the value of improved battery life

Improved device battery life could contribute to an increased market share if the end customer can perceive a better service than that delivered by competing operators. To estimate a value of the benefit, a small reduction in churn rate will be considered and assessed in the same way as the technique described in use case 1 above.

5.2.8 Reconfigurable Intelligent Surfaces

UC8	**Reconfigurable Intelligent Surfaces**	**WP3**
Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Cost savings through reduction in the number of radio access points required to provide defined coverage. 2. Increase in base station capacity from improved signal level in edge of coverage locations, reducing upgrade requirements.	1. Coverage increase in different environments from the deployment of RIS (ideally the three deployment scenarios of public, private and indoor networks), expressed in population, geographic area and/or dB increase. 2. Capacity improvement (in bit/s/Hz) from deployment of RIS in different environments.	**Input requirements** 1. (a) Baseline cost of improving coverage in public and private networks plus capacity cost (b) Calculated number of access points that could be saved by RIS in a deployment scenario (c) Estimated cost of RIS compared to access point 2. Calculated spectrum efficiency improvement from RIS **Output** Coverage and capacity deployment savings
Revenue		
Opportunity	**KPI**	**Input/Output**
1. Revenue associated with the manufacture of the RIS. 2. Planning services from organisations that determine the locations for RIS deployment (public, private and in-building networks).	6. Requirement for the number of RIS that would be deployed in public or private networks over time. 7. Number of RIS deployed split by deployment type (on-street, in-building) and frequency band.	**Input requirements** 1. Estimate of proportion of operator cost savings that RIS vendors could claim. 2. Estimate of planning service costs for on-street and in-building installations

3. Installation service revenue for organisations that acquire sites, install and initially configure RIS. 4. Optimisation algorithms for network architectures including RIS components bring revenue to algorithm developers. 5. Increase of market share and increased profit margin for installers and operators of B2B purpose-built networks	8. Number of RIS deployed split by deployment type (on-street, in-building) and frequency band. 9. Reduction in manpower requirements to configure and optimize RIS when operational compared to situation with automation. 10. Potential reduction of total cost of ownership of purpose-built network by using (more affordable) reflective equipment.	3. Estimate of installation service costs for on-street and in-building installations 4. Baseline estimate of optimization costs and percentage of cost reduction that could be claimed by the algorithm vendor **Output** RIS manufacturer, planning, installation and optimisation services revenues.

i) Baseline cost of improving coverage in public and private networks plus capacity cost

See use case 1.

ii) Calculated number of access points that could be saved by RIS in a deployment scenario

This information is being considered by analysis in WP3.

iii) Estimated cost of RIS compared to access point

The cost of a RIS will depend on the level of complexity, including whether this is a passive or active device. At the moment in WP2, the cost of a RIS is expressed as a fraction of the cost of a small cell.

iv) Calculated spectrum efficiency improvement from RIS

This information is considered in analysis by WP3.

v) Estimate of proportion of operator cost savings that RIS vendors could claim

In line with the assumptions in other aspects of this analysis, WP2 will assume that the RIS vendor could claim 50% of the calculated benefit of using a RIS relative to alternative coverage and capacity solutions.

vi) Estimate of planning service costs for on-street and in-building installations, **estimate of installation service costs for on-street and in-building installations**

A study by Analysys Mason has produced a view of how the component implementation and planning costs relate to the overall operating cost of a base station site (19). This is shown in Figure *15*.

An estimate of the relative opex cost of a RIS compared to a base station, depending on the type of the RIS (e.g. active, passive or hybrid) is derived by excluding the aspects of this opex estimate that are not relevant to a RIS, including backhaul, aspects of the operations and maintenance and, where relevant, power.

vii) Baseline estimate of optimisation costs and percentage of cost reduction that could be claimed by the algorithm vendor

In line with the estimates of the other aspects in AIMM WP2, it will be assumed that the RIS vendor can claim 50% of the value of any benefit arising from the deployment of a RIS, with the operator claiming the other 50%. For the purposes of AIMM it will currently be assumed that the algorithm is provide by the RIS vendor since the relative immaturity of the ecosystem makes any alternative assumption premature at the moment.

5.2.9 Self-organising In-Building Small Cells

UC9	Self-organising In-Building Small Cells	WP5

Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Avoidance of costs associated reducing the number of radio access points required in an indoor installation. 2. Potential avoidance of capacity upgrades through interference management and spectral efficiency improvement. 3. Automation of planning and configuration will reduce commissioning and optimisation costs. 4. Reduction of OPEX through the addition of automation, self-cognition and self-healing capabilities	1. Reduction in cost to augment coverage by additional site build enabled by beam optimisation. 2. Reduction in capacity upgrade costs from increase in spectral efficiency. 3. Reduction in optimisation costs from clearer cell boundaries. 4. Reduction of operational and maintenance costs though	**Input requirements** **1.** (a) Cost of access point installation (b) Percentage reduction in access point numbers through use of AIMM algorithms 2. (a) Baseline capacity upgrade requirements (b) Spectral efficiency improvement 3. Estimate of time and cost reduction from reduced manual optimization requirements **Output** Savings related to reduced deployment and efficient optimisation
Revenue		
Opportunity	**KPI**	**Input/Output**
1. Market share improvement for operator through the provision of lower total cost of ownership coverage. 2. Market share improvement for operator and market segment increase through better quality of service support for different customer sets (network slicing). 3. Value assigned to algorithm developer (which could be the network provider).	1. Market share improvement relative to coverage pricing leading to increased revenues. 2. Market share improvement and customer base increase through improved service differentiation. 3. New revenue for algorithm developer as a proportion of the cost savings made available to the operator.	**Input requirements** 1. Estimate of market share increase resulting from more competitive coverage pricing. 2. Estimate of market share and market size increase from improved service differentiation. **3.** Estimate of the proportion of market share improvement that could be claimed by the algorithm developer. **Output** 1. Value to operator of market share improvement. 2. Value to operator of market size improvement. 3. Value to algorithm developer.

i) Cost of access point installation

Data has been collected for a limited number of in-building installations. In several cases these are based on a turnkey solution where the installation is not identifiable as a single cost. WP2 estimates are based on this initial data.

ii) Percentage reduction in access point numbers through use of AIMM algorithms

This information is being considered by WP5.

iii) Baseline capacity upgrade requirements

See use case 1.

iv) Spectral efficiency improvement

Spectrum efficiency is not a primary expectation of this use case and so, for the purposes of WP2, it is assumed that spectral efficiency is not improved by automated base station placement.

v) Estimate of time and cost reduction from reduced manual optimization requirements

This aspect will be represented as proportion of the opex aspects of network site implementation as shown in Figure *15*.

vi) Estimate of market share increase resulting from more competitive coverage pricing. Estimate of market share and market size increase from improved service differentiation.

This particular item is more relevant to private network and neutral host installers than public network operators. An estimate of the total capex avoided by the reduction in the coverage sites will be derived from the figures given in items (i) and (ii). The price elasticity of the market for variations in input costs and services is difficult to establish in advance. For this reason, the same information relating the correlation of different aspects of service to satisfaction presented for use case 1 will be used to create a range of market share change.

vii) Estimate of the proportion of market share improvement that could be claimed by the algorithm developer.

In line with the other assumptions in AIM WP2, it will be assumed that the value of the benefits arising from the use of improved deployment algorithms will be split equally between the operator and the algorithm developer.

5.2.10 Massive MIMO Combined Pre-distortion Architectures for Power Amplifiers

UC10	**Massive MIMO Combined Pre-distortion Architectures for Power Amplifiers**	**WP3**
Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Reduction in power costs for operators by deploying RU with improved amplifiers.	1. Percentage reduction in overall network power achieved through the introduction of reduced power consumption amplifiers.	**Input requirements** 1. (a) Reduction in power consumption for improved amplifiers compared to existing commercial products. (b) Number of amplifiers deployed over time. (c) Amplifier power consumption as a percentage of total network power. **Output** Reduction in power requirements and therefore cost

		over time for public and private network operators.
Revenue		
Opportunity	**KPI**	**Input/Output**
1. Increased revenue from market share increase for amplifier vendors. 2. Premium price for lower power consumption amplifiers.	1. Estimate of the market share increase for low power consumption amplifiers. 2. Price premium for lower power consumption amplifiers (linked to 1)	**Input requirements** 1. (a) Estimate of the proportion of the power saving that the amplifier vendor can claim (b) The number of power amplifiers deployed over time 2. Estimate of value of power saving that could be allocated to the amplifier price premium. **Output** Estimate of increase in revenue for amplifier vendors

i) Reduction in power consumption for improved amplifiers compared to existing commercial products

This information is being considered by WP3.

ii) Number of amplifiers deployed over time

This information will be derived from the baseline network information described in use case 1.

iii) Amplifier power consumption as a percentage of total network power

See the information provided under use case 7.

iv) Estimate of the proportion of the power saving that the amplifier vendor can claim

As with other sections in AIMM WP2, it is assumed that the amplifier vendor can claim 50% of the total saving from the deployment of the higher capability devices.

v) Estimate of value of power saving that could be allocated to the amplifier price premium

Following a review of the data that might be available it has been determined that this information is difficult to separate from other components. For this reason the value of the improved amplifier performance will only be considered in terms of a share of the power cost reduction that will be realised.

5.2.11 AI for RAN Security

UC11	**AI for RAN Security**	**WP3**
Cost Saving		
Opportunity	**KPI**	**Input/Output**
1. Reduction in network monitoring costs related to security assessment.	1. Percentage reduction in manpower resource required for network monitoring.	**Input requirements** 1. (a) Baseline cost for network monitoring for interference events.

2. Reduction in reaction and mediation costs once an interference event has been identified. 3. Reduction, from initial design, of OPEX related with Security Operations through the implementation of preventative network designs and algorithms.	2. Percentage reduction in manpower resource required for response to jamming and similar occurrences. 3. Percentage reduction in overall costs associated with security operations.	(b) Estimate of proportion of monitoring costs that could be reduced by AIMM solution. 2. &3. (a) Baseline cost for network interference response. (b) Estimate of proportion of monitoring costs that could be reduced by AIMM solution. **Output** Operational cost savings from automated interference assessment and mitigation.
Revenue		
Opportunity	**KPI**	**Input/Output**
1. Increased revenue from operator market share increase through delivery of improved security features. 2. New security Service Level Agreements (SLAs) for specific customer sets (e.g. critical infrastructure owners), generating revenue for the network operator (public and private). 3. Increased sales for the security feature vendor.	1. Percentage market share increase arising from greater customer confidence. 2. Revenue from the support of higher value SLAs on a per organization basis (public and private networks). 3. New revenue and a proportion of cost savings from the operators that is passed to security feature vendor.	**Input requirements** 1. Estimate of operator market share improvement from improved security features. 2. Estimate of value of enhanced security SLAs for particular markets. **3.** Estimate of value to a security feature vendor that arises from operator revenue and cost saving. **Output** Operational cost savings and revenue improvements for operators and how these flow to new service providers.

i) Baseline cost for network monitoring for interference events

On existing networks, monitoring of interference is often inextricably linked with identifying trends in network performance or alarms registered on the network. As a result this aspect is considered as a proportion of the total operating and maintenance component of total opex, as illustrated in **Error! Reference source not found.**.

ii) Estimate of proportion of monitoring costs that could be reduced by AIMM solution

Through a consideration of the likely cost of identifying monitoring cost it has become clear that existing techniques are highly manual. With the expected growth in the number of additional private and neutral host networks and an increasing expectation for resilience of network connectivity, it is anticipated that current methods will be unsustainable and will require an automated process to replace them. With that perspective it is difficult to relate the value of the new automated process to the existing approach as requirements grow. For that reason, WP2 relates the value of the use of new interference management technologies to the potential revenue benefits.

iii) Baseline cost for network interference response

Information is available from live networks relating to the identification and rectification of interference events on the network. Typically these are expressed in terms of the number of days to reconcile a reported interference event and this can be converted to a financial figure based on personnel cost figures. From the experience of the WP2 members, an indicative figure for the response to an interference event is £1800. As mentioned above, the expected growth in the need for interference identification and management will require an automated procedure because existing approaches are not scalable.

iv) Estimate of operator market share improvement from improved security features

A proportion of the market share improvement associated with improved security is difficult to directly quantify. For this reason the expected output of AIMM WP2 for this aspect will be based on assigning a proportion of the churn reduction value described in use case 1.

v) Estimate of value of enhanced security SLAs for particular markets.

Reports by Ericsson (20) estimate the benefit in terms of ARPU that can be gained through growth in new service revenues. This is illustrated in Figure 26, showing the correlation between ARPU and service revenue for different classifications of operator types. Pacesetters are acknowledged as leading in terms of network capability and services. Aspirationals are market challengers with strong network performance but less focus on new services. Potentials primarily compete on price, as do Exploters, with different levels of focus on network capability.

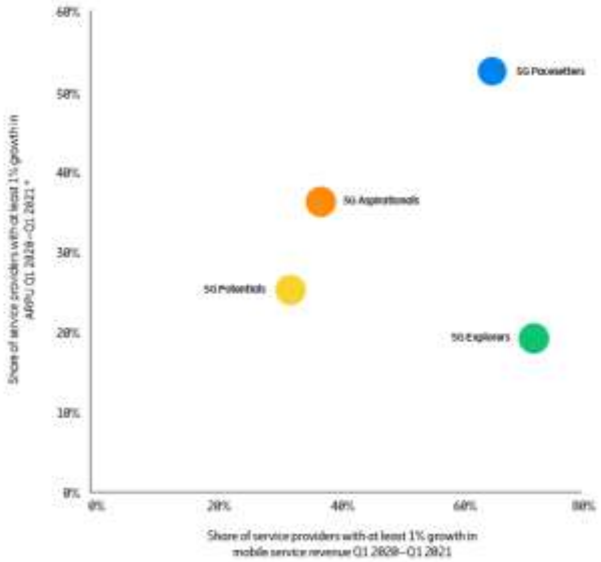


Figure 26. Correlation between ARPU growth and growth through new services *(20)*

Using this information in conjunction with the churn information in use case 1, a range of estimates of the total value of new service support has been derived.

vi) Estimate of value to a security feature vendor that arises from operator revenue and cost saving

In line with other use cases, AIMM WP2 will assume that 50% of the total benefit arising from AIMM solutions will be available to the system supplier. The remaining 50% benefit will be retained by the operator.

6 Models and Modelling

As described in Section 3.3, in most public network circumstances, a new feature or technology will be deployed into an environment where there is an existing network infrastructure and customer base. Most public network operators have a combination of 2G, 3G, 4G and 5G networks with established service portfolios. The benefits arising from a new feature are therefore dependent on the speed of rollout of the feature into the network, the ability of existing devices to access the features and the speed at which new devices will permeate the customer base and so increase the benefit of the feature. For new network installations with a closed user base, for example private networks in industrial installations where devices are supplied as part of the end to end service solution, this is less of an issue. However all network types do involve equipment going end of life either through technical obsolescence or financial depreciation to be replaced with new equipment. This point of replacement can be used to introduce new capabilities that are not backward compatible on the network.

Taking full account of legacy infrastructure and the timescales taken for a new deployment that can attract the traffic and ultimately the spectrum assets from an earlier network is a complex task. The development of such a modelling capability is beyond the primary focus and therefore the scope of the AIMM project. Fortunately a model does exist, developed by the UK telecommunications regulator Ofcom (6), to define the network cost of terminating voice calls originating in other networks. To do this a full analysis of network costs is required and a view of how the total voice and data traffic moves between different network nodes over time. The model was last updated in 2020 by consultants working for Ofcom and contains network node costs that have been reviewed by the UK mobile network operators.

The Ofcom model does provide a good starting point but does have a number of drawbacks:

- It is based on a 2G/3G/4G network deployments. 5G network capability has been added as part of AIMM WP2.
- The model is cost based. New revenue effects will require separate models. As described in Section 5, where new services have not been defined in advance, new revenue will be considered as the defence of market share through improved network capabilities.
- Power consumption, planning, operations and service management are all included in a single operating expenditure (opex) figure for each node. Separate supplementary models have been developed to consider the impact of energy efficiency and the attribution of energy efficiency improvements to different components within the RAN.
- The model is focused on the UK. Work with AIMM WP2 has retained this assumption but by altering the input parameters, other country networks could be considered.
- The network architecture follows the traditional approach of distributed base station equipment (eNodeBs) located at the antenna location. An early assumption established in WP2 is that the target RAN architecture is that being developed by O-RAN.

The Ofcom model has been taken as a starting point, the deficiencies of the model to meet the needs of AIMM have been identified and work carried out to address these deficiencies.

6.1 Defining the geography of network rollout

The coverage and capacity phases of network build phase have been described in Section 3.1. Within the AIMM network model, the coverage phase is determined by splitting the target area into a number of geotypes. For a national public network these have been set to be the values shown in Table 2.

Geotype
Urban
Suburban 1
Suburban 2
Rural 1
Rural 2
Rural 3

Rural 4
Highways
Railways

Table 2. National network geotypes

In this case the urban geotype is to define dense city areas. Two suburban types are included to differentiate suburban areas with high rise buildings from those with houses only. Three rural areas are provided to differentiate between mountainous areas, areas with villages and small towns and other sparsely populated areas. Highways and railways are identified as linear features that require coverage.

For a national rollout, the area of the whole country is partitioned into these geotypes as is the linear distance of highways and railways that require good coverage which might not be guaranteed by a general coverage deployment. Within the model the names of the geotypes can be changed appropriately and so a more limited set could be considered, for example, for a private network deployment in a campus.

Within AIMM, traffic is spread across different geotypes from an assessment of the proportion of the total market population that occur in each of the geotypes.

For each of the different geotypes, base station range figures are required for macrocells, microcells and picocells. The number of coverage base stations will be assessed by dividing the total geotype area by the base station type that is defined as the primary coverage technique. For a nationwide network the primary coverage technique will be macrocells. The rate of network build is specified year on year until the full target coverage for a geotype is achieved. For public networks rolling out a new generation onto existing infrastructure, eg a 5G rollout onto an existing 2G/3G/4G network, the model can be configured to allow for all existing physical infrastructure to used with a rate of rollout for a new generation or feature specified. The rate of feature rollout will impact the migration of end user devices and traffic onto the new network capability.

Once the primary coverage layer has been established, additional capacity sites are included in the geotype to meet a capacity demand which is growing with time

The equipment at each base station site has a specified lifetime, which is usually selected to coincide with the assumption of financial depreciation on the base station asset. At the expiry of the equipment lifetime it is assumed that the equipment is renewed, providing an opportunity for a new feature to be introduced if this is not backward compatible.

In some circumstances the base station, including the physical assets and electronics will be shared between multiple operators and therefore the costs, and any benefits, will be shared, resulting in a reduced cost for each operator.

Many of the use cases in AIMM consider mechanisms to improve the spectral efficiency of the system and so increase the capacity of the RAN infrastructure for a fixed amount of spectrum. The main financial benefit from this approach will be the cost reduction, in both capital and operational costs, that arises from requiring fewer capacity sites to be installed. An increase in spectral efficiency will also delay the time when an upgrade of capacity is required following potential migration of spectrum between different generations.

The Ofcom model does include the concept of the 2G Base Station Controller (BSC) and 3G RNC and so allows for connection between a radio site, an intermediate node and a core node. The model has been updated to re-use this same capability for the O-RAN architecture, allowing the connectivity requirements of fronthaul and midhaul, described in Section 3.2, to be considered. For the future architecture concepts, including cell-less architectures, the same approach to modelling is used but incorporating changes to the type and resulting change in cost of the fronthaul connection. At the moment it is expected that the cell-less approaches will be restricted to microcell and picocell deployment approaches. Some analysis in WP3 has shown that the relationship between the coherence time of the radio channel and delay in reporting and acting upon channel state information reports has a significant impact on the effectiveness of the cell-less approach. Fast moving mobile devices, which would typically be served by macrocells, have short channel coherence times and so the deployment of cell architectures would be less appropriate than for the slower moving devices encountered in small cell areas.

The model has been configured to enable both traditional and disaggregated RAN in 5G and to enable the sharing of different network assets to be accounted for in the overall network cost and value flow calculations.

6.2 Network node and connectivity cost and capability

The previous sections describe how the demand and basic RAN size and architecture will be considered from the perspective of network build. From this information the number of nodes, the capacity and connectivity requirements can be derived. To achieve an overall view of the cost of building and operating the network, an estimate of the cost of different nodes is required.

Within the Ofcom model there are capital and operational cost assumptions and price trend estimates for each of the active and passive network nodes for 2G/3G/4G implementations of traditional architectures for both RAN and core networks. These values have been reviewed by UK network operators and it is therefore these have been assumed to be the same for the AIMM WP2 analysis. For 5G deployment there is limited public domain information based on node level. The same is also true for the nodes required for an O-RAN defined architecture with different CU/DU/RU and RIC implementations. Fortunately the United States of America Federal Communications Commission (FCC) (15) produced a report in association with the requirement for operators to remove high risk vendors from their networks. This includes capital equipment 5G costs for both traditional and O-RAN architectures, defined as bundles of equipment for base station installations. This information is being used to populate the first instance of the model.

The cost of connectivity between nodes is highly dependent on the country that is being considered. Within the UK, the division of British Telecommunications that provides connectivity to other communications providers, Openreach, has a regulatory requirement to publish its product costs. Openreach usually provides layer 2 ethernet products and has no dark fibre alternative in its portfolio, although there are products that have the option for a customer to deploy wavelength division multiplexing on a fibre provided for an Openreach ethernet service. For the initial assessment within AIMM it will be assumed that the connectivity costs align with those of Openreach. The main issue will arise for future architectural solutions, including cell-less and some O-RAN compliant approaches, where the relationship between fronthaul bandwidth requirement and user data requirements is not linear. Under these circumstances a dark fibre approach, where the connectivity bandwidth is assumed to be in excess of that required by the AIMM technique will be more appropriate. This is described in more detail in Section 5.2.5.

For private networks, the cost of connectivity will be driven by the cost of installation and it is assumed that no bandwidth dependent operational expenditure charges are assigned.

6.3 Financial approach

As discussed above, the financial value of a benefit arising from an improvement to network performance does involve a timing element. On a large public network, even if a capability were available immediately from equipment vendors, if it requires a change to the handsets and network equipment it will take time for the benefit to become universally available. This delay will be a combination of the time to complete a network rollout, the time for existing network equipment to reach the end of its technical or financial life and, where necessary, the time for end user devices to penetrate the customer base.

To take full account of this timing issue in the assessment of benefit, a discounted cash flow approach has been adopted by AIMM WP2. Future costs and potential revenues are reduced year on year by a discount rate to reflect the time value of cash. The discount rate is typically used to account for a number of factors including the cost of borrowing money to make a network investment over time. This aspect of borrowing also sometimes includes an aspect of risk that the investment will not succeed in providing the expected benefit and so the discount rate is increased to include this element of risk. As an example, consider an investment in a competitive telecommunications network, where the actions of competitors might erode potential financial benefits. This is inherently more risky from a financial perspective than investment in electrical power distribution where, in many countries, the investment is made by a regulated monopoly company that has greater certainty over its ability to realise benefits. For the analysis in AIMM WP2, a discount rate of 10% has been applied for all financial considerations.

For single component items, for example power amplifiers, the period over which the value of any beneficial improvement is calculated is the expected lifetime of the component, starting at the point when it is deployed. The same approach is applied for private network installations which will sometimes be contracted for over an initial design lifetime. For wider public network benefits, where a new system or architecture is rolled out across the network, a longer timescale is considered. Within AIMM WP2 it is assumed that any new features will be deployed in 5G and future network evolutions. 2G and 3G legacy public networks are in the process of being switched off in many countries with the spectrum being released to 4G and 5G installations. 4G networks are continuing to grow in terms of carried traffic but most investment in new capability is associated with the rollout of 5G. Many network operators are also using the 5G rollout to refresh the 4G equipment on the large majority of sites that will carry both technologies, delaying the point at which a new 4G capability could be introduced until the time where the number of devices in the customer base that are not 5G capable will be reducing dramatically.

For the reasons outlined above, public network benefits analysis is considered over a 20 year period, starting in 2020 to coincide with the introduction of 5G in many markets and ending in 2040, which could be expected to be the point at which 5G is close to closure. However no provision is made for a 6G deployment which is currently in the very early stages of definition, with the expectation that AIMM improvements will be factored into 5G evolution. A twenty year period will also encompass two renewal periods for an assumed expected lifetime of eight years.

In Section 7, financial benefits are expressed in terms of the net present value (NPV), which is the sum of the annual discounted cash values over the calculation period. Where the calculation only considers cost and not a full treatment of associated revenue, this is expressed in terms of net present cost.

7 Benefits assessment results

The previous sections have discussed the large range of alternative input values, network deployments and features that have been considered as part of AIMM. Although the mechanism to calculate the financial benefits across a significant number of different alternative network configurations has been developed in AIMM, the number of scenarios that have been considered as part of the project have been relatively limited. The development of the approach, tools and capabilities is seen as one of the main contributions of AIMM that can be used in future benefit assessments.

In the following sections some example results of the output of analysis are presented. To make these results useful in themselves, these have been expressed in terms of the unit of benefit for a percentage change in a particular feature, for example a general improvement in spectral efficiency for the 5G aspects of a national public mobile network, rolled out as a new feature requiring some hardware renewal at a base station and taking full account of the migration of customers and therefore traffic from 4G to 5G.

7.1 Revenue improvement through market share growth

As described in Section 5, the mechanism for determining the value of an improvement that will increase the market share of a public network company has been considered in terms of the difference in churn between a high performing and a poor performing network. The difference in churn rate between the best performing and worst performing networks has been assessed at four percentage points. The proportion of this difference that is attributable to network performance is assumed to be 30% in the UK. This is an estimate based on a qualitative assessment of the information presented in Figure *18*. Correlation of different services features to the customer willingness to recommend across different countries and will vary by country. This level of influence of network performance in the willingness of end customers to recommend an operator will also be dependent on the network maintaining a benefit compared to other competitors. In reality, if there is a significant financial advantage to improved network performance it can be assumed that competitors will work to narrow this gap and so any benefit will be relatively short lived. In this study we have assumed that the benefits exists for two years before it must be augmented by another improvement.

Using the above analysis, the benefit of improved network performance is to reduce the churn rate by approximately 1.2%. Making the assumption that these customers would otherwise leave the network and so their lifetime value, based on a financial discount rate of 10%, would be lost rather than be retained for another contract period, it is possible to estimate an overall annual value of improved network performance in terms of operator revenue. Using figures from Ofcom, the annual revenue of a large UK operator is approximately £3.5bn. From the analysis above, the resulting annual value to revenue of enhanced network performance is £210m per annum, or approximately 6% of total revenue.

As discussed in Section 5.2.1, the customer perception of network performance is based on several different contributions. Some of these contributions will not be the responsibility of the network operator but will be determined by the application and server infrastructure operated by a content provider on the internet. Similarly, the core network features of the mobile operator, if perceived to be poor, will simultaneously affect larger proportions of the customer base than poor performance in areas of the radio network that are likely to be more localised in their effect. This is not true of private and in-building networks where all impacts will be localised and so the quality of the RAN will affect most users simultaneously.

Considering the discussions above, it is assumed that the quality of the RAN contributes to approximately one third of the customer perception of network performance. On that basis, for a large national network in the UK, it is assumed that the revenue value of any capability that is valued by end customers has a value of approximately £70m per annum, or 2% of total revenue.

7.2 Quantifying the value of RAN energy efficiency

In Section 5, input data is presented from various sources to apportion the energy consumption of different aspects of a mobile network. To create a quantifiable financial value for this information, the various aspects of the power distribution and cost have been mapped to provide an overall value for the cost of power saving in different aspects of the RAN.

Power saving is likely to be represented as a percentage reduction compared to an earlier benchmark. The results are shown in Table 3. A significant proportion of RAN power consumption is involved in the provision of air conditioning to cool the relevant components. For the purposes of this analysis, it is assumed that the power consumption of the air conditioning can be assigned to different components in the same proportion as the power consumption of the heat generating components themselves. This is a simplification but does include the effect that more power efficient components will reduce the requirement for cooling.

RAN Component	Percentage of RAN opex in component power consumption	Lifetime value of 1% power reduction (£m)
Power amplifier	8.3%	7.1
Small signal & digital if	4.9%	4.2
Baseband processing	2.3%	1.9
RAN main control	1.2%	1.0
Power supply	3.3%	2.8
Other	1.1%	0.9

Table 3. Assessment of the value of RAN power consumption by constituent components

The network-wide financial value of a 1% reduction in power consumption for a component is based on a large national mobile network operator. In this case the overall network opex figures are taken from BT in the UK. The lifetime of a component is assumed to be eight years and the financial discount rate is assumed to be 10%. For example, the figures should be in interpreted such that, for a power amplifier with a 10% reduction in consumption compared to current installation, the discounted component lifetime power savings for an operator in deploying that component throughout the network would be £71m. This saving would be realised over the eight years that the component is in the network before replacement.

In reality the components in a network would be upgraded over time so the long term view would differ from this and be represented as a net present cost reduction covering a number of years of rollout and replacement. That is the approach taken in the following section discussing spectral efficiency improvement

The input financial figures were derived from BT financial results presented in early 2022. As a result they do not include the escalation in energy prices that is occurring through the remainder of 2022 which will increase the potential benefit that can be achieved through improved energy efficiency.

7.3 The value of spectral efficiency improvement

Section 6 described the model that has been evolved from an original, developed by Ofcom, to include both 5G deployment and the option to alter the capabilities of different network nodes. A variety of different results can be derived from the model.

Figure 27 shows an undiscounted cash flow profile, separating capital expenditure (capex) and operational expenditure(opex) components. These represent the network costing for a baseline view that does not include any benefit from AIMM features. In this particular view includes the rollout of 5G onto base station sites that are already occupied by 4G and 2G/3G with site sharing arrangements in place between operators. The 5G rollout is assumed to start before the 2019/20 financial year to allow for launch of service in 2020 over a reasonable footprint. Rollout of the 5G network then continues for the early part of the decade but at a declining rate in the number of sites per year, resulting in a reduction in the overall capex requirement. At times when 5G and 4G equipment needs to be refreshed, there are further peaks in capex occurring at eight year intervals, which is the assumed lifetime of the equipment. An underlying growth in traffic leads to a growing background capex requirement that is indicated by the trend of the minima in the capex profile to grow over time.

The opex chart is dominated by the base station site rentals and backhaul connection charges. The rollout of 5G increases the power consumption on the sites and, in this particular example from the model, there is no power reduction benefit assumed in moving from 4G to 5G over time.

Although the capex cashflow profile has some variability, in reality this would be presented in the network operator's annual accounts with the depreciation rather than the cash in year presented as value leaving the company. In this way the profit and loss position of the company is smoothed relative to the cash flow profile.

The assumptions included in this chart pre-date the rises in power costs in Europe that have occurred during 2022.

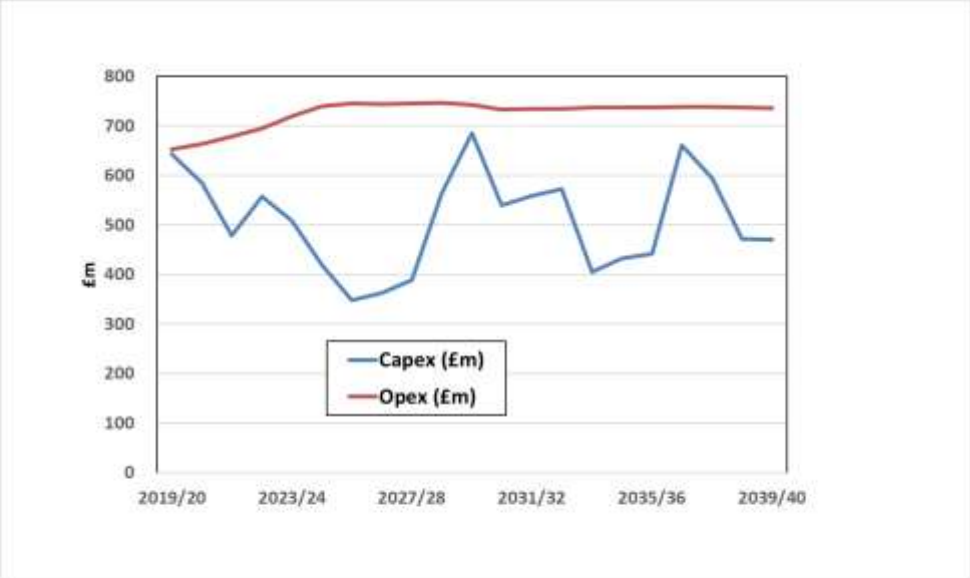


Figure 27. Capex and opex network cash flow over 20 years

By performing multiple examples of the cost benefit model to consider changes to the spectral efficiency of the network, the chart in Figure 28 has been produced. The figures here are expressed in terms of the difference in net present value (NPV) of capex, opex and Total Cost of Ownership (TCO = capex + opex) between a baseline that assumes no improvement in spectral efficiency and the value if different levels of spectral efficiency improvement are achieved.

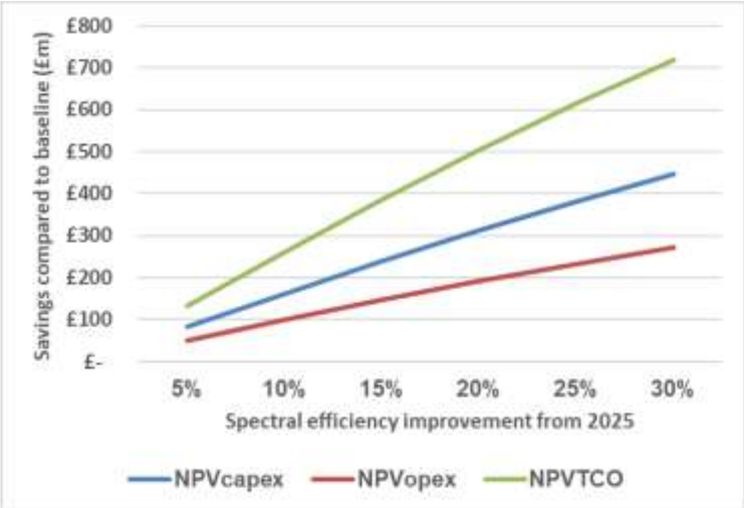


Figure 28.Savings resulting from different levels of spectral efficiency improvement

As discussed in Section 6, these values represent the sum of discounted cashflows for the period from 2020 to 2040. It is assumed that any improvements in spectral efficiency are only applied to 5G from 2025, starting three years from the completion of AIMM. In this version of the model the spectral efficiency improvement requires an update to the RAN hardware and so the rollout of the new capability is aligned to the eight year equipment replacement timescales.

The variation in traffic and handset penetration by technology are those presented in Section 5. The model has been configured to represent a UK national mobile network, with traffic, population and network build requirements calculated by geotype. As is normally the case in national networks, the greatest benefit from a reduction in capacity upgrades occurs in the urban areas where the larger traffic density leads to an increased requirement for new spectrum or additional base station sites. The level of improvement also increases over time although the nature of the financial discounting serves to reduce the impact of the later years.

In the example of Figure 28 there is no specific additional cost associated with the provision of the new capability that creates the increase in spectral efficiency. The presented cost savings represent the total value of the opportunity arising from the spectral efficiency improvement. With reference to the value flows described in Section 4, the increase in spectral efficiency will typically affect the flow of value from the mobile network operator to the RAN equipment supplier.

It appears that there is no benefit to the RAN equipment supplier to provide a new feature if it will result in them selling less equipment as traffic levels grow. This conclusion ignores the competitive pressure that is expected to occur in the RAN equipment supply market. Any equipment supplier who does not maintain efficiency in their products will ultimately lose market share and be replaced in networks as part of an end of life replacement. In Section 5.2, the split of benefit between supplier and end customer was considered with the introduction of a new feature. In reality the allocation of benefit will be based on the contractual arrangements, but the assumption is made that this value will be split equally between equipment vendor and network operator, effectively providing the equipment vendor with a price premium for their improved equipment.

Figure 28 represents the absolute value of benefit for different levels of spectral efficiency increase for a national public network. To provide more context regarding the impact on the network operator costs, Figure 29 recreates these results in terms of the percentage reduction in overall network costs as spectral efficiency increases.

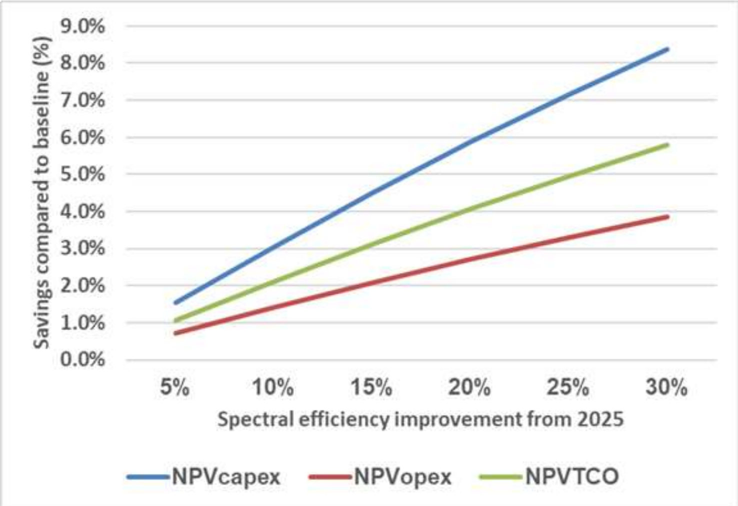


Figure 29. Percentage reduction in overall network costs relative to spectral efficiency improvement

This figure shows that the greatest percentage reduction occurs in the capex requirement in the network by reducing the need for additional equipment and some new sites. As stated above, the network opex is dominated by some fixed costs, including the ongoing rental of base station sites, and so the benefit from spectral efficiency improvement has a lesser effect on overall network opex.

8 Trust and ethical considerations

The use of AI and ML techniques is fundamental in the approach of AIMM. The use of AI in different applications has received widespread publicity and raised concerns surrounding the possibility of automated techniques to reinforce existing patterns in the data that they are based upon. The bias to reinforce these pre-existing data patterns may, in themselves, exhibit tendencies that are contrary to what society would consider as desirable or even acceptable. Historically these patterns could either go unnoticed, since the processes involved do not allow for the assessment of sufficient data for these to become apparent, or they will be accounted for by significant intervention of the analyst producing a model. It can also be the case that the analyst themselves introduces bias through their own perception.

WP2 considers the approach to the use of AI in AIMM that could be used to avoid bias in results or, where identification of bias is not possible before extensive deployment leads to new conditions that were not considered in the development, enable investigation of the issue.

The work of AIMM does not directly consider the use of personal information, such as that present in user subscription databases. However it has been recognised that the combination of a number of different data sources, including network data and additional contextual information such as geographic data, could introduce bias where a correlation of data with a particular personal characteristic is present. An example of this would be a correlation of a particular socio-economic group with particular geographic locations in the RAN coverage and the propensity of this group to access particular services.

This section reviews the work of other organisations to consider the ethical approach to the use of AI, before going on to identify an approach within AIMM to avoid bias and enable future investigation. A full assessment of all ethical approaches to the use of AI in network operation is beyond the scope of AIMM. The intention of this aspect of the work of WP2 is to provide guidance to other workpackages of aspects to consider in future deployment.

8.1 Public perception of automated networks

Given the wide range of data collection and analysis that AI has been applied to, much of the available literature considers the establishment of processes for an organisation to establish if it is to deploy AI in a sustainable way. As with any new technology, many of the risks surrounding the deployment of a technique that subsequently proves to have unacceptable features is a failure of the organisation to fully understand the way in which the technique will affect their customers. To avoid this, Table 4 is a summary of an approach proposed by Ericsson (21) for an organisation to provide safeguards.

Several of the steps identified in Table 4 fall outside the scope of AIMM and are more associated with the culture of the organisation that is introducing an AI approach. However, two principal areas, highlighted in red, are of direct relevance to AIMM. The first is the establishment of a baseline of technical and non-technical safeguards to be applied. The second is a process to monitor and control a deployed AI system to measure and encourage improvement.

Table 5 is a summary of work by ABI Research that considers the components of an ethical AI system. This recognises that any AI approach should be human centric. It is societal concerns that will dictate the general acceptance of AI techniques since a degree of scepticism will arise from the extensive use of routinely collected data. There is a limit to the extent to which technical solutions and safeguards will mitigate this scepticism since the fundamental purpose of the approach is often difficult to understand for society at large. Any application of AI approaches should therefore have a clearly stated societal benefit. A wider discussion of public concerns over the uses of AI is included in Section 8.2 and this indirect effect on the work of AMM is indicated by the dashed red box.

Other components of Table 5 are again concerned with the culture of an organisation and are of less direct relevance to AIMM. However the need to be able to explain and audit the operation of the system is directly applicable to the output of AIMM and has been outlined in red.

Table 4. Procedures to ensure the ethical use of AI (Ericsson)

Action for AI introduction	Description
1. Start at the top	Company leaders must be educated on the principles of trustworthy AI.
2. Conduct risk assessments	A risk assessment framework will be needed to map high risk activities and plan mitigation.
3. Roles and responsibilities	Employees responsible for ethics and compliance must agree roles and responsibilities with technology colleagues who have the understanding of AI.
4. Establish the baseline	Establish a combination of non-technical and technical safeguards. Non-technical measures include initiatives against discrimination and bias, technical measures involve ensuring compliant algorithms.
5. Drive company-wide awareness of ethics and AI	Provide training on ethics and AI rather than only guidelines to ensure understanding across the workforce.
6. Monitor and control	Provide continual assessment and improvement to support risk mitigation activities.
7. Onboard third parties	Ensure that third party components meet ethical requirements through supplier audits.
8. Create a speak-up culture	Ensure that channels are established to enable employees to speak up if they believe AI based solutions are having an impact on human rights.

Table 5. Key features of an ethical AI system (ABI Research)

Feature	Description
1. Human centricity	The development of AI must be in line with common societal values that are well accepted. The outcome of AI must serve as a betterment to human society as a whole.
2. Process orientation	Establish clear guidelines and mechanisms to support end to end audits.
3. Transparency and auditability	Ideally AI models should be capable of explaining to users the logic between the training and inference processes • AI models should be able to explain themselves, if possible. • AI developers should provide documentation relating to the decision making mechanism and processes which is made available for audit purposes
4. Specific requirements for AI applications	AI use cases must be subject to adequate safeguards. In particular description of the use case, how it impacts safety, end user rights and fundamental rights should be assessed.
5. Responsibility and accountability	Those responsible for the development of an AI system must be accountable for the decision making process.
6. Collaborative and inclusive	AI developers should encourage cross-disciplinary, cross-field, cross-regional and cross-border exchanges and cooperation.

The general approach of Table 4 and Table 5 is reflected in the work of the TM Forum who have created a set of checklists that an organisation should consider at different phases of the lifetime of a system that involves AI. This is presented in Table 6, where most aspects are directly relevant to the work of AIMM.

The early aspects of the TM Forum approach is to carry out an audit of potential ethical issues, so far as these can be identified in advance, and to establish clear definition of responsibilities between the different parties involved in the production. Following this, the establishment of a baseline model against which to test performance and actions with agreed test plans is required. The definition of limits of automated operation and the maintenance of records that can be audited for periodic review or investigation in the case of bias, follow the concepts in Table 4 and Table 5.

Table 6. Ethical aspects of AI Introduction checklist from TM Forum

Action for AI introduction	Description
1. Procurement	Establish a "chain of custody" to specify the responsibility for different stakeholders in the project. Carry out an ethical audit to ensure the solution is reasonable, proportionate and respects legal requirements. Identify performance metrics and targets. Define test procedures to identify bias and confirm data integrity.
2. Pre-development	Review existing AI components that are being used within the system to identify artefacts that might be carried through to the new system. Consider a simplified baseline model that can be used to test the results of the new system in terms of performance and validity.
3. Post-development	Complete test plans for verification. Identify known defects and inherent behavioural performance for potential audit by future reviews. Declare an "operational warranty" defining the limits of operation of the system within which valid result scan be obtained.
4. Deployment	Keep a record of all components necessary to reproduce the system in addition to records of test plans used in the development phase. Ensure fail safe procedures are in place to identify and respond to circumstances where the AI exceeds the operational limits. Establish an "AI contract" detailing performance expectations and establishing responsibilities between different parties producing the system. Define an end of life plan.
5. In-life	Establish regular reviews of performance, conformance with the operational warranty and adherence to the AI contract. This is particularly relevant when changes are made to the system while it is in operation.
6. End of life	Carry out an end of life review to ensure compliance with the end of life plan defined at the deployment stage.

From an assessment of the earlier literature it can be concluded that, in order to ensure an ethical approach to the use of AI, it is necessary to define procedures that include an early audit of potential issues around bias and also the avoidance of the use of compromised data. In terms of the operation of AI algorithms, a mechanism for defining boundaries for automated operation is necessary along with the collection of data and records during development and in operation to enable review and audit to take place. The possibility will always exist of bias being introduced inadvertently through the failure to recognise a correlation between patterns in the data being used to train and operate the AI algorithm and some group within society. The ability to investigate and explain the operation of any model should mitigate concerns and provide a route to resolve the unacceptable behaviour.

8.2 Public perception of automated networks

The concerns of the wider society with regard to AI systems are the result of a mixture of suspicion regarding the use of data that relates them, scepticism that sufficient control will be put in place to avoid misuse or unintended consequences, and lack of clarity of how such systems operate. If not addressed these concerns can be amplified into a distrust of new technology in general. An example of this is the online campaigns in many countries against 5G deployment on the basis of health concerns that are claimed to be specific to 5G.

One of the features identified in Section 8.1 is that AI solutions should be human centric in order to achieve acceptance. The aspects of this human centricity can be summarised as:

- **Safety** – will AI ever be a threat to society in general or human health in particular? These concerns are linked to who has control and if bias against specific groups will result.
- **Control** – how do we ensure that AI solutions will behave according to human values, preference and wellbeing? How do we audit these behaviours and take control when they are unacceptable?
- **Bias** – how do we incentivise the removal of bias (both algorithmic and data bias)? How do we audit, design and test AI solutions to mitigate bias?
- **Accessing the benefits of AI** – how do we make sure that society as a whole mostly benefits from the introduction of specific AI solutions? How do we ensure that benefits from AI are spread across the economy and remove the concern that AI will disadvantage the many and disproportionately benefit the few who can capitalise on investment in AI? How do we apply the benefits of AI to data poor environments where the collection of training data would be difficult and expensive?

If these factors are not addressed then the amplification of concern could result in a rejection of AI as a whole. Perceptions of AI as ultimately being capable of autonomous decision making, creating sentient beings that might have rights and some of whom will be bad actors, already appear in popular drama and publications. This can be illustrated in three example problems (22):

- Gorilla problem: is it a good idea to create a species that is more intelligent than you?
- King Midas problem: problems resulting with value alignment, where what we ask for has a number of inherent assumptions about what we don't want.
- 2001 Space Odyssey's HAL problem: where control is lost and the AI system pursues its goals to the detriment of its developers.

Human compatible AI, which could mitigate the concerns of society, can be described as a system where:

- The systems' objective is to maximise the realisation of human values, where human values are whatever human beings prefer to be achieved to ensure wellbeing.
- The system is initially uncertain about what those human values are and so avoids the single minded pursuit of a specific objective above everything else.
- The system learns from observing human behaviour, learning to predict what humans would prefer and not just copying behaviour but rather trying to understand the motivations behind it.

To resolve concerns in communicating the subject of AI, its benefits and risks, effort is required to uncover the ambiguity inherent in commonly used terms (23) and how these are used in different cultures, disciplines and sectors and build a consensus regarding their use. In this way a more rigorous evidence base can be developed for discussion of ethical and societal issues, drawing on a deeper understanding of what is technologically possible and more clearly explain the trade-offs between competing goals.

8.3 Potential sources of bias in AIMM

The previous sections outlined the requirements for the deployment of an ethical AI solution. To relate this to the work of AIMM, here we give an example of a use case in which the operation of AIMM can lead to a solution resulting in unintended bias. Use case (i), described in Section 2.1, is the use of AI to improve quality of service through management of interference. Within WP5, an approach has been taken to apply Reinforcement Learning (RL) to the issue of allocating frequency sub-bands from neighbouring cells to manage interference by learning the most appropriate allocation to maximise performance. Figure 30 is a schematic to illustrate this approach, where the state of the radio network following an action taken by the RL agent is rewarded or penalised depending on whether the radio environment improves towards a desired goal. Through trial the agent learns the approaches that lead to a reward and the performance of the system moves towards the desired goal.

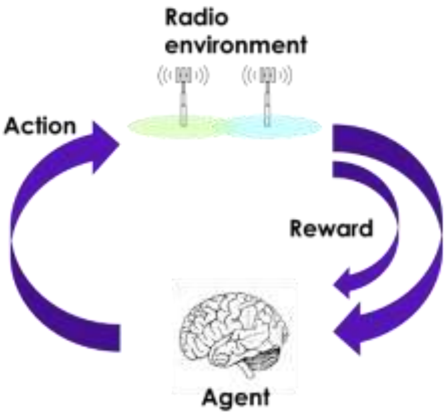


Figure 30. Schematic of the reinforcement learning for interference minimisation in WP5

Part of the design challenge of the RL technique is to define the target goal of the system. Following the discussion relating to cost in Section 5.2.1, the improvement of capacity as expressed in terms of spectral efficiency, would be an obvious goal to achieve the greatest financial benefit for the operator. It could be argued that if the data to be transferred to the algorithm from all users is given equal consideration then the resulting algorithm is without bias.

For a simple system it is possible to see how features that are inherent in the system could introduce bias. Achieving maximum spectral efficiency in a two base station system will tend to favour customers in locations close to the base station, where high signal to interference levels will provide better spectral efficiency per user. At the edge of the cell, where interference from the neighbouring base station is highest, spectral efficiency per user will be lower. At a system level therefore an algorithm with a target to maximise spectral efficiency will avoid carrying data from users at cell edge, preferring to reserve resources, in this example frequency sub-bands, for users at the cell centre.

Although the impact of such a simple approach to the selection of sub-bands could be considered undesirable because it would lead to significant changes in performance for mobile moving through the area, it remains the case that this treats all users in the same way. If the consideration of customers relates to those who are in their own home then the possibility of unintentional bias becomes apparent. If it is the case that there is a correlation between different socio-economic groups and where they live relative to base station coverage in an area then the algorithm might be treating these groups differently.

This simplistic view of bias is illustrated in Figure 31, where the residents of large houses in a one socio-economic group tend to be near the base station and, by the design of the RL algorithm, receive a much better performance than the residents of smaller houses who live near the cell edge.

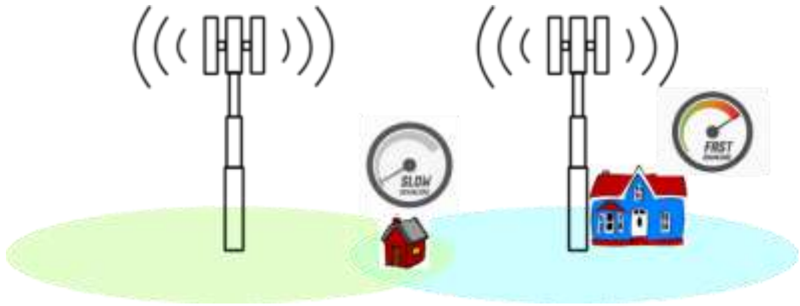


Figure 31. A simple example of unintentional bias arising from automated interference management

With reference to the ethical considerations of Section 8.1, the question arises as to if it would be possible in advance to avoid this type of bias being introduced. The choice of goal, although rational from the perspective of maximising throughput in the system, is not human centric because it does not directly consider the performance perceived by each end user. A more detailed review of performance requirements at the start of the process could have avoided this situation. Although that is obvious in this simplistic example, there will be circumstances where more complex algorithm definition and preliminary audit would still be unaware of a correlation between a feature of society and the decisions made by the algorithm.

When an outcome that is unforeseen in the design and early implementation stage of a system is identified as being unacceptable to society, this should be identified early in the routine review processes outlined in the TM Forum approach described in Section 8.1. To avoid suspicions that any identified bias is the result of intentional design, the data and processes to carry out the analysis should be recorded and suitable for audit. In the following section an approach to implementation of this approach in AIMM is considered.

8.4 Proposed methods to identify and avoid bias in AIMM

The primary purpose of AIMM is to study the use of advanced techniques in the evolution of radio access network, with particular focus on AI and Massive MIMO. A detailed consideration of the processes to be employed to ensure an ethical use of AI is outside the scope of AIMM. However it is within the scope of AIMM to propose how the necessary techniques of review, transparency and audit could be included in any solution and consider demonstrations of this approach.

As stated in Section 1, one of the assumptions of AIMM is that future solutions will be based on the O-RAN architecture. Within this architecture, the RAN Intelligent Controller (RIC) is where applications will be deployed to manage the RAN based on network measurements in addition to data made available from other sources. Part of the O-RAN concept is that these applications (xApps for near real-time and rApps for non-real time applications) can be supplied by third parties. The WP5 example of reinforcement learning for interference management, described in Section 8.3 has been implemented as an xApp.

Within this architectural framework, a potential audit function could be deployed as a separate xApp on the RIC. This would have access to all the data being used by the management xApps. The audit xApp could itself be implemented to identify when the performance of the RAN as a whole is operating outside some previously defined limits and commence data collection at that point.

9 5G Security

5G is the next generation of wireless technology and differs significantly from previous wireless networks. 5G expands upon previous generations through ubiquitous connectivity of things to people, services, the Internet, and things. To accomplish this, the network is re-architected to utilize software defined networking (SDN) for adaptability, network functions virtualization (NFV) for new services and enhanced capabilities and cloud-native architectures for scalability of resources. The novel network infrastructure enables disaggregation and virtualization, leading to a Control Plane and User Plane Separation with 5G Non standalone (NSA) and introduces capabilities like network slicing and multi-access edge computing with 5G Standalone (SA). Although the enhancements described above provide many benefits in improving spectral efficiency and flexible support of new types of services, it introduce new security domains requiring more complex protection and detection systems.

9.1 Security Concerns of Open Systems

The 5G architecture takes advantage of many new open system concepts such as:

Cloud-Native Architectures: like self-contained functions within or across data centers communicating in a micro-services environment with all elements working together to deliver services and applications. The cloud-native 5G architecture enables an elastic, automated environment where network, compute and storage services can expand, and contract as needed. Many telecommunications and mobility functions can now be hosted as software services and dynamically instantiated in different network segments. The overall 5G network needs to be pliable and is ultimately designed to be software configurable.

SDN & NFV: Software Defined Networking (SDN) and Network Function Virtualization (NFV) are the key pillars of future 5G communication systems that promise to support emerging new and applications such as enhanced mobile broadband, ultra-low latency, massive sensing type applications while improving the Time to Market and resiliency of the next generation wireless networks. Service providers and other vertical markets are poised to leverage SDN/NFV to provide flexible and cost-effective service without compromising the end user quality of service (QoS). While NFV and SDN allows for flexible networks and rapid service creation, these offer increased security concerns due to the introduction of additional complexities and the inherent open nature of this network deployment. To date, little attention has been given to the security aspects of virtualization. Although several standardization bodies have started looking into the many security issues introduced by SDN/NFV, additional work is needed with larger security community.

Dis-aggregation or User and Control Planes: Control plane and user plane separation is the concept of disaggregation that allows these two planes to exist on separate devices or at separate locations within the network. Separating the control plane from the user plane allows the two planes to scale independently, without having to augment the resources of one plane when additional resources are only required in the other. This separation allows the planes to operate at a distance from each other; they are no longer required to be co-located, further improving scalability.

Network Slicing: refers to the technique of isolating the end-to-end performance of a portion of the network compared to another. It involves all three domains of 5G: the RAN, the transport network and the 5G Core. This separation includes the user equipment (UE), whether that be a handset, fixed-wireless access point or IoT sensor. Network slicing delivers the stringent characteristics that 5G offers (eMBB, mMTC, URLLC) for specific subsets of users, operators, or applications.

Distributed Architecture: 5G will have a much more distributed architecture than previous wireless versions. A distributed architecture is needed because: the variety of frequency ranges used to deliver high bandwidth and the density of 5G end-points require many more base stations or small cells. 5G latency requirements for some services will require some application processing occurs much closer geographically to the user.

Complex Traffic Patterns: The nature of 5G traffic patterns will be unpredictable and ever changing. With a distributed architecture composed of virtualized elements, traffic will flow between services, elements, functions, and devices. Such a distributed architecture may introduce additional attack points but may also make attacks harder to execute. With a disaggregated packet core, parts in one location will be communicating with parts in another location. The packet core itself will become geographically diverse and more resilient, e.g., implemented by a primary packet core and a backup packet core. 5G enables “any-to-any” communication, for edge-component to edge component traffic flows.

Flexible Software Deployment: One of 5G’s goal is to make the network flexible and programmable. This can be achieved by the virtualization of many functions and services previously implemented in hardware. Software can take advantage of open-source software modules but may have some associated vulnerabilities. Vulnerabilities can be reduced when open-source communities thoroughly test software to discover and correct most of its flaws. Potentially greater vulnerabilities exist if a software supply chain does not have safeguards in place. If the origin, reliability and integrity of the open-source modules are unknown, it can introduce flaws and vulnerabilities. Dependency is another conflict; if a software module relies on other pieces of code, then the sudden lack of access to that dependent code could be problematic.

The new 5G cloud-based and distributed network will usher in a myriad of new and valuable services and lower cost that what was achievable with previous generation of wireless networks. However, the increase complexity as well as the “Openness” of the development and deployment architecture bodes well for nefarious and coordinated attacks of the new wireless environment. Some of these concerns will be addressed in the following section.

9.2 Security Concerns – Issues arising from more open systems

Supply Chain Threats
Hardware supply chain security is a well-known area for most operators, however the emerging software supply chains being developed for 5G deployments are mostly unknown and can introduce significant security threats. Both hardware and software supply chains need to be viewed with a zero trust approach from a Network Operations. The operators need to mature their Supplier Risk and Third-Party Risk Management processes to incorporate the new open-source platforms and open-source operating systems, and vendor software deliveries must be scanned and/ or routinely updated or hardened. The operators must require that the OEMs embrace a secure and reputable security auditor. Incorporating the security auditing by the OEMs via an independent 3rd party auditor will ensure that the OEMs follow best practices towards secure software development and securing the end-to-end supply chain.

Cloud Virtualization Threats
Since cloud computing systems comprise of various resources which are shared among users, it is possible that users can spread malicious traffic or consume more resources or stealthily access resource of other users impacting the performance of the whole system. Similarly, in multi-tenant cloud networks where tenants run their own control logic, interactions can cause conflicts in network configurations. Mobile Cloud Computing (MCC) migrates the concepts of cloud computing into the 5G eco-systems creating a number of security vulnerabilities that arise due to the architectural and infrastructural modifications in 5G. Therefore, the open architecture of MCC and the versatility of mobile terminals create vulnerabilities through which adversaries could launch threats and breach privacy of 5G based Cloud infrastructure.

SDN centralizes the network control platforms and enables programmability in communication networks. These features create opportunities for hacking the network. For example, the centralized control can be vulnerable to DoS attacks and exposing the critical Application Programming Interfaces (APIs) to unintended software which can severely impact the network performance. Also, NFV platforms do not provide proper security and isolation to virtualized telecommunication services. One of the main challenges persistent to the use of NFV in mobile networks is the dynamic nature of Virtual Network Functions (VNFs) that can lead to configuration errors and thus security lapses.

IoT Device Threats
With the drastic increase in IoT technology, the risk of security threats and challenges are also increasing rapidly. Not only the technology but also threats are getting smarter the threats can impact many mission critical use-cases and industries that are deploying IoT based networks as part of their critical infrastructure. This security concern needs to be resolved with the ability to detect threats with proposed IoT solutions as the threat landscape differs across IoT applications.

Network Slice Vulnerability

Network slicing is central to realizing many of 5G’s more ambitious capabilities because it enables individual access points or base stations to subdivide networks into multiple logical sections—slices effectively providing entirely separate networks for multiple uses. The slices can be used for different purposes such as mobile broadband for end-users and massive IoT connectivity—at the same time, without interfering with each other. Recently, vulnerabilities have been discovered that, if exploited, can enable an attacker on one slice to gain access to data being exchanged on another or in some circumstances gain access to the 5G provider’s core network.

Adversarial Attacks on the “Open” Air Interface

Many characterize threats on the “Open” Air Interface and RAN Spoofing or data collection (privacy) threats. But there is more to the Security of the Open Air interfaces relating to the Wireless Physical layer. These will be defined below.

RAN Related Security Threats

In recent years, a large body of literature has revealed numerous security and privacy issues in 4G mobile networks. Most of the published attacks at the 4G RAN layer involve IMSI catchers to target IMSIs during the UE’s initial attach procedure to the network, or paging attacks using the IMSI paging feature. In such attacks, the obtained information about particular IMSIs may be used later for other types of attacks. Fortunately, the 5G technology and standards are expected to address the known threats at this layer at all access types since 5G does not transmit an unencrypted IMSI.

Also, the data and signaling transmitted and received at the radio layer is expected to be appropriately encrypted and integrity protected at higher layers, whenever possible. However, Rogue Base Station (RBS) threat where the RBS masquerades as a legitimate base station to facilitate a Man-in-The-Middle (MiTM) attack between the mobile user equipment (UE) and the mobile network. An attacker can use the RBS to launch different attacks on mobile users and networks. These attacks include stealing user information, tampering with transmitted information, tracking users, compromising user privacy or causing DoS for 5G services. Although 5G has taken measures to reduce the RBS threat, 5G networks could still be a target to RBS-based threats using recently discovered threat vectors (details not covered here).

Open Air – Physical-Layer Threats

As we have seen with LTE, despite being designed for commercial communications, the latest cellular technology is often utilized for mission-critical applications such as public safety and military communications. Just as we have become dependent on LTE, over the next decade we will likely become dependent on 5G NR, which is why we must ensure it is secure and available when and where it is needed. Unfortunately, like any wireless technology, disruption through deliberate radio frequency (RF) interference, or jamming, is possible.

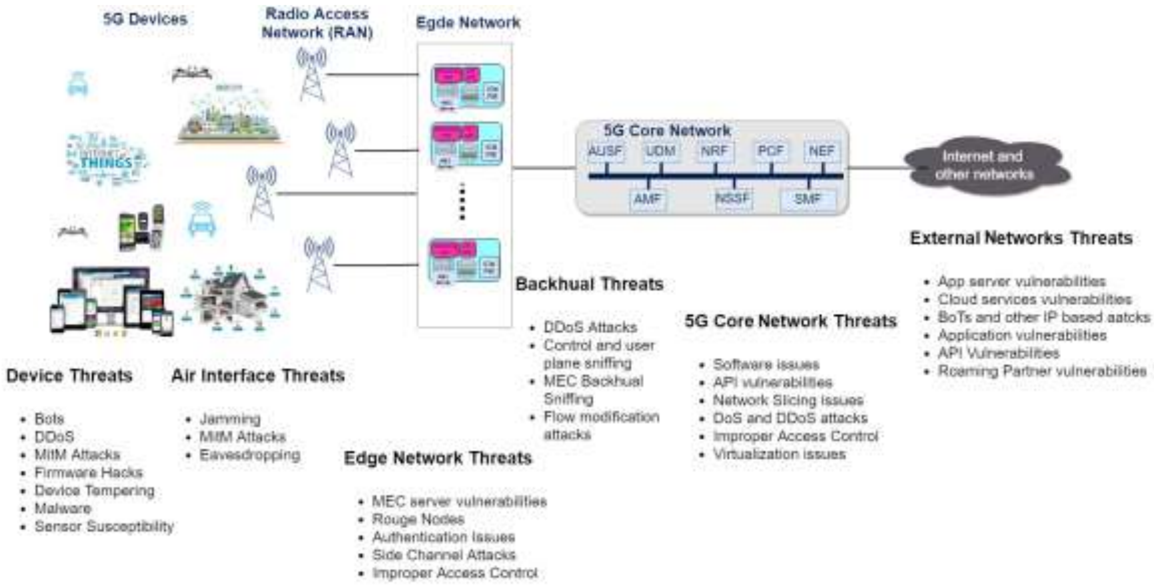


Figure 32. An overview of threats in 5G networks *(24)*

9.3 Approaches to security through AIMM

Security Vulnerabilities introduced within the Trust-Chain by De-coupled functions including the use of Open-Source code. Cyber supply chain risks touch sourcing, vendor management, supply chain continuity and quality, transportation security and many other functions across the enterprise and require a coordinated effort to address.

Cyber Supply Chain Security Principles:

1. Develop your defenses based on the principle that your systems will be breached. The question becomes not just how to prevent a breach, but how to mitigate an attacker's ability to exploit the information they have accessed and how to recover from the breach.

2. Cybersecurity is a people, processes and knowledge problem. Breaches tend to be less about a technology failure and more about human error. IT security systems won't secure critical information unless employees throughout the supply chain use secure cybersecurity practices.

3. Holistic Security. There should be no gap between physical and cybersecurity.
Key Cyber Supply Chain Risks that should be addressed

Cyber supply chain risks covers a lot of territory. Some of the concerns include risks from:

- Third party service providers or vendors with physical or virtual access to information systems, software code, or IP.
- Poor information security practices by lower-tier suppliers.
- Compromised software or hardware purchased from suppliers.
- Software security vulnerabilities in supply chain management or supplier systems.
- Counterfeit hardware or hardware with embedded malware.
- Third party data storage or data aggregators.

Intelligent Controllers potential Attack Surface
5G networks are poised to become very complex systems in order to meet key spectral efficiency metrics while providing new levels of service quality across a number of network service types from IoT and ultrawide bandwidth. A communications systems capability of operating in these objectives will require the use of very sophisticated intelligent control systems using the latest analytic and predictive algorithms and technologies. With any advanced technology integrated in large scale deployments and potential across multi-vendor configurations it will be inherently difficult to ensure system behaviour is well understood and manageable under all system boundary conditions. Thus, it is important to ensure that in addition to the traditional security policies and controls such as supply-

side verifications and 3rd party SW vulnerability assessments are conducted, expected system behaviour is also well understood. Harmful or malice code could inadvertently be introduced into intelligent controllers that could impact 5G services under certain conditions. Testing of intelligent control software under all expected conditions will be required to ensure biased or malware and related impacts are detected before being put into service.

Management Interfaces provide a significant attack surface to potential hackers and may not always align with best security practices. Some of the key practices that should be adopted are:

- Isolate the management interface on a dedicated management VLAN.
- Limit the source IP addresses allowed into the management network to those of your dedicated management devices,
- Use Authentication Policy with multi-factor authentication (MFA) to require administrators to successfully authenticate before you allow them to continue to the firewall web interface login page or CLI login prompt.
- Limit access to users in your security admin, network admin, or IT user groups, as appropriate for your organization.
- If you must enable remote access to the management network, require access through a VPN.
- Only use Admin Accounts for Administration usage. Never use personal or corporate user accounts to access Admin functions.

RF / Wireless (PHY level) Disruptions & Attacks

Wireless attacks as defined above can be unintended in the case of interference or nefarious acts with intentions to disrupt, destroy or gain from the network impairment caused as well as to collect and extricate private data. Both disruptions caused by interference and physical layer attacks should be monitored.

AIMM project – Work item WP5.4 is developing a Wireless Monitoring capability for the Next Generation RF based communications (5G, WIFI) to enable businesses and critical infrastructures owners with access to real time information to monitor and detect interference and attacks on their networks in order to optimize quality of Service (QoS), network performance and security. The Wireless interface detection system is based on deep learning outlier detection algorithms which learn the characteristics of a 5G (high Density) signal and is able to discern between a normal signal with that of a signal with narrowband interference & attacks as well as smart Jamming threats that are able to evade traditional interference detection systems.

10 Future Work Beyond AIMM

The intention of AIMM is to consider the improvement in the performance of the RAN using a variety of techniques from improvement of the air interface (WP4), to changes in the network architecture (WP3) and evolution of techniques in the management of the RAN (WP5). All of these approaches involved the use of advanced techniques, including AI/ML, that make use of the data increasingly available from the RAN. This is particularly the case for the O-RAN architecture, with defined interfaces for the exchange of data with the RAN Intelligent Controller (RIC). Many of the approaches and use cases in AIMM have involved increasing the efficiency of providing and capacity and coverage in the RAN and so reducing cost of installation and operation.

Through the course of AIMM, and particularly in the work of WP2, awareness has grown that a more holistic approach to the operation of the RAN is required for the future. This is particularly true for the vision of 5G where the number of mobile networks will grow with the emergence of private networks and the support of multiple groups of users with differing service requirements enabled by the concept of network slicing. Network slicing will reduce the current separation between public and private networks, enabling a combination of both on shared infrastructure.

In addition to the provision for multiple service requirements across multiple, geographically collocated networks which share functions at different levels in the architecture to meet the user needs, additional considerations are growing in importance. Power consumption has historically been regarded as a lower priority than the need to improve spectral efficiency and coverage on a radio network. With the move away from fossil fuels for environmental reasons and increasing uncertainty in the global supply chain for energy with its resulting impact on energy prices, energy efficiency is an area of increasing priority. Increasing dependence on the use of mobile devices for societal inclusion and reliance on mobile networks as a whole for the provision of critical national infrastructure has also provided an increased focus on resilience and security.

One of the challenges for the future will be to balance the different and sometimes competing requirements for access to networks from different users while improving efficiency and reducing the impact of network node failure or localised impaired performance. This impaired performance could be unintentional or the result of malicious activity. Performing the balance of different requirements in near-real time could not be performed manually or by systems performing off-line analysis. It is also the case that a centralised control system performing this activity would ultimately become cumbersome, difficult to maintain and itself represent a potential point of failure to disrupt communication. A more distributed approach of multiple AI/ML agents optimising within a federated hierarchy of other AI/ML agents would provide both flexibility and resilience.

For the reasons above it is proposed that future work in the area of AI/ML should concentrate on a holistic approach to RAN management, operation and optimisation. This could define the necessary boundaries and interactions for individual AI agents in a wider federation. The following section gives some examples of the areas where management and optimisation are required and how the inter-connected nature of these actions require an overriding layer of management.

10.1 Efficiency

Traditionally efficiency in mobile networks has focused on cost reduction through improvements in coverage and capacity of network nodes to maintain an acceptable and generally uniform quality of service. Additional efficiencies arise from reductions in the level of human intervention required in the planning and operation of the network. Going forward, the achievement of efficiency will include other factors, only some of which could be expressed in financial terms.

- Coverage and capacity
    - This will continue as a significant area of consideration, with increasing attention paid to the variation of the network response to different traffic patterns and user distributions. The expansion of mobile networks from predominantly public access national networks serving a general consumer requirement to a combination of public, private and hybrid networks will change the balance of spectral efficiency and service requirements.
- Power efficiency
    - Earlier work on power efficiency through sleep modes have included a focus on the handset and impacts on battery life. Increasingly, power consumption is a societal and financial concern and work has extended to powering off aspects of the RAN at

times of low traffic, generally in response to longer term, time of day trends. These changes could be made more dynamic, relevant to the instantaneous coverage requirements of individual base station sector user plane connections.

- Operational efficiency
    - As the number of different service types and different user communities grow on mobile networks, management and operation of services will become increasingly complicated. The introduction of a new service on an infrastructure shared between different communities, ensuring that all SLAs can be met, would involve significant delays using current techniques. It is expected that the definition of the required control and optimisation techniques will require ML approaches, federated between the control functions serving the different user communities.
- Maintenance of service level agreements
    - Multiple user communities operating in the same network will have different service level agreements (SLAs). Achieving the SLAs for users is generally a function of the scheduler based on some prioritisation between the users, with the SLA being exceeded in times where there is reduced network traffic.
    - In future, meeting the SLA could also be balance against other factors, including power reduction and the control of interference between different sites across the network. Rather than a traditional approach where the network is defined by its capability to meet an SLA in the local network busy hour, the capability of the network could be adjusted to reduce power consumption to the minimum required to continue to satisfy the SLA in light traffic periods.

10.2 Resilience

In recent years, the ability to have access to a personal mobile device has become increasingly important to end users. In the UK, Ofcom reported in 2019 that of all devices they own, half of adults would miss their mobile phone the most (25). In response, expectations of the availability of service have increased and so place additional requirements on network resilience in times of equipment failure or unforeseen natural events, including storms.

Here, resilience is taken to include the response to all factors that might impair the service to the end user. Different end user communities will have different levels of resilience assured in their SLA and so the network should intelligently trade off the capability of the network in times of fault or loss of power or connectivity. Aspects of resilience include:

- Node failure
    - Rapid automated identification of the failure of a network node.
        - Automated assessment of the timescale of the failure in terms of its impact on customers and potential time to rectification (e.g. minutes, days, months)
    - Automated determination of the expected response of a node experiencing failure and any other surrounding or impacted nodes in the area.
        - The response could follow an earlier, pre-tested response to an expected failure mode. Planning and algorithm training for this approach could be through the use of a network digital twin.
        - For longer term failures the system could learn a new set of solutions, trading off the impact on surrounding nodes with meeting SLAs for the affected user communities.
    - Trade-off the impact on power consumption and SLAs
        - Determine the appropriate response in terms of increasing transmit power on surrounding sites (in the case of a base station failure) or on carriers on the same site (in the case of a low frequency carrier failure). This will need to take account of both the capability of the equipment and other regulatory requirements, including licence obligations and public exposure regulations.

    - For longer term outages, learn the response to provide efficiency in the new network configuration.

- Performance degradation
  - Anomalous detection
    - Automated identification that the behaviour of the network has altered in some way that is impacting the network efficiency (e.g. failure to meet SLA, capacity reduction)
    - Identify the timescale of the effect (e.g. repetitive short term, continuous long term)
    - Identify the root cause of the network degradation both in location and the nature (e.g localised interference)
  - Trade-off the effects of mitigating the performance degradation with the impact across an area involving surrounding nodes. It is anticipated that this will involve the use of federated AI, with one control function proposing a potential change to operation that other control functions will need to assess the impact on their own operations.

10.3 Security

Security in this sense is taken to be effects detectable on the RAN that could point to a malicious attack on the network or on the integrity of a user's communications. It is assumed that non-malicious interference events will be covered by the resilience considerations above.

- Anomalous interference detection
  - Potential correlation with anomalous traffic patterns on the local radio network.
  - Correlation with anomalous cyber activity acting against the network as a whole.
  - Correlation with individual user identification at the core network or radio level (e.g. by rf fingerprinting, identifying specific devices through the effects of manufacturing variability in the device components).
- Continuing assessment of response to mitigation actions (similar to the activities in the "resilience" section).

These security features have, by their nature, a range of personal data and other ethical considerations that do not occur to the same extent in the automated management of efficiency and resilience. These ethical concerns are not limited to the radio network and are expected to form part of a wider study concerned with the protection of the end to end service across the network.

11 References

1. **document, AIMM WP2 internal.** *All Use Case proposals.* November 2020.
2. **document, AIMM WP2 Internal.** *AIMM Deployment models.* October 2020.
3. **Oughton, E.j. and Frias, Z.** The Cost, Coverage and Rollout Implications of 5G Infrastucture in Britain. *Telecommunications Policy.* 2018, Vol. 42.
4. **Vodafone.** O2, Three and Vodafone agree new deal to enhance rural coverage. *www.vodafone.co,uk.* [Online] 27 January 2021. [Cited: 27 April 2022.] https://newscentre.vodafone.co.uk/press-release/o2-three-and-vodafone-agree-new-deal-to-enhance-rural-coverage-srn/.
5. **Mason, Analysys.** *Wireless network data traffic: worldwide trends and forecasts 2021–2026.* 2021.
6. **Ofcom.** Ofcom MTR model. [Online]
7. **Mason, Analysys.** What are key considerations for 5G sites? *www.analysysmason.com.* [Online] 2 October 2019. [Cited: 27 April 2022.] https://www.analysysmason.com/research/content/perspectives/5g-key-considerations-rma18/#:~:text=%22Analysys%20Mason%20forecasts%20that%20there,increasing%20levels%20of%20data%20traffic.).
8. **ACIA, 5G.** 5G Non-Public Networks for Industrial Applications. *5g-acia.org.* [Online] September 2021. [Cited: 27 April 2022.] https://5g-acia.org/wp-content/uploads/5G-ACIA_5G_Non-Public_Networks_for_Industrial_Scenarios_09-2021.pdf.
9. **Ofcom.** Telecommunications Market Data Update Q4 2021. *www.ofcom.org.uk.* [Online] 28 April 2022. [Cited: 29 April 2022.] https://www.ofcom.org.uk/research-and-data/telecoms-research/data-updates/telecommunications-market-data-update-q4-2021.
10. **Mason, Analysys.** Sample: Connected Consumer Survey . *www.analysysmason.com.* [Online] April 2021. [Cited: 3 May 2022.] https://www.analysysmason.com/contentassets/9e3d634d9c08457ba94c6d98f97400e6/analysys_mason_customer_satisfaction_europe_usa_sample_apr2021_rdmm0.pdf.
11. **3GPP.** *www.3gpp.org.* [Online] 31 March 2022. [Cited: 6 May 2022.] https://www.3gpp.org/ftp/Specs/archive/38_series/38.901/. TR 38.901 v17.0.0.
12. **Openreach.** Ethernet Access Direct (EAD). *www.openreach.co.uk.* [Online] 29 April 2022. [Cited: 6 May 2022.] https://www.openreach.co.uk/cpportal/products/ethernet/ethernet-access-direct.
13. **Deloitte.** The Open Future of Radio Access etworks. *www2.deloitte.com.* [Online] 2021. [Cited: 11 May 2022.] https://www2.deloitte.com/content/dam/Deloitte/pt/Documents/technology-media-telecommunications/TEE/The-Open-Future-of-Radio-Access-Networks.pdf.
14. **Mason, Analyys.** OpenRAN could deliver up to 30% TCO savings. *analysysmason.com.* [Online] 23 February 2022. [Cited: 11 May 2022.] https://www.analysysmason.com/research/content/perspectives/open-ran-tco-rma18-rma16/.
15. **FCC.** DA-21-355A1.pdf (fcc.gov). [Online]
16. **NGMN.** *Green future Networks: Network Energy Efficiency v1.1.* s.l. : NGMN, 2021.
17. **GSMA.** Going green: Benchmarking the enery efficiency of mobile. *data.gsmaintelligence.com.* [Online] June 2021. [Cited: 20 May 2022.] https://data.gsmaintelligence.com/api-web/v2/research-file-download?id=60621137&file=300621-Going-Green-efficiency-mobile.pdf.
18. —. 5G energy efficiencies: green is the new black. *data.gsmaintelligence.com.* [Online] November 2020. [Cited: 20 May 2022.] https://data.gsmaintelligence.com/research/research/research-2020/5g-energy-efficiencies-green-is-the-new-black.
19. **Mason, Analysys.** What are key considerations for 5G sites? *www.analysysmason.com.* [Online] 2 October 2019. [Cited: 22 May 2022.] https://www.analysysmason.com/research/content/perspectives/5g-key-considerations-rma18/#:~:text=%22Analysys%20Mason%20forecasts%20that%20there,increasing%20levels%20of%20data%20traffic..
20. **Ericsson.** www.ericsson.com. *5G Pacesetters: winning in the eyes of consumers and growing revenues.* [Online] November 2021. [Cited: 22 May 2022.] https://www.ericsson.com/en/reports-and-papers/consumerlab/reports/5g-pacesetters/overview.
21. —. Ethics and AI: 8 steps to build trust in technology - Ericsson. [Online]
22. **YouTube.** https://www.youtube.com/watch?v=EBK-a94IFHY . [Online]
23. **Foundation, Nuffield.** https://www.nuffieldfoundation.org/sites/default/files/files/Ethical-and-Societal-Implications-of-Data-and-AI-report-Nuffield-Foundat.pdf . [Online]
24. **IEEE.** [Online]

25. **Ofcom.** Adults: Media use and attitudes report 2019. *ofcom.org.uk.* [Online] 30 May 2019. [Cited: 22 May 2022.] https://www.ofcom.org.uk/__data/assets/pdf_file/0021/149124/adults-media-use-and-attitudes-report.pdf.
26. **Forum, TM.** AI checklist. [Online]
27. **GSMA.** GSMA | 5G-era Mobile Network Cost Evolution - Future Networks. [Online]
28. **Mason, Analysys.** [Online]

BT

CEMWorks

IMST

interdigital

Loughborough University

NOKIA

thinkRF

University of BRISTOL

Universität Stuttgart

Vilicom
signalling the future

Bundesministerium für Wirtschaft und Energie

UKRI Innovate UK

Canada
NRC·CNRC