ITUNews MAGAZINE

No. 5, 2020

AI and Machine Learning in 5G

Lessons from the ITU Challenge

ITU

AI and machine learning in 5G – the ITU Challenge 2020

By Houlin Zhao, ITU Secretary-General

In February this year, the International Telecommunication Union (ITU) set the first ITU AI/ML in 5G Challenge in motion – a global competition that will culminate in an online prize-winning event on 15–17 December, 2020.

Through the Challenge, ITU encourages and supports the growing community driving the integration of artificial intelligence (AI) and machine learning (ML) in networks and at the same time enhances the community driving ITU standardization work for AI/ML.

The ITU Challenge enables the collaborative culture necessary for success in emerging and future networks such as 5G and creates new opportunities for industry and academia to influence the evolution of ITU standards.

As the UN specialized agency for ICTs, ITU plays a central role in ensuring that these networks are rolled out widely and follow the highest quality standards. Most recently, we announced the approval by our 193 Member States of an ITU Radiocommunication Sector (ITU–R) Recommendation: "Detailed specifications of the radio interfaces of IMT-2020."

IMT-2020 specifications for the fifth generation of mobile communications (5G) will be the backbone of tomorrow's digital economy, leading industry and society into the automated and intelligent world and promising to improve people's lives on a scale never seen before.

In this edition of the ITU Magazine you will learn all about the ITU AI/ML in 5G Challenge and also find ample insight articles from industry and academia.

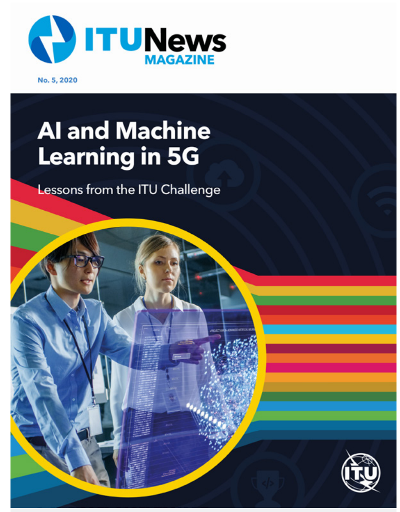
The Grand Challenge Finale will feature keynotes by Professor Vincent Poor of Princeton University, United States, Chih-Lin I of the China Mobile Research Institute, and Wojciech Samek of Fraunhofer HHI, Germany. It will also launch the Challenge 2021. Enjoy!

> *"Through the Challenge, ITU encourages and supports the growing community driving the integration of artificial intelligence and machine learning in networks."*
>
> Houlin Zhao

Contents

AI and Machine Learning in 5G

Lessons from the ITU Challenge



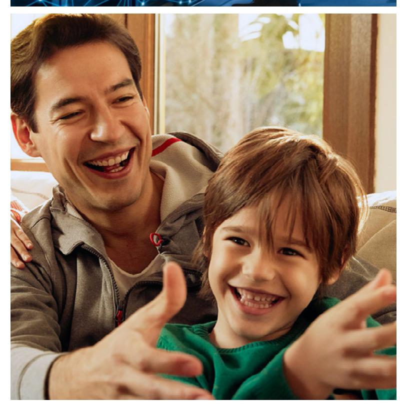
Cover photo: Shutterstock

ISSN 1020-4148
itunews.itu.int
Six issues per year

Editorial Coordinator & Copywriter: Nicole Harper
Art Editor: Christine Vanoli
Editorial Assistant: Angela Smith

Editorial office:
Tel.: +41 22 730 5723/5683
E-mail: itunews@itu.int

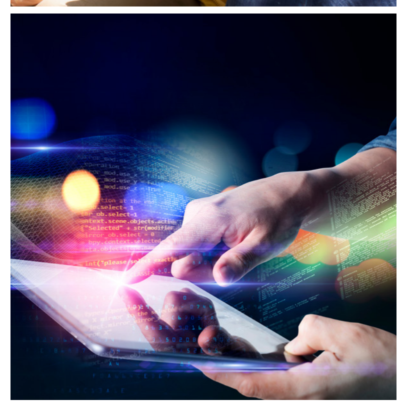
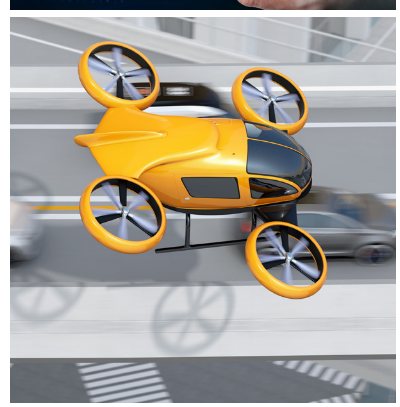
Mailing address:
International Telecommunication Union
Place des Nations
CH-1211 Geneva 20 (Switzerland)



Contents

Insights from academia




ITU thanks the sponsors of the 2020 AI/Machine Learning in 5G Challenge

Gold Sponsor

TRA, United Arab Emirates

TRA

TRA هيئة تنظيم الاتصالات

TELECOMMUNICATIONS REGULATORY AUTHORITY

Bronze Sponsors

Cisco Systems and ZTE

CISCO

CISCO™

ZTE

ZTE

Building community and trust on the ITU platform

ITU News spoke with Chaesub Lee, Director of the ITU Telecommunication Standardization Bureau, to learn more about the context for the ITU Challenge on AI and Machine Learning in 5G and its connection with the strategic priorities of ITU.

This edition shares experiences from the ITU Challenge. How would you describe the aims of the ITU Challenge?

The ITU Challenge provided a platform for participants to apply ITU's Machine Learning Toolkit in solving practical problem statements. The ITU Challenge allowed participants to connect with new partners in the ITU community – and new tools and data resources – to achieve goals set out by problem statements contributed by industry and academia in Brazil, China, India, Ireland, Japan, Russia, Spain, Turkey and the United States. It offered participants an opportunity to showcase their talent, test their concepts on real data and real-world problems, and compete for global recognition.

How does the ITU Challenge factor into the strategic priorities of ITU?

Building community and trust is at the heart of all that we do at ITU. We are a global membership of 193 Member States and over 900 companies, universities, and international and regional organizations. ITU standards are developed in a community, building the mutual understanding that enables the community to advance together. ITU standards are significant feats of international collaboration. They represent voluntary commitments to common approaches to technology development, application and supporting business relationships. The value of ITU standardization, just like the value of the ITU Challenge, lies in the community that it creates.

> *"The ITU Challenge provided a platform for participants to apply ITU's Machine Learning Toolkit."*
>
> Chaesub Lee
>
> Director, ITU Telecommunication Standardization Bureau

How do ITU standards connect with the ITU Challenge and how might this connection evolve?

New ITU standards for AI/ML provide toolsets to enable AI/ML integration in 5G and future networks as these networks evolve. The ITU Telecommunication Standardization Sector (ITU–T) Y.3172 architecture – derived from the study of use cases published in ITU–T Y.Supplement55 – introduced the basic toolsets in relation to the underlying network: ML Pipeline for model optimization and serving; ML Sandbox to trial models before deployment; and ML Function Orchestrator (MLFO) to control AI/ML integration. ITU–T Y.3173 (intelligence evaluation), ITU–T Y.3174 (data handling) and ITU–T Y.3176 (marketplace integration) all build on the ITU–T Y.3172 architecture. The ITU Challenge aimed to demonstrate and validate these ITU standards and create new opportunities for industry and academia to influence their evolution.

> *5G represents major advances in networking to meet the needs of a very diverse set of applications, across industry sectors.*
>
> Chaesub Lee

Why are AI/ML and supporting standards important to 5G and future networks?

Companies in the networking business are introducing AI/ML as part of their innovations to optimize network operations and increase energy and cost efficiency. 5G represents major advances in networking to meet the needs of a very diverse set of applications, across industry sectors. Networks are growing in sophistication and complexity. AI/ML will be key in managing this complexity. The ITU–T Y.317x standards provide versatile toolsets to support AI/ML integration in tune with network evolution.

Standard "toolsets", built to be adapted to evolving user requirements and a broad scope of use cases, are also found in ITU standards in fields such as multimedia, security, blockchain and quantum information technology.

The ICT industry evolves very rapidly. How have recent years' evolutions impacted ITU standardization?

ITU's standardization arm (ITU–T) has seen a very strong increase in new members in the past four years, topping over 50 last year. We are addressing exciting new subjects, but the role of the ITU platform has remained unchanged for over 150 years – we build community and trust to enable information and communication technology (ICT) advances on a global scale. The ITU standardization platform – for many years central to building mutual understanding within the ICT sector – is now helping the ICT sector to build mutual understanding with its many new partners. We see new partners collectively advancing ITU standardization work in fields such as smart cities, energy, health care, finance, automotive, and AI/ML.

How has ITU approached this need to support a more diverse set of ICT applications?

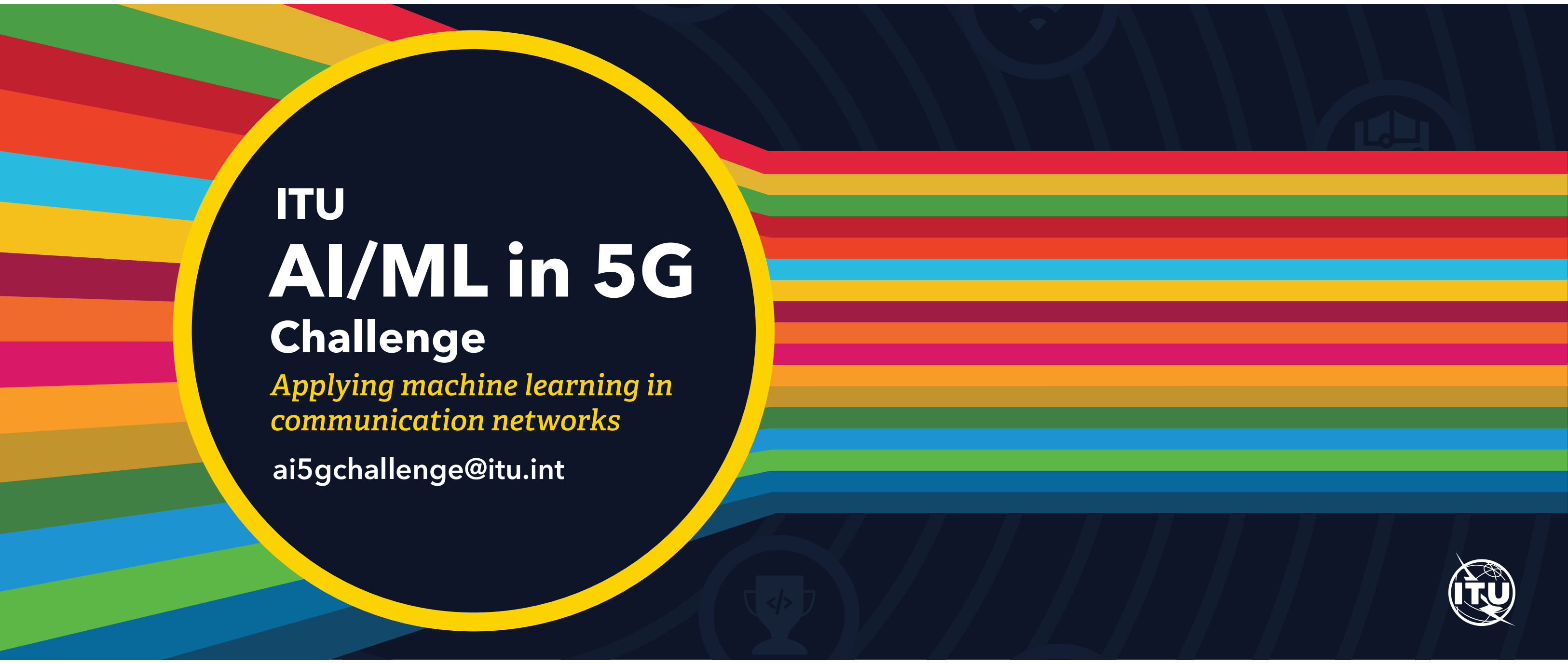
Although ITU's role in building community and trust remains unchanged, we have entered a new era of standardization in need of new approaches to continue building this community and trust. We have spent many years bringing ICT decision-makers together with decision-makers in other sectors. This inclusive dialogue has helped us to create the conditions necessary to deliver influential standards in fields of innovation given life by new partnerships; fields such as digital health, digital finance, intelligent transport systems and AI/ML. Here we see the value of open platforms such as ITU focus groups or the AI for Good Global Summit. These open platforms help to build community and trust. They help to clarify the contributions expected of different stakeholders, including the contribution of ITU standardization.

Where is the influence of AI/ML most pronounced in ITU standardization work and what are the opportunities to participate?

AI/ML is playing a key part in ITU standardization work in fields such as network orchestration and management, multimedia coding, service quality assessment, digital health, environmental efficiency, and autonomous driving. And the concept of a truly autonomous network – enabled by the Level 5 intelligence described by ITU–T Y.3173 – has sparked considerable discussion in ITU. We welcome you to join us.

ITU continues to grow in inclusivity. This year we introduced a reduced membership fee option for start-ups and SMEs. Academia have benefitted from reduced fees since 2011. Companies of all sizes in "low income" developing countries also benefit from reduced fees.

> *"The concept of a truly autonomous network — enabled by the Level 5 intelligence described by ITU–T Y.3173 — has sparked considerable discussion in ITU."*
>
> Chaesub Lee



Message from the organizers

By Thomas Basikolo, AI/ML Consultant

The ITU AI/ML in 5G Challenge rallied like-minded students and professionals from around the globe to study the practical application of artificial intelligence (AI) and machine learning (ML) in emerging and future networks. The Challenge was a first for ITU, but with many valuable lessons learnt, it looks to be the first of many.

> *"The ITU AI/ML in 5G Challenge rallied like-minded students and professionals from around the globe."*
>
> Thomas Basikolo

The Challenge welcomed over 1300 participants from 62 countries, forming 911 teams, and we are looking forward to the Grand Challenge Finale, 15–17 December online, where outstanding teams will compete for a share in a prize fund totalling 20 000 CHF and a range of other prizes offering global recognition.

Partnerships made the ITU Challenge possible, and partnerships were also the name of the game.

> *The Challenge welcomed over 1300 participants from 62 countries, forming 911 teams.*
>
> Thomas Basikolo

The ITU Challenge enabled participants to connect with new partners in industry and academia — and new tools and data resources — to solve real-world problems with AI/ML, showcase their talent and share new experiences. Twenty-three problem statements were contributed by industry and academia in Brazil, China, India, Ireland, Japan, Russia, Spain, Turkey and the United States, and these "regional hosts" offered resources and expert guidance to support participants in addressing their challenges.

We would like to thank the community that gave life to the Challenge, our participants and regional hosts; our promotion partners LF AI & Data, NGMN and SGInnovate; and our Gold sponsor, the Telecommunications Regulatory Authority (TRA) of the United Arab Emirates; and Bronze sponsors Cisco and ZTE.

Mapping solutions to ITU standards

New ITU standards for AI/ML provide toolsets that, when integrated, form an end-to-end pipeline for AI/ML integration in networks. The ITU Challenge aimed to demonstrate and validate these ITU standards. In mapping solutions to ITU standards, the ITU Challenge contributes to the growth of the community able to support the iterative evolution of these ITU standards.

The ITU–T Y.3172 architecture — derived from the study of use cases published in ITU–T Y.Supplement55 — introduced the basic toolsets in relation to the underlying network: ML Pipeline for model optimization and serving; ML Sandbox to trial models before deployment; and ML Function Orchestrator (MLFO) to control AI/ML integration. ITU–T Y.3173 (intelligence evaluation), ITU–T Y.3174 (data handling) and ITU–T Y.3176 (marketplace integration) all build on the ITU–T Y.3172 architecture.

The problem statements of this first ITU Challenge offered a variety of opportunities to apply the ITU–T Y.317x techniques, and one problem statement demonstrated ML function orchestrator capabilities via reference implementations.

In future editions of the ITU Challenge, we aim to provide a reference implementation of an end-to-end ML pipeline as defined by ITU–T Y.3172. Such reference implementations could include notebooks for ML coding and integration; tools for data processing and management; and tools for ML model selection, training, optimization and verification.

We also aim to enable access to ITU-standard toolsets for initiatives such as plugfests and hackathons and to set the stage for collaboration in open-source projects and standardization work.

A learning experience for all

Data availability is a key challenge to be navigated when bringing together a global community to innovate with AI/ML.

Fifteen problem statements were open to all participants. Eight were limited to participation under conditions set by their hosts. And fourteen remain "under development" without the necessary tools or data resources for this first ITU Challenge. We hope to see new partners coming together to address these fourteen problem statements in future editions of the ITU Challenge.

The data sharing guidelines of the ITU Challenge incorporate a wide range of perspectives from industry and academia on access to real network data, synthetic data and open data. The guidelines describe measures to enable data sharing in view of different classifications of datasets, pre-processing steps (including anonymizing) and secure hosting of data.

We also saw the best outcomes achieved in close collaboration. The Challenge highlighted that problem statements are best positioned for success when supported not only by the necessary tools and data resources, but also by close collaboration between participants and regional hosts.

Our priority was to create community value in the field of AI/ML. In our work to offer participants a level playing field, ITU and our partners developed tailored workflows delivering participants a unique, customized Challenge experience.

ITU engaged participants in technical roundtables and webinars to provide expert guidance in addressing problem statements and the value of new ITU standards in support. Together with our regional hosts, we reached out in local languages, connected participants with mentors and maintained interactive discussions on our Slack channel.

Up to the challenge in 2021?

Preparations for the ITU Challenge 2.0 are in motion, driven by a core team of challenge management board members, judges, promotion partners and sponsors.

We will continue to encourage new partnerships in AI/ML and establish guiding principles for the sharing of tools and data resources necessary to enact these partnerships. We are welcoming new partners and new problem statements, and new tools and data resources. We are creating new opportunities for industry and academia to solve problems together, and new opportunities to influence the direction of ITU standards development and application. Contact us to participate in the problem-solving, judge some of the interesting submissions, promote the challenge, sponsor a prize, or mentor a few students.

We thank you for your support and look forward to seeing you soon in Challenge 2.0. ■

> *"Preparations for the ITU Challenge 2.0 are in motion, driven by a core team of challenge management board members, judges, promotion partners and sponsors."*
>
> Thomas Basikolo

Follow the ITU AI/ML in 5G Challenge

26 partners (telecom operators, vendors, and academia) hosted **23** problem statements

1300+ participants from

60+ countries from **6** regions

45% industry – **55%** academia

26 webinars

4 technical tracks: Networks, Enablers, Verticals, Social Good

20 000 CHF in cash prizes

Certificates – **5** categories

See Challenge website

Don't miss the Grand Challenge Finale winner announcements 15–17 December 2020 online here

Timeline

Problem statements

Title	Host entity
ML5G-PHY-beam-selection: Machine learning applied to the physical layer of millimeter-wave MIMO systems	Federal University of Pará (UFPA), Brazil
Improving the capacity of IEEE 802.11 WLANs through machine learning	Pompeu Fabra University (UPF), Spain
Graph Neural Networking Challenge 2020	Barcelona Neural Networking Center (BNN-UPC), Spain
Compression of deep learning models	ZTE
5G+AI (smart transportation)	Jawaharlal Nehru University (JNU), India
Improving experience and enhancing immersiveness of video conferencing and collaboration	Dview
5G+ML/AI (dynamic spectrum access)	Indian Institute of Technology Delhi (IITD)
Privacy preserving AI/ML in 5G networks for healthcare applications	Centre for Development of Telematics (C-DOT)
Shared experience using 5G+AI (3D augmented + virtual reality)	Hike, India
Demonstration of Machine Learning Function Orchestrator (MLFO) capabilities via reference implementations	Letterkenny Institute of Technology (LYIT), Ireland
ML5G-PHY-channel estimation: Machine learning applied to the physical layer of millimeter-wave MIMO systems	North Carolina State University, United States
Network state estimation by analysing raw video data	NEC, RISING Committee, Telecommunication Technology Committee (TTC)
Analysis on route information failure in IP core networks by NFV-based test environment	KDDI, RISING Committee, Telecommunication Technology Committee (TTC)
Using weather info for radio link failure (RLF) prediction	Turkcell
Traffic recognition and Long-term traffic forecasting based on AI algorithms and metadata for 5G/IMT-2020 and beyond	St. Petersburg State University of Telecommunications (SPbSUT)
5G+AI+AR	China Unicom (Zhejiang Division)
Fault localization of loop network devices based on MEC platform	China Unicom (Guangdong Division)
Configuration knowledge graph construction of loop network devices based on MEC architecture	China Unicom (Guangdong Division)
Alarm and prevention for public health emergency based on telecom data	China Unicom (Beijing Division)
Energy-saving prediction of base station cells in mobile communication network	China Unicom (Shanghai Division)
Core network KPI index anomaly detection	China Unicom (Shanghai Division)
Network topology optimization	China Mobile
Out of service (OoS) alarm prediction of 4/5G network base station	China Mobile

The Grand Challenge Finale – Tuesday, 15 December 2020

Time (CET)	Challenge Title	Team Members	Affiliation
12:15	5G+AI+AR	Jiawang Liu, Jiaping Jiang	CITC and China Unicom
12:30	Analysis on route information failure in IP core networks by NFV-based test environment	Fei Xia, Aerman Tuerxun, Jiaxing Lu, Ping Du	The University of Tokyo
12:45	Analysis on route information failure in IP core networks by NFV-based test environment	Takanori Hara, Kentaro Fujita	Nara Institute of Science and Technology, Japan
13:00	Analysis on route information failure in IP core networks by NFV-based test environment	Ryoma Kondo, Takashi Ubukata, Kentaro Matsuura, Hirofumi Ohzeki	The University of Tokyo
13:15	Fault localization of network devices based on MEC Platform	Zhang Qi, Lin Xueqin	Guochuang Software Co. Ltd
13:30	Network topology optimization	Han Zengfu, Wang Zhiguo, Zhang Yiwei, Wu Desheng, Li Sicong	China Mobile Shandong
13:45	Network topology optimization	Gang Zhouwei, Rao Qianyin, Feng Zezhong, Xi Lin, Guo Lin	China Mobile Guizhou
14:00	Break	Break	Break
14:15	Energy-Saving Prediction of Base Station Cells in Mobile Communication Network	Wei Jiang, Shiyi Zhu, Xu Xu	AsiaInfo Technologies Ltd
14:30	Out of Service (OoS) alarm prediction of 4/5G network base station	Zhou Chao, Zheng Tianyu, Jiang Meijun	Nankai University
14:45	Demonstration of Machine Learning Function Orchestrator (MLFO) capabilities via reference implementations	Abhishek Dandekar	Technical University Berlin
15:00	ML5G-PHY-beam-selection: Machine learning applied to the physical layer of millimeter-wave MIMO systems	Mahdi Boloursaz Mashhadi, Tze-Yang Tung, Mikolaj Jankowski, Szymon Kobus	Imperial College London
15:15	ML5G-PHY-beam-selection: Machine learning applied to the physical layer of millimeter-wave MIMO systems	Batool Salehihikouei, Debashri Roy, Guillem Reus Muns, Zifeng Wang, Tong Jian	Northeastern University, Brazil
15:30	ML5G-PHY-beam-selection: Machine learning applied to the physical layer of millimeter-wave MIMO systems	Zecchin Matteo	Eurecom, Brazil
15:45	Improving the capacity of IEEE 802.11 WLANs through machine learning	Ramon Vallès	Pompeu Fabra University, Spain
16:00	Improving the capacity of IEEE 802.11 WLANs through machine learning	Paola Soto, David Goez, Miguel Camelo, Natalia Gaviria	University of Antwerp, Belgium
16:15	Improving the capacity of IEEE 802.11 WLANs through machine learning	Mohammad Abid, Ayman M. Aloshan, Faisal Alomar, Mohammad Alfaifi, Abdulrahman Algunayyah, Khaled M. Sahari	Saudi Telecom

Note: The above teams have been selected to make presentations at the Grand Challenge Finale (Finale Conference). (Each team has 8 minutes for its presentation, followed by a 7-minute Q&A with the judges and the audience).

Take a look at the list of best teams.

Don't miss the Final Conference! Register here.

The Grand Challenge Finale – Wednesday, 16 December 2020

Time (CET)	Challenge Title	Team Members	Affiliation
12:00	Network state estimation by analysing raw video data	Yuusuke Hashimoto, Yuya Seki, Daishi Kondo	Osaka Prefecture University, Japan
12:15	Network state estimation by analysing raw video data	Yimeng Sun, Badr Mochizuki	The Kyoto College of Graduate Studies for Informatics, Japan
12:30	Network state estimation by analysing raw video data	Fuyuki Higa, Gen Utidomari, Ryuma Kinjyo, Nao Uehara	National Institute of Technology, Okinawa College, Japan
12:45	Compression of deep learning models	Yuwei Wang, Sheng Sun	Institute of Computing Technology Chinese Academy of Sciences
13:00	Compression of deep learning models	Satheesh Kumar Perepu, Saravanan Mohan, Vidya G, Thrivikram G L, Sethuraman T V	Ericsson Research India
13:15	5G+AI (smart transportation)	Atheer K. Alsaif, Nora M. Almuhanna, Abdulrahman Alromaih, Abdullah O. Alwashmi	Saudi Telecom Company
13:30	Privacy preserving AI/ML in 5G networks for healthcare applications	Mohammad Malekzadeh, Mehmet Emre Ozfatura, Kunal Katarya, Mital Nitish	Imperial College London
13:45	Shared experience using 5G+AI (3D augmented + virtual reality)	Nitish Kumar Singh	Easyrewardz Software Services
14:00	Break	Break	Break
14:15	Graph Neural Networking Challenge 2020	Loïck Bonniot, Christoph Neumann, François Schnitzler, François Taiani	InterDigital; Inria/Irisa
14:30	Graph Neural Networking Challenge 2020	Nick Vincent Hainke, Stefan Venz, Johannes Wegener, Henrike Wissing	Fraunhofer HHI, Germany
14:45	Graph Neural Networking Challenge 2020	Martin Happ, Christian Maier, Jia Lei Du, Matthias Herlich	Salzburg Research Forschungsgesellschaft
15:00	Using weather info for radio link failure (RLF) prediction	Dheeraj Kotagiri, Anan Sawabe, Takanora Iwai	NEC Corporation
15:15	Using weather info for radio link failure (RLF) prediction	Juan Samuel Pérez, Amín Deschamps, Willmer Quiñones, Yobany Díaz	Santo Domingo Institute of Technology (INTEC)
15:30	Traffic recognition and long-term traffic forecasting based on AI algorithms and metadata for 5G/ IMT-2020 and beyond	Ainaz Hamidulin, Viktor Adadurov, Denis Garaev, Artem Andriesvky	Ufa State Aviation Technical University (USATU) University, Russia
15:45	ML5G-PHY-Channel Estimation: Machine Learning Applied to the Physical Layer of Millimeter-Wave MIMO Systems	Dolores Garcia, Joan Palacios, Joerg Widmer	IMDEA Networks
16:00	ML5G-PHY-Channel Estimation: Machine learning applied to the physical layer of millimeter-wave MIMO systems	Emil Björnson, Pontus Giselsson, Mustafa Cenk Yetis, Özlem Tugfe Demir	Linköping University and Lund University, Sweden
16:15	ML5G-PHY-Channel Estimation: Machine learning applied to the physical layer of millimeter-wave MIMO systems	Chandra Murthy, Christo Kurisummoottil Thomas, Marios Kountouris, Rakesh Mundlamuri, Sai Subramanyam Thoota, Sameera Bharadwaja H	Eurecom, France, Indian Institute of Science, India Communications, Canada

Note: The above teams have been selected to make presentations at the Grand Challenge Finale (Finale Conference). (Each team has 8 minutes for its presentation, followed by a 7-minute Q&A with the judges and the audience).

Take a look at the list of best teams.

Don't miss the Final Conference! Register here.

The Grand Challenge Finale – Thursday, 17 December 2020

Time (CET)	Programme
11:30–12:00	Join session to test connection
12:00–12:30	**Opening ceremony**
	Welcome remarks Houlin Zhao, ITU Secretary-General Chaesub Lee, Director, ITU Telecommunication Standardization Bureau United Arab Emirates Telecommunications Regulatory Authority
	Overview of the 2020 Challenge Thomas Basikolo, ITU
12:30–12:55	**Keynote – Recent advances in federated learning for communications** Wojciech Samek, Head of Machine Learning Group, Fraunhofer HHI
12:55–13:40	**Special Session: Vision for the future – AI/ML in 5G roadmap**
	Regulator perspective Telecommunications Regulatory Authority, United Arab Emirates
	Industry perspective Cisco
	Industry perspective Wei Meng, Director of Standard and Open Source Planning, ZTE Corporation
13:40–14:05	**Keynote – The unfinished journey of network AI** Chih-Lin I, Chief Scientist, Wireless Technologies, China Mobile Research Institute
14:05–14:30	**Keynote – Learning at the wireless edge** H. Vincent Poor, Professor of Electrical Engineering, Princeton University, United States
14:30–15:15	**Winners' presentations**
15:15–15:30	**Award announcements** Prizes and certificates
15:30–15:35	Call for papers to special issue of ITU Journal on Future and Evolving Technologies (ITU J-FET): "AI/ML Solutions in 5G and Future Networks" Ian Akyildiz, Editor-in-Chief, Georgia Institute of Technology, United States
15:35–15:45	**2021 outlook for Challenge 2.0** Vishnu Ram, Independent Researcher
15:45–16:00	**Closing ceremony**
	Closing remarks Hosts of the ITU AI/ML in 5G Challenge 2020 Chaesub Lee, Director, ITU Telecommunication Standardization Bureau

Don't miss the Final Conference!
Register here.

Winning prizes and certificates

Teams from various problem statements will compete for the ITU AI/ML in 5G Challenge Champion title, and several awards will be presented to winning solutions at the Grand Challenge Finale taking place 15–17 December 2020.

Winners' certificate: Awarded to winning teams in the following categories:

1st prize:
ITU AI/ML in 5G Challenge Gold Champion
Cash prize: 5000 CHF

2nd prize:
ITU AI/ML in 5G Challenge Silver Champion
Cash prize: 3000 CHF

3rd prize:
ITU AI/ML in 5G Challenge Bronze Champion
Cash prize: 2000 CHF

Three *Runners up* will receive 1000 CHF each.

Judges Prize certificates: Awarded to winners of each problem statement as recommended by the host (excluding those under *Winners certificate*). Each winner receives 300 CHF.

Honorable mention certificate.

Encouragement/Community award certificate: Awarded to teams that were active during the mentoring programme and successfully submitted a solution.

Certificate of completion: Awarded to teams that completed the challenge by submitting a solution.

Shutterstock

A guide to AI/ML challenges for the next-generation CTx

By Vishnu Ram OV, Independent Research Consultant

The new CTx* at FutureXG analysed the reports on the screen.

(x+1)G specification delayed. xG deployments yet to be justified. Research & development lost in a maze of acronyms, old and new. New architecture diagrams every few weeks. New use cases to support in every market. Applying and integrating AI/machine learning (ML) in the network was nothing smooth. The open source repository that CTx was banking on was being pulled in a zillion directions. And the buzz around autonomous networks meant that every part of the network was working on its own brand of autonomy.

Will CTx survive this challenge?

New ITU standards describe concepts to enable AI/ML integration in 5G and future networks as these networks evolve.

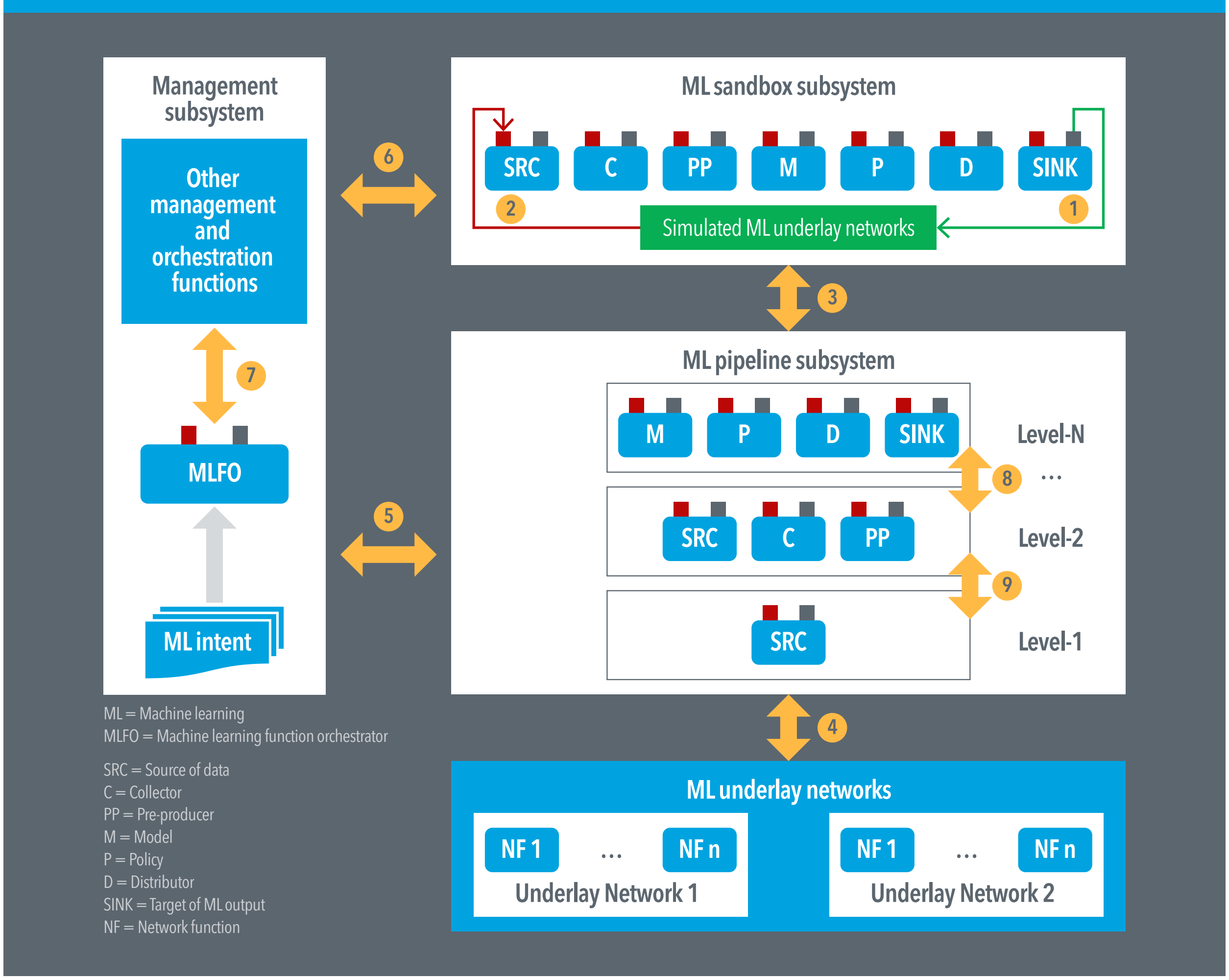
"And the buzz around autonomous networks meant that every part of the network was working on its own brand of autonomy."

**Any resemblances or similarities with real-life CTOs are purely futuristic.*
Disclaimer: This article contains some fictional information which may be referred to as forward-looking statements.

The ITU–T Y.3172 architecture, derived from the study of use cases published in ITU–T Y.Supplement55, introduced basic toolsets including the ML Pipeline, ML Sandbox and ML Function Orchestrator (MLFO) in relation with the underlying network. ITU–T Y.3173 (intelligence evaluation), ITU–T Y.3174 (data handling) and ITU–T Y.3176 (marketplace integration) all build on the ITU–T Y.3172 architecture.

Together these ITU standards provide powerful toolsets — standard toolsets — for operators to monitor and adapt to changes in the underlying network. Using the concepts described by the ITU–T Y.317x standards, even as the underlying network architecture changes from generation to the next, it will remain possible to specify AI/ML integration using the common terminology provided by ITU.

High level architecture for integration of AI/ML in networks (ITU–T Y.3172)



> *New alerts pop up on the screen from the MLFO monitor. What? A network-update alert!*

The details of a new use case arrive in the message-box. CTx runs it through the Intent-parser tool. Interesting, but how to implement it? CTx finds an ITU webinar on MLFO orchestration for managed AI/ML integration. A few API calls later, CTx is ready with a tentative ML pipeline.

CTx kicks off simulations in the ML Sandbox while waiting for approval to access real network data. Digital "twins" rev into action; data is generated based on previous patterns and models are trained in the ML Sandbox, all while the Approval Authority takes its time. CTx sends the results from the ML Sandbox trial models. This has the desired effect. The approval arrives in the message-box.

Real network data adds to the accuracy of the models. "[ML-usecase-1xx::status::ready]" CTx inputs to the message-box.

The MLFO described by ITU–T Y.3172 is a logical node that manages and orchestrates the nodes in an ML pipeline. ITU–T Y.3173 (intelligence evaluation) describes a key architecture scenario for the evaluation of network intelligence levels by the MLFO. ITU–T Y.3174 (data handling) describes the sequence diagrams corresponding to the instantiation of various components of the ITU–T Y.317x toolsets, based on the incoming ML Intent from the operator.

In combination with MLFO, ML Sandbox provides a managed environment for operators to train, test and validate ML models before they are deployed in the live network. The data handling mechanism defined in ITU–T Y.3174 further allows the addition of new sources of data and other scenarios.

New alerts pop up on the screen from the MLFO monitor. What? A network-update alert! As usual, an unscheduled virtualized network function upgrade by the vendor. Do we need to rework the whole ML pipeline?

The ITU–T Y.317x concept of ML Pipeline and ML Sandbox managed by MLFO enables operators to decouple the underlying network from the AI/ML integration.

At reference point 7, the ITU–T Y.3172 architecture allows the tracking of changes in the underlying network and the application of optimizations and configurations in the ML pipeline by the MLFO. The architecture scenario described by ITU–T Y.3173 (intelligence evaluation) also includes monitoring the intelligence level of each node of an ML pipeline by the MLFO.

The draft ITU standard Y.ML-IMT2020-MODEL-SERV aims to provide an architectural framework supporting the efficient optimization of ML models for heterogeneous hardware environments, flexible deployment of ML models for different use-case scenarios, and effective interfaces in the ML pipeline when a serving model is deployed.

CTx parses a new message in the message-box "[ML-usecase-1xx::Evaluate::partner.edu::-model.url]". Pioneering algorithm work at a partner university had produced a wrapped model suitable for the use case. But the Approval Authority needs an evaluation of the model. Hopefully the external ML marketplace complies to ITU–T Y.3176! CTx pulls the model from the marketplace.

ML marketplace integration can help network operators to follow the ML innovation curve.

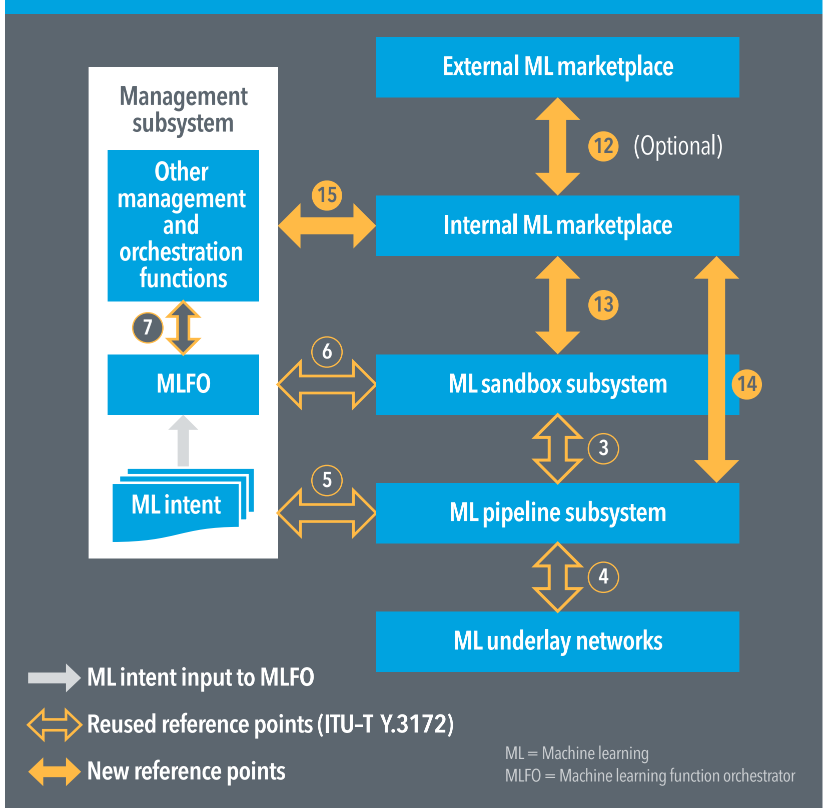
The ML model metadata, ML marketplace requirements and the architecture reference points defined in ITU–T Y.3176 (marketplace integration) enable the efficient exchange and deployment of ML models using standard interfaces. Not only can this method help solve networking problems using ML techniques, but also has the potential to share and monetize ML techniques.

ITU–T Y.3176 supports the administration of different types of ML marketplaces, internal or external, and ML marketplace federation. The APIs defined in ITU–T Y.3176 enable marketplaces to find and select ML models in other marketplaces and pull from federated marketplaces. And they enable marketplaces to exchange updated ML models and interact with ML Sandboxes.

Done! A new ML pipeline in place in the ML Sandbox, tested and verified for the new use case. "[status::ready]" CTx inputs to the message-box. "[status::approved]" the Approval Authority responds. CTx schedules an update of the network.

Meanwhile, unknown to CTx, a CTx-software-update package had arrived in the message-box. It was time for evolution and a new CTx-agent to take over.

Architecture for ML marketplace integration in network (ITU–T Y.3176)



On top of the ability to adapt and improve network management and control, an autonomous network could self-evolve through online experimentation, enabling better compositions of controllers and controller hierarchies.

Vishnu Ram OV

About ITU AI/ML in 5G Challenge

ITU AI/ML in 5G Challenge provided a platform for participants to apply the ITU–T Y.317x techniques in solving practical problem statements. A varied selection of topics including beam selection, WLAN capacity analysis, network state analysis, network slicing and traffic forecasting, radio link failure prediction, the optimization of deep learning models, and MLFO reference implementations, were offered in the Challenge. Different types of data, including data from real networks, were provided in some cases for developing solutions to these problems.

The concept of a truly autonomous network – enabled by the Level 5 intelligence described by ITU–T Y.3173 – sparked considerable discussion in the ITU–T Focus Group on "machine learning for future networks including 5G" and this discussion continues in the ITU's standardization expert group for "future networks and cloud", ITU–T Study Group 13.

Autonomous networks would display the "self" properties: the ability to monitor, operate, recover, heal, protect, optimize and reconfigure themselves. On top of the ability to adapt and improve network management and control, an autonomous network could self-evolve through online experimentation, enabling better compositions of controllers and controller hierarchies.

CTx.v2 scanned the environment.

ML Pipelines, Sandboxes and ML marketplaces are in place and MLFO reports are green, but issues remain. Divergent data formats impacting latency between the ML pipelines in the network. A multitude of open-source toolkits to integrate. More challenges in the demand mapping towards Level 5 intelligence. The disaggregation of network components, rapid DevOps and better and better AI/ML models meant more work in the AI/ML integration.

CTx.v2 searches the context for solutions.

Perhaps time for another ITU AI/ML in 5G Challenge? CTx.v2 logs into the Geneva Sandbox of ITU and triggers "[AI-ML-Challenge::v2::init]", but that's another day, another story (for CTx.v3).

A standards round-up on autonomous networks

By Xiaojia Song, Researcher, **Xi Cao,** Senior Researcher, **Lingli Deng,** Technical Manager, **Li Yu,** Chief Researcher, and **Junlan Feng,** Chief Scientist, China Mobile Research Institute

Mobile networks are evolving into the intelligence era with multiple application scenarios, features, services and operation requirements. Technologies including artificial intelligence (AI) are expected to enable autonomous networks in areas such as network planning, deployment, operation, optimization, service deployment, and quality assurance.

> *Mobile networks are evolving into the intelligence era with multiple application scenarios, features, services and operation requirements.*

Most of the standards development organizations (SDOs), e.g. the ITU Telecommunication Standardization Sector (ITU–T), 3GPP, ETSI, and CCSA, are actively developing standards for autonomous networks.

Industry bodies such as GSMA, TM Forum, and the Global TD-LTE Initiative (GTI) are working to promote autonomous networks. GSMA stated that the automatic network operation capability will become the indispensable 4th dimension of the 5G era together with enhanced mobile broadband (eMBB), massive machine type communications (mMTC) and ultra-reliable low-latency communications (URLLC), and become one of the most important driving forces for 5G service innovation and development.

There are discussions among SDOs about the level of autonomous capabilities in networks (see framework approach in Table 1).

The study of autonomous network levels (ANL) can provide reference and guidance to operators, vendors and other participants of the telecommunication industry for autonomous networks, standardization works and roadmap planning.

Since industrial convergence is the key for reducing the cost for any single vendor or single network operator, building an open collaboration platform (see Figure 1) for cohesively developing both a reference implementation for case-agnostic functional architecture and standardized external or internal interfaces would be the easy way for communication service providers (CSPs) to kick off and stay in the converged direction towards network autonomy.

For example, a rule-based policy engine could be one of the common functional modules to support both timed control tasks in Level 1, imperative closed loops in Level 2, and adding intent-to-rule translation modules in Levels 3 and 4.

The main activities on autonomous networks in the SDOs and industry bodies are presented in Figure 2 and briefly introduced below.

Table 1 – Framework approach for classification of autonomous network intelligence level (source: ITU-T Y.3173)

Network intelligence level		**Dimensions**				
		Action implementation	**Data collection**	**Analysis**	**Decision**	**Demand mapping**
L0	Manual network operation	Human	Human	Human	Human	Human
L1	Assisted network operation	Human and System	Human and System	Human	Human	Human
L2	Preliminary intelligence	System	Human and System	Human and System	Human	Human
L3	Intermediate intelligence	System	System	Human and System	Human and System	Human
L4	Advanced intelligence	System	System	System	System	Human and System
L5	Full intelligence	System	System	System	System	System

NOTE 1 – For each network intelligence level, the decision process has to support intervention by human being, i.e., decisions and execution instructions provided by a human being have the highest authority.
NOTE 2 – This table may be used to only determine the network intelligence level for each dimension (and not the overall network intelligence level).

Figure 1 – Open industrial collaboration towards autonomous networks

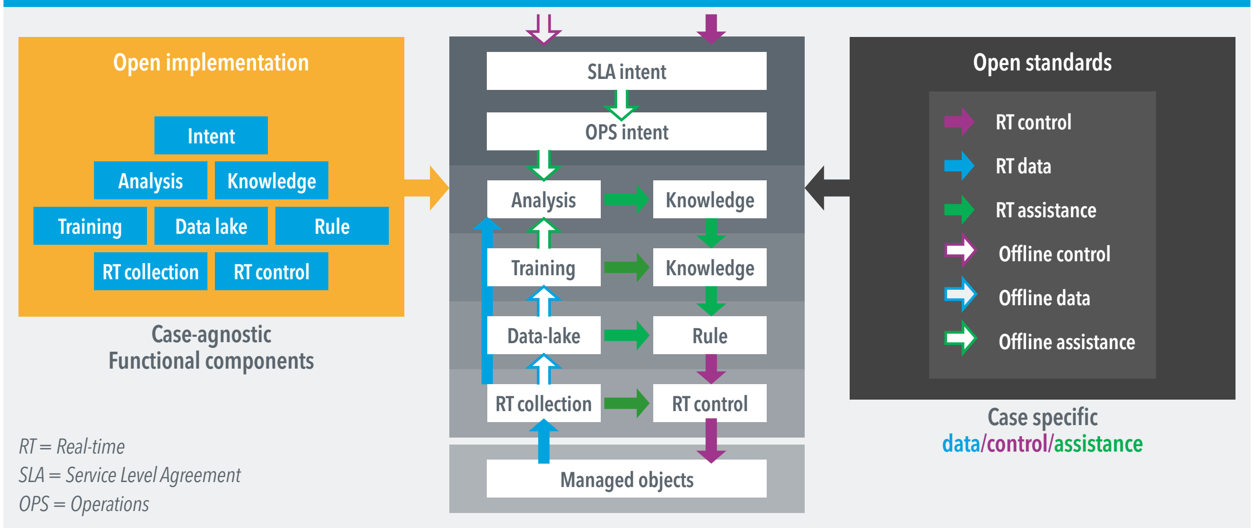
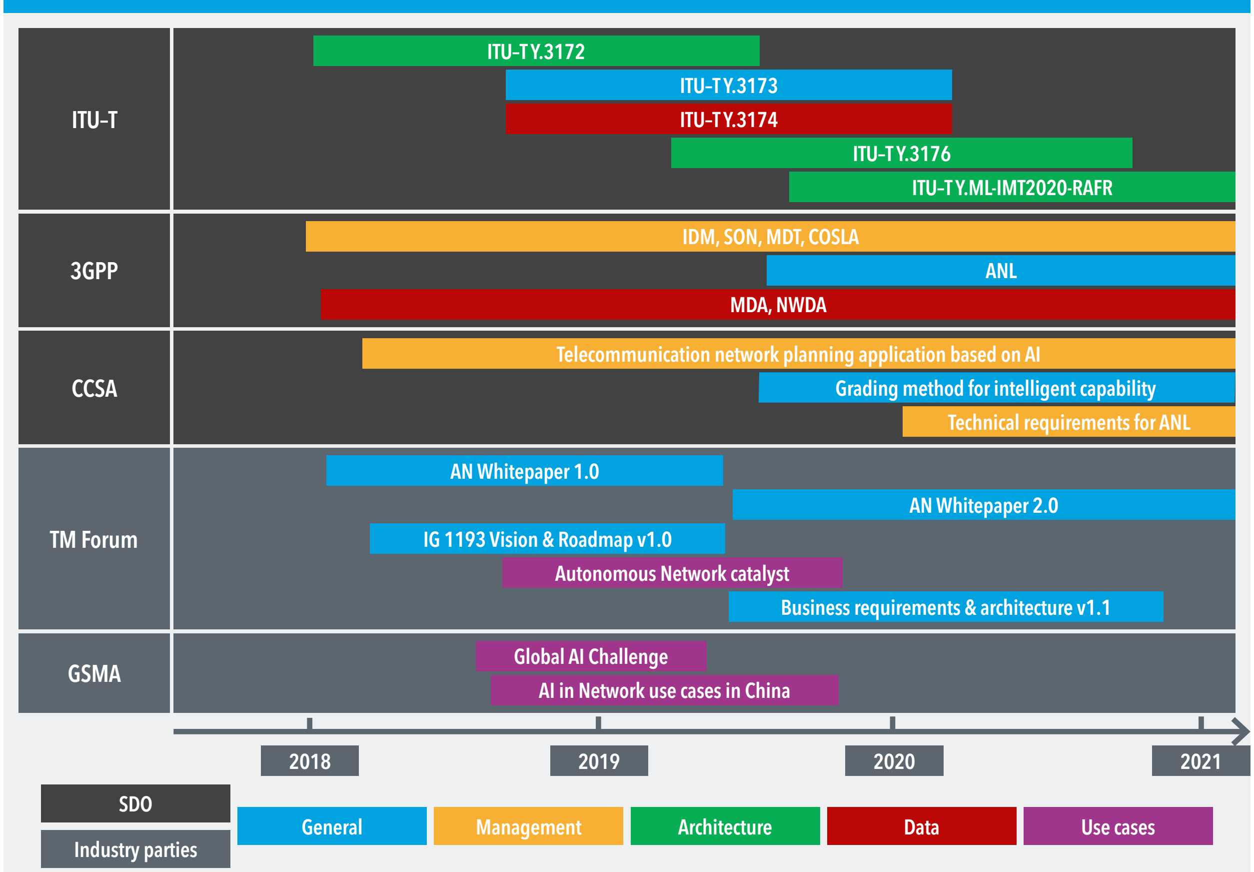


Figure 2 – Main activities of SDOs and industry bodies



Activities in ITU

In ITU–T, Study Group 13 focuses on future networks and network aspects of mobile telecommunications. The Focus Group on Machine Learning for Future Networks including 5G (FG-ML5G), active from January 2018 until July 2020, had been set up to study interfaces, network architectures, protocols, algorithms and data formats. Of FG ML5G's ten technical specifications, four have already been turned into ITU Recommendations (standards), one into a Supplement, and the other five are in the process of being turned into ITU standards. Recommendations on the AI-based autonomous network, e.g. ITU–T Y.ML-IMT2020-RAFR, are currently in draft stage (see Table 2).

Table 2 – ITU-T standardization activities on AI/ML and AI-based autonomous networks

Reference number	Title
Supplement 55 to Y.3170 Series	Machine learning in future networks including IMT-2020: use cases
ITU-T Y.3172	Architectural framework for machine learning in future networks including IMT-2020
ITU-T Y.3173	Framework for evaluating intelligence level of future networks including IMT-2020
ITU-T Y.3174	Framework for data handling to enable machine learning in future networks including IMT-2020
ITU-T Y.3176	Machine learning marketplace integration in future networks including IMT-2020
FG-ML5G spec	Requirements, architecture and design for machine learning function orchestrator
FG-ML5G spec	Machine Learning Sandbox for future networks including IMT-2020 requirements and architecture framework
FG-ML5G spec	Machine learning based end-to-end network slice management and orchestration
FG-ML5G spec	Vertical-assisted Network Slicing Based on a Cognitive Framework
Draft ITU-T Y.ML-IMT2020-RAFR	Architecture framework for AI-based network automation of resource adaptation and failure recovery for future networks including IMT-2020

Activities in 3GPP

Autonomous networks came in sight of 3GPP in the 4G era. The topics mainly focused on Self-Organizing Networks (SON) and Minimization of Drive Tests (MDT). In the 5G era, 3GPP undertakes standardization efforts to enable autonomous networks:

- 3GPP RAN: RAN data collection (TR 37.816), SON/MDT (TS 38.314, TS 38.300, TS 37.320, TS 38.306, TS 38.331, etc.) (see Table 3).

Table 3 – Standardization activities in 3GPP RAN

TS/TR	Title
3GPP TR 37.816	Study on RAN-centric data collection and utilization for LTE and NR
3GPP TS 38.314	New Radio (NR); Layer 2 measurements
3GPP TS 38.300	NR; Overall description; Stage 2
3GPP TS 37.320	Minimization of Drive Tests (MDT); Overall description; Stage 2
3GPP TS 38.306	NR; User Equipment (UE) radio access capabilities
3GPP TS 38.331	NR; Radio Resource Control (RRC); Protocol specification

TS=Technical Specification, TR=Technical Report.

- 3GPP SA2: Network Data Analytics (NWDA) (TR 23.791, TR 23.288, TR 23.700-91) (see Table 4).

- 3GPP SA5: Management Data Analytics (MDA) (TR 28.809), autonomous networks levels (TR 28.810, TS 28.100, intent driven management (TR 28.812, TS 28.312), closed loop SLS assurance (TR 28.805, TR 28.535, TR 28.536, etc.), SON (TR 28.861, TS 28.313) and MDT (TS 28.313, TS 32.42X series) (see Table 5).

Activities in ETSI

ETSI is actively studying autonomous networks and has several groups working on the following relevant topics:

- ENI (Experiential Networked Intelligence).

- NFV (Network Functions Virtualization).

- OSM (Open Source MANO (Management and Orchestration).

- MEC (Multi-access Edge Computing).

- F5G (Fifth Generation Fixed Network).

Table 4 – Standardization activities in 3GPP SA2

TS/TR	Title
3GPP TR 23.791	Study of Enablers for Network Automation for 5G
3GPP TS 23.288	Architecture enhancements for 5G System to support network data analytics services
3GPP TR 23.700-91	Study on Enablers for Network Automation for 5G – Phase 2

TS=Technical Specification, TR=Technical Report.

Table 5 – Standardization activities in 3GPP SA5

TS/TR	Title
3GPP TR 28.809	Study on enhancement of Management Data Analytics (MDA)
3GPP TR 28.810	Study on concept, requirements and solutions for levels of autonomous network
3GPP TS 28.100	Management and orchestration; Levels of autonomous network
3GPP TR 28.812	Telecommunication management; Study on scenarios for Intent driven management services for mobile networks
3GPP TS 28.312	Intent driven management services for mobile networks
3GPP TR 28.805	Telecommunication management; Study on management aspects of communication services
3GPP TS 28.535	Management and orchestration; Management services for communication service assurance; Requirements
3GPP TS 28.536	Management and orchestration; Management services for communication service assurance; Stage 2 and Stage 3
3GPP TR 28.861	Study on the Self-Organizing Networks (SON) for 5G networks
3GPP TS 28.313	Self-Organizing Networks (SON) for 5G networks
3GPP TS 32.42X series	
3GPP TS 32.421	Telecommunication management; Subscriber and equipment trace; Trace concepts and requirements
3GPP TS 32.422	Telecommunication management; Subscriber and equipment trace; Trace control and configuration management
3GPP TS 32.423	Telecommunication management; Subscriber and equipment trace; Trace data definition and management
3GPP TS 32.425	Telecommunication management; Performance Management (PM); Performance measurements Evolved Universal Terrestrial Radio Access Network (E-UTRAN)
3GPP TS 32.426	Telecommunication management; Performance Management (PM); Performance measurements Evolved Packet Core (EPC) network

TS=Technical Specification, TR=Technical Report.

> *"SDOs should continue their relevant standardization work and play a leading role in enabling autonomous networks."*

- TC INT AFI (Technical Committee Core (TC) Network and Interoperability Testing (INT) working group Autonomic Management and Control Intelligence for Self-Managed Fixed and Mobile Integrated Networks).
- ZSM (Zero-touch network & Service Management).

TC INT AFI is studying Generic Autonomic Networking Architecture (GANA), and ZSM is discussing closed-loop automation in the ZSM framework, optimized for data-driven machine learning and AI algorithms.

In November 2019, ETSI published the report "Experiential Networked Intelligence (ENI): ENI Definition of Categories for AI Application to Networks" (ETSI GR ENI 007), which defines various categories for the level of application of AI techniques to the management of the network, going from basic limited aspects to the full use of AI techniques for performing network management.

ENI is developing a general-purpose architecture for enhanced network intelligence, and a mapping on NFV policy management is under discussion as NFV started policy modelling work for automating NFV management and VNF CI/CD.

Activities in CCSA

As one of the most influential SDOs in the field of communication in China, China Communications Standards Association (CCSA) began standards work on autonomous networks from 2010, and the items are mainly set up in Technical Committee TC 1, TC5 and TC7, including use cases, architecture, data handling, levels of autonomous network, management requirements, etc.

Activities in industry

Industry bodies such as GSMA, TM Forum and GTI are exploring and promoting the collaboration of autonomous network topics among SDOs, operators, vendors and other industry participants.

In GSMA, AI and Automation is one of the topics of the "Future Network". In June 2019, the first GSMA Global AI Challenge was held, investigating three specific areas: connectivity in rural areas, mobile energy efficiency and enhanced services in urban areas.

At its workshop in June, at the AI in Network Seminar of the Mobile World Congress Shanghai 2019, GSMA called on the entire industry to focus on and contribute to the key applications of AI in the mobile networks, and jointly build the 5G era for intelligent autonomous networks. In October 2019, GSMA published AI in Network Use Cases in China".

TM Forum has held several workshops on autonomous networks since 2019, and the Autonomous Networks Project (ANP) was established in August 2019.

Three whitepapers have been published: Autonomous Networks Whitepaper 1.0, IG1193 Vision and Roadmap v1.0, and IG1218 Business requirement and architecture v1.0. In 2020, Autonomous Networks Whitepaper 2.0, business requirements and architecture v1.1, technical architecture, demo of Catalyst projects, user stories/use cases, etc. are underway.

The Global TD-LTE Initiative has established a project dedicated to intelligent networks. Under the 5G eMBB program, it will specify use cases and requirements for intelligent networks in the fields of levels, architecture, network elements and network management.

Activities in open source communities

The End User Advisory Group (EUAG) of Linux Foundation Networking (LFN) is drafting a survey to solicit CSPs' deployment status, requirements and strategies of network autonomy, with the hope to abstract a converged vision and serve as input to corresponding technical groups.

Since its earliest releases, the Open Network Automation Platform (ONAP) has been building and enhancing use cases for radio, core and transport automation on top of its policy-driven close-loop framework. Intent-based edge-to-edge (E2E) slicing automation is on its way. One ETSI official proof-of-concept shows that ONAP would provide the basis for constructing a reference stack for autonomous networks.

The importance of standards for autonomous networks

As technologies and networks evolve, autonomous networks will be an important enabler in future. In order to implement network autonomy, a staged path is envisioned, where the key differentiator to ensure convergence in later stages is to build a common architecture across layers, standardize cross-layer interfaces and drive SDO alignment in earlier stages.

SDOs should continue their relevant standardization work and play a leading role in enabling autonomous networks.

ITU AI/Machine Learning in 5G Challenge webinars

19/06/2020

Graph Neural Networking Challenge 2020

José Suárez-Varela
Researcher, Barcelona Neural Networking Center, Universitat Politècnica de Catalunya (BNN-UPC), Spain

See webinar

26/06/2020

Beam Selection – Machine Learning Applied to the Physical Layer of Millimeter – Wave MIMO Systems

Aldebaro Klautau
Professor, Federal University of Pará (UFPA), Brazil

See webinar

03/07/2020

Channel Estimation – Machine Learning Applied to the Physical Layer of Millimeter – Wave MIMO Systems

Nuria González Prelcic
Associate Professor, North Carolina State University, United States

See webinar

10/07/2020

ITU AI/ML in 5G Challenge: Improving the Capacity of IEEE 802.11 WLANs through Machine Learning

Francesc Wilhelmi
Researcher, Universitat Pompeu Fabra, Spain

See webinar

17/07/2020

ITU AI/ML in 5G Challenge: DNN Inference Optimization Challenge

Liya Yua
Open Source & Standardization Engineer, ZTE

See webinar

22/07/2020

Radio Link Failure Prediction Challenge

Salih Ergüt
5G R&D Senior Expert, Turkcell

See webinar

24/07/2020

5G + AI + Immersive + Assistive Services in Telecommunications

Brejesh Lall
Professor, Indian Institute of Technology, Delhi

See webinar

27/07/2020

AI Techniques for Privacy-Preserving Remote Medical Diagnosis + Spectrum and Network Resource Sharing in 5G Networks

Brejesh Lall
Professor, Indian Institute of Technology, Delhi

See webinar

29/07/2020

Machine Learning for Wireless LANs + Japan Challenge Introduction

Akihiro Nakao
Professor, University of Tokyo

Koji Yamamoto
Associate Professor, Kyoto University

Tomohiro Otani
Executive Director, KDDI Research, Inc.

Takanori Iwai
Research Manager, NEC Corporation

See webinar

31/07/2020

LYIT/ITU-T AI Challenge: Demonstration of Machine Learning Function Orchestrator (MLFO) Via Reference Implementations

Shagufta Henna
Lecturer, Letterkenny Institute of Technology (LYIT), Ireland

See webinar

07/08/2020

ITU AI/ML in 5G Challenge: Lecture on Machine Learning and Participation in Japan Challenge

Akihiro Nakao
Professor, University of Tokyo

Koji Yamamoto
Associate Professor, Kyoto University

Tomohiro Otani
Executive Director, KDDI Research, Inc.

Takanori Iwai
Research Manager, NEC Corporation

See webinar

07/08/2020

An Overview of the ITU-ML5G-PS-012 "ML5G-PHY [Beam Selection]"

Aldebaro Klautau
Federal University of Pará (UFPA), Brazil

See webinar

17/08/2020

Applying Knowledge Graph and Digital Twin Technologies to Smart Optical Network

Anran Xu
Researcher, China Information and Communication Technologies Group Corporation (CICT)

See webinar

19/08/2020

ITU/AI/ML in 5 Challenge Open House and Roundtable No. 2

Prerana Mukherjee
Assistant Professor, School of Engineering, Jawaharial Nehru University, Delhi, India

See webinar

21/08/2020

A Universal Compression Algorithm for Deep Neural Networks

Wojciech Samek
Head of the Machine Learning Group, Fraunhofer Heinrich Hertz Institute, Germany

See webinar

26/08/2020

ITU AI/ML in 5G Challenge: China Mobile Network Topology Optimization Competition Question Analysis

Wang Xing
Researcher, China Mobile Research Institute

See webinar

31/08/2020

Traffic Recognition and Long-Term Traffic Forecasting Based on AI Algorithms and Metadata for 5G/IMT-2020 and Beyond

Artem Volkov
Researcher

Ammar Muthanna
Associate Professor, St. Petersburg State University of Telecommunications, Russia

See webinar

01/09/2020

Milvus: An Open Source Vector Similarity Search Engine

Jun Gu
Partner, Zilliz

See webinar

04/09/2020

How to Bring AI into 5G Radio Access Network

Qi Sun
Senior Researcher, China Mobile Research Institute

See webinar

28/09/2020

Wireless 2.0: Towards a Smart Radio Environment Empowered by Reconfigurable Intelligent Metasurfaces and Artificial Intelligence

Marco Di Renzo
Professor, CNRS & Paris-Saclay University, France

See webinar

16/11/20

Harnessing Deep Learning for Mobile Service Traffic Decomposition to Support Network Slicing

Alexis Duque
Research Associate, Net AI

See webinar

18/11/2020

The Road Towards an AI-Native Air Interface for 6G

Jakob Hoydis
Head, Research Department on Radio Systems and AI, Nokia Bell Labs

See webinar

27/11/2020

Leveraging AI & Machine Learning to Optimize Today's 5G Radio Access Network Systems and to Build the Foundation of Tomorrow's 6G Wireless Systems

Tim O'Shea
Co-Founder/CTO, DeepSig

See webinar

02/12/2020

Toward Effective Network Traffic Analytics of Mobile Apps via Deep Learning

Domenico Ciuonzo
Assistant Professor, DIETI, University of Naples, Federico II, Italy

See webinar

04/12/2020

Scaling CNN Inference for Extreme Throughput

Michaela Blott
Distinguished Engineer, Xilinx

See webinar

08/12/2020

Towards Open, Programmable, and Virtualized 5G Networks

Michele Polese
Associate Research Scientist, Northeastern University, United States

See webinar

Capability evaluation and AI accumulation in future networks

By Jun Liao, Artificial Intelligence Director, **Tengfei Liu, Yameng Li,** and **Jiaxin Wei,** Artificial Intelligence Engineers, China Unicom Research Institute

The rapid development of 5G networks has brought many new challenges – networking becomes more complex, the services are more diverse, and the connections are increasing on a large scale.

Using traditional ways of network operation and maintenance will make meeting new requirements for network development difficult. As an important problem-solving method, network intelligence has become the focus of the information and communication technology industry and a predominant trend in future network development.

China Unicom believes that future network intellectualization will ensure safe and reliable network services and provide customers with a fast and ultimate experience. For network operation and maintenance (O&M), it has the capabilities of self-configuration, self-monitoring, self-healing, and self-optimization.

> *Network intelligence has become the focus of the information and communication technology industry and a predominant trend in future network development.*
>
> Jun Liao, Tengfei Liu, Yameng Li, and Jiaxin Wei

> *The test bed will evaluate the intelligence level of networks in a quantitative manner and will help realize the rapid implementation of network intelligences.*
>
> Jun Liao, Tengfei Liu, Yameng Li, and Jiaxin Wei

By using advanced automation and intelligent technology comprehensively, we can reconstruct the existing network architecture, O&M mode, enable services innovation, and create extreme user experience.

Test bed for evaluating network intelligence capabilities

Evaluating and sharing network intelligence capabilities are key to ensuring the effectiveness of network intelligence. To achieve rapid innovation in the future, we need to build a platform with capability accumulation and open data. When demands change, we can flexibly integrate the existing services and resources to respond quickly – improving the overall efficiency and operation ability.

China Unicom is in the process of building a test bed for evaluating network intelligence capabilities. Based on the ITU Telecommunication Standardization Sector's (ITU-T's) Y.3173 Recommendation: *Framework for evaluating intelligence levels of future networks including IMT-2020*, it will provide professional evaluation methods and services which will include computing power, a machine learning (ML) model, and data and network intelligence capability.

The test bed will evaluate the intelligence level of networks in a quantitative manner and will help realize the rapid implementation of network intelligence. The test content includes the following items:

- **Computing power evaluation:** For different types of accelerated AI chips, e.g. training chips or serving chips from different companies are evaluated in terms of throughput capacity, delay, power consumption, etc. An objective and accurate report will be provided by the test bed.

- **ML model evaluation:** This is conducted to solve the problem of estimating performance of the ML model, especially for the ML model used in telecommunication networks, including: accuracy, security, robustness, etc. The applicability of different frameworks and algorithms in different scenarios will be evaluated. The testing of the operating efficiency and flexibility of various deep learning frameworks (such as tensorflow, paddle paddle) is supported, and the results of comparative analysis are provided. The efficiency and accuracy of different AI algorithms in the same application scenario are compared.

- **Data handling:** Network data is collected and pre-processed to form a network data set, which can be used to train ML models. The fields covered by the dataset include network data and image data. Network data is mainly text data, involving five aspects: wireless, core network, transmission network, bearer network, and access network. Image data is mainly based on target detection and semantic segmentation and annotation.

CubeAI 智立方

About CubeAI ★

CubeAI is an open source AI platform completely independently developed by the China Unicom Research Institute. It currently includes sub-platforms and functional modules such as AI online training, automated model release and deployment, and visualization of AI capabilities.

Its core role is to break through the barriers between AI model development and actual production applications, accelerate the process of AI innovation and application, and promote the rapid iteration and evolution of the entire life cycle of AI applications from design and development to deployment and operation.

- **Network intelligence capability evaluation:** Network intelligence capability is evaluated by the test bed from various dimensions: demand mapping, data collection, analysis, decision-making and action implementation. Each application is evaluated according to the ITU–T Y.3173 standard and is divided into L0~L5 levels. The test process includes five steps: 1) determining the evaluation object; 2) dividing the evaluation dimension; 3) analysing the evaluation object; 4) scoring the evaluation dimension; and 5) obtaining the evaluation result. At the same time, the test results of stability, ease of use, accuracy, and throughput of each application are given.

Machine learning marketplace integration in future networks

Another important exploration is ML marketplace integration in future networks, including IMT-2020 (commonly known as 5G), which refers to network ML capability accumulation. Nowadays, ML models may be hosted in varied types of ML marketplaces e.g. The Linux Foundation's Acumos AI, China Unicom's CubeAI, AWS Marketplace, and Huawei's Network AI Engine.

Sometimes, the latest advances in predictive analysis or algorithms have no dependency on network architecture evolution in ML underlay networks. Cloud-hosted ML marketplaces may attract developers of innovative ML mechanisms and algorithms to host their solutions.

Telecommunication operators integrate their own or third-party ML marketplace into the future network and push the application of AI on the telco network and promote the level of network intelligence.

While designing the ML application, network operators need interoperable mechanisms for identification of the ML marketplace which may be used as a source for ML models. Lack of standard mechanisms to exchange ML models and related metadata between ML marketplaces and the network operator's ML deployment environments limits interoperability.

The ITU–T Y.3176 Recommendation: Machine learning marketplace integration in future networks including IMT-2020, which was edited by China Unicom, China Mobile and ZTE, provides the architecture and reference points for integrating ML marketplaces in future networks. Also, the interaction process of model search, model selection and push, model discovery, model training and model deployment are provided. ■

Accelerating deep learning inference with Adlik open-source toolkit

By Liya Yuan, Open Source and Standardization Engineer, ZTE

Machine learning, and deep learning in particular (ML/DL), have gained great popularity in many areas, including machine translation, computer vision and natural language processing.

With ML/DL frameworks such as Tensorflow, Pytorch and Caffe, we can build and train ML/DL models to learn knowledge from data, and ultimately use their inference ability to generate business value in a production environment.

Models may perform well in training, but trained models face new challenges in a production environment.

> *"Models may perform well in training, but trained models face new challenges in a production environment."*
>
> Liya Yuan

In production environments, models can be deployed on various hardware platforms (e.g. central processing unit (CPU), graphics processing unit (GPU), field-programmable gate array (FPGA)) with different performance requirements for computing cost, memory footprints and inference latency in different scenarios.

Challenges in deploying ML/DL models

The result is that there are still some challenges to be addressed when deploying ML/DL models in production environments, even when the models converge well during training stages:

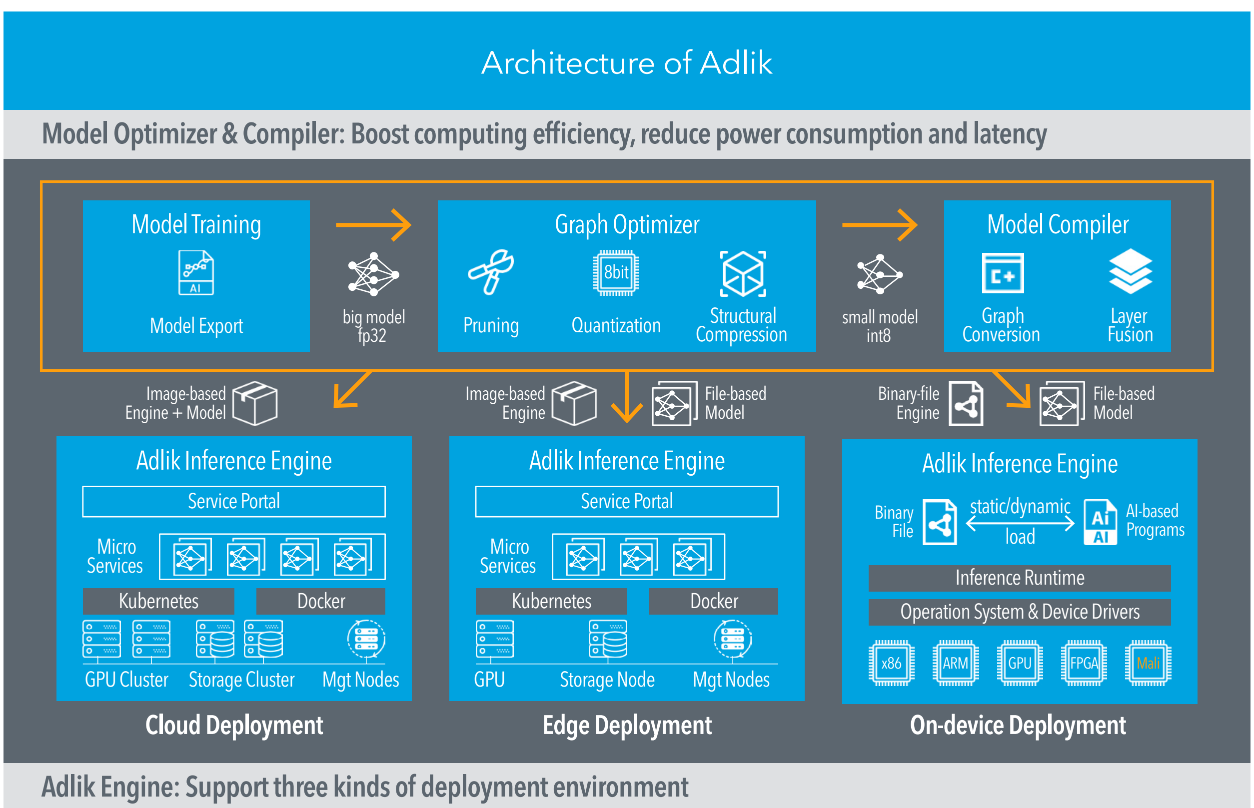
- There is a steep learning curve for users to decide the inference framework best suited to a specific hardware.

- Users require their own solutions to deploy ML/DL models, e.g. deploying models as containers or integrating models into applications on embedded hardware.

- Models need to be optimized to meet different performance requirements in different scenarios.

Adlik open-source project

These challenges are the focus of the open-source project Adlik initiated by ZTE and now incubated by the Linux Foundation AI Foundation (LF AI).



Adlik is an end-to-end optimizing framework for deep learning models with the goal of accelerating the deep learning inference process in cloud, edge, and embedded environments. It is composed of a model optimizer, model compiler and inference engine.

The model optimizer optimizes DL models trained with different frameworks for better inference performance. The model compiler then compiles the models into a format that the inference engine supports. The inference engine then loads the compiled models to provide inference ability on cloud, edge or embedded devices.

The model optimizer plays a key role in Adlik, especially when ML/DL models need to be deployed in environments with strict constraints in computing cost, memory footprints or inference latency – environments such as 5G edge computing scenarios.

Adlik is an open-source project and we are eager to incorporate as many good solutions as we can in the Adlik model optimizer.

Tackling DNN inference optimization – ITU AI/ML in 5G Challenge

That's why we invited the competitors of the ITU AI/ML in 5G Challenge to tackle our problem statement on "Deep Neural Network (DNN) inference optimization" – our challenge to construct a general model-optimization algorithm that can help to achieve model acceleration.

There are many technologies to be explored to optimize ML/DL models for better execution, in fields including model-targetted optimization, system-communication optimization and hardware-specific optimization.

The Adlik model optimizer currently supports model-targetted optimization methods including model pruning and quantization, and our team has focused its recent work on knowledge distillation. This model-targetted optimization focuses mainly on the compression of ML/DL models, with other promising methods including kernel sparseness and low-rank decomposition.

> *"There are many technologies to be explored to optimize ML/DL models for better execution."*
>
> Liya Yuan

System-communication optimization methods such as Deep Neural Network partition can accelerate model inference by optimizing communication between different computing nodes or layers. And hardware-specific optimization with toolkits such as TensorRT can optimize inference operations based on the characteristics of the involved hardware.

The second release of Adlik will come out soon. We welcome you to try it and contribute to it on Github. ■

Turkcell

Challenges and opportunities for communication service providers in applying AI/ML

By Salih Ergüt, 5G R&D Senior Expert, Turkcell

Artificial intelligence (AI) has transformed many industries. With new technical developments in computer vision, where AI trains computers to interpret and understand the visual world, natural language processing (NLP), which involves the interaction between data science and human language, time series forecasting to predict the future, etc., it is hard to keep up with the progress in this field. The science fiction of yesterday is indeed fast becoming today's reality.

However, AI applications have not lived up to their potential for communication networks, considering the vast amount of data available and the difficulty in managing infrastructure, that is increasing in complexity.

5G deployments are slowly picking up all around the world and many standardization organizations have already started to work on 6G; the ITU Telecommunication Standardization Sector (ITU–T) Focus Group on "Technologies for Network 2030" (FG NET-2030) and NGMN's 6G Task Force, to name a few.

> *It is hard to imagine how we can deploy and operate future networks without AI assistance.*
>
> Salih Ergüt

It is hard to imagine how we can deploy and operate future networks without AI assistance.

Service providers – investigating and experimenting

It is common for communication service providers (CSP) to implement non-network related use cases using AI. Churn prediction, customer segmentation, community heatmaps, up-sale/cross-sale, and fraud prediction are popular examples. Many operators have started deploying and experimenting with AI techniques in networking, though it is not an easy task due to limited resources in this field. Also, developing unbiased and accountable models is necessary.

Like many CSPs, Turkcell is committed to using AI responsibly and ethically and announced its commitment publicly as "AI Intelligence Principles".

Service-based architecture is getting more traction and end-to-end network slicing is the key technology for CSPs to provide services with different quality-of-service (QoS) requirements on their 5G infrastructure. In the RELIANCE project of which Turkcell is a partner, while we are assessing the benefits and overhead of slicing under several use cases – ranging from video conferencing to smart trains to smart buildings – we are also investigating the prediction of future performance metrics, which helps devise and optimize resource allocation strategies.

As opposed to focusing on periodic QoS reports and statistical distributions, providing a QoS forecast instead is becoming more valuable for mission-critical services. For example, an autonomous car promising a high reliability score of 99.9999…% is no longer sufficient, as it needs to be warned about a coverage problem in advance, to be able to take precautions. 3GPP TR 22.886 defines use cases on information sharing for partial and fully automated driving and platooning scenarios.

To ensure the availability and quality of services, customer experience centric network management is becoming increasingly important for CSPs. In the 5G-PERFECTA project we are developing radio key performance indicator (KPI)- based models that predict the quality of experience (QoE) of TV streaming services.

Such models allow near real-time customer experience measurements for this service and results can be extended to model other internal and external services through transfer learning.

The challenges of building AI applications

The challenges of building AI applications in an operator network in a multi-vendor and multi-technology setup are outlined in an ITU–T Focus Group on "Machine Learning for Future Networks Including 5G" (FG-ML5G) contribution.

Implementing a real-time application is difficult on a centralized database architecture as it takes a lot of processing power and introduces delays.

About the ITU Focus Group ML5G

The ITU Standardization Sector (ITU–T) Focus Group on Machine Learning for Future Networks including 5G was established by ITU–T Study Group 13 in 2017. The Focus Group drafted ten technical specifications for machine learning (ML) for future networks, including interfaces, network architectures, protocols, algorithms and data formats. FG ML5G was active from January 2018 until July 2020.

Learn more here.

> *The ITU–T Focus Group on 'Machine Learning for Future Networks including 5G' contributed significantly.*
>
> Salih Ergüt

When implementing a real-time special event handling module we had to use non-standard backdoors to network equipment, to access the real-time statistics, and to control the nodes through remote commands. Non-standard interfaces mandate close assistance from the vendor to manipulate the proprietary interfaces and some of those actions may cause security vulnerabilities.

Since these interfaces differ from one vendor to another, AI applications need to be customized and sometimes redesigned for each vendor. The problem with KPIs is not limited to real-time access hurdles, but also the vendor-specific definition of certain KPIs or even a lack of those KPIs which makes AI application development for networks difficult. For example, timing advance (TA) is reported directly as a KPI in the equipment from one vendor, while trace logs need to be parsed to extract TA for another vendor's equipment.

The value of an AI-ready architectural design

An AI-ready architectural design that supports interoperability, data handling mechanisms and tools to evaluate network maturity in this journey is very valuable for CSPs when transforming their network into an intelligent network.

The ITU–T Focus Group on "Machine Learning for Future Networks including 5G" contributed significantly to this cause through its specifications during its lifetime.

Examples include: "Architectural framework for machine learning in future networks including IMT-2020" (ITU–T Y.3172), "Framework for data handling to enable machine learning in future networks including IMT-2020" (ITU–T Y.3174) and "Framework for evaluating intelligence level of future networks including IMT-2020: use cases" (ITU–T Y.3173).

Simpler models for mission-critical industries

There are still some gaps between AI research and experimenting AI in a live operator network. Deep learning in mobile and wireless networking: A survey, "*IEEE Communications Surveys & Tutorials*", vol. 21, no. 3, pp. 2224–2287, 2019, by C. Zhang, P. Patras and H. Hamed, gives an excellent survey of deep learning techniques for mobile and wireless networks.

For mission-critical and highly regularized industries such as telecommunications and health care, trust and validation of the autonomous decisions are of utmost importance. Achieving high predictive accuracy is not sufficient for such systems; the results should be obtained from proper problem representation rather than flawed input data.

Therefore, simpler models such as linear regressions or decision trees are preferred over complex models in these industries due to their interpretability. However, recent research has shown that complex models can also be designed to provide explainability.

Providing the best prediction

Another shortcoming of most AI algorithms is providing the best prediction based on the previously observed data. When they encounter a new scenario, or the underlying data statistics have changed, rather than reporting a lack of confidence in their decision, algorithms still provide their best guess. In telecom networks, it is essential to adopt human-in-the-loop AI mechanisms for robust and resilient operation.

Transfer learning, which provides learning with fewer examples by utilizing knowledge from a previously solved related problem, is an active research area with positive implications for networks. Since the AI agent would be able to adapt to changes to its environment, a model of a well-known service could be extended to a similar service whose inner-workings were unknown.

Federated learning – easing the burden of big data processing

Finally, federated learning, which allows jointly training a model from multiple datasets, is another future research direction that will ease the burden of collecting and processing huge amounts of data from all nodes into a central data warehouse. In addition to increasing loads on the system, centralized approaches also suffer from extended delays and need to be handled with great care to prevent any privacy concerns. Federated learning also allows multiple parties to develop a better model without compromising data privacy.

To conclude, it is hard to imagine network operations without AI-assistance for 5G and beyond networks. There are some challenges but also big opportunities for CSPs to adopt AI into their networks. ■

Rakuten Mobile

Autonomous networks: Adapting to the unknown

By Paul Harvey, Research Lead, and **Prakaiwan Vajrabhaya,** Research Outreach and Promotion Lead, Innovation Studio, Rakuten Mobile

Modern pedagogy has shifted away from traditional rote learning towards critical thinking. This pendulum swing of mindset creates people who can autonomously problem-solve when a yet-to-be-seen situation is encountered, which is essential for the workplace of tomorrow. Like these workplaces, telecommunication networks are difficult, ever-changing environments, consisting of new technologies and services, as well as, complex traffic patterns. This begs the question: if we are training talents to be autonomous; then, is it perhaps time we trained networks to be autonomous too?

Virtualization

Before we dive into autonomy, it's important to ground ourselves in its enabler: virtualization. Virtualization is the process by which hardware in the network is abstracted by software, decoupling an application from the hardware upon which it operates. This new abstraction enables telco operators to unify and simplify their infrastructures, presenting a mechanism to manage their network. The issue here is the control of this mechanism, which is still mainly performed by humans or well-defined automated processes. Future networks, however, are full of challenges, many of which are never-before-seen. Therefore, for future networks to function properly, we must embrace an autonomous mindset, in the same way that teachers now train students to think critically and solve novel problems on the spot.

> *For future networks to function properly, we must embrace an autonomous mindset.*
>
> Paul Harvey, Prakaiwan Vajrabhaya

Automation is not autonomy

First and foremost: automation is not autonomy.

- **Automation** is operating within well-defined parameters or with pre-defined constraints.
- **Autonomy** is the independence to reflect about and adapt behaviour – to go beyond well-defined parameters or pre-defined constraints.

Automation is a powerful tool. It is targeted at a particular problem or set of problems; for example: applying deep learning techniques to find anomalies, identifying specific web traffic or applications, or even performing intrusion detection.

Indeed, various machine learning technologies have been shown as effective ways to automatically address many use cases within the network, such as those identified by the ITU Telecommunication Standardization Sector (ITU–T) Focus Group on Machine Learning for Future Networks including 5G (FG-ML5G), ITU–T Y.3170-series – Machine learning in future networks including IMT-2020: Use cases.

Although automation brings great benefits, it also presents an inherent challenge due to its pre-defined nature. Consider when there is an unforeseen change, for example, in the problem domain, the technology domain, or a new application class, human engineers must intervene and make modifications.

Autonomy, on the other hand, requires little or no human intervention as it is intended – by design – to be self-adaptive and go beyond its well-defined parameters or pre-defined constraints.

Exploring new technologies autonomously

Imagine giving a young child a marker and then walking away. It's likely that by the time you return they will have used it on every surface they could see: paper, face, wall, or even the dog. Looking at ourselves, we, as humans, instinctively experiment with the world around us, receiving feedback (praise or scolding, for instance) that guides but does not dictate future actions.

The same is true with future technologies. Autonomous networks must learn to incorporate new technologies without being explicitly told their purpose and/or learn to apply existing technologies to address changing problem domains. In this way, we learn how to use new tools and their effects upon the world.

How to write down creativity

How does trial and error experimentation actually work? Given the large number of potential technology combinations and configurations, this requires an efficient mechanism to search the space to find potentially useful selections. Again, we look at ourselves for inspiration.

Evolution depends on the semi-arbitrary recombination and modification of small building blocks (genomes) and the application of reward "through survival of the fittest" or increased mating chances for stronger offspring. Evolution-based approaches have been shown as effective mechanisms to search large spaces and find (optimal) solutions to problems. Antenna design in NASA (Hornby, G., Globus, A., Linden, D. and Lohn, J., 2006. Automated antenna design with evolutionary algorithms. In Space 2006 (p. 7242)), and autonomous "rediscovery" of machine learning technologies at Google (Real, E., Liang, C., So, D.R. and Le, Q.V., 2020. AutoML-Zero: Evolving Machine Learning Algorithms From Scratch. arXiv preprint arXiv:2003.03384), are some examples of the concept having been implemented.

Autonomous networks will be no different, where biological genomes are replaced by the modular software blocks that are now standard practice in all software disciplines. As such, evolution is a codifiable mechanism to drive the creativity that is required in addressing the unknown challenges of future networks.

Knowledge is power

Despite promising results, not everything will (or should) be a semi-arbitrary exploration of possible technology combinations. Humans accumulate knowledge diachronically and apply this gained knowledge to shape the decision-making process. Autonomous networks should be able to benefit from this collective knowledge, as well as gain their own.

For this, we incorporate ontology and taxonomy to represent the relationships that exist between the relevant entities found within the network. This representation of human knowledge, combined with collected information from the network and ongoing knowledge gained through trial and error experimentation, enables a codifiable way to guide autonomy. This decreases the likelihood of certain evolutionary combinations being chosen, shrinks the space to be searched, and makes the entire process faster. Existing efforts, such as TM Forum's Telecom Application Map (TAM), serves as a starting point for this process.

Is that everything?

This is, without a doubt, not a simple task. Just as decisions without actions don't solve problems, the ability to autonomously adapt to the unknown requires many different considerations including the creation of small building blocks, the construction of different use case ontologies and taxonomies, the specification of languages to describe these elements, or the construction of simulation and canary testing environments. And the list goes on and on.

The challenge of achieving an autonomous network is one that can only be met by a concerted effort. Not only to address the above, but also to ensure that autonomy can be realized as an interoperable platform between operators, for the practical benefit of all. This approach is already being pursued in standardization organizations for the new wave of edge computing platforms arriving with the 5G era of telecommunications.

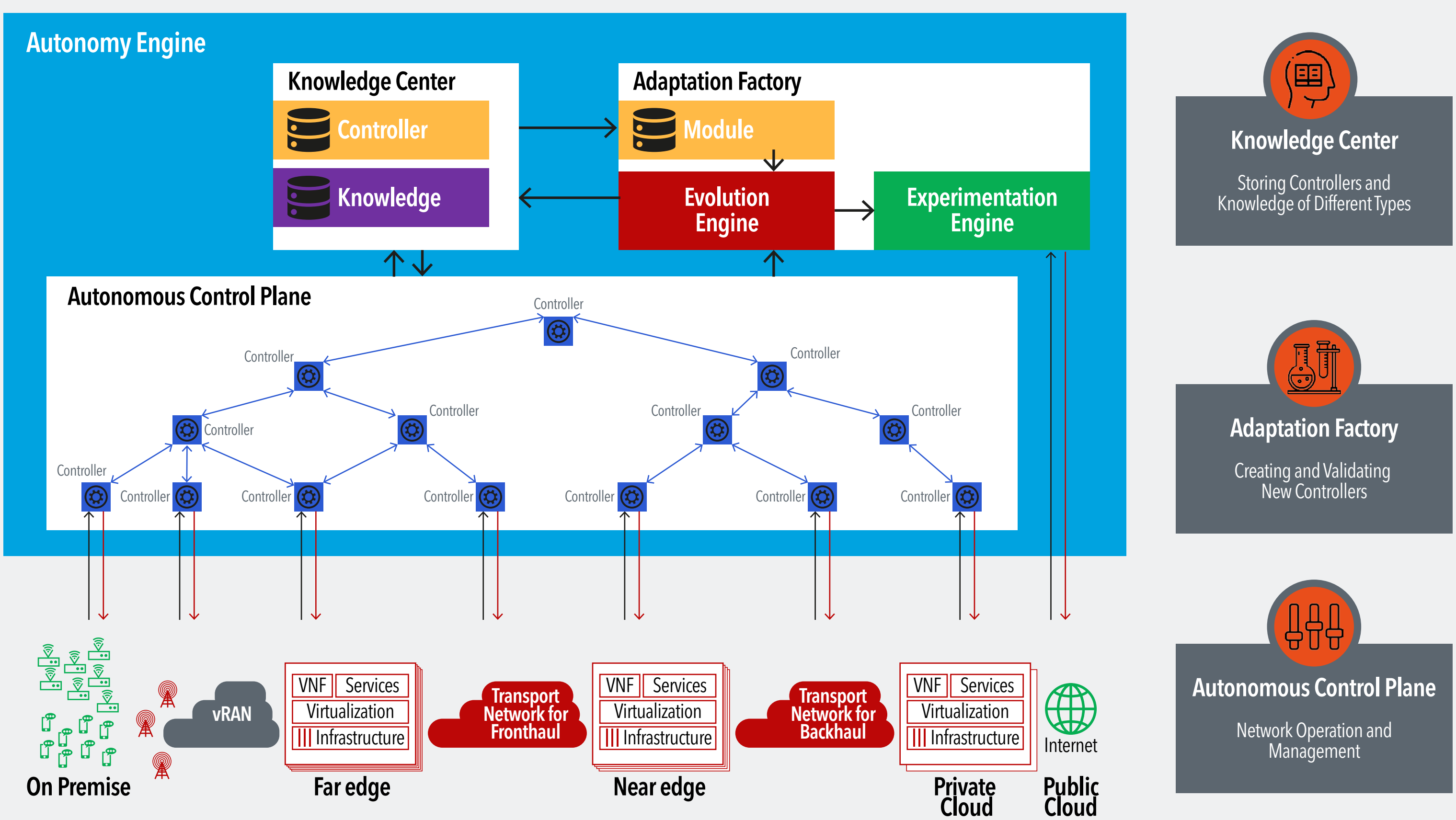
In this regard, ITU and its members have the opportunity to take a leading role in bringing together the necessary community and bridging the gap.

> *In this way, autonomous networks present the opportunity to free telco engineers from the mundane, to focus on the extraordinary.*
>
> Paul Harvey, Prakaiwan Vajrabhaya

Autonomy: Freedom from the mundane

Just as our children are being prepared for the unknown workplace of the future, we too are preparing our networks to autonomously adapt to the unknown challenges of the future network. We are replacing hand-crafted automatic approaches with evolutionary-driven, experimentally verified, machine-crafted approaches. In this way, autonomous networks present the opportunity to free telco engineers from the mundane – to focus on the extraordinary. ■

High-level architecture for evolutionary-driven autonomous network adaptation



vRAN = Virtualized radio access network

VNF = Virtual network function

Source: ©2020 Mobile Rakuten Innovation Studio.

Quality of Experience testing in mobile networks

By Arnd Sibila, Technology Marketing Manager,
Mobile Network Testing, Rohde & Schwarz

Mobile network operators need to test the stability and performance of their networks in order to ensure good service. Due to the enormous amounts of data involved, this is hardly possible with manual methods. This is where artificial intelligence (AI) comes to the rescue.

Network testing in the 5G era

With the advent of the fifth generation of mobile communications, network testers are confronted with a novel situation. Many aspects of 5G – diverse frequency bands, network operators' different rollout programmes, the breadth of applications such as the Internet of things (IoT), conventional mobile communications, traffic networking, etc. – lead to highly differentiated networks and test data.

"Mobile network operators need to test the stability and performance of their networks in order to ensure good service."

Arnd Sibila

Analysing this data in the usual aggregated form quickly leads to distorted results and incorrect interpretations. AI offers a good solution to this dilemma. While algorithm-based methods reflect specific theories, AI methods, such as pattern recognition, are able to evaluate data sets without preconceptions, and discover relationships that would remain hidden to human analysts.

Big data needs machine learning

The term "machine learning" is a bit more specific than the term "artificial intelligence". With machine learning (ML) the goal is to automatically derive general rules from a large volume of data. In our ML applications on drive testing data, a typical deep learning pipeline covers two stages: training and inference.

At the **training stage**, we collect the training data from numerous tests performed under different configurations and environments. This data goes through a computing-intensive training process, normally using graphics processing units (GPUs) owing to their ability to perform numerous simple operations in parallel. The output of the training stage is a model that contains the knowledge of all the training data needed to accomplish our specific objective function.

At the **inference stage**, we take the learned model and apply it to new data to generate a prediction or insight otherwise hidden in the structure of the data. This stage requires less computing power than the training stage, thus making it suitable for standard server central processing units (CPUs) or even edge computing, which avoids sending sensitive data to a server.

After the intensive training phase, the model is able to correctly interpret new measurement data almost spontaneously.

Opting for machine-learning methods

We use ML methods for applications such as simplifying the optimization of mobile networks or improving the assessment of qualitative differences between providers. The Data Intelligence Lab established in 2018 tackles these issues and supports Rohde & Schwarz research and development (R&D) departments with data-based analysis methods. These approaches are especially promising for testing mobile networks where particularly large amounts of data are generated, so that manual analysis and rule formulation are no longer practical. ML makes it possible to use the information hidden in large data sets, for example to derive new assessment metrics. An example is the call stability score (CSS).

Smart – a new generation platform for mobile network testing

Watch this video to get a glimpse of the new paradigm in mobile network testing that allows to reduce complexity for Quality of Experience (QoE) centric and targeted network quality and performance improvements.

Rohde & Schwarz and ITU standards

Rohde & Schwarz is an ITU Sector Member active in the work of the ITU Standardization Sector (ITU-T) Study Group 12 (SG12: Performance, Quality of Service and Quality of Experience) for driving AI and machine learning.

AI and machine learning are today widely used in developing models to assess the quality of speech, audio and video, for example in ITU standards for the quality assessment of audiovisual streaming, in particular ITU P.1203 (progressive-download and adaptive-bitrate AV) and ITU P.1204 (video streaming services up to 4K).

New ITU quality-assessment standards address intelligent network analytics and diagnostics (ITU E.475) and the creation and validation of machine learning based models to assess media quality (ITU P.565).

Based on the experience of developing these ITU standards, Study Group 12 is giving further guidelines on how to apply AI and machine learning in ITU standardization in an upcoming ITU Technical Report and Supplement.

Rohde & Schwarz expects a significant increase in studies and recommendations using AI and machine learning techniques and supports these activities in ITU very actively.

> *"A suddenly dropped phone call is an annoying experience."*
>
> Arnd Sibila

Call stability score: A new assessment metric for reliable communications

A suddenly dropped phone call is an annoying experience. That is why mobile network operators have been testing voice quality and connection stability for many years. A popular statistic is the call drop rate (CDR). But since the number of dropped calls is very low in mature networks, it is necessary to make a large number of calls in order to obtain a statistically significant value. Consequently, drive test campaigns are long and expensive.

Therefore, we use a method to replace the binary call status (either successfully completed or dropped) with a finely graduated analog value. This is done by creating a statistical AI-generated model that links the transmission conditions with the call status.

The CSS derived from the model allows the reliability of the mobile connection to be measured over the entire call duration and classified based on quality. The diagnostic also includes unstable calls that were successfully completed but the data proves they were not far away from being dropped. In conventional CDR statistics, those unstable calls would be assessed positively as successful calls, distorting the network quality assessment.

The CSS value is based on information gathered from millions of test calls and incorporated in the model during the learning process. The assessment is conclusive right from the first call. The network call quality is registered more accurately and with less test effort. The model assesses the data based on the learned rules, and outputs a number between 0 and 1. The higher the number, the lower the likelihood of a drop occurring in the observed interval.

The CSS measurement is part of the SmartAnalytics analysis platform from Rohde & Schwarz (see Screenshot 1).

Time-based anomaly detection

Another AI-driven function in this software suite is anomaly detection using unsupervised learning – the neural network is trained to learn information that is hidden in the data without labels.

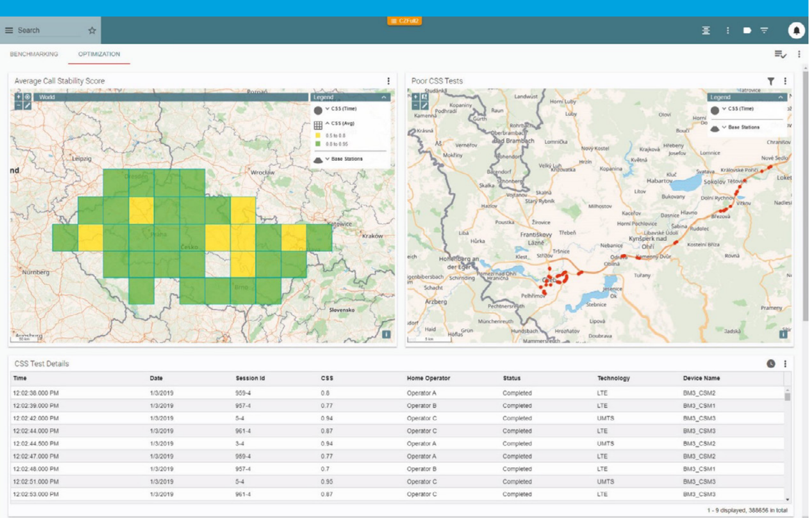
Anomaly detection is a much-appreciated tool by data scientists. It aims to find data samples that do not conform to the regular distribution of the dataset to which they belong. Finding anomalous samples, also known as distribution outliers, provides valuable insight that often correlates with defects or errors in the data collection process (e.g. faulty or misconfigured equipment).

Detecting anomalies in data transfers based on variable-length time series benefits mobile network operators by detecting deviations and identifying problematic areas instantaneously that are otherwise masked by key performance indicator (KPI) averages.

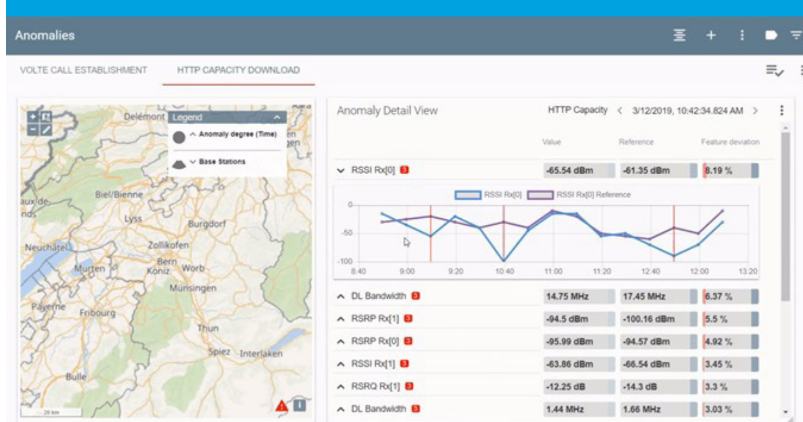
The visualization of the feature in SmartAnalytics from Rohde & Schwarz enables users to quickly see which phases of a test deviate from the model. The overall effect of time-based anomaly detection is a more efficient methodology for drive tests and optimization.

Anomaly detection, in particular time-based anomaly detection, is applicable to various test cases.

Screenshot 1: Call stability score visualization in SmartAnalytics from Rohde & Schwarz



Screenshot 2: Time-based anomaly detection visualization in SmartAnalytics from Rohde & Schwarz




Shutterstock

A network operator's view of the role of AI in future radio access networks

By Chih-Lin I, Chief Scientist, and **Qi Sun,** Senior Researcher, Wireless Technologies, China Mobile Research Institute

5G is under commercialization all around the world and the speed of its development is amazing. China, for example, is expected to have more than 600 000 5G base stations at the end of this year.

5G, however, is an extremely complex system, which cannot be accomplished by simply building hundreds of thousands of base stations. There are still many challenges on how to achieve commercial success. For example, 5G power consumption, high costs, and 5G's high flexibility lead to difficulties in operation and maintenance optimization. How to quickly and effectively serve the diverse vertical industrial application demands is another challenge.

> *We need to rethink the traditional telecommunications paradigm and embrace new technologies.*
>
> Chih-Lin I, Qi Sun

To tackle these issues, we need to rethink the traditional telecommunications paradigm and embrace new technologies, such as cloud, big data analytics and machine learning (ML). Deep integration of information technologies, data technologies and communication technologies are envisioned for 5G and beyond.

Intelligence is expected to be introduced into wireless networks for all the domains and at all levels, from local, to edge, to the cloud. Data analytics, ML and artificial intelligence (AI) have been identified as key drivers of the intelligence evolution and revolution in wireless networks.

Network capabilities with data analytics and machine learning

Data analytics and ML will empower networks with the following capabilities:

- **Reliable prediction:** The vast multi-dimensional data collected from the network enables network traffic as well as the prediction of network anomalies, service pattern/type, user trajectory/position, service quality of experience, radio fingerprint, interference, etc.

  These predictions will undoubtedly empower proactive network management and control, leading to significantly improved network resource allocation and energy efficiency while ensuring a customized user experience.

- **Advanced network optimization and decision-making:** Driven by the collected data from the real network, data analytics and ML can help to efficiently solve massive problems in the 5G network. These are usually hard to model or entail great computational complexity due to extremely high dimensions or non-deterministic polynomial-time (NP)-hardness.

Focusing on the radio access network (RAN), the use cases can be roughly categorized into four types: 1) Intelligent network management and orchestration; 2) Intelligent mobile edge computing; 3) Intelligent radio resource management; and 4) Intelligent radio transmission technology.

Use cases may also further extend to optimize the radio frequency area, i.e., AI-assisted digital pre-distortion.

Four use case studies from the operator's perspective

The following four typical use cases are detailed below, from the operator's perspective and based on the commercialized and test-trial experience.

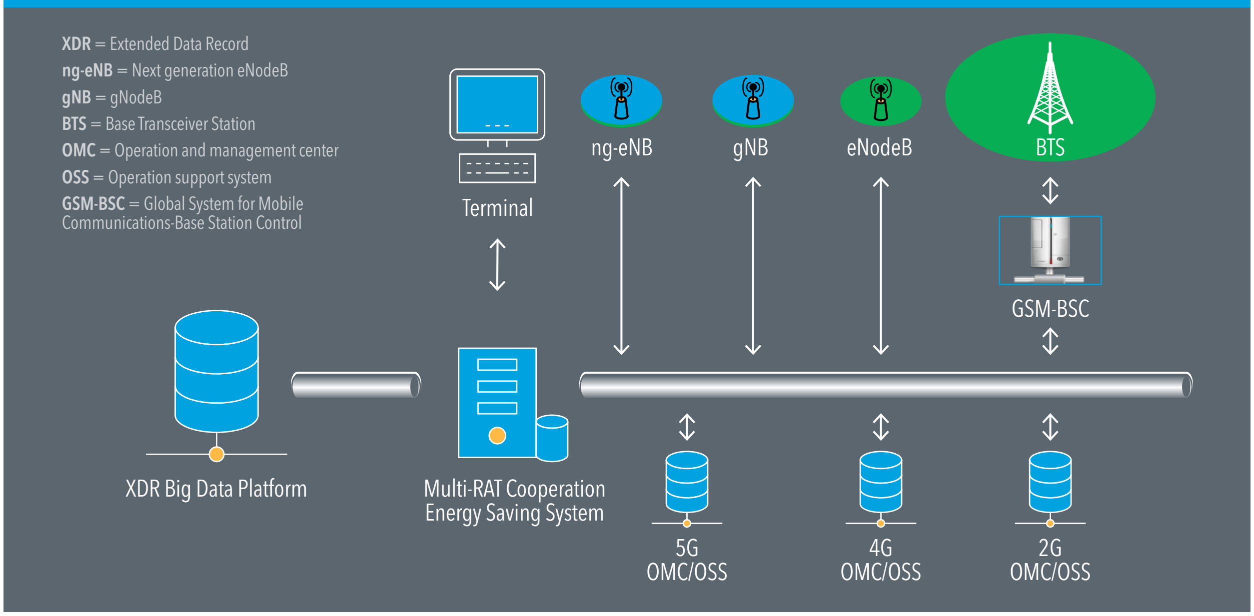
1 Energy saving

To cope with the energy challenges caused by mobile network expansion, the multiple radio access technology (RAT) cooperation energy-saving system (MCES) was developed by China Mobile to improve the energy efficiency of mobile networks. MCES interacts with the radio access network in real-time, and can support 2G/3G/4G RAN equipment from multiple vendors. Specifically, the MCES system has three major technical features:

- Network-level energy saving.
- Energy-saving cell discovery function based on big data.
- Cell turn-off/on at a lower time-granularity.

MCES has been deployed in 18 provinces across China, including 970 000 cells. In 2019, the total energy saved amounted to over 40 million KWh. Now the MCES is also evolving to incorporate a 5G system to enable coordinated energy saving in both 4G and 5G networks.

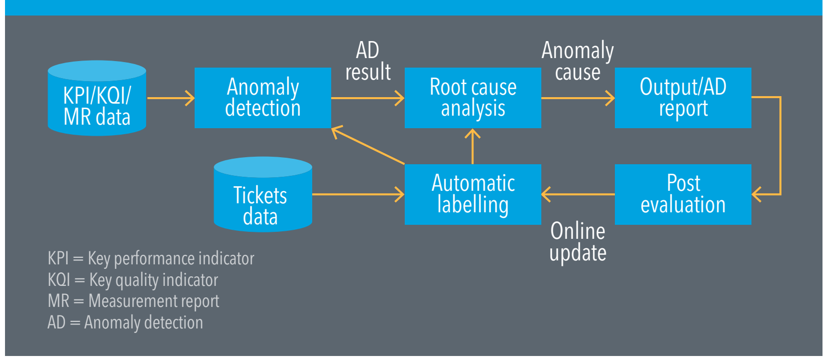
Topology of multi-RAT cooperation energy-saving system (MCES)



2 Automatic anomaly analysis

Anomaly detection (AD) and analysis has always been an important part of the operation and management (OAM) system. By introducing a dynamic machine-learning-based AD algorithm, the number of manual rules can be reduced and more accurate root cause information can be revealed. Meanwhile, the AI-based root cause analysis (RCA) algorithm directly replaces part of the human effort involved in the anomaly analysis process.

Workflow of the automatic anomaly analysis



3 Quality of experience optimization

The 5G business model is changing from "volume" to "value". User quality of experience (QoE) is playing a key role in the commercialization of 5G, thus network optimization targets are shifting from key performance indicator (KPI) to key quality indicator (KQI) related QoE.

The Radio Intelligent Controller is positioned as a data-driven platform to provide customized RAN capabilities and exposure, especially for vertical industries and over-the-top (OTT) services. High-definition (HD) video streaming, cloud virtual reality (VR), and cloud gaming are expected to be among the most popular services of the 5G era.

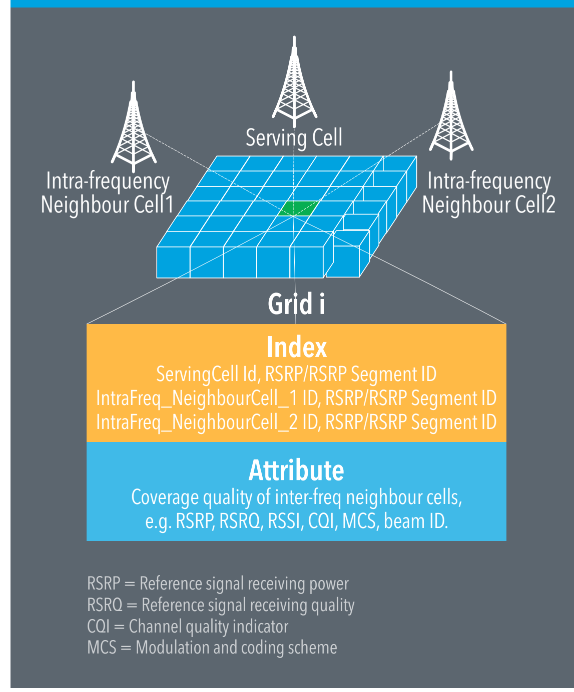
In 2019, China Mobile conducted trials in the Shanghai 5G network and the following features have been verified: 1) AI/ML-based QoE prediction and guarantee for cloud VR; and 2) Radio bandwidth estimation for cloud VR adaptive coding selection.

4 Radio fingerprint-based traffic steering

Traffic steering, also termed mobile load-balancing, is a widely used network solution to distribute the traffic load among cells, or to transfer traffic in order to improve network performance. This solution aims to enhance the performance of traffic steering by constructing a radio fingerprint which divides the cell into grids by the serving cells and neighbouring cells radio signal levels, in order to locate the user equipment's (UE's) grid and perceive UE's coverage information. This can largely reduce the number of UE inter-frequency measurements and speed up traffic steering.

China Mobile and partners also carried out trial tests of radio fingerprint-based traffic steering in the commercial network. Based on the optimization of radio fingerprint-based traffic steering, the test results showed that compared with traditional load balancing, the duration of high loading is reduced by 13 per cent.

Virtual grid of radio fingerprint



Also, the measurement reconfiguration from base station and the measurement report signalling overhead from UE are reduced by 54 per cent and 83 per cent, respectively. Moreover, with the radio fingerprint-based load balancing, the average Internet protocol (IP) delay of the tested cells is reduced by 20 per cent.

Standards progress

Various studies on network intelligence standardization, from management to control planes, including use cases and requirements, AI functional framework, procedures and architecture have been actively explored in standardization organizations including the ITU Telecommunication Standardization Sector (ITU–T), 3GPP, O-RAN and ETSI.

3GPP has already introduced the data analysis and AI/ML related function of the core network and management plane with the service-based core network and management architecture. While for the radio access network, considering the distributed architecture and more stringent timing and reliability characteristics, it faces more challenges when we bring in embedded AI.

The industry is working actively towards the open and smart radio access network to fully ensure the commercial success of 5G.

Future work

Although great progress has been achieved, 5G specifications barely offer mobile operators any guidance on how to make their 5G networks truly AI/ML-enabled. Lots of further work is envisioned on better supporting AI/ML-enabled networks, and to incorporate fundamental concepts of AI and ML as the core fabric of the network, which includes:

- Customized fine granularity data collection.
- RAN capability exposure for customized network and service optimization.
- RAN programmability, service-based architecture in RAN to enable AI/ML.
- Decoupling the communication and data analytics/AI/ML to enable efficient innovation.
- Wireless open data sets to accelerate the pace of wireless AI algorithm and application innovation. ■

"The industry is working actively towards the open and smart radio access network to fully ensure the commercial success of 5G."

Chih-Lin I, Qi Sun


Shutterstock

AI and open interfaces: Key enablers for campus networks

By **Günther Bräutigam,** Managing Director, Airpuls; **Renato L.G. Cavalcante,** Research Fellow, and **Martin Kasparick,** Research Associate, Fraunhofer HHI; **Alexander Keller,** Director of Research, NVIDIA; and **Slawomir Stanczak,** Head of Wireless Communications and Networks Department, Fraunhofer HHI, Germany

Modern communication is the foundation of successful digitization. Thanks to the 5G standard, completely new applications are emerging across industry verticals, and there is a strong demand for new wireless technologies that address industry-specific problems in campus networks, also known as private networks.

To address this demand, for example, the Federal Government in Germany has laid the foundations for the development of campus networks by allowing their operation in the 3.7–3.8 GHz band, which is particularly suitable for covering large areas and supporting high mobility scenarios.

> *“Modern communication is the foundation of successful digitization.”*
>
> Günther Bräutigam, Renato L.G. Cavalcante, Martin Kasparick, Alexander Keller, and Slawomir Stanczak

This frequency band is expected to be complemented with the millimeter-wave band, which offers many advantages in terms of transmission bandwidth, security against tapping, and robustness against jamming, to name a few.

Similar to conventional mobile networks, campus networks have a core network containing the central elements for network control, and a radio access network (RAN) managing the wireless connections from base stations to mobile terminals.

Many vendors offer software for the core network. Only a few vendors have been supplying RAN technologies. The vertically integrated solutions hardly allow for interoperability. Due to the high barrier to entry, industry-specific technologies targeting niche markets are not being developed. This lack of competition impedes innovation in RAN technology.

Fostering innovation in campus networks

To foster innovation in campus networks, a new business model has been advocated. It is comprised of system design, optimization, and the integration of open and secure wireless technologies.

The technologies are based on disaggregation, virtualization, openness, and AI. In this new approach, the RAN may be disaggregated into modules performing network functions in software. Open interfaces between such virtualized RAN modules will be crucial in enabling interoperability among vendors.

While the interfaces are standardized, standards do not address the implementation of the network functions, enabling vendors to differentiate to their own advantage and the advantage of their customers.

This high degree of softwarization of RAN functions, in combination with the use of commercial off-the-shelf (COTS) software and hardware leads to positive economies of scale, and reduced costs.

The above evolution from closed to open and programmable systems has two important consequences.

It reduces the investment required to develop technologies for radio access networks, enabling small and medium-sized companies to develop industry-specific solutions for campus networks.

This lower barrier to entry also expands the market by creating the conditions for niche technologies from other markets to enter campus networks. For example, existing data mining algorithms may be customized for wireless systems, such as the analysis of RAN data streams to detect attacks on radio interfaces.

If based on flexible multi-vendor solutions, campus networks would meet the needs of complex industry-specific environments.

Moving beyond 5G networks with AI

As we move beyond 5G networks, a holistic approach to manage this complexity requires AI as an integral part of the overall system design.

When integrating new hardware and software components into a network, the integration can be tested in a "digital twin" environment where simulation is accelerated by AI.

Novel AI tools are needed in support of network functions such as scheduling, beam management, interference coordination, localization, symbol detection, and channel estimation, to name just a few prime examples.

These tools may be fundamentally different from existing tools applied in fields such as speech recognition and computer vision.

However, the success of open and programmable systems will mainly be decided on the lower layers of the communication stack, where the environment is highly dynamic and uncertain.

The challenges

Acquiring datasets for training data-driven machine learning algorithms becomes a challenge as a result. By the time sufficient data has been collected, the environment may have changed significantly enough to render the training data out of date.

Pure model-based methods are now widely applied but may also face serious challenges. As current wireless models often ignore beam-squinting effects and rely on further approximations such as waves in the far field region, they may become too crude with increasing operating frequency, rates, and number of antennas.

These issues may be addressed by hybrid model and data-driven methods, where the data is used to mitigate the uncertainty of the models and the coarse models help to reduce the amount of training data required by the learning tools. Crucially, some of these operations must be performed within timeframes ranging from microseconds to milliseconds, making it increasingly important to develop scalable algorithms that can be highly parallelized on commercial off-the-shelf (COTS) hardware.

Call to AI action

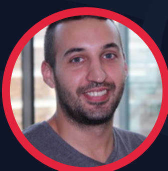
To meet the strict requirements of mobile applications in campus networks, the industry needs to devote a great effort to the development and application of novel AI methods.

This will create opportunities for countries to develop their digital sovereignty and for companies to provide solutions for niche markets that are waiting to be served.

"To meet the strict requirements of mobile applications in campus networks, the industry needs to devote a great effort to the development and application of novel AI methods."

Günther Bräutigam, Renato L.G. Cavalcante, Martin Kasparick, Alexander Keller, and Slawomir Stanczak

Quotes from hosts of the ITU AI/ML in 5G Challenge problem statements



The ITU AI/ML in 5G Challenge has provided a great opportunity to bring together standardization activities with academia by allowing students and professionals to solve important problems in communications. In particular, the challenge has led to obtaining very interesting results that may open the door to revolutionize the way communications are understood. For the Dynamic Channel Bonding problem, the application of Deep Learning models accounts for a significant breakthrough in the field.

Francesc Wilhelmi

PhD Student, Wireless Networking Research Group, UPF, Spain

First of use cases in 2030?

Andrey Koucheryavy

Chair and Professor of Telecommunication Networks and Data Transmission Dept., SPbSUT, Chief Researcher NIIR, Russia, and Chairman of ITU–T SG11

To ensure network coverage density even in sparsely populated areas while guaranteeing high security and satisfying quality of service, an integration of intelligence edge computing and blockchain technology will play an important role in future 2030 networks.

Ammar Muthanna

Deputy Head in Science, Telecommunication Networks and Data Transmission Dept., SPbSUT, and Head, SDN Laboratory, Russia

We are grateful to the "ITU AI/ML in 5G Challenge" for giving us the opportunity to connect with AI researchers around the globe in solving an operational challenge for Turkcell. Our challenge topic, "Radio link failures", has critical implications for an operator network. Modelling link failures based on weather forecasts is a daunting task due to the unreliable nature of forecasts and rare occurrences of radio failures. Challenge preparation and communication with contestants have already helped us better formulate the problem, resolve inconsistences in the data, and consider alternative approaches. We are expecting to receive high-quality solution proposals for this challenging task and we are looking forward to applying the learnings from this challenge to implement mechanisms that take precautions before radio link failures occur, to improve customer experience.

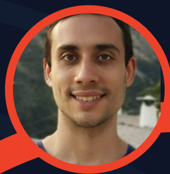
Salih Ergut

5G R&D Senior Expert, Turkcell, Turkey

The main important task for AI in 5G is traffic recognition "on the fly" without making delays for new services traffic management such as the Tactile Internet, medical networks, and autonomous vehicles.

Artem Volkov

Researcher and PhD Student, Telecommunication Networks and Data Transmission Dept., SPbSUT, Russia



Hosting a competition within the AI/ML in 5G Challenge has been a completely worthwhile experience. ITU organizers have made it very easy to prepare and run the whole competition, and have done a great job in attracting participants. In our problem statement, the participation numbers were far beyond what we could expect at the beginning. As members of academia, we see this Challenge as a pathway to disseminate the research we do. This kind of competition is very interesting for us to connect with a different public from both the academia and industry. Moreover, we were impressed by the proposed solutions, some of them pushing the state of the art. We are looking forward to the next edition in 2021.

José Suárez-Varela
PostDoctoral Researcher, BNN-UPC, Spain

We are very honoured to be a part of the ITU AI/ML in 5G Challenge — a great opportunity to accelerate the pace of applying AI/ML algorithms in telecommunication networks. We proposed a problem statement "DNN model inference optimization" in the Enabler Track. The topic is quite important when considering deploying ML models in the network, especially where there are strict requirements for inference performance. We have received some great submissions so far, which we believe will definitely enlighten our work in the Adlik open source project.

Liya Yuan
Open Source and Standardization Engineer, ZTE, China

The ITU AI/ML in 5G Challenge helps to depart from the "small-data regime" and properly assess AI/ML algorithms such as deep learning, with large datasets and reproducible experiments.

Aldebaro Klautau
Professor of Electrical and Computer Engineering, UFPA, Brazil

The Challenge provides a stage to showcase the potential of applying machine learning to enable network intelligence. It attracts more researchers and engineers to put effort into the research and development of the intelligent network and has gathered the operators, suppliers, researchers, college students and related communities together to build a more flexible, efficient, green, and resilient network. By integrating AI into the network services, the network is being evolved to support better social, commercial and technology services. The Artificial Intelligence Industry Alliance will continuously support ITU in organizing challenges to drive the convergence of AI technologies and network infrastructure and services.

Qiang Cheng
Artificial Intelligence Industry Alliance (AIIA), China

We are very honored to be part of this very exciting ITU AI/ML in 5G Challenge and to provide a problem set. AI/ML adoption is still at the very early stage in the telecommunication industry and innovation will definitely be brought by very unique ITU communities consisting of different countries, organizations, and backgrounds. Let's keep innovative and collaborative in the 5G era and beyond.

Tomohiro Otani
Executive Director, KDDI Research, Japan

Running the ITU challenge on MIMO channel estimation at millimeter wave has been a great way to kick off my new research program at North Carolina State University. It has made me realize that putting together the AI and wireless communities, so they can speak the same language, is a huge challenge. I have enjoyed exposing the participants to one of my favourite problems in millimeter wave MIMO communications.

Nuria Gonzalez Prelcic
Associate Professor, NCSU, United States

The Ministry of Internal Affairs and Communications in Japan recognizes that research and development on network autonomy using AI/ML is extremely important for the further development of 5G and the early realization of B5G (Beyond 5G). By using AI/ML and competing to solve various 5G network problems, engineers will greatly contribute to the development of information and communications, such as the improvement of AI/ML technology and human resource development. The global ITU AI/ML in 5G Challenge is a very meaningful activity.

Ministry of Internal Affairs and Communications (MIC), Japan

We posit that one of the promising research and development directions for Beyond 5G/6G telecommunications is enabling "autonomic intelligence" in networking, as outlined by the "Beyond 5G Strategic Board" held by the Japanese Government's Ministry of Internal Affairs and Communications. We observe that the ITU AI/ML in 5G Challenge is in perfect alignment with the strategy and has been quickly and globally encouraging young researchers to participate in competitions. It is an indication that ITU–T's activity is proving to be an effective means for putting forth the strategic direction of R&D for enabling super-intelligence in networking. In Japan, RISING (Cross-Field Research Association of Super-Intelligent Networking) has led the regional competition of ITU's AI/ML in 5G Challenge, fully supported by MIC, 5GMF, TTC, and industry partners, KDDI and NEC. I really appreciate ITU's global promotion of applying AI/ML in telecommunications and I am looking forward to a successful outcome of the activities.

Akihiro Nakao
Professor, UTokyo, Japan

Quotes from the ITU AI/ML in 5G Challenge participants

My participation in the ITU AI/ ML in 5G Challenge has allowed me to gain hands-on experience of relevant topics to build and design future wireless technologies. Specifically, it has been exciting to apply different ML methods to the particularities of future spectrum access mechanisms that allow even better performance of what we have today.

Paola Soto-Arenas
PhD Researcher, University of Antwerp, Belgium

Participating in this challenge was a golden chance to test and improve our ML skills in the context of newly introduced technology such as 5G. During this competition, we have managed to overcome some challenges and meet new people from around the world.

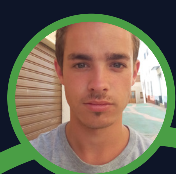
Khalid Al-Bagami
Telecom Engineer, Ericsson, Saudi Arabia

The challenge has given us the opportunity to rethink classical channel estimation approaches from an ML perspective. I am excited to see the results of other participants.

Dolores Garcia
PhD Student, IMDEA, Spain

I feel ecstatic to say that my participation in the ITU AI/ ML in 5G Challenge has been a great learning experience. It's been a good opportunity to work on the exploratory problems in wireless communication and trying to find solutions to them using AI/ ML approaches.

Megha Gururaj Kulkarni
Student, PES University, India



At UC3M we trained an NN solution to forecast 802.11 WLANs throughput. The organizers (UPF) provided both datasets, and support during the challenge. It has been a very fulfilling experience.

Jorge Martín Pérez
PhD Student, UC3M, Spain

I'm a health tech enthusiast and my UNDP R42 group's work is on an AI in 5G mobile chatbot providing COVID-19 information. My work spans Africa and Europe, and I aim to reach more regions to influence global markets and society, young and old. ITU has been fantastic in bringing together such dynamic participants to take part in innovations for 2020 challenges.

Mahlet Shimellis
UNDP, Ethiopia/United States


Shutterstock

AI/machine learning for ultra-reliable low-latency communication

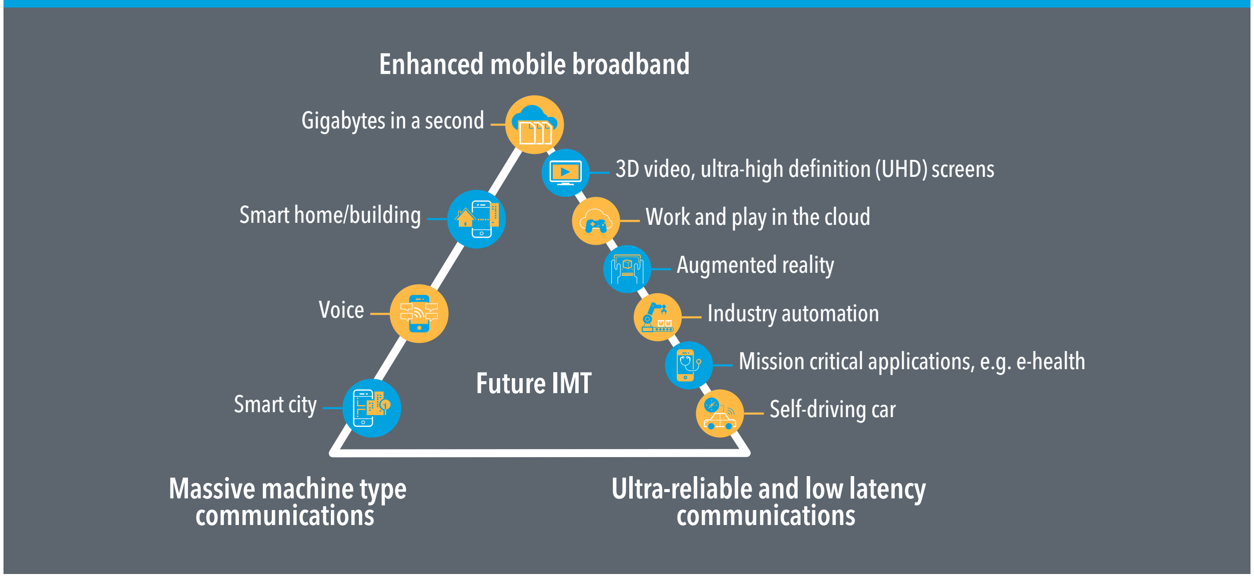
By Andrey Koucheryavy, Chaired Professor, Telecommunication Networks and Data Transmission Department, The Bonch-Bruevich Saint Petersburg State University of Telecommunications (SPbSUT), Chief Researcher, NIIR, and Chairman, ITU–T SG11; **Ammar Muthanna,** Deputy Head, Science, Telecommunication Networks and Data Transmission Department, SPbSUT, and Head, SDN Laboratory; **Artem Volkov,** Researcher and PhD Student, Telecommunication Networks and Data Transmission Department, SPbSUT, Russia

5G networks are designed to integrate all the achievements of mobile and fixed communication networks to provide ultra-high data speeds, enabling a range of new services with new cloud computing structures such as fog and edge.

The International Mobile Telecommunications (IMT) Vision in ITU Recommendation ITU–R M.2083-0, describes three 5G usage scenarios: 1) enhanced mobile broadband (eMBB); 2) massive machine-type communications (mMTC); and 3) ultra-reliable and low latency communications (URLLC).

Massive machine-type communications, the concept of the Internet of Things (ITU Recommendation ITU–T Y.2060/Y.4000), refers to trillions of things connected and uniquely identified, and requires a fundamental shift from traditional ideas about the number and volume of device databases in the network.

5G usage scenarios from the ITU-R IMT-2020 Vision Recommendation



The Internet of Things, as a driver for the development of network technologies, has put forward a whole layer of new services covering all spheres of life and society.

But ultra-reliable and low latency communication creates the most significant challenges for the scientific and technical community.

While it could give life to "Tactile Internet" applications in areas such as telemedicine, autonomous vehicles and industrial robotics, it also generates stringent new quality of service (QoS) requirements.

Reducing latency in 5G networks and beyond

We can reduce latency drastically in 5G networks and beyond, building on innovations in software-defined networking (SDN), network function virtualization (NFV) and edge computing.

And working towards the ambitious goal of latencies lower than 1 millisecond to enable tactile applications, the scientific and technical community is looking to fog and edge computing technologies as key to architectural approaches in future networks.

This represents a movement towards a greater number and variety of network and computing technologies, resulting in more complex networks and more complex network management.

This calls for a review of the established principles in network management.

And this introduces artificial intelligence (AI) — new capabilities well-positioned to support the high programmability and automated provisioning possible with SDN-enabled orchestration systems.

> *A Tactile Internet application in 2020 may have been remote-controlled robots assisting patients in hospitals ill-prepared for outbreaks of COVID-19.*
>
> Andrey Koucheryavy, Ammar Muthanna, Artem Volkov

AI is a class of mathematical machine-learning algorithms and models and Big Data processing algorithms. The latest processor technology makes it possible to implement AI algorithms quite effectively.

Volumes of traffic are growing, and the heterogeneity of traffic is also growing. The Internet of Things and ultra-reliable and low latency communication introduce a wide variety of new requirements, and this will call for much greater efficiency in decision-making relevant to quality of service (QoS).

Tactile Internet

Here are the key characteristics of Tactile Internet applications:

- Decentralized network architecture to enable Tactile Internet services at the network edge. (Can network decentralization reduce digital inequalities?)
- Real-time tactile interaction matching human sensations requires latency lower than 1 millisecond.
- The under 1 millisecond target creates stringent new requirements for system-network solutions.

A Tactile Internet application in 2020 may have been remote-controlled robots assisting patients in hospitals ill-prepared for outbreaks of COVID-19.

Existing QoS tools cannot perform at the required level. And most solutions already need to have load forecasts for certain services – forecasts that take into account geographical and dynamic factors such as subscriber movements, including at high speeds.

Operators also need full-fledged system forecasts of infrastructure development that take into account the speed at which new technologies are introduced, to enable new services over the Internet, as well as the related changes expected in people's lifestyles.

New AI approaches to traffic identification with software defined networking

AI can perform two tasks of key importance to QoS assessment: unambiguous identification of traffic, and subsequent forecasting.

The AI task of identifying traffic includes the need to recognize a large number of traffic types without introducing additional delays (taking into account ultra-reliable and low latency communication services), and the need to expand and adjust the AI algorithm to different geographical locations of the network and services.

The software defined networking (SDN) capabilities of 5G networks make new approaches to traffic identification possible.

The machine learning model developed by St. Petersburg University (SPbSUT) as part of the ITU AI/ML in 5G Challenge analyses the metadata of SDN network streams to identify and predict traffic.

This method of traffic identification and forecasting introduces no delays in the transmitted traffic at the data plane level, ensures the portability of analytical modules across SDN controllers, and can expand in types of recognized traffic.

AI algorithms – in conjunction with new technologies for building networks and cloud services calculations – can offer valuable support to the movement towards ultra-reliable and low latency communications.

But the main breakthrough to achieve the full promise of tactile, ultra-reliable and low latency communication applications is expected with the next generation of physical-layer technologies – namely, quantum communications. ■

> *"The main breakthrough to achieve the full promise of tactile, ultra-reliable and low latency communication applications, is expected with the next generation of physical-layer technologies."*
>
> Andrey Koucheryavy, Ammar Muthanna, Artem Volkov

AI/ML integration for autonomous networking – a future direction for next-generation telecommunications

By Akihiro Nakao, Professor, The University of Tokyo, Japan

Recently, commercial 5G services have been globally deployed and utilized all over the world. At the same time, research and development (R&D) strategies aimed at beyond 5G, i.e. 6G, are already in progress.

In Japan, 5G services started in the spring of 2020, and the discussion body organized by the Ministry of Internal Affairs and Communications (MIC) for defining the R&D directions towards 6G was initiated prior to that, in January 2020.

In June 2020, the MIC Beyond 5G Strategic Board outlined the strategic proposal for 6G R&D.

> *"Emergent strategy proposals for beyond 5G/6G stipulate new key performance indicators."*
>
> Akihiro Nakao

The emergent strategic proposals for beyond 5G/6G stipulate new key performance indicators (KPIs) that should improve target values of existing 5G wireless technologies, such as large bandwidth, low latency, and a large number of connections, by orders of magnitude, as well as new elements such as ultra-low power, security, autonomy and deployability.

These last two new elements are especially interesting ones, as the goal is to realize self-operating networks and to extend telecommunications to include systems that have been considered difficult to deploy, such as High-Altitude Platform Stations (HAPS) and underwater, etc.

In fact, autonomous networking using machine learning (ML) and artificial intelligence (AI) has been actively discussed recently. "Autonomy" is the design and implementation of a communication infrastructure that has the functions of autonomous operations without human intervention. It is "zero touch" networking that involves the construction of optimal network infrastructure that goes beyond wired and wireless.

NTT has recently organized a global forum called IOWN (Innovative Optical and Wireless Network). Although the concept includes many insightful building blocks such as all photonics, data centric, low power, and low latency networking, the primary aim seems to build infrastructure that can be operated automatically by smart algorithms beyond human intelligence and experience.

Advanced communication research in Japan

In Japan's academia, the Institute of Electronics, Information and Communication Engineers (IEICE) is playing a central role in putting forth advanced communication research.

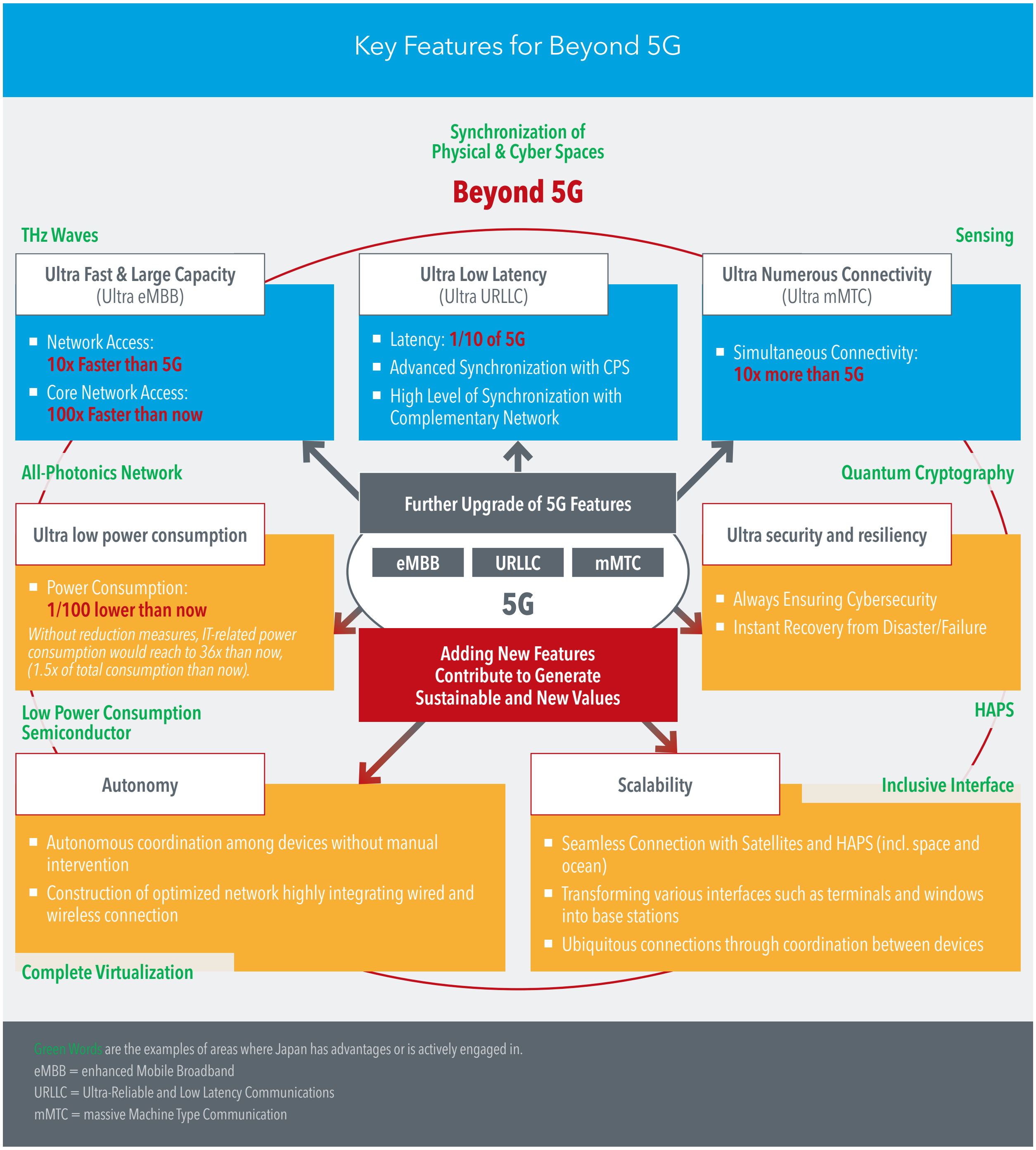
We recognize the need for research on autonomous networking to automate operations and automatically detect and predict failures in information and communication technology infrastructures. This prompted us to organize RISING – a cross-field study group on super-intelligent networks – of which I am Chair.

> *"We recognize the need for research on autonomous networking."*
>
> Akihiro Nakao

The first RISING symposium included the presentation of 111 posters and panel discussions to enlighten many young researchers and students. It attracted great interest from researchers in all fields of wireless and wired communication infrastructure technology.

The ITU and AI and machine learning for 5G

Many stakeholders from Japan, including the University of Tokyo, as well as industry partners have participated in the ITU Focus Group on Machine Learning for Future Networks including 5G (FG-ML5G).


Key Features for Beyond 5G
Synchronization of Physical & Cyber Spaces
Beyond 5G
THz Waves
Sensing
Ultra Fast & Large Capacity (Ultra eMBB)
Network Access: 10x Faster than 5G
Core Network Access: 100x Faster than now
Ultra Low Latency (Ultra URLLC)
Latency: 1/10 of 5G
Advanced Synchronization with CPS
High Level of Synchronization with Complementary Network
Ultra Numerous Connectivity (Ultra mMTC)
Simultaneous Connectivity: 10x more than 5G
All-Photonics Network
Quantum Cryptography
Further Upgrade of 5G Features
eMBB
URLLC
mMTC
5G
Ultra low power consumption
Power Consumption: 1/100 lower than now
Without reduction measures, IT-related power consumption would reach to 36x than now, (1.5x of total consumption than now).
Ultra security and resiliency
Always Ensuring Cybersecurity
Instant Recovery from Disaster/Failure
Adding New Features Contribute to Generate Sustainable and New Values
Low Power Consumption Semiconductor
HAPS
Inclusive Interface
Autonomy
Autonomous coordination among devices without manual intervention
Construction of optimized network highly integrating wired and wireless connection
Scalability
Seamless Connection with Satellites and HAPS (incl. space and ocean)
Transforming various interfaces such as terminals and windows into base stations
Ubiquitous connections through coordination between devices
Complete Virtualization
Green Words are the examples of areas where Japan has advantages or is actively engaged in.
eMBB = enhanced Mobile Broadband
URLLC = Ultra-Reliable and Low Latency Communications
mMTC = massive Machine Type Communication

> *Many stakeholders from Japan, including the University of Tokyo, as well as industry partners have participated in the ITU Focus Group on Machine Learning for Future Networks including 5G.*
>
> Akihiro Nakao

The Ministry of Internal Affairs and Communications, KDDI, NEC, Hitachi, and NICT, are all Japanese entities that are conducting research and development of innovative AI network integration and have also proposed many contributions to ML5G.

Since July 2020, the ITU AI/ML in 5G Challenge has been taking place – a problem-solving competition related to the application of AI/ML in 5G networking.

Countries including China, Brazil, Turkey, Ireland, India and the United States have hosted contests. In our region a contest is being held by a coalition of partners such as TTC and NEC/KDDI, centered on the aforementioned RISING community.

In the regional ITU Challenge in Japan, KDDI and NEC presented research problem sets, respectively. About 20 teams of four persons registered to take part in working out solutions to each of the problem sets.

RISING came up with another problem set regarding wireless communication.

We closed the contest in mid-October 2020, and we will select three winners per problem set, and hopefully send the researchers to present their results at the global conference taking place online at ITU from 15 to 17 December, 2020.

In the light of a series of research and development activities and a great deal of engagement by many stakeholders in the ITU Challenge, we posit that the integration of AI/ML in networking is a promising direction for defining next generation telecommunications, such as Beyond 5G/6G.

We indeed intend to contribute to the research and development, as well as standardization, in the integration of AI/ML into telecommunications in the coming decade. ■

Shutterstock

Realistic simulations in Raymobtime to design the physical layer of AI-based wireless systems

By Aldebaro Klautau, Professor, Federal University of Pará, Brazil; and **Nuria González-Prelcic,** Associate Professor, North Carolina State University, United States

In innovation towards 5G and beyond, there is a clear trend towards learning from experience.

Simulated network environments are helping us to investigate a range of open questions of fundamental importance to AI's contribution to wireless communications.

Complex mathematical models have provided the basis for the design of the current wireless systems. Human intelligence was the main driver of this design. But systems are moving towards wireless networks based on AI, with machine learning algorithms playing a key role.

To derive AI-based solutions, large and diverse datasets are needed to achieve good generalization performance, especially when deep-learning approaches are considered.

"In innovation towards 5G and beyond, there is a clear trend towards learning from experience."

Aldebaro Klautau and Nuria González-Prelcic

Raymobtime datasets and the ITU AI/ML in 5G Challenge

Datasets and sensible benchmarks are essential in guiding research efforts. Raymobtime focuses on the physical (PHY) layer, but datasets for many other applications are now available via the timely ITU AI/ML in 5G Challenge.

There are currently ten Raymobtime datasets available here, totalling 290K communication channels.

A subset of these datasets are supporting contestants in the ITU Challenge on AI/ML in 5G in two challenges:

- **The ML5G-PHY [beam-selection] challenge** assumes a millimeter wave (mmWave) system in a vehicle-to-infrastructure network using an analog MIMO architecture. The machine learning (ML) model to be developed takes input features such as the ones indicated in Figure 2 and outputs the indices of the best pair of beams.
- **The ML5G-PHY [channel estimation] challenge** attacks one of the most difficult problems in the 5G physical layer: acquiring channel information to establish a mmWave link considering a hybrid MIMO architecture. It consists of estimating the frequency-selective channel from a small number of received training pilots.

The two challenges look at key 5G challenges where ML can offer valuable solutions, with other examples including user selection, link adaptation, and joint positioning/communication.

Some applications of deep learning in 5G networks – such as anomaly detection based on performance indicators from routers and other network components – rely on data that is abundant.

The use of machine learning (ML) in designing the physical layer, however, depends on the available channel data.

Building the wireless infrastructure and prototypes to make extensive channel measurement campaigns to store these channel data is infeasible, due to cost and time constraints. Not to mention that some design decisions need to be made prior to building the system.

In this context, realistic simulations deliver essential value.

They do not completely substitute prototypes and measurements, but they support rapid innovation by providing efficient environments for the assessment of new algorithms.

How Raymobtime datasets support the ITU AI and Machine Learning in 5G Challenge

The Raymobtime methodology was conceived to provide realistic simulations of communication channels using **ray**-tracing, taking into account the **mob**ility of transceivers and radio scatterers and their evolution over **time**.

By adopting sensible values for the sampling intervals between distinct scenes (snapshots), ray-tracing leads to simulated communication channels with consistency over time, frequency and space. These channels can enable, for instance, the assessment of ML-based channel-tracking methods for multi-user-multiple-input-multiple-output (MIMO) algorithms.

And Raymobtime datasets are not restricted to communication channels.

Motivated by the increasing usage of sensing information to aid communication systems, additional software is used to compose multimodal datasets.

Figure 1 indicates how Raymobtime pioneered joint simulations of ray-tracing and 3D computer graphics software. The result is "paired" information regarding the communication channel and the respective visual information obtained from a camera and/or the point cloud from a LIDAR (Light Detection and Ranging) sensor. The open-source Blender and BlenSor packages are used to simulate cameras and LIDAR sensors, respectively. Commercial solutions Wireless InSite from Remcom and Altair's Winprop are used for ray-tracing.

Raymobtime also allows the development of data-driven algorithms for specific sites in a city.

To obtain realistic data for outdoor scenarios, Raymobtime uses Cadmapper and OpenStreetMap to create 3D simulations of a site's buildings, streets and other immovable objects. Moving objects such as vehicles, pedestrians and drones are also modeled in 3D, and their positions in different scenes are controlled by the open-source simulator SUMO (Simulator of Urban Mobility). SUMO allows imposing realistic traffic statistics and facilitates studying seasonal aspects, such as fluctuations in number of users and data traffic in specific regions over a day.

The interaction among all software packages is orchestrated via Python, a widely favoured programming language for AI – which can also enable the stages of feature extraction and AI modelling to be integrated, as depicted in Figure 1.

Figure 1 – Block diagram indicating the software involved and how Raymobtime datasets are used for developing ML-based solutions to communication problems

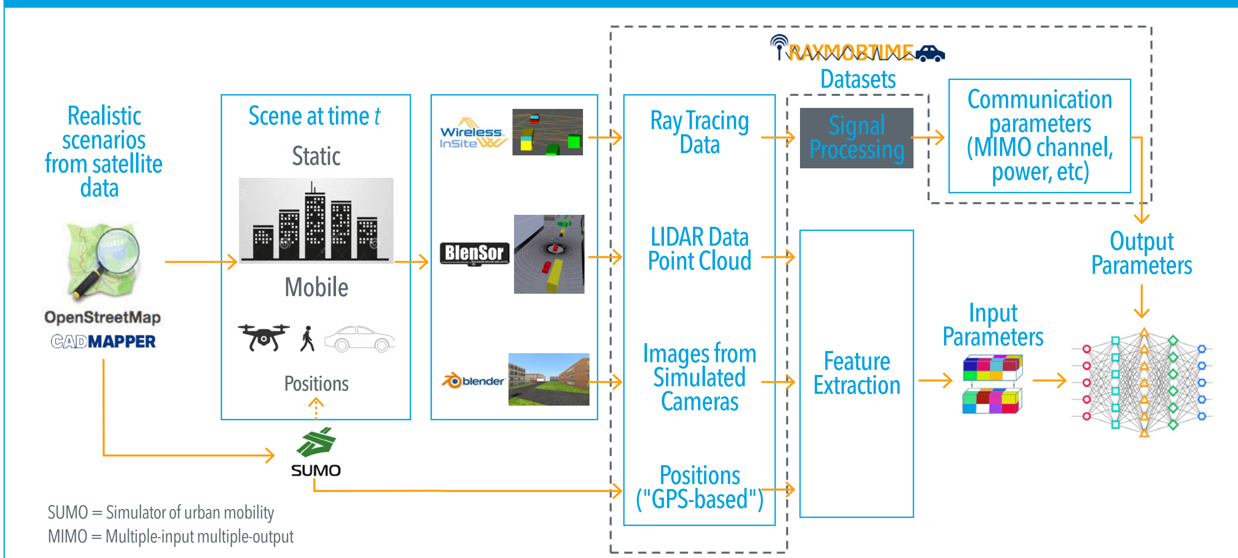
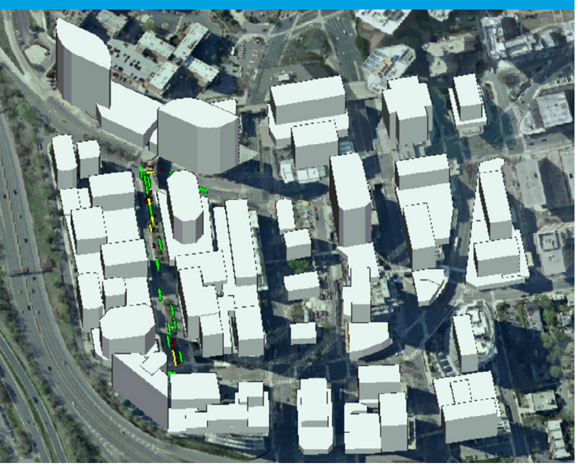


Figure 2 – Example of 3D scenario in which the number of faces was reduced to decrease the ray-tracing execution time. The vehicles are highlighted in vibrant colors in the street under study



Research to improve Raymobtime

Computational cost can become a challenge when repeatedly invoking ray-tracing and other simulation techniques because the time scales differ significantly.

A wireless channel may change in tens of milliseconds even while objects like vehicles barely move.

The time to execute ray-tracing depends on the complexity of the 3D scene, which is related to its total number of faces (simulated renderings of objects).

The number of faces in a 3D scene can be reduced to speed up ray-tracing and create smoother simulations, such as in Figure 2 where faces were reduced to create smoother simulations of the buildings superimposed over satellite imagery of the city scene.

Research continues in optimizing the tradeoff between the speed of ray-tracing and the accuracy of 3D scenes for AI-enabled communications.

Another challenge to improve Raymobtime is the assignment of a material to each face or object that composes a 3D scene.

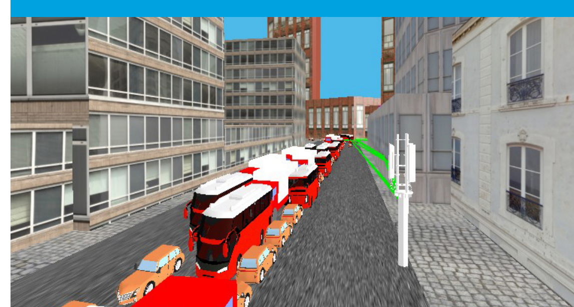
Each material has electromagnetic properties that impact ray-tracing, and consequently, the communication channel. Raymobtime evolved from simpler simulations with two materials, as shown in Figure 3.

Version 2 of Raymobtime will incorporate automatic assignment of electromagnetic properties to materials. It will also be based on improved 3D engines used to generate datasets in a larger number of sites to enable studies in the context of transfer learning.

This blending of communication systems and virtual reality will also be leveraged by research in hybrid channel models, combining ray-tracing and statistical channel models. ■

Figure 3 – Evolution of realism in describing 3D scenarios for wireless simulations

Static ray-tracing with rectangular shapes



Raymobtime

Raymobtime v2.0

Building reliability and trust with network simulators and standards

By Francesc Wilhelmi, Post-Doctoral Researcher, Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Spain

Network simulators can play a key part in building reliability and trust in machine learning (ML). But for simulators to deliver this value, innovation towards ML-enabled 5G and 6G networks must include the integration of interoperable testing environments.

Truly autonomous, secure, and reliable ML-aware communication systems can be achieved in the near future. And network simulators will contribute to this achievement. But the successful adoption of simulators will depend on the definition and standardization of interoperable components.

The integration of simulators into autonomous ML-aware networks is now becoming possible with the softwarization and virtualization of networking functions.

> *"Network simulators can play a key part in building reliability and trust in machine learning."*
>
> Francesc Wilhelmi

As for network simulators, we find a plethora of proprietary and open-source tools (e.g. ns-3, OMNET++, OPNET) for characterizing multiple types of scenarios, technologies, and networking functionalities.

But bringing all these tools together forms a very significant challenge – one that can only be addressed through the definition and implementation of standardized interfaces.

Interoperability in future ML-aware networks will enable entities in different domains such as network functionality, machine learning repositories, and data providers, to interact seamlessly in addressing optimization challenges.

Cross-domain interoperability will be one of the main enablers of fully autonomous future networks, and the conditions for this interoperability must be established soon.

Why ML5G?

The increasing application of machine learning in communication networks is motivated by large amounts of **unexploited data** and **inherent complexity** of novel use cases such as Vehicle to Everything (V2X) communications, massive Machine Type Communications (mMTC), and extended reality and high-quality video communications.

These use cases vary widely in mobility requirements, the numbers of implicated devices, and bandwidth and latency requirements. In complex 5G networks, **substantial performance gains** may result from machine learning's ability to learn **complex patterns** and adapt to multiple contexts and domains.

Neural networks, for example, are fast becoming very popular in signal processing thanks to their ability to characterize unknown channel models.

Overcoming barriers to reliability and trust with the help of simulators

Machine learning mechanisms can produce non-linear outputs, prediction functions for example, leading to questions around the trustworthiness of the outputs produced by these "**black boxes**".

The perhaps most notorious black boxes are deep learning models for problems with high dimensional spaces, where the accuracy of a model is typically tied to its complexity. The greater the complexity of the dataset, the higher the number of neurons and hidden layers required by the deep learning model, thereby hampering interpretability and explainability.

Moreover, depending on the characteristics of the use case, training data may be scarce, noisy, and even non-stationary, thus compromising the **reliability** of ML models and questioning their sustainability. Highly complex models can also require large computational resources that are not always available.

Consider V2X communications, where security is critical, and where connected vehicles and mobile devices will create complex radio frequency environments. Machine learning can help to address such complexity, but shortfalls in data and computational resources can lead to model misbehaviour and degrade the key performance indicators informing machine learning's adoption in networking.

Research towards explainable AI shows promise in building trust in the outputs of neural networks and other complex AI methods, but in the near-term network simulators can provide powerful tools to enhance trust in machine learning. These simulators can also be integrated into ML-aware communication systems.

Network simulators are a **cost-effective** tool to reproduce the behaviour of communications systems, ranging from communication protocols to physical phenomena related to signal propagation.

Network simulators can be integrated into ML-based networks to provide the following functionalities:

- **Validate** the output of machine learning models before being applied to an operating network.
- **Generate** synthetic data for training machine learning models, which can be used to address the lack of data, or to extend training data sets.
- **Train** machine learning models in a simulation domain, which is particularly useful to avoid the effects of exploration in online learning.
- Act as **experts** to assist the operation of machine learning models, thus providing advice in certain situations (e.g., initialization, narrow exploration, break ties).

To illustrate the potential simulators in future communications systems, consider the example of transmitter-receiver implementations based on neural networks.

Aiming to improve the accuracy of this kind of solution, network simulators can generate synthetic data characterizing human behaviour and thus enrich the datasets used for training. This is motivating the design of novel standardized datasets and tools, including network simulators, and the resulting synthetic datasets have played an important part in the ITU AI/ Machine Learning in 5G Challenge.

> *"ITU standards are providing a toolkit to introduce machine learning methods into 5G networks."*
>
> Francesc Wilhelmi

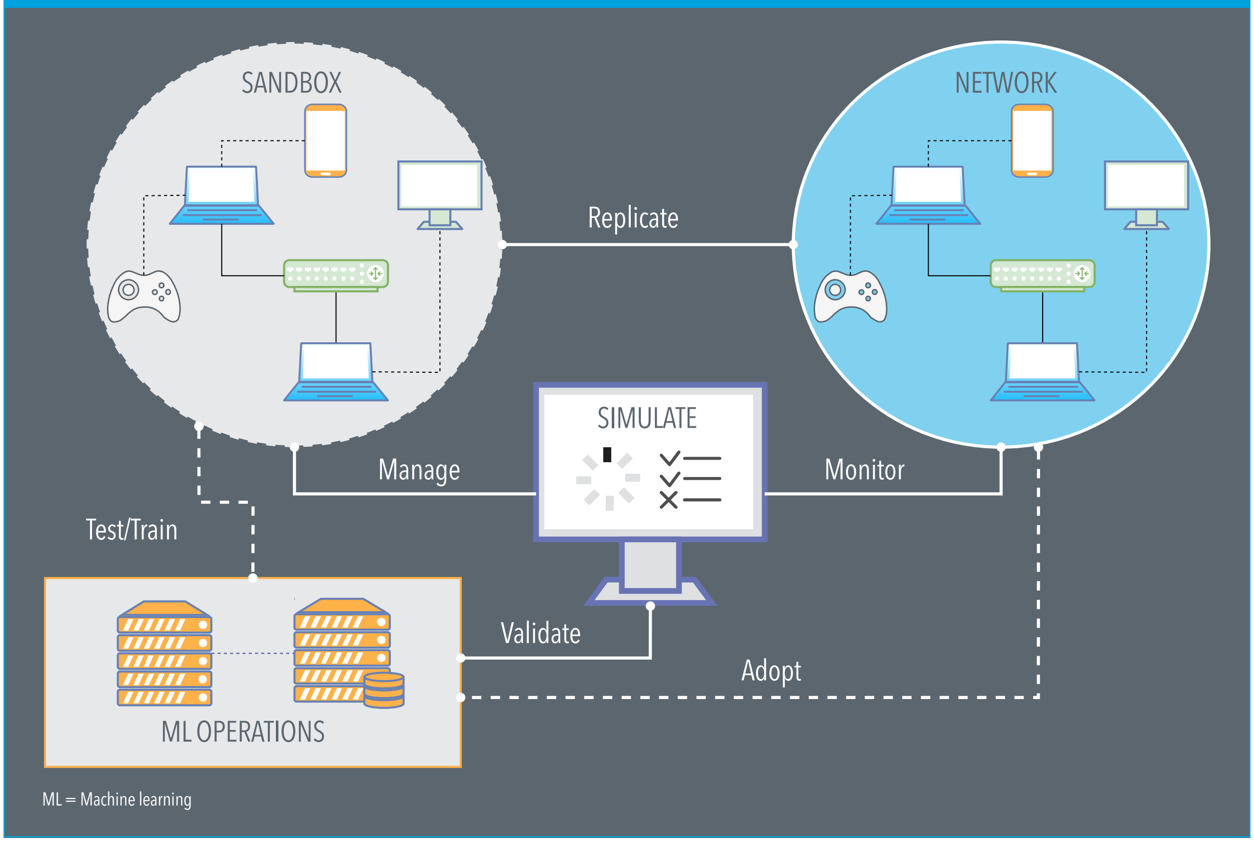
ITU's flexible ML-aware architecture and proposed ML sandbox

ITU standards are providing a toolkit to introduce machine learning methods into 5G networks.

Particularly relevant is ITU–T Y.3172, which defines an architectural framework for the flexible integration of machine learning into networks.

This ITU–T Y.3172 ML-aware architecture proposes the ML sandbox as a solution for enhancing confidence in machine learning application.

Proposed machine learning sandbox



An ML sandbox enables the testing, training, and evaluation of machine learning models in a safe environment before their application to operating networks.

Network simulators are an important part of the ML sandbox, effectively and flexibly reproducing different networking behaviours and scenarios.

Ongoing ITU studies are defining the requirements, architecture, and interfaces for the proposed ML sandbox.


Federal University of Technology, Minna

Research projects in Nigeria advancing education and speech recognition

By James Agajo, Associate Professor and Head of WINEST Research Group, Department of Computer Engineering, **Abdullahi Sani Shuaibu** and **Blessed Guda,** Students, Federal University of Technology, Minna, Nigeria

As part of our participation in the ITU Focus Group on Machine Learning for Future Networks including 5G (FG-ML5G), in March 2019, our WINEST (Wireless Networks and Embedded Systems Technologies) Research Group launched a study on "use cases and solutions for migrating to IMT-2020/5G networks in emerging markets".

Our focus was to determine how machine learning could help emerging markets to leapfrog technology generations to take advantage of emerging and future networks while optimizing energy consumption, network coverage, and communication overheads.

Improving education in Africa

This study led us to propose the "AI-Based Classroom" project, designed to improve education for young pupils in Africa.

Students are driving the project, under the guidance of James Agajo, Head of WINEST Research Group at the Federal University of Technology, Minna, Nigeria.

> *This study led us to propose the 'AI-Based Classroom' project, designed to improve education for young pupils in Africa.*
>
> James Agajo, Abdullahi Sani Shuaibu, and Blessed Guda

With AI-based natural language processing (NLP), classroom conversations between pupils and teachers are processed at the network edge to extract keywords while maintaining speaker anonymity.

These keywords are transmitted to a trained classifier in the central server, which is able to recommend captivating media content, providing students with intuitive examples, supporting a teacher's explanations. The media content is then shared on a digital display in the classroom.

The system is designed to augment the efforts of elementary school teachers rather than attempting to replace them.

An efficient speech recognition library is a critical prerequisite for the development of an AI-based classroom.

This proved very difficult to find.

Automatic speech recognition for Africa

We were in search of a speech recognition library which was able to function locally, meet users' privacy concerns, and was freely available. In view of the extraordinary number of languages spoken across Nigeria, and Africa at large, we also needed a library able to perform well in processing the English language accented in many different ways.

We evaluated many software libraries, but none of them succeeded in meeting all of these requirements.

This led to the launch of a new WINEST Research Group project in February 2020 to develop a new speech recognition framework able to meet the unique requirements of the AI-Based Classroom project.

The project evolved from the discussions sparked by our presentation of the AI-Based Classroom project at the 7th Regional Workshop on "Standardization of future networks towards building a better-connected Africa" in Abuja, Nigeria, 3–4 February 2020, convened by ITU's standardization expert group for future networks and cloud computing, ITU Telecommunication Standardization Sector (ITU–T) Study Group 13.

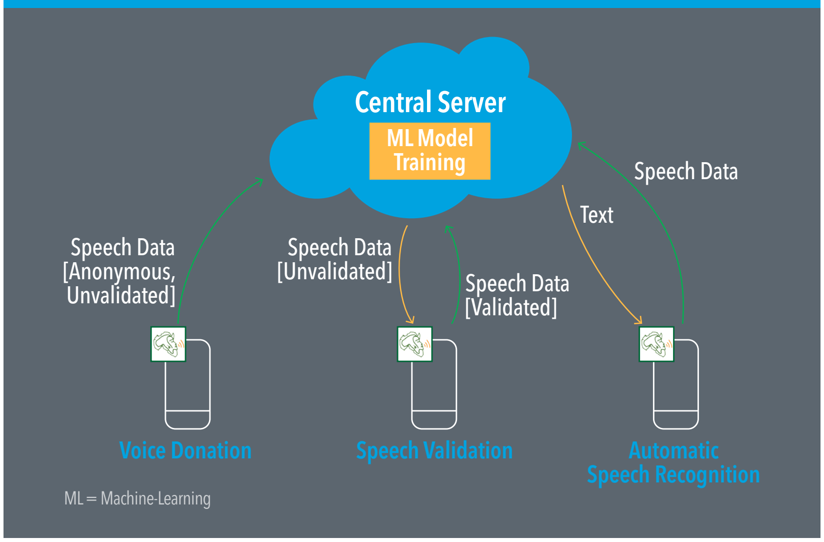
The expert feedback provided by the Abuja workshop motivated our launch of a pilot project in Nigeria to develop an African automatic speech recognition (ASR) system.

We are collecting speech data and developing the ASR engine to deliver a prototype able to guide the development of a system ready for market deployment.

We have developed the "Wazobia" mobile application, to support the necessary data collection, where Nigerian "voice donors" read displayed text aloud and donate the recording – anonymously.

"Wazobia" is an amalgam of three words meaning "come" in Yoruba (wa), Hausa (zo) and Igbo (bia), Nigeria's three largest linguistics groups.

Architecture of the African automatic speech recognition project



The speech data is stored on our server as "unvalidated" by default, pending the crowd-sourced validation of this data by volunteers via the Wazobia mobile app. This validation results in Boolean evaluations of the accuracy of the ASR engine's transcriptions of recorded speech.

To date, the project has collected over three hours of speech corpus from over 170 voice donors.

The African ASR system development phase includes data pre-processing, training and software design. The project uses the Wav2letter++ ASR toolkit and looks to Facebook's AI research article as a reference implementation.

We are progressing with the segmentation and pre-processing of the collected data for supervised and semi-supervised machine learning settings, but thus far the African ASR project only accepts English as an input language.

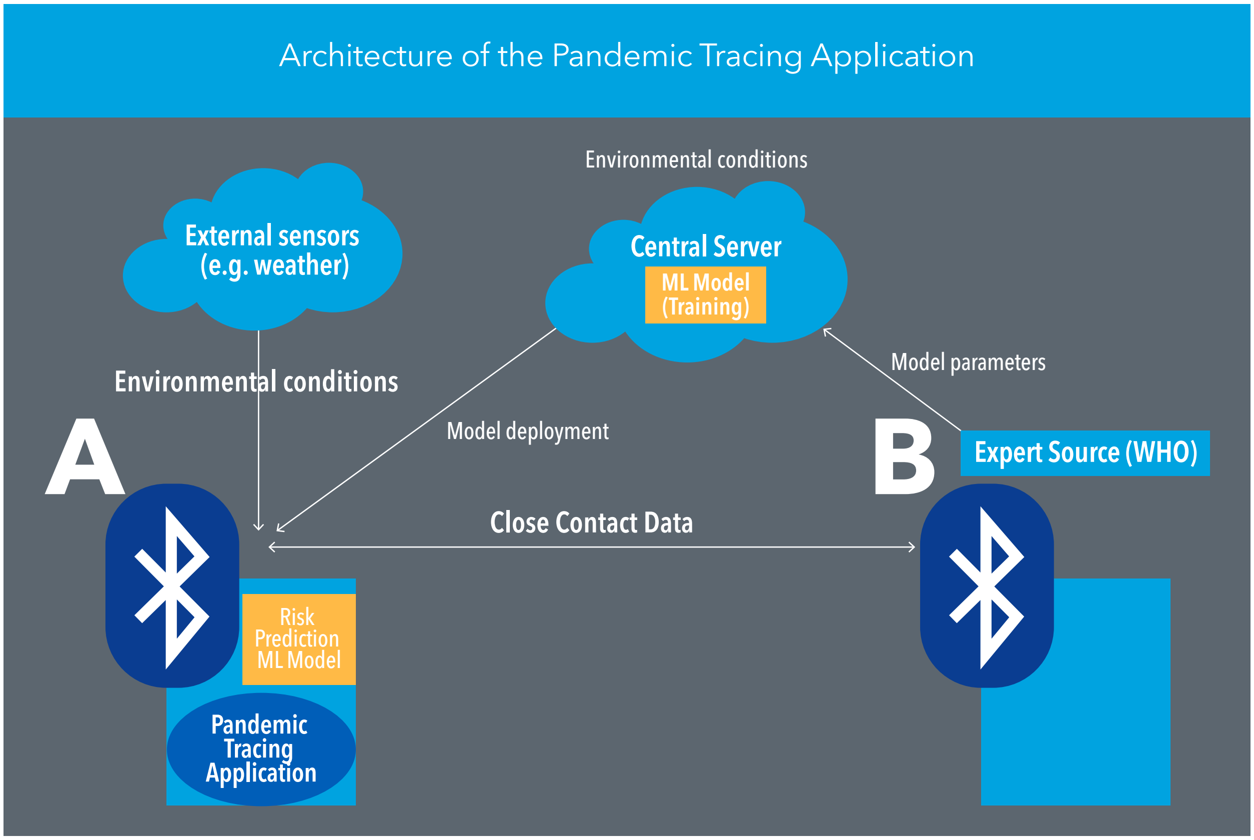
We aim to introduce African languages as inputs as the ASR project develops, and we plan to stimulate this key avenue of innovation by submitting our speech corpus of African languages to future ITU challenges on AI and Machine Learning in 5G and beyond.

AI and machine learning to help Africa manage pandemics

In future ITU AI/machine learning in 5G challenges we also plan to propose a new Bluetooth®-enabled contact-tracing application supported by machine learning.

This pandemic tracing application (PTA) project aims to build exposure-risk prediction models trained from anonymized user data [see table].

Data collection for pandemic tracing application contact detection
Straight line distance between user equipment
Bluetooth signal strength
User equipment model
Operating system version
Indoor/outdoor (based on ambient light)
Radio-frequency interference (wireless local area network)



The proposed PTA deployment scenario would incorporate data collected from users with Bluetooth® discoverable, not requiring all users to install the application. However, a more detailed picture of the environment can be achieved by incorporating data from mobile devices' gyroscopes and accelerometers when two Bluetooth®-connected devices both have the PTA installed.

The development of the proposed PTA will follow these guiding principles:

1. The contact tracing will be generic with configurable parameters to accommodate future pandemics.
2. It will reuse relevant features of existing frameworks but customize these features for application in Africa.
3. It will include privacy-preserving mechanisms by design.
4. The application will determine the appropriate scope of data sharing beyond a user's device-based user-indicated preferences with respect to privacy.

The training data will be specific to pandemics, as recommended by the World Health Organization (WHO) and other health authorities. The resulting exposure-risk prediction models will also be specific to pandemics (trained and deployed from a central server).

We are now focused on collecting the required data and we plan to submit this data to future ITU challenges on AI and Machine Learning in 5G and beyond. ■

Shutterstock

Why we need new partnerships for new data

By Ignacio Rodriguez Larrad, PostDoc, Wireless Communication Networks, Aalborg University, Denmark

The ITU AI/Machine Learning (ML) in 5G Challenge attracted over 20 problem statements proposed by different industry and academic organizations. Participants from all over the world are competing to find the neural network-based algorithms that lead to the optimal solution.

The available problem statements are organized into different tracks related to diverse themes such as network, security, operators, or vertical markets.

The number of statements is certain to increase significantly in future editions of the ITU AI/ML in 5G Challenge. The application of AI and machine learning to networks is gaining strong momentum in industry and academia, and the ITU AI/ML in 5G Challenge will engage more and more students, researchers and engineering professionals.

> *"The Grand Challenge Finale in December is an exciting prospect."*
>
> Ignacio Rodriguez Larrad

The Grand Challenge Finale in December is an exciting prospect – sure to shed new light on how AI and machine learning will help us to optimize networking.

New partnerships to expand access to data

Access to high-quality data remains a key challenge facing innovation in AI and machine learning.

Synthetic datasets built with simulations have played an important part in the ITU AI/ML in 5G Challenge in view of the difficulties associated with access to data from operating networks. And some of the problem statements were very relevant and well-formulated but lacked any source of data – participants needed to find and provide their own datasets as part of their solutions.

While most of the problem statements were open to international participants, some were restricted to national competitions. However, one has to understand that the data belongs to the institution that provides it and is therefore subject to internal/national export regulations and national general data protection regulations.

Regulation, business and the nature of data all factor considerably into data availability, and data availability factors considerably into the success of a global competition like the ITU AI/ML in 5G Challenge.

This also highlights the importance of the ITU AI/ML in 5G Challenge in encouraging data availability.

ITU is working to attract and engage new institutions who are able to share problem statements and data with the international community.

To achieve accurate functional ML/AI algorithms, high-quality input data is essential.

Extensive datasets from operating networks would be ideal inputs to AI/ML algorithms. However, at present it is very difficult to obtain such collections of data from commercial network operators. It requires some effort to collect specific data and treat and anonymize it. Operators need to protect the privacy of their customers, and they are also hesitant to share data containing business-critical information about the operational status or performance of their networks.

> *"To achieve accurate functional ML/AI algorithms, high-quality input data is essential."*
>
> Ignacio Rodriguez Larrad

Research networks – real-world operational networks managed by public research institutions – will continue to play a key role in the collection of traces for ML/AI processing in the coming years.

Collecting data from research networks also requires some effort, and typically also the treatment and anonymization of the collected data, but research networks come with fewer complexities when it comes to sharing datasets with other institutions.

Research institutions from all over the world are part of the open data movement, with research communities advocating for open access to experimental research data and associated scientific findings and articles.

But research networks are not easy to build.

While it is quite common to see research networks operating in unlicensed spectrum such as Wi-Fi networks or wireless Internet of Things (IoT) networks such as long range (LoRa), it is rare to see research networks in 4G or 5G cellular systems.

These networks come at a much higher cost and research institutes will depend on the sponsorship and close collaboration of vendors and operators.

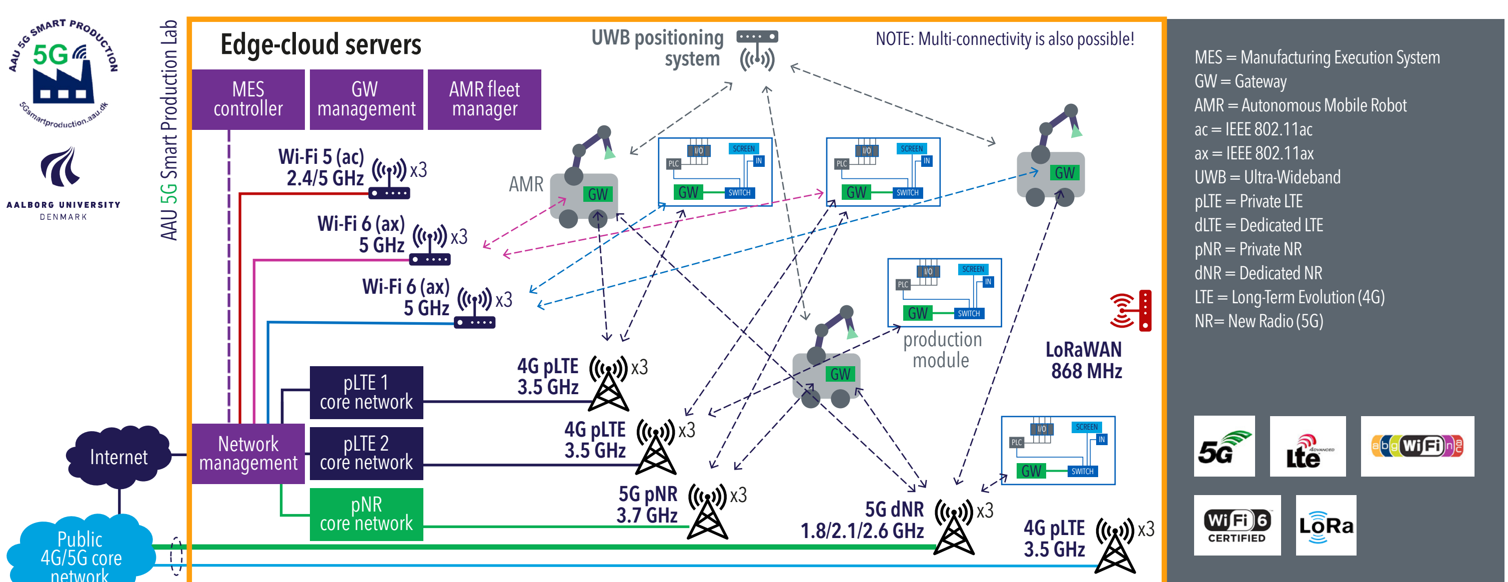
How Aalborg University collaborates with industry on its AAU 5G Smart Production Lab

As an example of a research network with big ML/AI potential established at a public research institute in collaboration with vendors and operators, Aalborg University (AAU), Denmark, hosts one of the most advanced Industry 4.0 wireless playgrounds in Europe.

The recently inaugurated AAU 5G Smart Production Lab is a small factory industrial research lab with access to a wide range of operational industrial-grade manufacturing and production equipment from different vendors, including production line modules, robotic arms, and autonomous mobile robots.

In collaboration with Nokia Bell Labs and the telecommunications operator Telenor Denmark, the research lab is equipped with 4G (two private networks), 5G (one private, standalone network), and Wi-Fi 6 (two networks). This makes it possible for AAU researchers and industrial partners to work together in the integration and testing of advanced Industrial IoT systems for the factories of the future.

AAU 5G Smart Production Lab



Source: Aalborg University, Denmark, 2020.

The lab is also equipped with a research LoRa network sponsored by Cibicom Denmark, and an ultra-wideband positioning system.

All wireless research networks in the lab are interconnected via a central dedicated network management interface which allows not only for the control and configuration of the different networks but also the monitoring and collection of network data traces in dedicated edge-cloud servers.

Having the possibility of recording unrestricted network information, simultaneously with measurements from the different user devices, opens new research possibilities for AAU and the international research community working on AI and machine learning applications in networking.

Let's expand our work together

Research networks built together by research communities and industry can generate extensive and relevant sets of multi-domain and multi-technology performance data. This data could be very useful for the design and optimization of advanced wireless solutions targeting the different communication requirements of real-world operational industrial IoT use cases. ■

Industry 4.0

Source: techstartups.com, 2020.

Recently inaugurated AAU 5G Smart Production Lab

Source: Aalborg University, Denmark, 2020.

Equipment in the AAU 5G Smart Production Lab

Source: Aalborg University, Denmark, 2020.


Shutterstock

Machine learning function orchestration for future generation communication networks

By Shagufta Henna, Lecturer in Computing, Letterkenny Institute of Technology, Ireland

Fifth generation and beyond network operators are interested in exploiting the potentials of machine learning (ML) to solve intractable problems using a large amount of data.

However, these network operators struggle to integrate ML in their networks and often rely on data scientists to create an ML pipeline, i.e., from data collection to model deployment. An ML pipeline, if not managed and orchestrated appropriately, is subject to bottleneck. Further, lack of standardized ML orchestration mechanisms can result in highly complex and expensive ML pipelines.

“Fifth generation and beyond network operators are interested in exploiting the potentials of machine learning to solve intractable problems using a large amount of data.”

Shagufta Henna

Other challenges in the context of future generation networks include ML model updating, model optimization, ML pipeline chaining, ML pipeline performance monitoring, evaluation, pipeline splitting, policy-based ML pipeline deployment, and management and coordination of multiple ML pipeline instances across the network.

These challenging characteristics call for the requirements of big-data-enabled tools that can collect, store, and pre-process data to train ML models. Currently, there are efforts to address some of the above issues by the big players including Uber, Netflix, Google, Facebook, and Airbnb, with the help of custom ML orchestration platforms.

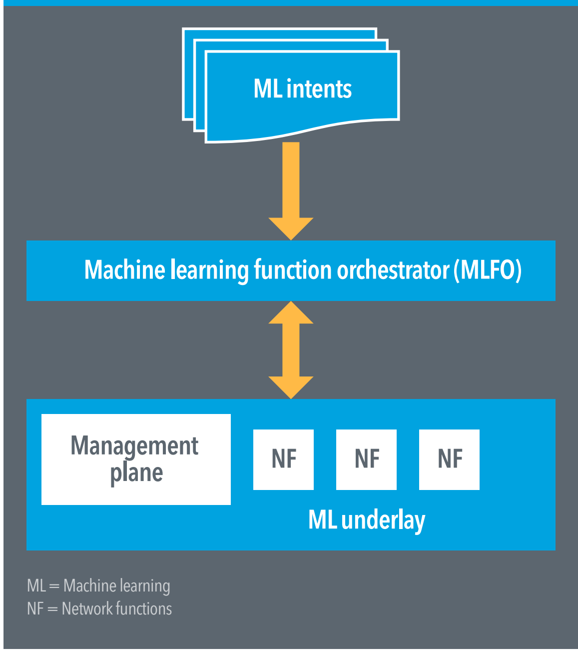
However, these solutions aim at in-house ML pipeline management, and do not address the requirements of diverse 5G and beyond use cases. In-house ML orchestration solutions require significant investments without the actual realization of their benefits within an industry setting.

Overcoming the challenges with machine learning function orchestrator

The main objective of machine learning function orchestrator (MLFO), as per the draft ITU Telecommunication Standardization Sector (ITU–T) Recommendation: "Requirements, architecture and design for machine learning function orchestrator", is to overcome the aforementioned challenges of integration, orchestration, and management of pipeline nodes and dependencies in an ML pipeline, while reducing operational costs. Machine learning function orchestrator offers a unified architecture to facilitate the orchestration of end-to-end ML workflows, i.e. data collection, pre-processing, training, model inference, model optimization, and model deployment. It can monitor and evaluate ML pipeline instances for performance optimization.

It aims to hide the underlying complexities of orchestrating ML pipeline nodes by providing an abstraction to the users and developers with the help of high-level application programming interfaces (APIs) as illustrated in the figure. Its architecture provides flexibility, reusability, and extension of the ML pipeline to accommodate the rapid pace and development in ML pipeline nodes.

Machine learning function orchestrator



Future aims will be to consider its distributed implementation with minimum complexity and overhead. Further, it will be interesting to test the machine learning function orchestrator-specific concepts in various 5G and beyond use cases.

Sponsorship opportunities for 2021

What are the benefits of sponsoring next year's ITU AI/Machine Learning in 5G Challenge?

Brand visibility for an entire year

- on the Challenge website
- at the successful ITU AI ML in 5G webinar series
- at the Grand Challenge Finale

Program opportunities

Media opportunities

Mentorship

Tailored hands-on workshops

If you are interested in sponsoring the 2021 AI/Machine Learning in 5G Challenge, please contact:
ai5gchallenge@itu.int

Stay current // // Stay informed

ITU

Sign up to:

// Key ICT trends worldwide // Insights from ICT thought leaders //
// The latest on ITU events and initiatives //

ITU News Weekly

// *Every Tuesday* //

ITU News Regular articles

// *Regular blogs* //

ITU News Magazine

// *Six editions per year* //

ITU Podcasts

// *Follow podcasts* //

Press releases and media advisories

// *Receive latest news* //

Join ITU's online communities on your favorite channel