The Role of Cognitive Autonomy in "5G and Beyond" Communication Networks

Christian Mannweiler, Henning Sanneck, Stephen Mwanje

Architecture & Management Group
Nokia Standards
München

Tutorial

The Role of Cognitive Autonomy in “5G and Beyond” Communication Networks

Outline

- Cognitive Autonomy
  - Definitions
  - Classic AI (for Reasoning) vs. Machine Learning (for Cognition)
  - Taxonomy
- Case Studies (Network Automation in 4G/5G RAN)
  - Anomaly Detection
  - Predictive Location-Aware Network Automation
  - (AI in Campus Networks)
- AI enabling areas for Cognitive Autonomy
- Outlook: Towards (beyond 5G) Cognitive Autonomous Networks (RAN)
- Takeaways

The Need for Cognitive Autonomy in Communication Networks

Definitions

Autonomous - able to act on its own, without dictation or rules from anyone else

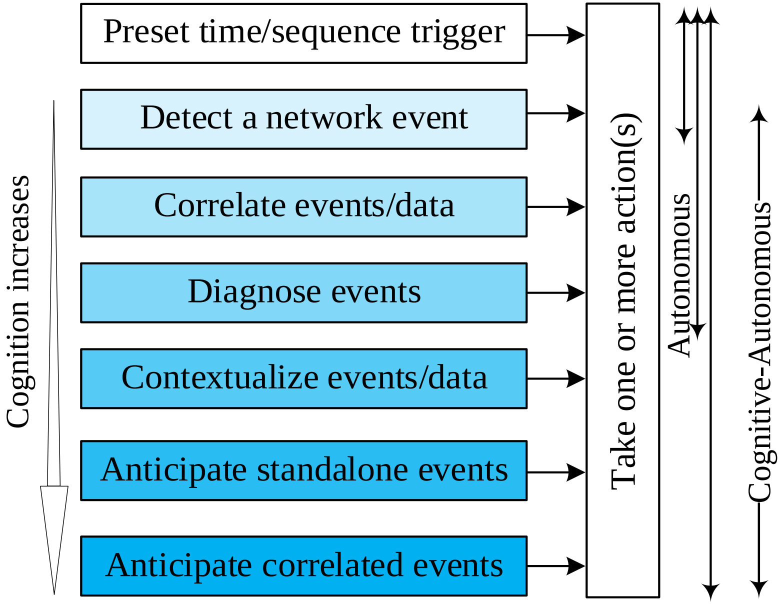
- not necessarily able to reason based on its environment or even smart enough to make the best static decisions

Cognitive - able to reason and formulate recommendations for subsequent behaviour

- may however require the operator's approval

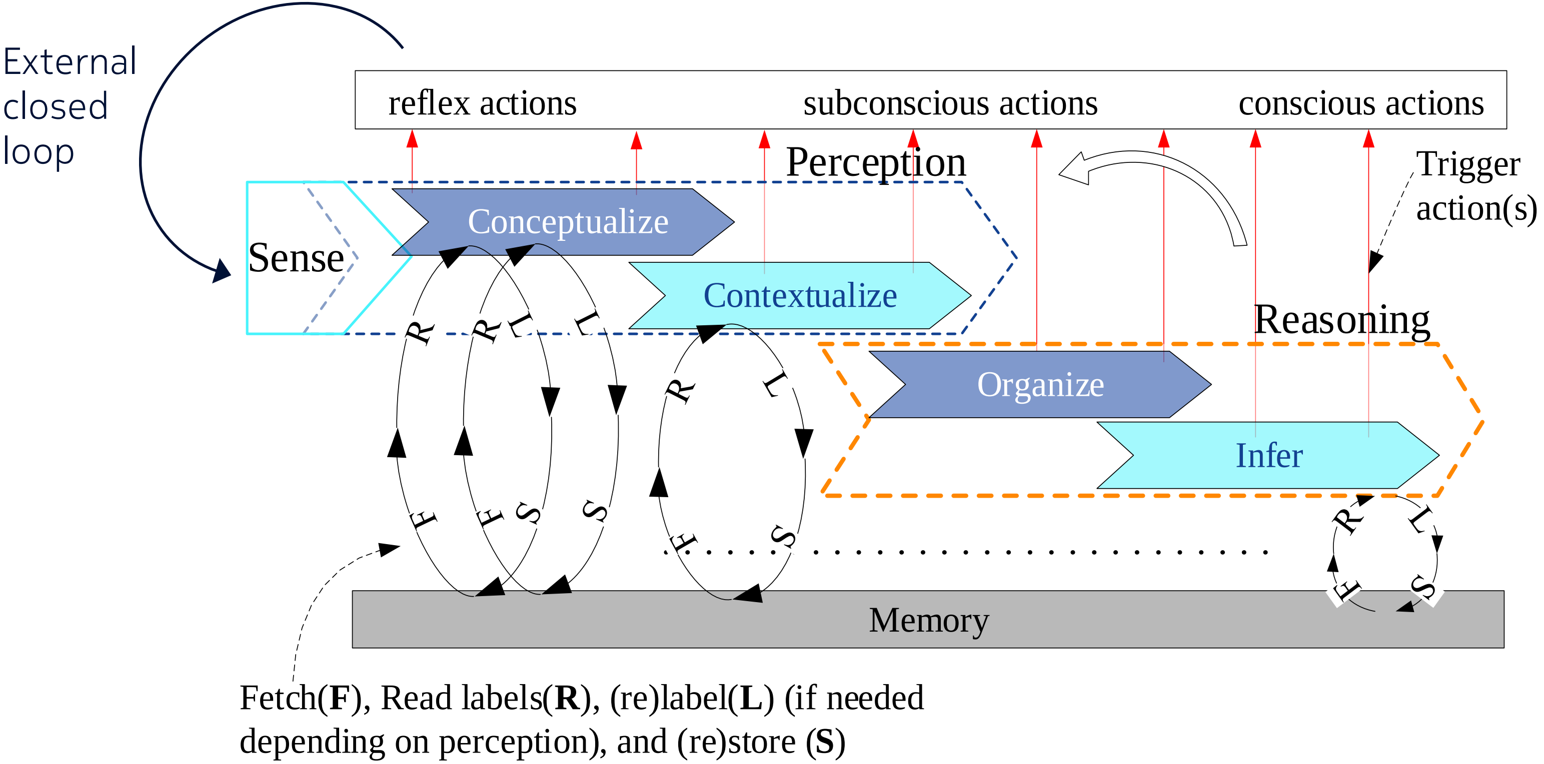
Self-Organizing – Achieve steady state without external control

- automates the selection and execution of actions
- interprets events & context to determine the cause-effect relations.



Modelling Cognitive Decision Making

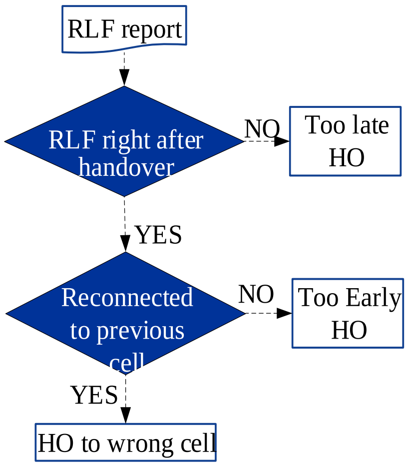
A Perception-Reasoning pipeline



  Public NOKIA

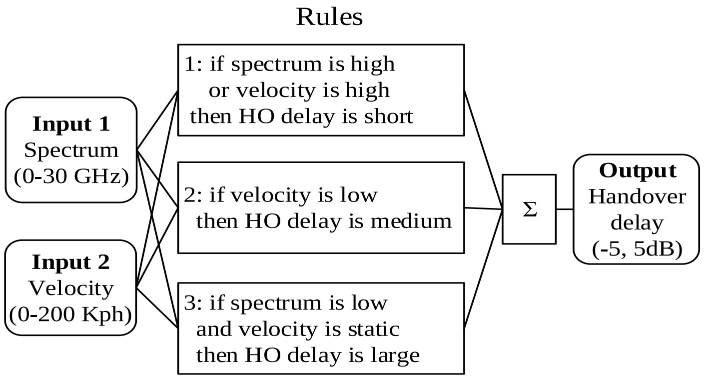
"Classic" AI: Tools for Autonomous Reasoning
From discrete to probabilistic rules

Expert system to diagnose RLFs



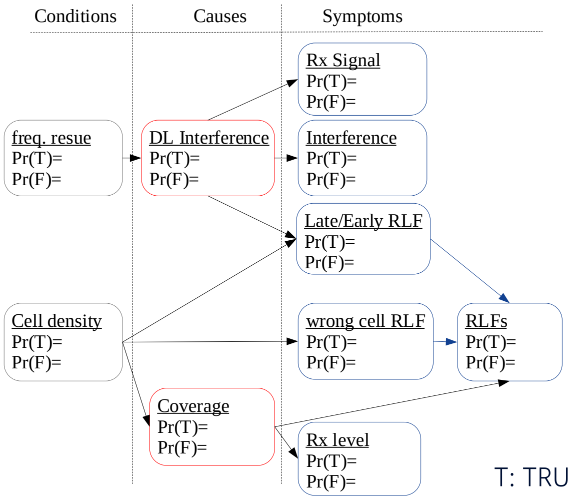
Discrete rules on discrete events

Fuzzy Inference analysis for Handover



Rules on continuous valued logic

Bayesian network on radio events



T: TRUE

F: FALSE

Conditional probabilities define paths through the graph

Rules focus on decisions, perception is “built-in” at design/implementation

Cognition in Communication Networks
Static model vs. learning from data

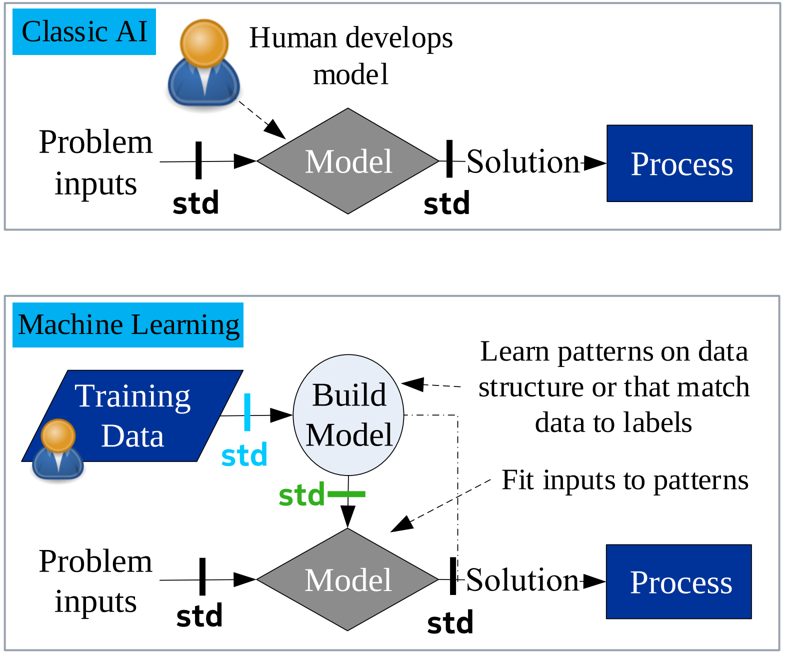
Classic AI: Model is available

- Static – fixed rule set on discrete or continuous valued logic
- Adaptive to deployment – by probabilities

Learning: Happens if performance at task T as measured by performance measure P, improves with experience E [1]

- Parameterise a model (θ) with loss function J → P=J(θ)
- Iteratively update model with experience E from data

$$\theta := \theta - \alpha \frac{\partial}{\partial \theta} J(\theta)$$



Expanded & New interfaces for standardization – data & model transfer

[1] adapted from Tom Mitchell (1997))

Cognition vs. Autonomy Taxonomy

Autonomy

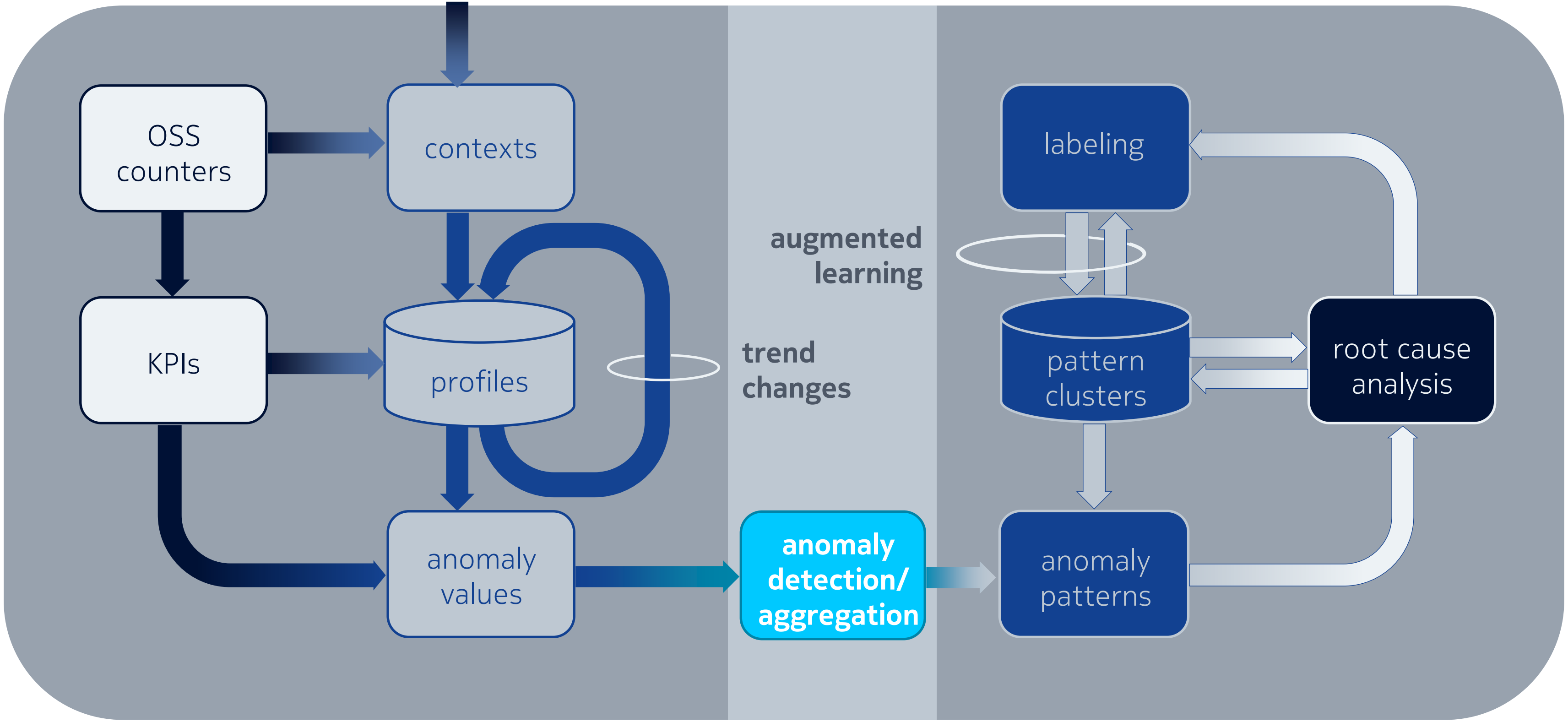
Cognition	Manual	Assisted	Partially automated	Automated	Partially autonomous	Autonomous
	Machine: None Human execution & supervision	**Machine: assisted execution & supervision Human: partial execution**	**Machine: partial execution; Human: supervision via policy**	**Machine: execution; Human: supervision via policy**	**Machine: execution & partial supervision Human: policy & intent**	**Machine: execution & supervision; Human: intent-only**
+ Anticipate correlated events	Human experience		+ Machine prediction	+ Machine auto	Machine	+ Machine reasoning
+ Anticipate individual events			+ Automatic pro-action	pro-action selection	ediction of new policies	
+ Context-ualize		Machine2huma visualization	**e profiling** atic re-act	**+ Machine automatic re-acti selection**	**chine learning of new policies**	**+ Machine reasoning + General learning; +Trustworthiness**
+ Diagnose events	Human diagnosis	Machine2human selective exposure	**+ Machine mapping to causes (rules)** +Automatic re-act	**+ Human labelling of causes identified by machine**	**+ Model-free (Reinf- Learning) + Transfer learning**	**+ Machine explanation**
+ Correlate events	Human correlation	Machine correlation	+Automatic re-action	(n.a. – due to limited - scalability of automation - feasibility of machine supervision in a system with low cognitive capability)		
Detect an event	**Human** detection	**Machine** detection	**+Automatic re-action**			

KICK Case Study

PLANAR Case Study

Anomaly Detection Case Study

Anomaly Detection and Diagnosis
Overview

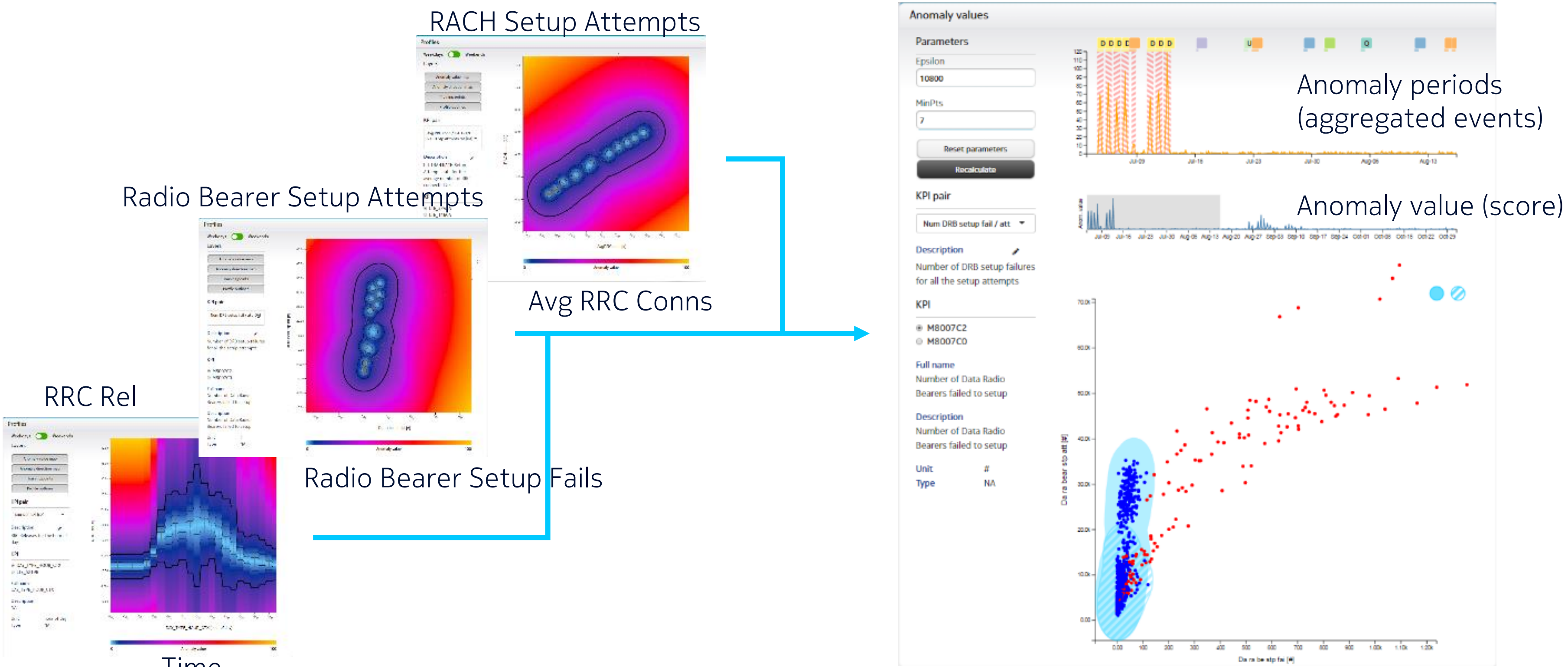


H. Sanneck, Janne Ali-Tolppa, Levente Bodrog, Szilard Kocsis, Benedek Schultz,
Addressing 5G NM Challenges with ML, Industry keynote, 23. ITG Fachtagung Mobilkommunikation, 2018

NOKIA

Anomaly Detection Procedure

Context-aware learning of normal states, measuring anomalousness in diurnal behavior and correlation



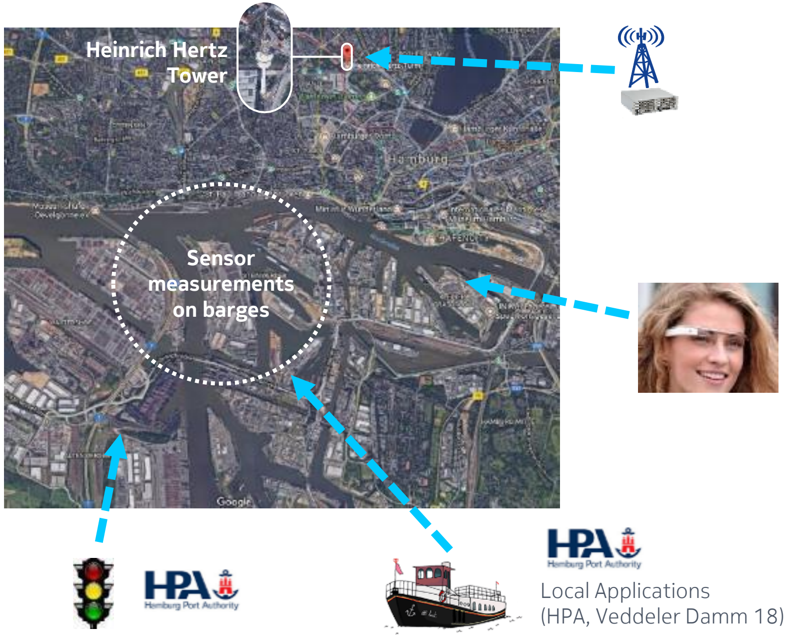
PLANAR Case Study

Predictive Location-Aware Network Automation for Radio
The 5G network slicing testbed at Hamburg Seaport

A live testbed demonstrating 5G slicing at the Hamburg Harbor:

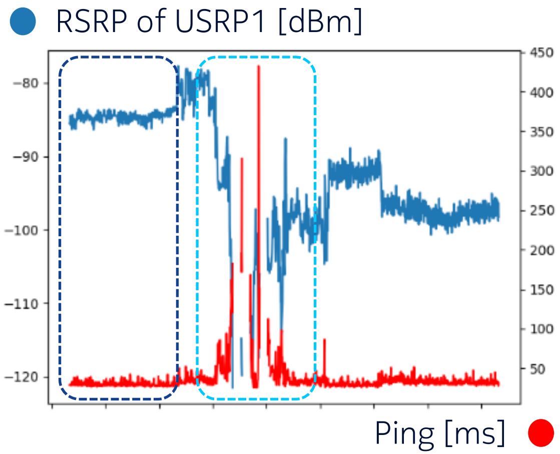
- Three slices
    - eMBB: Local applications in the harbor
    - URLLC: Traffic light control
    - **IoT: Emission sensor readings from barges**

- Data collection
    - **Slice-specific BTS KPIs**: PRB usage, throughput, latency etc.
    - **UE measurements** from up to three ships including position (by GPS), RSRP, RSRQ, ping etc.
    - Collected for 6 months every 5 seconds **~3M records**



Predictive Location-Aware Network Automation for Radio
Problem statement

- IoT requires high reliability
- In certain areas of the testbed, coverage and mobility issues are observed (for devices served by the IoT slice)
  - Shadowing effects and/or
  - Long distances from the base station
- Reliable service must be guaranteed, but without overprovisioning of resources or compromising the performance of the other slices

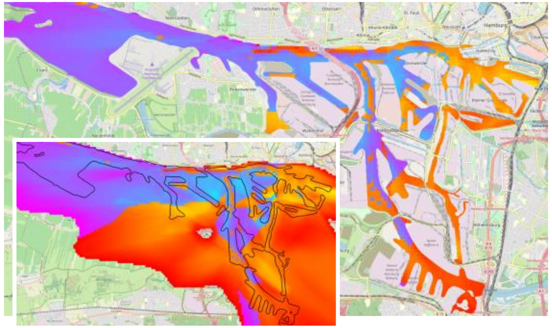


Predictive Location-Aware Network Automation for Radio

Prediction of Mobility and QoS/RSRP

Radio Propagation Map:

- Created based on UE measurements (reported GPS position, RSRP)
- Using an FNN

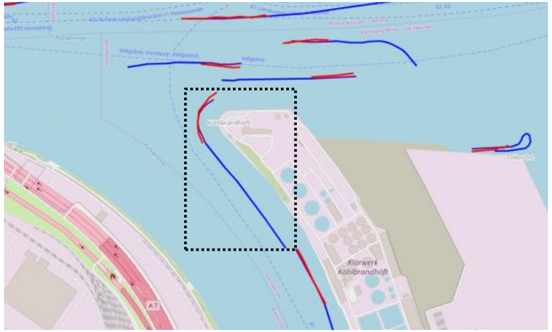


RSRP, serving cell: Cell_1

RSRP, serving cell: Cell_2

Mobility Pattern Prediction (MPP):

- Positions reported by the barges
- Prediction of barge movement using a convolutional neural network



Input sequence

Ground truth

Prediction

Combining the mobility prediction with the coverage model, of 62200 sequences in a validation set, we were able to predict up to **90%** of the low-RSRP events and RLFs **40 seconds** ahead

Technische Universität München

NOKIA

Predictive Location-Aware Network Automation for Radio

Closed-Loop Automation Evaluation with Simulation

A **digital twin** of the testbed setup is **mirrored** in a simulator

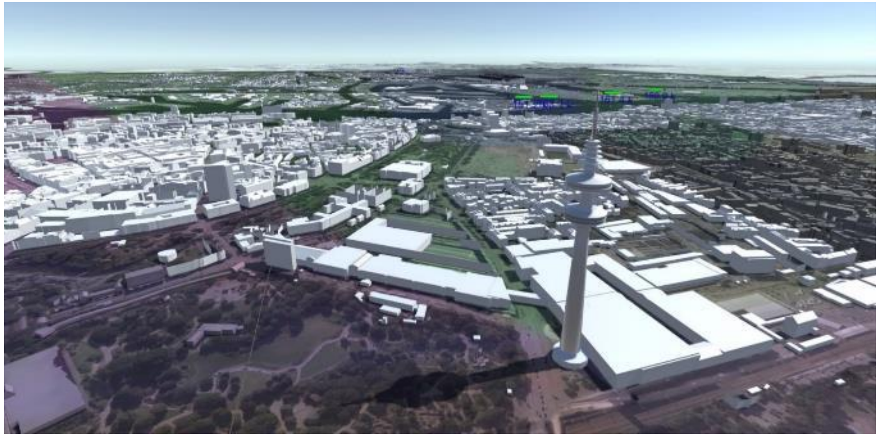
- Full 3D model of the city of Hamburg and especially the harbor area
- Network topology and configuration as in the real testbed
- Traces of the movement of the real barges are collected from the testbed and imported into the simulation scenario

Predictive Location-Aware Network Automation for Radio

Closed-Loop Automation Evaluation with Simulation

The coverage issues of the real testbed can be **reproduced**:

- The simulator is connected to a cognitive network management experimental system, which does the ML inference and implements the closed-loop automation functions
- The mobility predictions from the MPP model are created for each ship, simulation's radio propagation model is used for modeling the RSRP
- Prediction 40 simulated seconds ahead

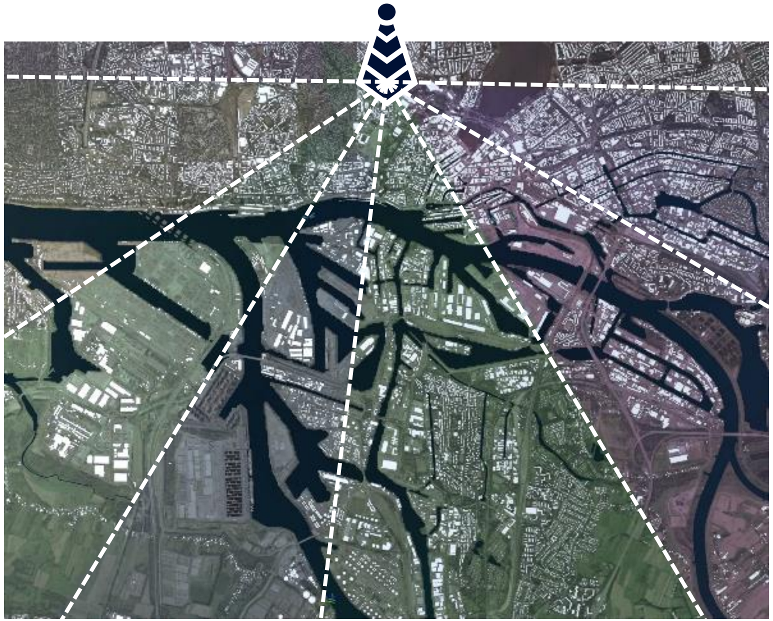


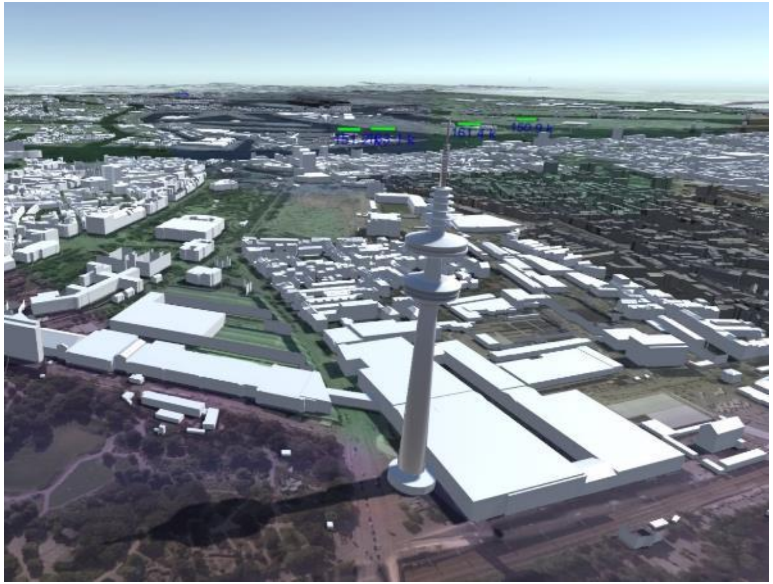
Predictive Location-Aware Network Automation for Radio

Closed-Loop Automation Evaluation with Simulation

The **digital twin is extended** to simulate NR with **beam forming**

- Four beams per each of the two cells, with two antenna elements for two simultaneous beams



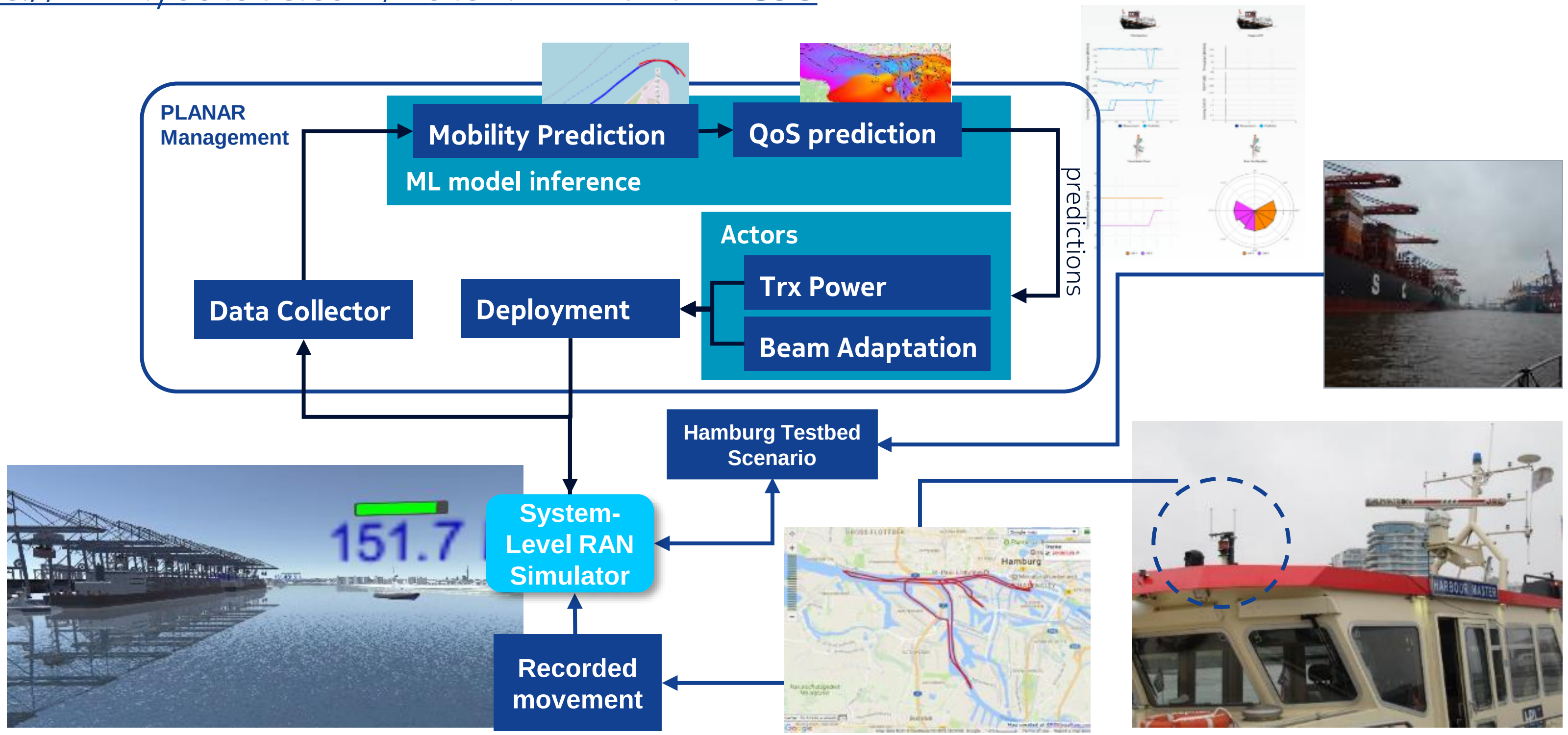


Public

Predictive Location-Aware Network Automation for Radio

Demo @ IEEE NOMS (Janne Ali-Tolppa, Marton Kajo)

https://www.youtube.com/watch?v=nMdBbLv2G98

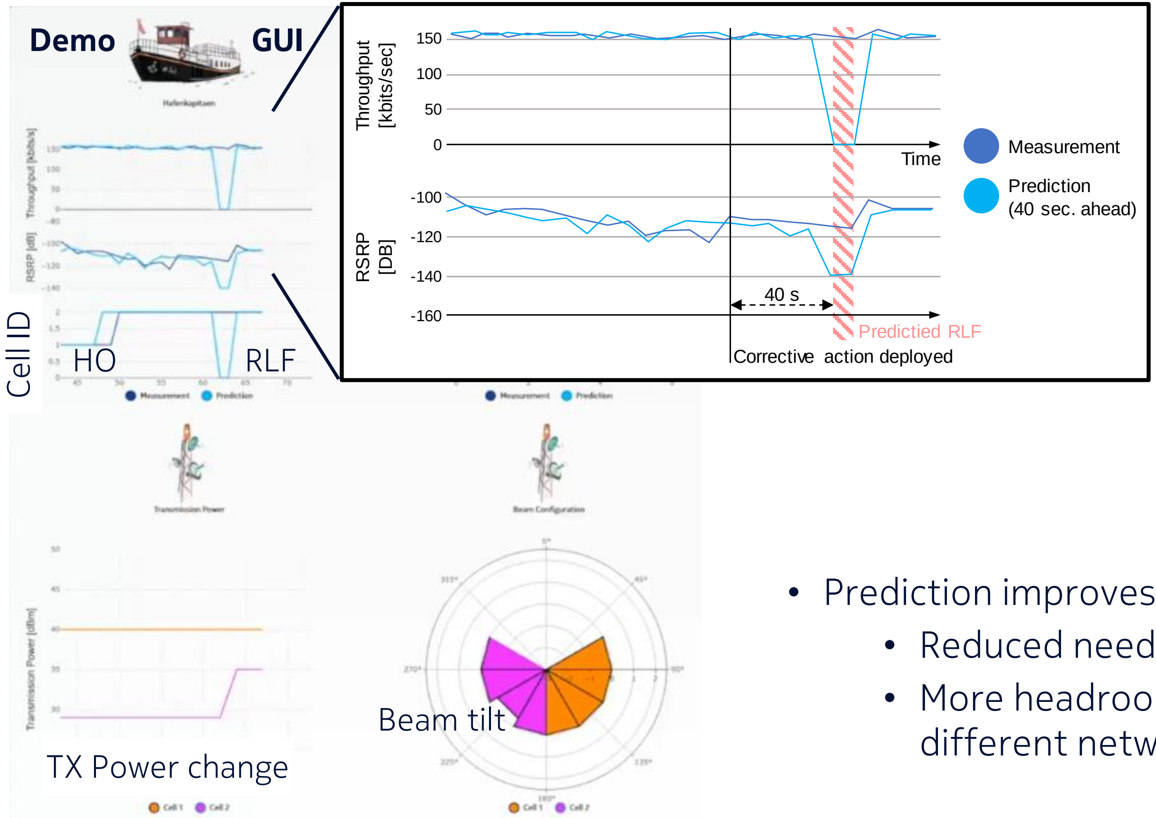


  Public

Technische Universität München

Predictive Location-Aware Network Automation for Radio (PLANAR)

Preventive Closed-Loop Optimization



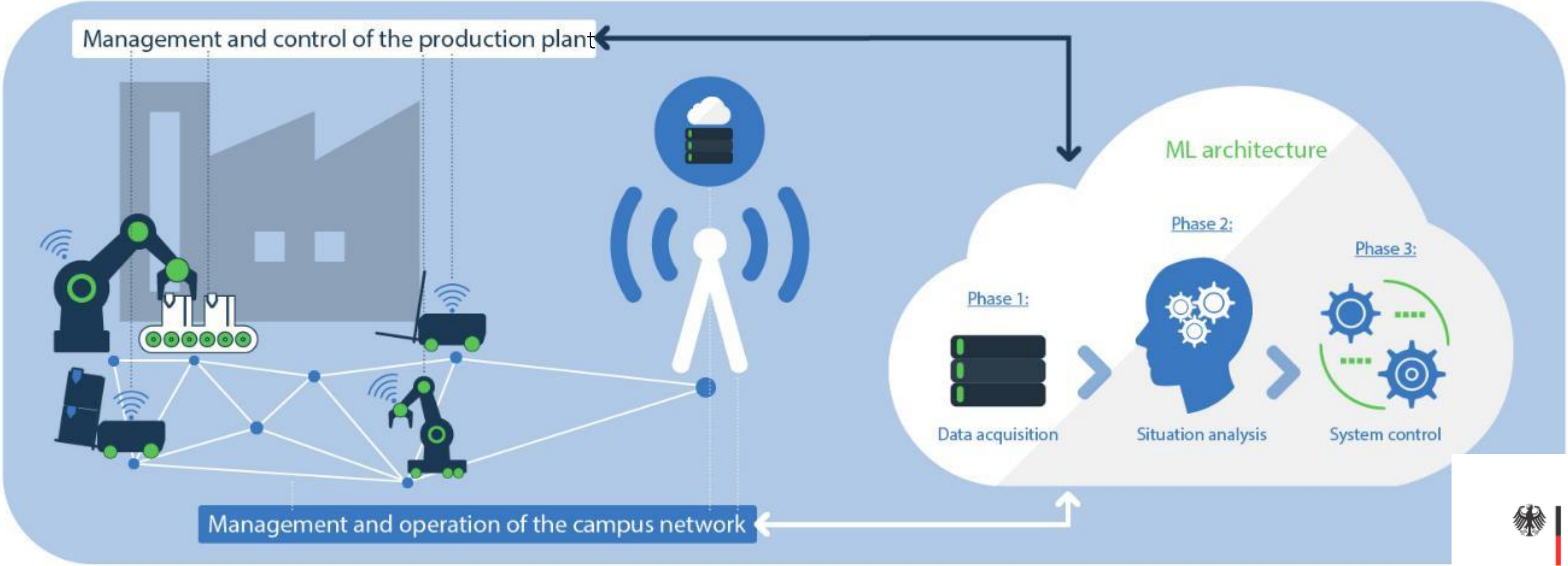
QoS and RLFs can be predicted and prevented by:

- Optimizing the transmission power
- Beam forming
  - Activate or de-active beams
  - Tilting of individual beams

- Prediction improves robustness
  - Reduced need for overprovisioning of resources
  - More headroom for resource sharing between different network slices

KICK Project (kick-project.de)
Künstliche Intelligenz für Campus-Kommunikation

NOKIA Stuttgart Munich
DFKI Deutsches Forschungszentrum für Künstliche Intelligenz GmbH
GHMT
BOSCH
TRUMPF
Fraunhofer HHI
atesio
SIEMENS
infosim

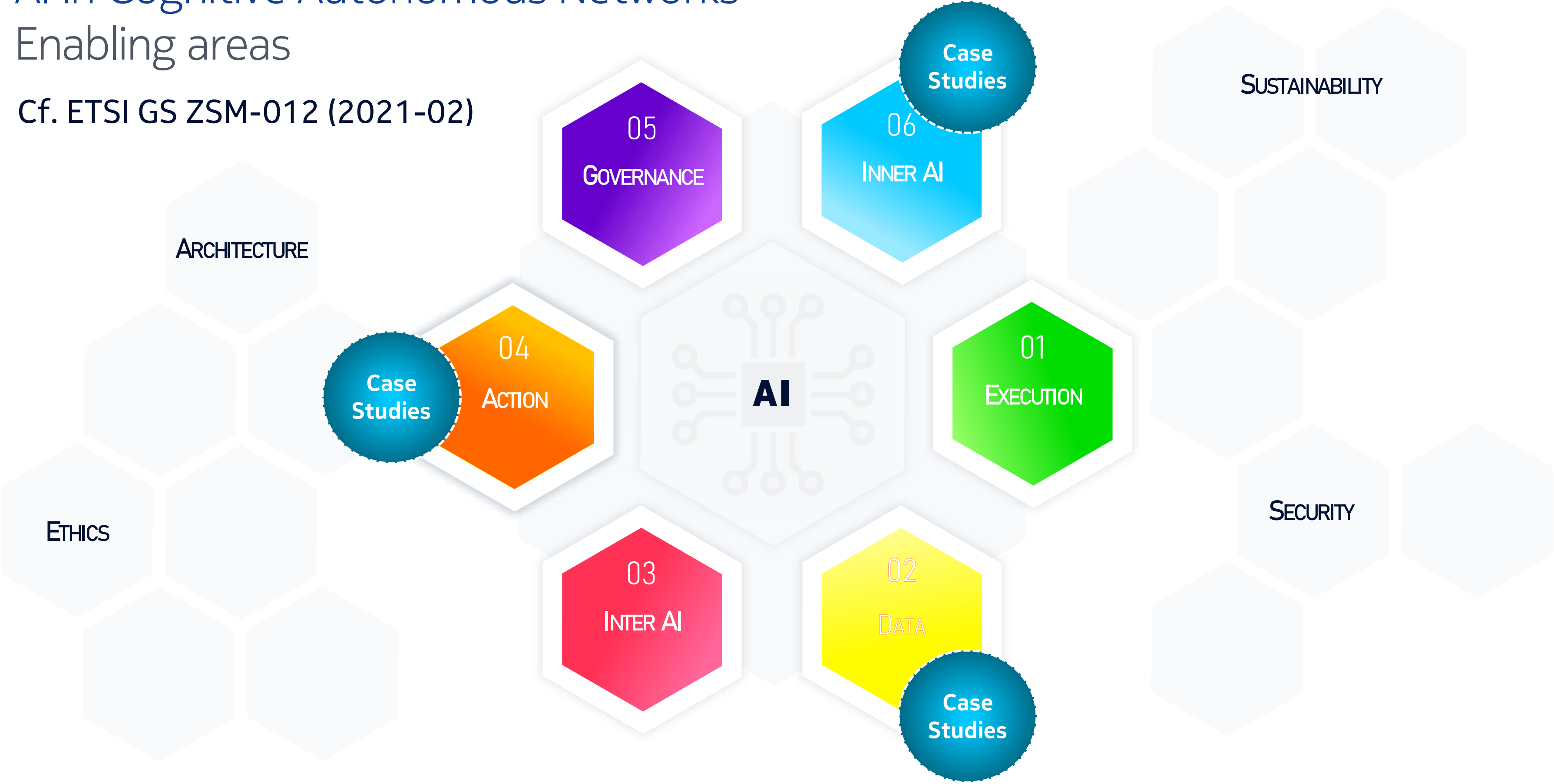


SPONSORED BY THE
Federal Ministry of Education and Research

Application of AI/ML methods to campus network automation to enable dynamic and autonomous industrial wireless networks and a flexibly reconfigurable production environment

AI in Cognitive Autonomous Networks
Enabling areas

Cf. ETSI GS ZSM-012 (2021-02)

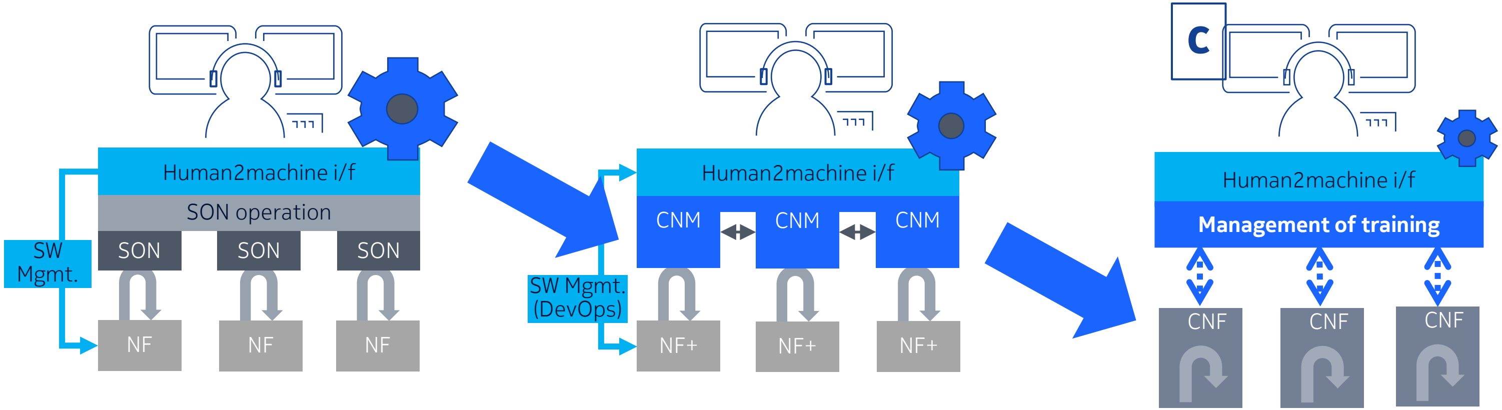


  Public NOKIA

Towards Cognitive Autonomous Networks

5G and Beyond (RAN)

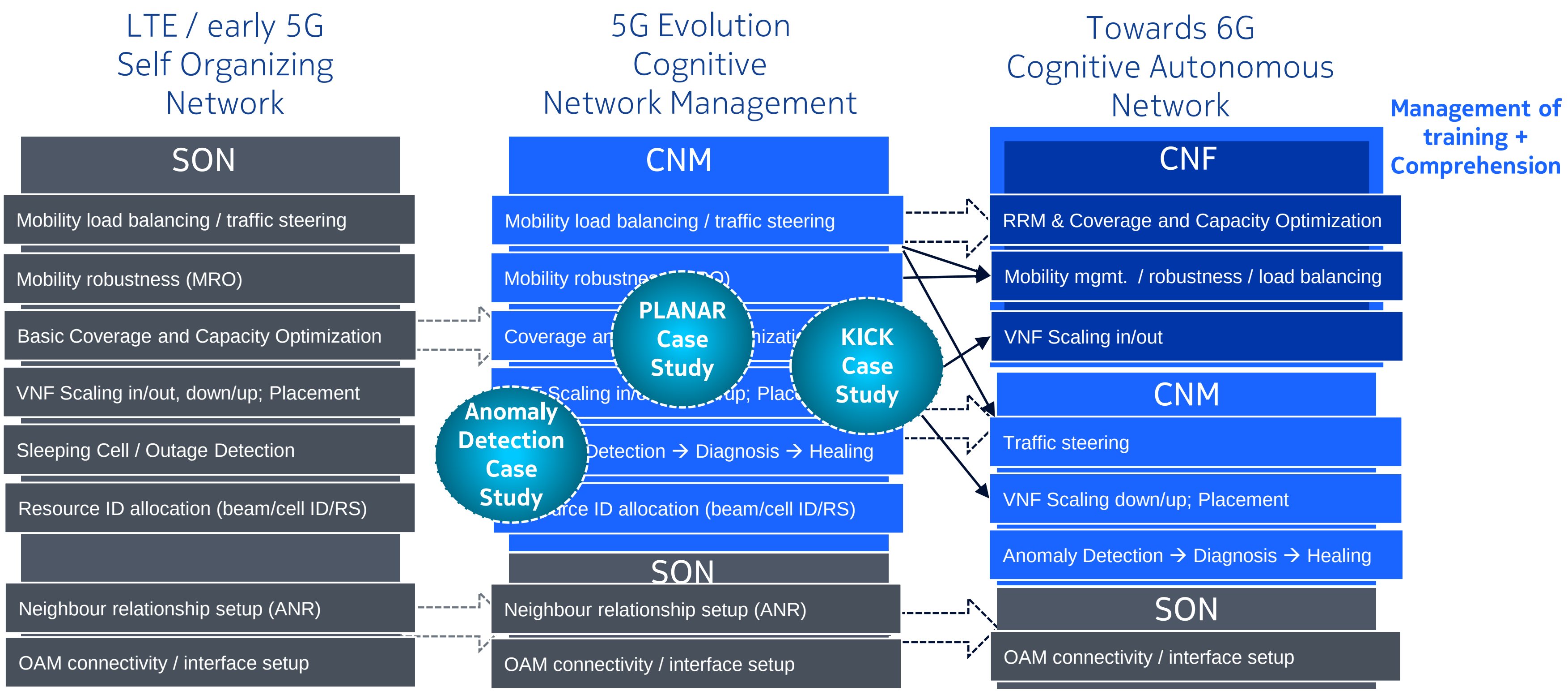
LTE / early 5G Self Organizing Network	5G Evolution Cognitive Network Management	Towards 6G Cognitive Autonomous Network
Rule-based management functions, Software Management	Learning management functions, Software Management (DevOps)	Management of training + Comprehension (C)



Rule-based network functions	Rule-based / data-driven network functions	Learning network functions (embedded management)

(C)NF: (Cognitive) Network Functions

Towards Cognitive Autonomous Networks
5G and Beyond (RAN)



The Role of Cognitive Autonomy in “5G and Beyond” Communication Networks

Takeaways

- Cognition & Autonomy are crucial, separable properties
  - Understand system concepts and contexts to enable decision making at machine level
  - Actions can be taken at any step in the cognitive process

- Anomaly Detection case study
  - Leverage the potential of data, by comprehensive analysis → high quality detection
  - Motivate Network Operations Experts to combine their knowledge with the machine-level → (semi-)supervised learning (augmented diagnosis)

- PLANAR case study
  - Separated (but chained) modeling of the network (QoS) and application (barges) mobility
  - Increased robustness by prediction of radio QoS degradation
  - Replica of physical testbed (digital twin) enabling to evaluate features (beamforming) not (yet) present in the testbed

NOKIA

The Role of Cognitive Autonomy in “5G and Beyond” Communication Networks

Takeaways

- AI requires expanded (training data) and new (model management) standardized interfaces
  - AI Enabling Areas: Data → Execution, Inner-AI → Action; inter-AI & Governance
  - ETSI ISG ZSM WI12 defines the AI enablers within the ZSM architectural framework

- LTE/early 5G: SON → 5.5G: Cognitive NM (CNM) → 6G: Cognitive Autonomous Network (CAN)
  - Cognition & Autonomy penetrate the u- and c-plane
    - AI is the technology area to enable this
  - m-plane takes a new role (managing the AI on the u- and c-plane)
    - Process (integration / replacement of legacy functions)
    - Some SON and CNM functions remain

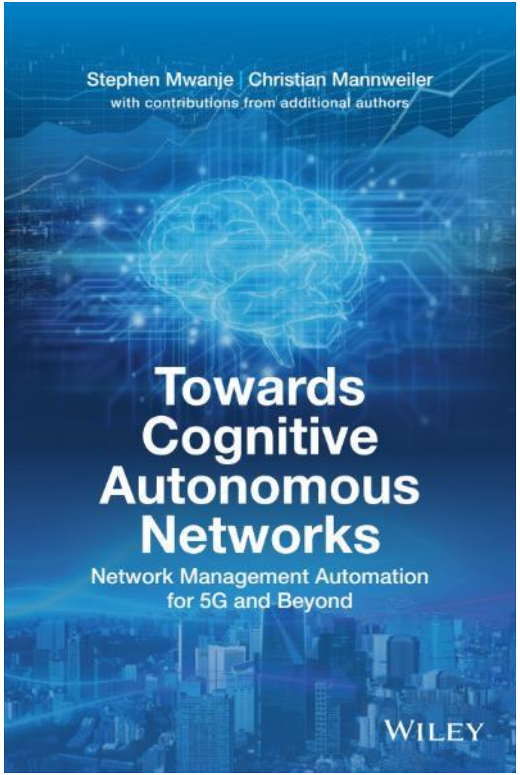
→ Performance improvement of existing network (management) functions & enabling of new functions

→ Human operator: shift from low value “execution” to high impact “supervision” tasks (simple, recurring problems are not visible any more → “simplification”)

Towards Cognitive Autonomous Networks
Recent journal / conference publications & tutorials

- S. Mwanje, M. Kajo, J. Ali-Tolppa, Environment Modeling and Abstraction of Network States using Deep Clustering, IEEE Network special issue "AI for mobile networks“ 2020
- M. Kajo, B. Schultz, Deep clustering of mobile network data with sparse autoencoders, IEEE NOMS 2020
- M.L. Mari-Altozano, S. Mwanje, S. Luna-Ramirez, M. Toril, H. Sanneck, C. Gijon, A Service-centric Q-learning Algorithm for Mobility Robustness Optimization in LTE, IEEE TNSM, 2021.
- A. Masri, T. Veijalainen, H. Martikainen, J. Ali-Tolppa, S. Mwanje, M. Kajo, Machine Learning Based Predictive Handover, IEEE IM 2021
- A. Banerjee, S. Mwanje, G. Carle, Optimal configuration determination in cognitive autonomus networks, IEEE IM 2021
- A. Banerjee, S. Mwanje, G. Carle, RAN Cognitive Controller, 2nd KuVS Fachgespräch on Machine Learning and Networking, 2020
- A. Banerjee, S. Mwanje, G. Carle, Game theoretic Conflict Resolution Mechanism for Cognitive Autonomous Networks, IEEE SPECTS 2020
- P. Kalmbach, P. Diederich, W. Kellerer, R. Pries, A. Blenk, sfc2cpu: Operating a Service Function Chain Platform with Multi-Agent Reinforcement Learning, IEEE IM 2021
- S. Mwanje, Towards Cognitive Autonomous Networks, Tutorial, IEEE NoF 2020
- S. Mwanje, Towards Cognitive Autonomous Networks, Tutorial, IEEE CNSM 2020

Towards Cognitive Autonomous Networks
Network Management Automation for 5G and beyond



Book structure and outline

Background	Core argumentation	Solutions/Application	Open challenges
2. Network Evolution	1. Introduction	7. Cognitive Auto-Configuration	11. System Challenges of CANs
3. SON in pre-5G Networks	4. Modeling Cognition	8. Cognitive Autonomy in Optimization	12. Towards realizing CANs
	5. Classical AI for reasoning	9. Cognitive Self-Healing	
	6. ML for cognitive decision making	10. Cognitive Self-Operation	

https://www.wiley.com/en-us/Towards+Cognitive+Autonomous+Networks+%3A+Network+Management+Automation+for+5G+and+Beyond-p-9781119586388

NOKIA

Cognitive NM functions: inputs → ML algorithm / rules → outputs

Cognitive NM Function	Examples for applicable ML algorithm
Load balancing / traffic steering	Metaheuristics (PSO*) / decision tree
Coverage and Capacity Optimization	Reinforcement learning (QL*)
Mobility robustness (MRO)	Reinforcement learning (QL*)
Scaling in/out, down/up	Reinforcement learning; Graph Neural Nets
Anomaly Detection → Diagnosis → Healing	Topic Modeling → MLN* → Utility Theory
Network Function Chaining	Metaheuristics (genetic algorithms)
LCM / (re-)configuration, (re-)placement	Metaheuristics (harmony search)
Neighbour relationship setup (ANR)	n/a (rules)
Resource ID allocation (beam/cell ID/RS)	Unsupervised learning (clustering)
OAM connectivity / interface setup	n/a (rules)

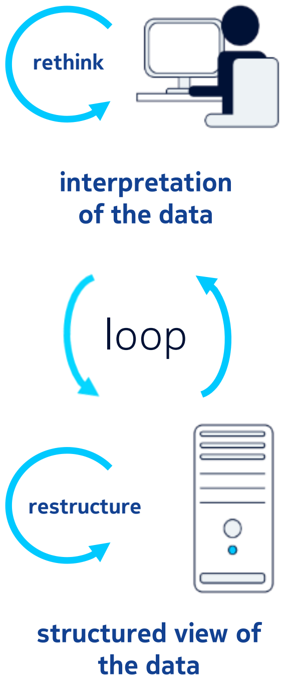
* PSO: Particle Swarm Optimization, QL: Q(uality) Learning, MLN: Markov Logic Networks

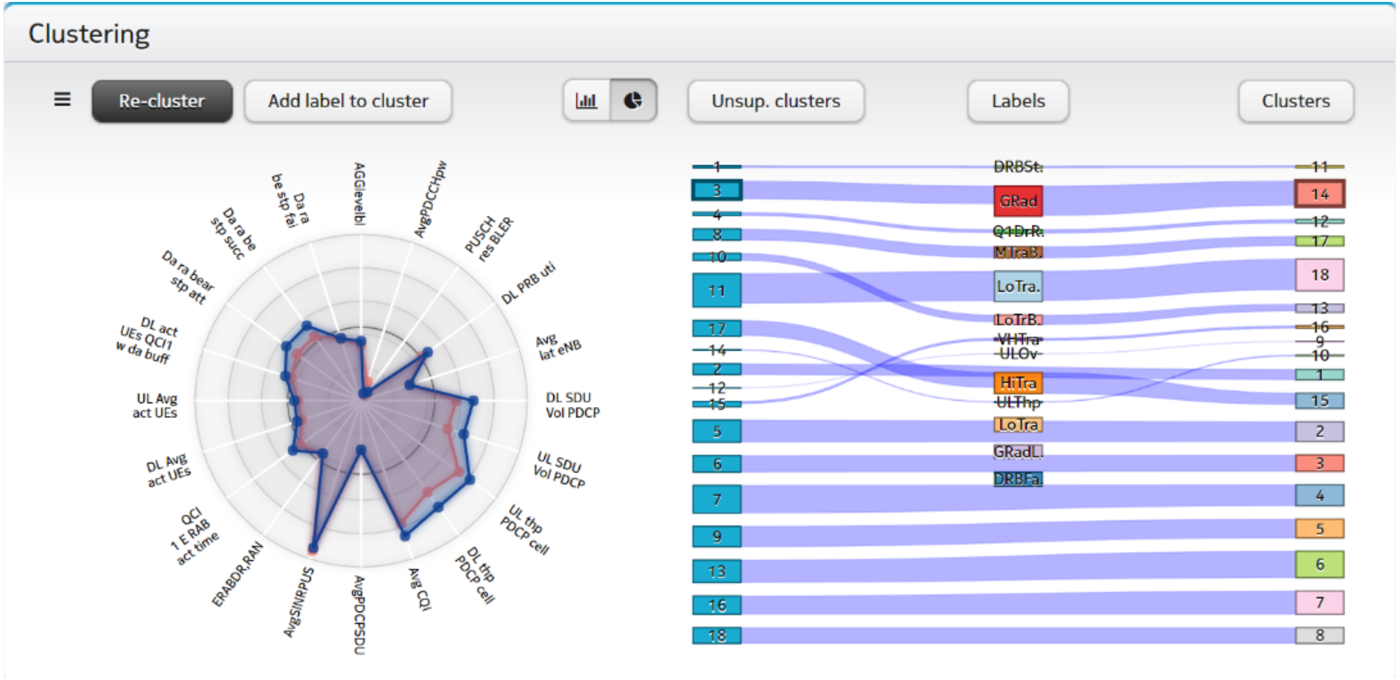
Example: Cell anomaly detection and diagnosis – ML Algorithms

Cognitive NM Function	Examples for applicable ML algorithm
Input selection	Genetic algorithms, PCA
Multi-dimensional non-normally distributed profiling	Clustering: k-NN, SOM, GNG
Anomaly level calculation	Multi-dimensional probabilistic distributions
Anomaly event aggregation	DBSCAN
Diagnosis	Decision theory, rulebases, different distance measures: Mahalanobis, Kullback-Leibler divergence or Hellinger
Augmented learning	Active learning, DBSCAN, k-NN

Augmented diagnosis

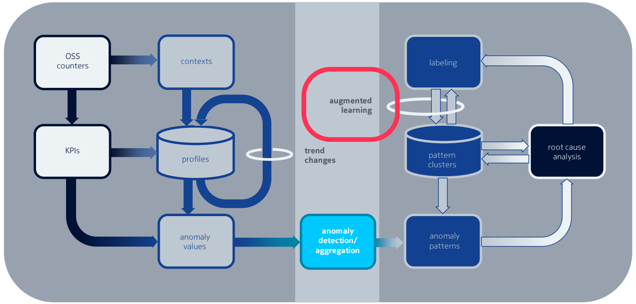
Synergetic exploitation of human-machine capabilities for fast and efficient analysis





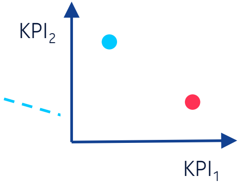
NOKIA Bell Labs

Augmented diagnosis



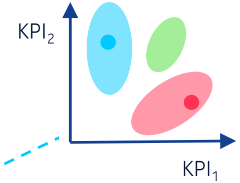
Human

1. Labels a few examples of the major expected anomaly groups

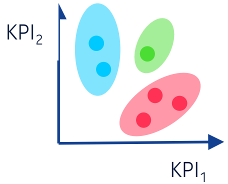


Machine

2. Creates a clustering taking the labeled points into account



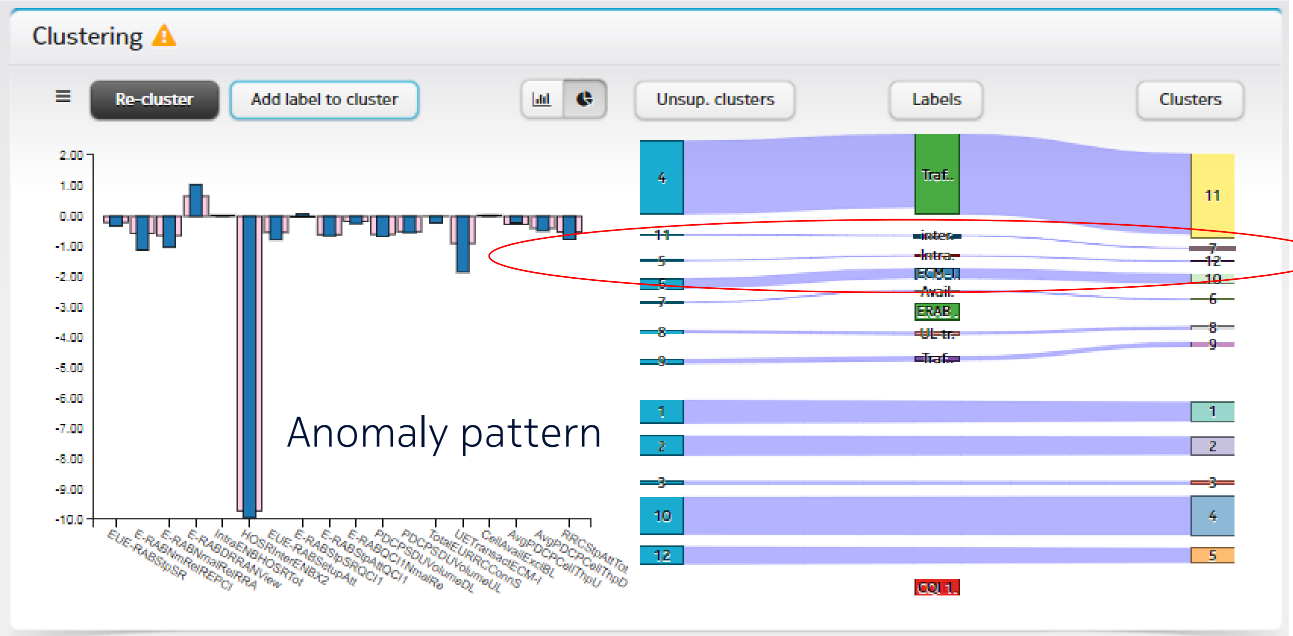
3. Expands and refines the labeling:
   - Labels the previously unlabeled anomalies in the expected groups
   - Labels the anomalies in the unexpected groups



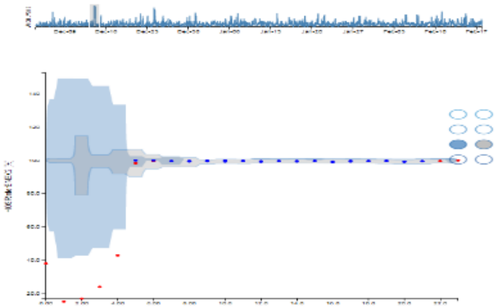
4. Any newly arriving anomalies can be **automatically labeled** with this information

Use Case Example from a Major Operator

- Augmented learning → new cluster of intra- and inter-eNB Handover problems (distinct classes of anomalies that had not been discovered / analyzed yet)



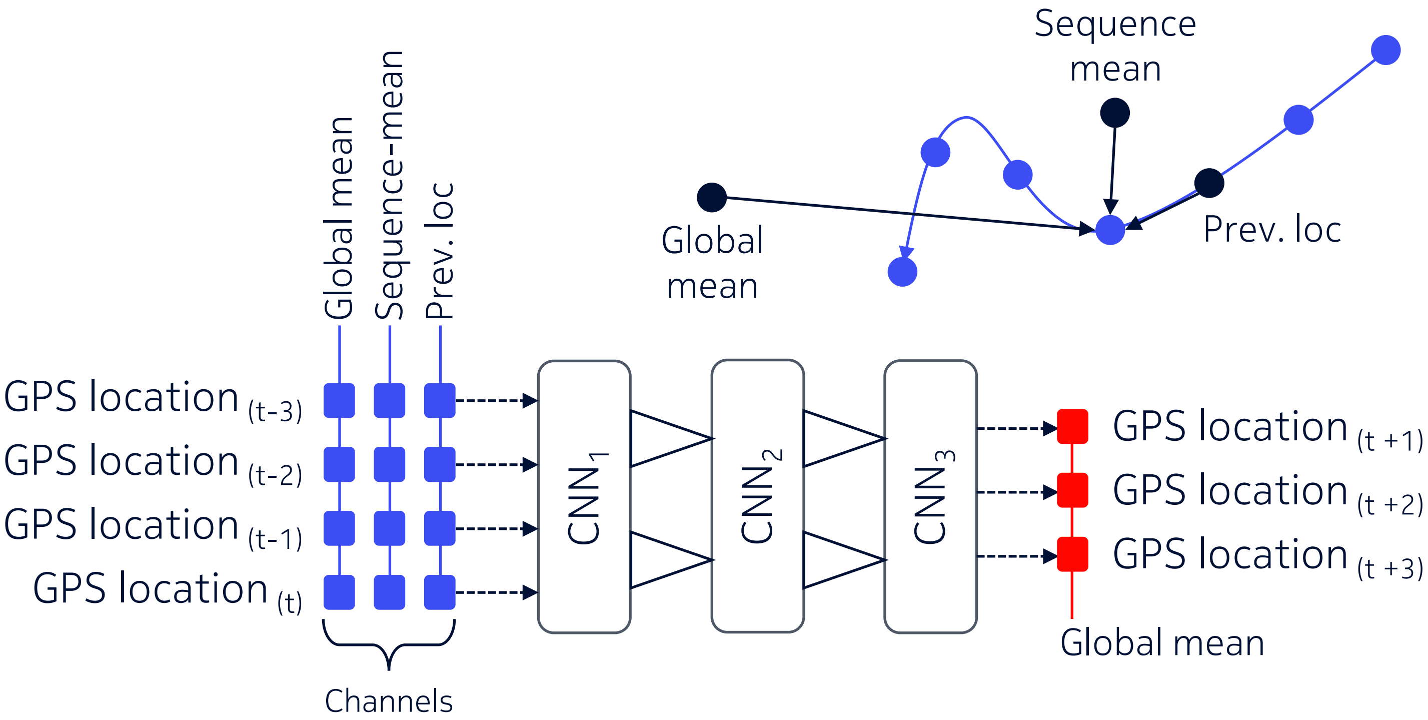
New identified problem cluster → new „golden configuration“



NOKIA Bell Labs

Predictive Location-Aware Network Automation for Radio (PLANAR)

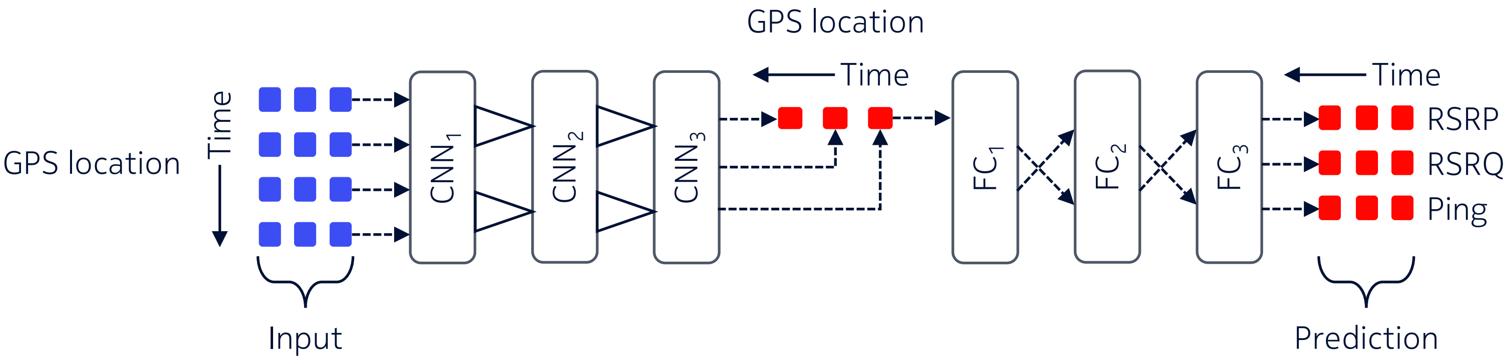
Mobility prediction



Technische Universität München
NOKIA

Predictive Location-Aware Network Automation for Radio (PLANAR)

Evaluation with location and radio environment combined



- 62200 sequences in the validation set
- Predicting RLFs **40 seconds** ahead
- True positives: **97.6%**
- False positives: **16.7%** of predicted events
- We optimized for high true positive rate, since the corrective actions were non-intrusive
- In many of the examples classified as false positive, the actions were still justified

Technische Universität München

NOKIA

Predictive Location-Aware Network Automation for Radio

Example Deployment Architecture