Design and Analysis of a 5G Digital Twin

Piotr Olesinski
Department of Systems and Computer Networks
Wrocław University of Science and Technology,
Wyb. Wyspianskiego 27, 50-370 Wrocław, Poland
piotr.olesinski@pwr.edu.pl

Róża Goścień
Department of Systems and Computer Networks
Wrocław University of Science and Technology,
Wyb. Wyspianskiego 27, 50-370 Wrocław, Poland
roza.goscien@pwr.edu.pl

***Abstract*—Digital twin (DT) is one of the key technologies to enable intelligent mobile networks and the future sixth-generation (6G). This PhD thesis is going to design, test, analyze, and implement a cell-level digital twin for a mobile network management to address the practical need for the evaluation of new network functionalities and configuration options before their deployment in real-world networks. The research proposes simulation-assisted assessment of individual cells and nodes under both observed and previously unseen configurations.**

***Index Terms*—Digital Twin; Artificial Intelligence; Machine Learning; Mobile Network; 5G; 6G**

I. Introduction

Mobile networks have been rapidly evolving over the last few decades towards complex and autonomous systems, driven by growing traffic demand and the introduction of advanced radio and configuration features [1]. Alongside, Digital Twin (DT) emerges as a promising technology enabling the creation of a virtual representation of a physical object, process, or system that is continuously updated with data from its real counterpart. It allows us to mimic, understand, and explain the behaviors of complex objects, processes, and systems, and, in turn, optimize them and improve their maintenance [2], [3]. The DT is also considered for mobile networks due to the increasing complexity of radio systems, which makes expert-driven assessments insufficient for reliable predictions of network behavior under new operating conditions. For mobile network applications, DTs are expected to support monitoring, simulation, optimization, and decision support [4]. However, current trend studies focus either on large-scale Digital Twin Networks (DTNs) or specific DTs for network components, such as Reconfigurable Intelligent Surfaces (RIS), antenna arrays, or other radio elements [1], [5].

From the perspective of network management, one of the most difficult practical problems is the deployment of new functionalities or configuration options [6]. Existing DT approaches are commonly oriented towards continuous operation, such as anomaly detection or optimization over already observed network settings [7]. In contrast, operators need methods to assess the impact of configuration changes before activation in live networks, without relying exclusively on physical test networks or extensive manual testbed preparation [8]. The individual cell and node levels are highly important for this challenge, as performance depends on local traffic patterns, radio conditions, and configuration settings [6].

II. Problem Statement and Importance

The PhD thesis is going to address the identified gaps. Its main goal is to design, test, analyze, and implement a cell-level digital twin for mobile network management that can reconstruct cell behavior from operational data and support the simulation evaluation of unseen functionalities and configuration changes before deployment in live networks. The thesis is realized in the implementation mode in collaboration with Nokia, which shares real data from an operational mobile network and physical infrastructure to build, validate, and implement DT. The problem is motivated by several competing objectives: modeling accuracy, computational complexity, and the ability to generalize across cells with different traffic patterns, radio conditions, and deployment characteristics. The scope of the PhD thesis therefore covers individual cells and network nodes within a local cluster, with emphasis on practical applicability in operational environments.

The main challenges addressed in the thesis are as follows:

Operational Inefficiencies: With the high complexity of the system, it is difficult to ensure consistent and scalable execution. Currently, network development requires manual interventions and demanding tools, which reduce overall efficiency and are susceptible to delays and errors.

Lack of Real-World Validation: Limited availability of alternatives to real-world network environments in which new features can be verified under realistic operational conditions.

Methodological Gaps: There is a lack of reproducible methods for assessing the effectiveness of the proposed approach in real-world conditions. In many cases, existing studies or implementations focus on high-level concepts without providing sufficient practical validation or guidance for deployment in operational environments.

The importance of the PhD thesis lies in the need for solutions that are not only technically effective but also practical and aligned with organizational goals. Addressing these challenges can improve efficiency and reduce operational risk.

III. PhD Scope and Work Plan

The PhD thesis is organized around the development and validation of a cell-level DT framework for simulation-assisted network management. The research plan includes five main stages. In the first stage, real datasets from Nokia are acquired and processed to identify the relevant Key Performance Indicator (KPI) set, feature activations, and configuration parameters

979-8-3195-0167-7/26/$31.00 ©2026 IEEE

are identified, and the dimensionality of the configuration space is analyzed. In the second stage, the first hierarchical DT prototype is built from operational measurements and validated on forecasting and reconstruction tasks. In the third stage, the DT is extended with knowledge obtained from controlled experiments and simulation environments to enable reasoning about unseen configurations. In the fourth stage, the complete framework is assessed in terms of prediction quality, uncertainty, transferability across cells, and usefulness for simulation-assisted optimization. In the last stage, the final DT version is implemented in a real network environment.

This plan is expected to lead to a methodology for practical cell-level DT construction that can be used for pre-deployment evaluation in mobile network engineering.

IV. Methodologies

The main methodological component of the PhD thesis is a Multi-Level Hierarchical Digital Twin (MLHDT), developed to reconstruct the state of an individual cell by modeling the relationships between load, radio conditions, and final performance reflected in throughput. The methodology follows three main steps. First, load-related KPIs, including the number of active users and Physical Resource Block (PRB) utilization, are forecast independently. Second, radio indicators such as Block Error Rate (BLER), Modulation and Coding Scheme (MCS), and Signal-to-Interference-plus-Noise Ratio (SINR) are modeled separately. Third, these intermediate outputs are combined to estimate throughput within the virtual environment. This hierarchical formulation makes throughput dependent on the reconstructed virtual state rather than on direct forecasting alone.

The study is based on operational data from an operator test network over 3 months, with 1-hour granularity, for approx. 200 TDD-related cells in 1 cluster. Data preprocessing includes cleaning, temporal alignment of measurements, handling of missing values, and construction of consistent feature sets for forecasting and reconstruction tasks. The available data include KPIs, feature activation status, and configuration parameters. Dedicated models are trained for each KPI, while throughput is estimated using a regression model. The main modeling approach uses boosting methods, in particular eXtreme Gradient BOOSTing (XGBoost), which has been compared against time-series and machine learning baselines incl. Auto-Regressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), Random Forest, Light Gradient Boosting Machine (LightGBM), and CATegorical BOOSTing (CatBoost). To control the complexity of the configuration space, the methodology includes correlation-based filtering, importance ranking, and reduced configuration representations. Validation is performed using train/test splits, rolling-horizon forecasting, cross-cell validation, and tests on unseen cell types and unseen configurations. Performance is assessed using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and coefficient of determination R^2.

A second methodological component addresses the limitation that operational data describe only configurations already observed in the network. To support the evaluation of unseen functionalities, the PhD thesis extends the MLHDT with knowledge obtained from controlled Nokia testbeds and simulation environments. The methodology consists of identifying the KPI impact of selected functionality activations or configuration changes under controlled conditions and transferring this knowledge into the DT as an additional simulation capability. This enables the assessment of previously unseen configurations while preserving the link to operational data. An important methodological aspect of this stage is the development of confidence-aware mechanisms for transferring knowledge from testbeds and simulators to the DT.

V. Progress and Future Work

Initial progress of the thesis includes the work on stages 1 and 2 – the identification of key configuration parameters and KPI groups relevant for DT construction, as well as the implementation of the first MLHDT prototype for active users, PRB utilization, BLER, MCS, SINR, and uplink throughput. The current research focus is on validating the hierarchical DT against operational measurements and quantifying the performance gap between using measured and forecasted intermediate KPIs. The next steps include termination of stage 2 and following tasks of stages 3-5, the refinement of KPI selection, further validation of forecasting models, integration of testbed and simulation knowledge, and uncertainty-aware assessment of the DT for simulation. At present, the PhD thesis considers an offline DT and does not yet include real-time closed-loop integration with the live network.

Acknowledgment

Piotr Olesiński and Róża Goścień was supported by the Polish Ministry of Science and Higher Education under the grant DWD/8/0116/2024.

References

[1] M. Sheraz *et al.*, "A comprehensive survey on revolutionizing connectivity through artificial intelligence-enabled digital twin network in 6G," *IEEE Access*, vol. 12, pp. 49 184–49 215, 2024.

[2] Y. Pan *et al.*, "A survey on digital twin networks: Architecture, technologies, applications, and open issues," *IEEE Internet Things J.*, vol. 12, pp. 19 119–19 143, 2025.

[3] S. Mihai *et al.*, "Digital twins: A survey on enabling technologies, challenges, trends and future prospects," *IEEE Commun. Surv. Tutorials*, vol. 24, pp. 2255–2291, 2022.

[4] L. Kundu *et al.*, "AI-RAN: Transforming RAN with AI-driven computing infrastructure," *IEEE Coomun. Mag.*, vol. 64, pp. 168–174, 2026.

[5] M. S. Akbar *et al.*, "On challenges of sixth-generation (6G) wireless networks: A comprehensive survey of requirements, applications, and security issues," *J. Network Comput. Appl.*, vol. 233, p. 104040, 2025.

[6] G. Liu *et al.*, "Native design for 6G digital twin network: Use cases, architecture, functions, and key technologies," *IEEE Internet Things J.*, vol. 12, pp. 29 135–29 151, 2025.

[7] K. Duran *et al.*, "Toward digital twin-as-a-service (DTaaS) platforms: A survey on architecture, design requirements, and performance metrics," *IEEE Commun. Surv. Tutorials*, vol. 28, pp. 1845–1878, 2026.

[8] R. Poorzare *et al.*, "Network digital twin toward networking, telecommunications, and traffic engineering: A survey," *IEEE Access*, vol. 13, pp. 16 489–16 538, 2025.