Internship Plan: Survey on AI for Mobile Networks

Internship Period: 25 June – 24 August, 2025

Phase 1: Analysis of Development Trends

(Timeline: Mid May – June 24, 2025)

- **Objectives:**
    - Analyze the current landscape, key standards (3GPP, O-RAN, AI-RAN), and research trends to build foundational knowledge.
- **Task:**
    - Review technical reports from standards bodies like 3GPP and O-RAN.
    - Survey industry white papers to identify key use cases and strategies.
    - Analyze academic research on AI-native concepts and federated learning.
    - Categorize AI application models (automation, optimization, services).
- **Deliverables:**
    - Written study notes summary of AI development trends and standards evolution.
    - Organized reference list.

Phase 2: Toolkit Evaluation and Application Potential

(Timeline: June 25 – July 13, 2025 / Weeks 1-3)

- **Objectives:**
    - Gain hands-on experience with AI toolkits to assess their features, feasibility, and constraints for network applications.
- **Task:**
    - Install and test key AI development toolkits (e.g., NVIDIA Aerial SDK, Sionna) within mobile network scenarios.
    - Complete tutorials and run a sample use case to assess workflows.
    - Document the capabilities, limitations, and system requirements for each platform.
    - Propose suitable application concepts and integration approaches based on toolkit capabilities.
- **Deliverables:**
    - A technical memo summarizing the features and evaluation of the toolkits.
    - Working code or scripts from experiments and simulations.

Phase 3: Modeling Deployment and Computational Requirements

(Timeline: July 14 – August 3, 2025 / Weeks 4-6)

- **Objectives:**
    - Analyze the trade-offs of deploying AI models at different points in the mobile network (e.g., RAN, Edge, Core).

- o Model the resource demands associated with these deployment scenarios.
- **Task:**
  - o Model and evaluate latency tolerance, bandwidth constraints, data collection types, and inference cycles for AI models in the RAN vs. Edge vs. Core.
  - o Analyze trade-offs in latency, bandwidth, and data availability for each model.
  - o Model the computational resource demands (CPU/GPU cycles, memory) for different deployment scenarios to illustrate key differences.
- **Deliverables:**
  - o An analysis report detailing different AI deployment models.
  - o Comparative tables and graphs illustrating resource demands for each model.
  - o Recommendations for model deployment conditions and resource allocation.

Phase 4: System Integration Recommendations and Future Directions

(Timeline: August 4 – August 17, 2025 / Weeks 7-8)

- **Objectives:**
  - o Synthesize previous findings into a coherent system integration proposal.
  - o Identify key challenges, mitigation strategies, and future research directions.
- **Task:**
  - o Design a high-level architecture diagram for integrating AI into a mobile network, showing data flows and control loops.
  - o Identify potential deployment challenges such as model management, interoperability, and scalability.
  - o Propose practical mitigation strategies for the identified challenges and outline future research directions.
- **Deliverables:**
  - o An integration plan featuring a proposed demonstration architecture.
  - o A document detailing potential deployment challenges and their mitigation strategies.

Phase 5: Final Report Drafting and Presentation

(Timeline: August 18 – August 24, 2025 / Week 9)

- **Objectives:**
  - o Consolidate all analyses and recommendations into a comprehensive survey report to serve as a key reference for future planning.
- **Task:**
  - o Combine all previous deliverables (memos, reports, architecture documents) into a single, structured internship report.
  - o Prepare a final presentation that summarizes the key findings and recommendations from the survey.
- **Deliverables:**
  - o A final, comprehensive internship report.
  - o A final presentation deck.

Reference List:

Phase 1: Analyis of Development Trends

[1] 3GPP, "Study on Artificial Intelligence (AI) / Machine Learning (ML) for NR Air Interface," TR 38.843 V18.0.0, Mar. 2025. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.843/

[2] AI-RAN Alliance, "AI-RAN Alliance Whitepaper: Revolutionizing Radio Access Networks with AI," Dec. 2024. [Online]. Available: https://ai-ran.org/wp-content/uploads/2024/12/AI-RAN_Alliance_Whitepaper.pdf

[3] J. Montojo et al., "Overview of AI in 3GPP's RAN Release 18," IEEE ComSoc Tech News, 2024. [Online]. Available: https://www.comsoc.org/publications/ctn/2024/june/overview-ai-3gpps-ran-release-18

[4] Ericsson, "AI in RAN - Enhancing Network Performance of CSPs," White Paper, 2024. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/white-papers/ai-in-ran

[5] O-RAN Alliance, "Principles and Methodologies for AI/ML Testing in Next-Gen Networks," Research Report, 2024. [Online]. Available: https://mediastorage.o-ran.org/specifications/o-ran-public/O-RAN.WG1.Principles-and-Methodologies-for-AI-ML-Testing-v01.00

[6] 5G Americas, "Artificial Intelligence and Cellular Networks," White Paper, Dec. 2024. [Online]. Available: https://www.5gamericas.org/wp-content/uploads/2022/12/5G-Americas-AI-in-Cellular-Networks-WP-final.pdf

Phase 2: Toolkit Evaluation and Application Potential

[1] NVIDIA, "Aerial Omniverse Digital Twin Boosts Development of AI-Native Wireless," *NVIDIA Technical Blog*, Mar. 2025. [Online]. Available: https://developer.nvidia.com/blog/aerial-omniverse-digital-twin-boosts-development-of-ai-native-wireless/

[2] NVIDIA, "Sionna - An Open-Source Library for 6G Research," *NVIDIA Developer Portal*, 2025. [Online]. Available: https://developer.nvidia.com/sionna

[3] NVIDIA, "NVIDIA Aerial cuPHY Documentation: Overview," *docs.nvidia.com*. [Online]. Available: https://docs.nvidia.com/aerial/aerial-cuphy/current/text/overview.html

[4] Intel, "FlexRAN™ Reference Architecture for Wireless Access," 2024. [Online]. Available: https://www.intel.com/content/www/us/en/products/platforms/details/flexran.html

[5] OpenAirInterface, "OAI 5G RAN Roadmap - Sept 2024," 2024. [Online]. Available: https://openairinterface.org/oai-5g-ran-project-group/

Phase 3: Modeling Deployment and Computational Requirements

[1] ETSI, "Network Data Analytics Services (NWDAF); Stage 3," *TS 29.520 V17.9.0*, Jan. 2023. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/129500_129599/129520/17.09.00_60/ts_129520v170900p.pdf

[2] Y. Zhang *et al.*, "6G EdgeAI: Performance Evaluation and Analysis," *arXiv:2504.16529*, 2025. [Online]. Available: https://arxiv.org/abs/2504.16529

[3] M. Varela *et al.*, "Latency-Aware and Proactive Service Placement for Edge Computing," *IEEE Access*, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/8687355

[4] Rohde & Schwarz, "On the Verge of 6G - Will AI/ML Revolutionize Wireless Communication?," White Paper, May 2024. [Online]. Available: https://www.rohde-schwarz.com/us/knowledge-center/white-papers/on-the-verge-of-6g---will-ai/ml-revolutionize-wireless-communication--white-paper_254972.html

Phase 4: System Integration Recommendations and Future Directions

[1] Ericsson, "SMO Enabling Intelligent RAN Operations," White Paper, 2025. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/white-papers/a-guide-to-service-management-and-orchestration

[2] O-RAN Alliance, "O-RAN Service-Based RAN for 6G," Research Report, Nov. 2024. [Online]. Available: https://static1.squarespace.com/static/5ad774cce74940d71151f8d4/t/63863773173a195438883701/1669740405101/O-RAN.WG1.Research-Report-on-Service-Based-RAN-for-6G-v01.00.pdf

[3] Ericsson, "Intelligent RAN Automation - Managing 5G Complexity," 2024. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/white-papers/intelligent-ran-automation

[4] 3GPP, "Study on AI/ML Management Enhancements (SA5)," *S5-252558r3*, May 2025. [Online]. Available: https://www.3gpp.org/ftp/tsg_sa/WG5_TM/TSGS5_150e/Docs/S5-234677.zip