universidade de aveiro

instituto de telecomunicações

INTELLIGENCE IN ACTION: AI-DRIVEN NETWORKS

Checkpoint 2

Projeto em Informática 2024/2025

Hugo Ribeiro - 113402
Rodrigo Abreu - 113626
Eduardo Lopes - 103070
Jorge Domingues - 113278
Joao Neto - 113482

May 2025

TABLE OF CONTENTS

1. WORK DONE

WORK DONE ML

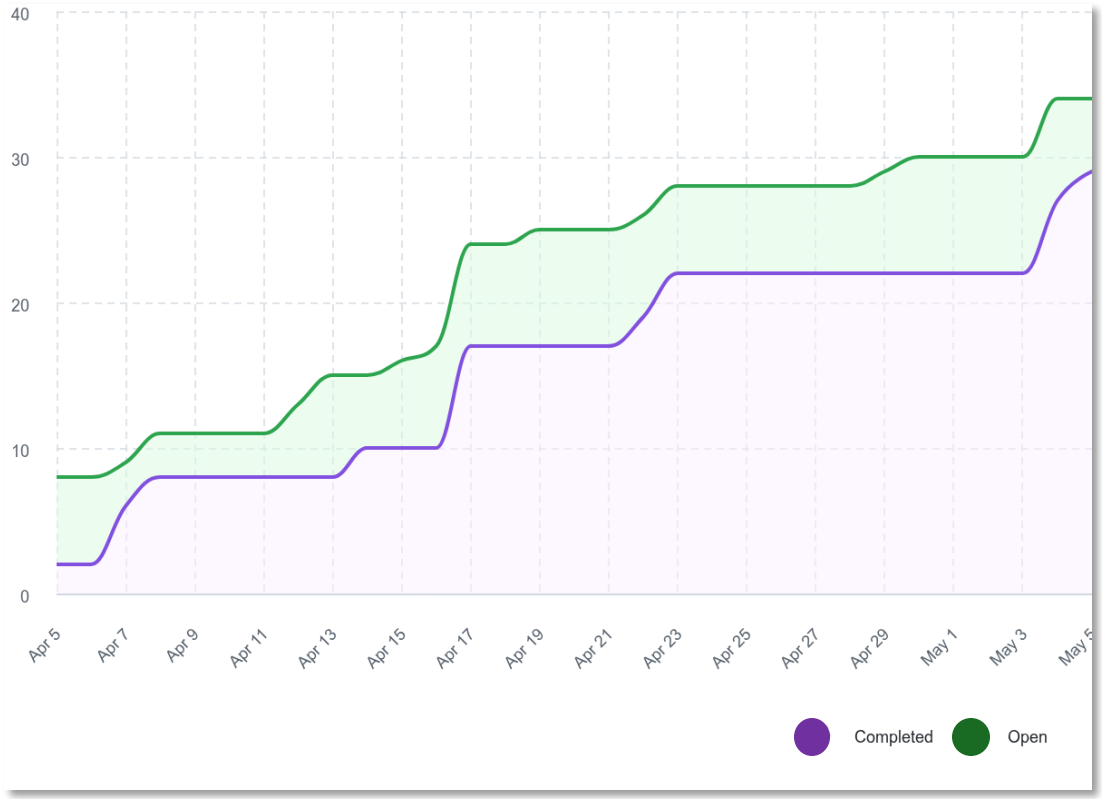
- Massively increased the pipeline speed so we have more readily available data.
- Developed the binary classification component statically.
  - Preprocessed the data.
  - Trained multiple model types systematically and ranked them.
    - **Random Forest, Gradient Boosting, MLP (Neural Networks), XGBoost.**
  - Produces the model file (Pickle) useful for model deployment.
    - The best model is the one with the highest F1-Score
  - Applies balancing strategies (SMOTE and Undersampling).

WORK DONE ML

- Developed the attack classification component statically.
  - Same achievements.
  - Works only with attack sample data.

- Integrated the binary classification component into pipeline.

- Developed the model deployment component statically.
  - The model can receive flows and it predicts whether its an attack.



2. PROBLEMS

ML Problems

- Takes much time to get enough data to train the models.
- Even when we would have enough data for binary classification, we still would have insufficient knowledge to tell the attack types apart.
- The differentiation of the Label feature is not a deterministic process and it 'broke' the model fitting upon deployment.

3. STRATEGIES

ML STRATEGIES

- Systematic training but only when a new batch of data is ready.
- Dynamically train various types of models and keep the best one (by f1-score).
- Dynamically deploy the best model to accommodate data drift.

ML STRATEGIES

- Train 2 models, binary classification (attack or benign) and attack classification.
- Deploy said models and get inferences.
- Save important inferences (attacks) for displaying on Chronograf.
- Remove TensorFlow and Keras from our plans.

4. PLAN

ML PLAN

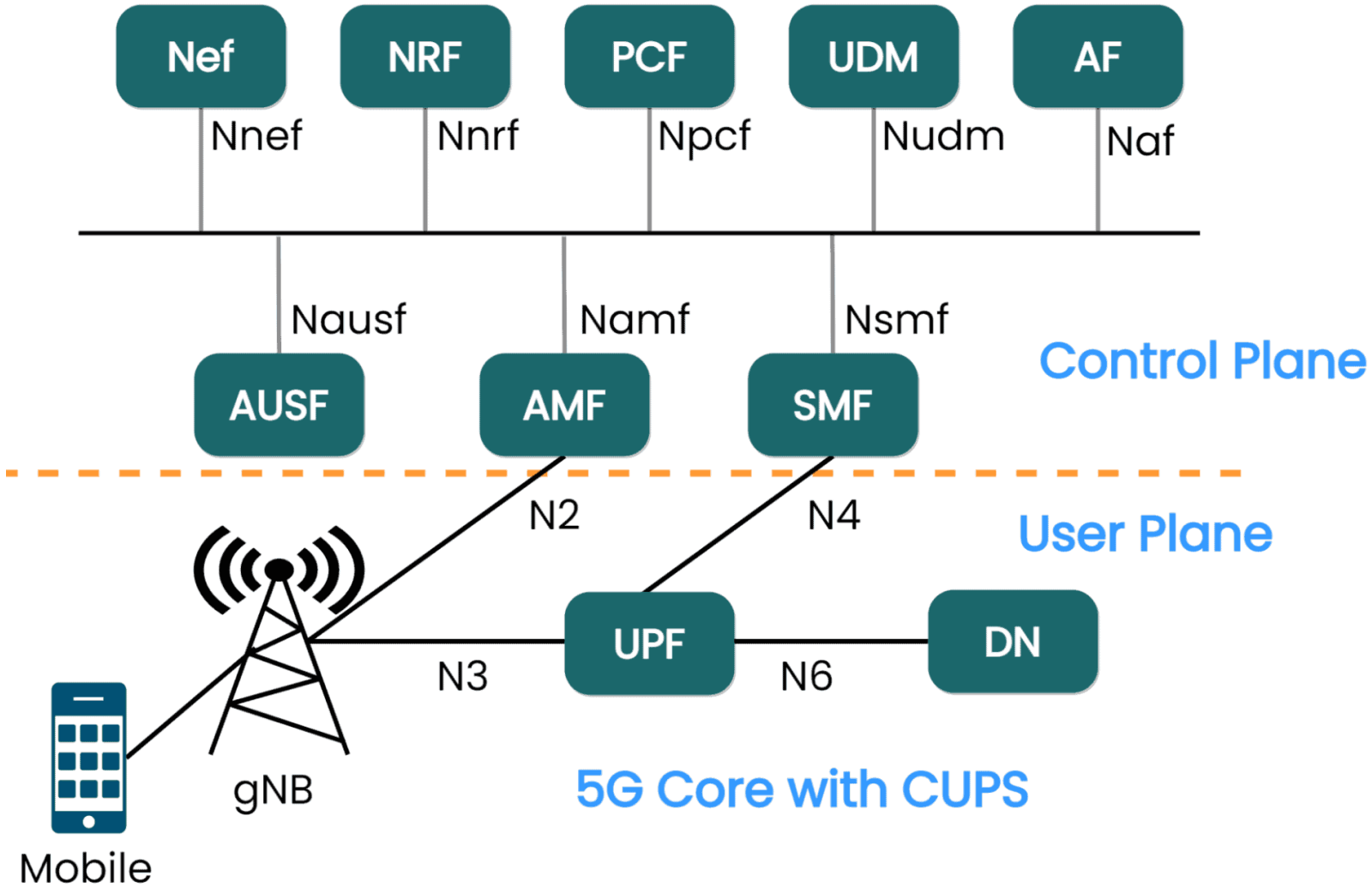
- Complete the model(s) deployment and Integrate them in the pipeline.

- Assess whether it's feasible to have a good attack categorizing model. (Will we ever have enough data?)

- Assess whether to cold start the system or prepopulate the ML databases with sample data.

5. 5G Core Integration

5G Integration



Scan the QR code to check our documentation website.

Or click here.

THANK YOU