Machine Learning Based Propagation Loss Module for Enabling Digital Twins of Wireless Networks in ns-3

Eduardo Nuno Almeida 1, Mohammed Rushad 2, Sumanth Reddy Kota 2, Akshat Nambiar 2, Hardik L. Harti 2, Chinmay Gupta 2, Danish Waseem 2, Gonçalo Santos 1, Helder Fontes 1, Rui Campos 1, Mohit P. Tahiliani 2

1 INESC TEC, Portugal
2 University of Porto, Portugal

India
OUTLINE

Introduction

ML-based Propagation Loss (MLPL) Module

Validation of the MLPL Module

Conclusions & Future Work
INTRODUCTION

Next-generation wireless networks require validation & performance evaluation.

<table>
<thead>
<tr>
<th>SIMULATION</th>
<th>EXPERIMENTAL TESTBED</th>
</tr>
</thead>
<tbody>
<tr>
<td>✗ Medium Accuracy</td>
<td>✓ Perfect Accuracy</td>
</tr>
<tr>
<td>✓ Repeatability</td>
<td>✗ Cost &amp; Availability</td>
</tr>
<tr>
<td>✓ Simplicity</td>
<td>✗ Complexity</td>
</tr>
</tbody>
</table>

DIGITAL TWIN

- ✔ Reproduction of experimental environment in simulation
- ✔ Accuracy, simplicity and repeatability
NS-3 TRACE-BASED SIMULATION APPROACH

- Propagation loss model based on experimental network traces
  - Replicate experimental environment conditions in simulation at PHY layer
  - Repeatable & reproducible
  - Single and multiple access, SISO, MIMO and Wi-Fi channel occupancy

- Simulation setup = Experimental setup
  - Network traces collected and applied per packet
  - Can not change topology, traffic or duration
## EXISTING APPROACHES FOR VALIDATION

Existing approaches for **extreme scenarios**
(e.g., crowded scenarios, dynamic traffic demands)

<table>
<thead>
<tr>
<th></th>
<th>Pure Simulation</th>
<th>Experimental</th>
<th>Trace-Based</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>Low</td>
<td>Excellent</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Existing models are generic</td>
<td>Real results</td>
<td>Assuming setup matches</td>
</tr>
<tr>
<td><strong>Repeatability &amp; Reproducibility</strong></td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variable environment conditions</td>
<td></td>
</tr>
<tr>
<td><strong>Fast-Fading</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Simulation Setup</strong></td>
<td>Any</td>
<td>Any</td>
<td>Exact match</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Simulation setup = Experimental setup</td>
</tr>
<tr>
<td><strong>Complexity &amp; Cost</strong></td>
<td>Low</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Limited testbed availability</td>
<td></td>
</tr>
</tbody>
</table>
# EXISTING APPROACHES FOR VALIDATION

Existing approaches for **extreme scenarios**
(e.g., crowded scenarios, dynamic traffic demands)

<table>
<thead>
<tr>
<th></th>
<th>Pure Simulation</th>
<th>Experimental</th>
<th>Trace-Based</th>
<th>ML Trace-Based</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>Low</td>
<td>Excellent</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Existing models are generic</td>
<td></td>
<td>Real results</td>
<td>Assuming setup matches</td>
<td>Assuming similar conditions as traces</td>
</tr>
<tr>
<td><strong>Repeatability &amp; Reproducibility</strong></td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Variable environment conditions</td>
<td></td>
<td></td>
<td></td>
<td>Controlled by RNG seeds</td>
</tr>
<tr>
<td><strong>Fast-Fading</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Simulation Setup</strong></td>
<td>Any</td>
<td>Any</td>
<td>Exact match</td>
<td>Any</td>
</tr>
<tr>
<td>Simulation setup = Experimental setup</td>
<td></td>
<td></td>
<td></td>
<td>Any setup can be used</td>
</tr>
<tr>
<td><strong>Complexity &amp; Cost</strong></td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Limited testbed availability</td>
<td></td>
<td></td>
<td></td>
<td>ML model training</td>
</tr>
</tbody>
</table>
CONTRIBUTION

- **ML-based Propagation Loss (MLPL) module**
  - Propagation loss model for ns-3 (path loss + fast-fading)
  - ML model trained with experimental network traces

- **Digital twin of experimental wireless network environment**
  - Repeatable and reproducible
  - Any network topology, mobility pattern and duration of simulation
  - Network traces represent environment dynamics
MLPL MODULE
MLPropagationLossModel

Deterministic Path Loss
- Calculated according to distance
- Deterministic value

Stochastic Fast-Fading
- Random value according to CDF
  - Using ns-3 RNG
  - Repeatable & reproducible simulations controlled by ns-3 seed

ML models trained with experimental network traces

MLPL MODULE

HELPER SCRIPTS

**train_ml_propagation_loss_model.py**

- **Train** ML model with dataset
  - Train ML model with external ML framework
  - Save ML model in files

**run_ml_propagation_loss_model.py**

- **Run** trained ML model
  - Start external ML framework and load ML model
  - Start ns3-ai module
  - Wait for ns-3 simulation to start

**Using ns3-ai module**

- Allows using existing ML frameworks
- Avoids complex integration of ML models directly in ns-3

---

MLPL MODULE

**INTRODUCTION**

**MLPL MODULE**

**MLPL VALIDATION**

**CONCLUSIONS & FUTURE WORK**

---

**MLPL MODULE**

MLPropagationLossModel → MIPredicted ▪ Path Loss (dB)

Distance (m) → PATH LOSS

ns-3 RNG

FAST-FADING CDF

PATH LOSS

ML External Framework

**CONCLUSIONS & FUTURE WORK**

**MLPL MODULE**

MLPropagationLossModelNs3AIDL → MIPredicted ▪ Path Loss (dB)

**MLPL VALIDATION**

**INTRODUCTION**

**MLPL MODULE**

**CONCLUSIONS & FUTURE WORK**
**MLPL DATASET FORMAT**

### Simple Data Format
- Distance (m)
- Propagation loss (dB)
  - Path loss + fast-fading

### Raw Data Format
- Nodes coordinates (m)
- Tx power (dBm)
- Antenna gains (dBi)
- Channel frequency (MHz)
- SNR (dB)

### Data pre-processing
- **Isolate** path loss from fast-fading
- Assuming fast-fading modelled as Normal distribution with $\mu = 0$

- **Conversion** to Simple Data Format
MACHINE LEARNING BASED PROPAGATION LOSS MODULE FOR ENABLING DIGITAL TWINS OF WIRELESS NETWORKS IN NS-3

Eduardo Nuno Almeida et al. | Workshop on ns-3 (WNS3) 2022

INTRODUCTION

MLPL MODULE

MLPL VALIDATION

CONCLUSIONS & FUTURE WORK

MLPL MODULE ACCURACY

Compare Propagation Loss Values

- MLPL Module
- MLPL Prediction
- MLPL’s Accuracy
- Existing Propagation Loss Model’s Accuracy
- Real Value
- Baseline Value
- Distance
- Network Traces
- Existing Propagation Loss Model
MLPL MODULE ACCURACY

TRAINING STRATEGIES

Extrapolation Training Strategy

Interpolation Training Strategy

Full Set Training Strategy

Different datasets for training and testing, collected on the same environment

MLPL MODULE ACCURACY
EXPERIMENTAL SET-UP

- **Wireless Network**
  - IEEE 802.11a
  - Tx Power: [0 dBm, 12 dBm]
  - Antenna Gain: -7 dBi
  - Channel: 5220 MHz (20 MHz)
  - Warehouse Environment

- **Traffic Generated**
  - Distance: [2.07 m, 24.09 m]
  - 54 Mbit/s UDP Constant Bitrate
  - Packet Size: 1400 Bytes

- **Nodes & Models**
  - 1 Fixed Node + 1 Mobile Node
  - ML Models: SVR and XGB
  - Existing Models: Friis and Log-dist. + Jakes fast-fading
MLPL MODULE ACCURACY
EXTRAPOLATION TRAINING STRATEGY

**XGB Most Accurate Model**

**SVR too Optimistic**

---

**INTRODUCTION**

**MLPL MODULE**

**VALIDATION**

**CONCLUSIONS & FUTURE WORK**
MLPL MODULE ACCURACY

INTERPOLATION TRAINING STRATEGY

Accurate Models Despite Training Gaps

Requires Less Data

INTRODUCTION

MLPL MODULE

MLPL VALIDATION

CONCLUSIONS & FUTURE WORK
MLPL MODULE ACCURACY
FULL SET TRAINING STRATEGY

Most Accurate Training Strategy

XGB Able to Predict Spikes

-55
-60
-65
-70
-75

Path Loss (dB)

5 10 15 20 25

Distance (m)

-55
-60
-65
-70
-75

Path loss (dB)

5 10 15 20 25

Distance (m)
MLPL MODULE EFFECTIVENESS

Compare Network Performance

- MLPL Module
- MLPL Prediction
- MLPL’s Performance
- Real Value
- Existing Propagation Loss Model
- Baseline Value
- Existing Propagation Loss Model’s Performance

Distance

Network Traces

Existing Propagation Loss Model
MLPL MODULE EFFECTIVENESS

NS-3.35 SIMULATION PARAMETERS

<table>
<thead>
<tr>
<th>Wireless Network</th>
<th>Traffic Generated</th>
<th>Nodes &amp; Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEEE 802.11a</td>
<td>Distance: [2.07 m, 24.09 m]</td>
<td>1 Fixed Node + 1 Mobile Node</td>
</tr>
<tr>
<td>Tx Power: 7 dBm</td>
<td>54 Mbit/s UDP Constant Bitrate</td>
<td>ML Training Strategy: Full Set</td>
</tr>
<tr>
<td>Channel: 5220 MHz (20 MHz)</td>
<td>Simulation Duration: 404 s</td>
<td></td>
</tr>
<tr>
<td>Preamble Threshold: -90 dBm</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

INTRODUCTION

MLPL MODULE

MLPL VALIDATION

CONCLUSIONS & FUTURE WORK
MLPL MODULE EFFECTIVENESS

MLPL NETWORK PERFORMANCE

- **XGB**
  - Most accurate ML model
  - Reproduce data spread
  - Optimistic for longer distances

- **SVR**
  - Follow general trend
  - Too optimistic

- **Friis and Log-distance**
  - Too optimistic
  - Do not reproduce goodput spread
CONCLUSIONS

- ML-based Propagation Loss (MLPL) Module for ns-3
  - Digital twin of experimental wireless environment
  - Trained with experimental network traces
  - Repeatable, reproducible and flexible

- More accurate than existing models
  - Especially in highly dynamic scenarios
FUTURE WORK

▪ Improve ML model accuracy
▪ Consider more parameters

▪ Publish in ns-3 App Store
  – Module already available on GitLab
  – Finish user API of ML helper scripts
  – ETA: Few weeks after WNS3 2022
Machine Learning Based Propagation Loss Module for Enabling Digital Twins of Wireless Networks in ns-3

Eduardo Nuno Almeida | INESC TEC & FEUP, Portugal
eduardo.n.almeida@inesctec.pt


Acknowledgments
This article is a result of the project “DECARBONIZE – DEvelopment of strategies and policies based on energy and non-energy applications towards CARBON neutrality via digitalization for citIZEns and society” (NORTE-01-0145-FEDER-000065), supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF).