Dataset Description for AI-Driven Mobility Management

1. Overview

This dataset is generated using a system-level simulation platform which we build based on QuaDRiGa, designed to support research on AI-driven mobility management in high-speed railway (HSR) communications. The dataset is associated with the paper "AI-Driven Mobility Management for High-Speed Railway Communications: Compressed Measurements and Proactive Handover," which is currently under major revision in IEEE Vehicular Technology Magazine. The dataset consists of two experimental groups corresponding to different sections of the paper:

1. Beam-Level Mobility Management Dataset (Filename: dataset\_beam\_level\_mobility\_management): This dataset corresponds to the experiment described in Section II.B of the paper. For each mobility speed, the dataset contains:

* Training set: 41,857 samples
* Validation set: 5,232 samples
* Test set: 5,232 samples
* Sample dimension: 32 × 8 × 5 × 2
* 32: Number of beams at the transmitter
* 8: Number of beams at the receiver
* 5: Number of sampling time steps
* 2: Real and imaginary components of the complex-valued channel

1. Cell-Level Mobility Management Dataset (Filename: dataset\_cell\_level\_mobility\_management): This dataset corresponds to the experiment described in Section III.B of the paper. For each mobility speed, the dataset contains:

* Training set: 30,000 samples
* Validation set: 10,000 samples
* Test set: 10,000 samples
* Sample dimension: 10 × 32 × 21, RSRP (Reference Signal Received Power) recorded for each beam
* 10: Sampling time steps
* 32: Number of beams per cell
* 21: Number of cells

The dataset is valuable for research in AI/ML-based mobility management, including beam tracking, handover optimization, and proactive decision-making in dynamic HSR environments.

2. System-Level Simulation Platform Configuration

The dataset is generated using a QuaDRiGa-based system-level simulation platform which we build, with the following parameter settings:

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Channel Type | UMa |
| Bandwidth | 80MHz |
| Subcarrier Bandwidth | 120KHz |
| Number of Subcarriers | 64 |
| Carrier Frequency | 30GHz |
| Base Station Antenna Configuration | （M, N, P, Mg, Ng）=（4, 8, 1, 1, 1） |
| Base Station Antenna Downtilt | 12° |
| Base Station Height | 25m |
| User Equipment (UE) Antenna Configuration | （M, N, P, Mg, Ng）=（2, 1, 1, 1, 1） |
| UE Height | 1.5m |
| UE Antenna Downtilt | 0° |
| UE Distribution | 100% outdoor |
| UE Speed | 60km/h, 120km/h, 350km/h, 500km/h |
| Base Station Transmit Power | 40dbm |
| Number of Sites | 7 |
| Number of Sectors per Site | 3 |
| Number of UEs | 1 |
| Sampling Interval | 40ms |
| Inter-Site Distance | 200m |

3. Dataset cloud storage link

Dataset Download Link: [https://pan.quark.cn/s/f11a09a732e0](https://pan.quark.cn/s/f11a09a732e0" \t "_new)  
Access Code: X8Xq

4. Usage and Citation

Researchers using this dataset are encouraged to cite the associated paper (if accepted). This dataset can be applied to various AI/ML models for improving HSR communication reliability, predictive mobility management, and proactive handover mechanisms.

For further inquiries or contributions, please refer to the GitHub repository where this dataset is hosted.