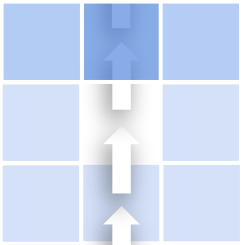
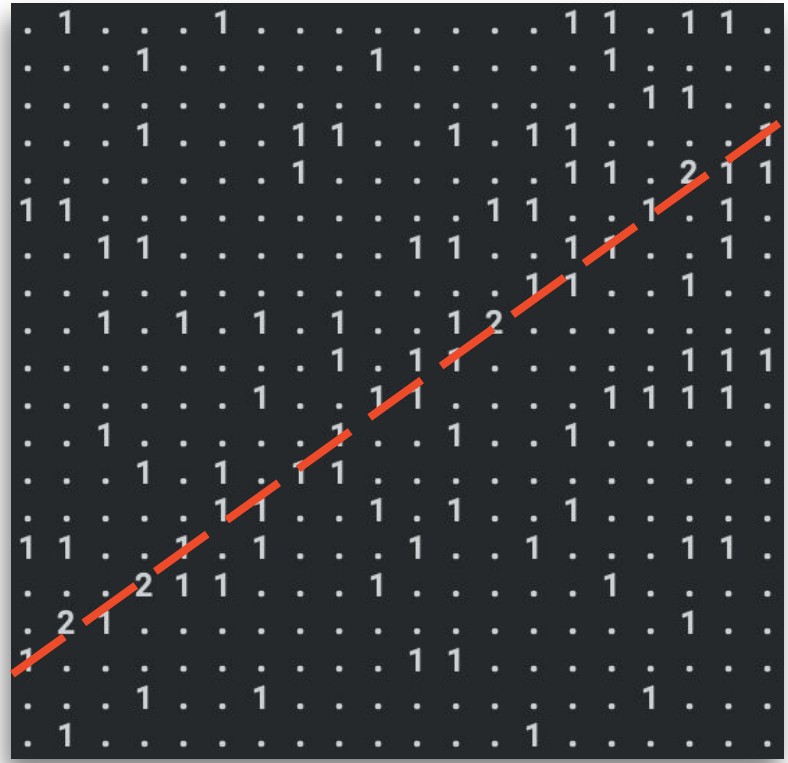


# Noise-Tolerant Track Finding with Language Models

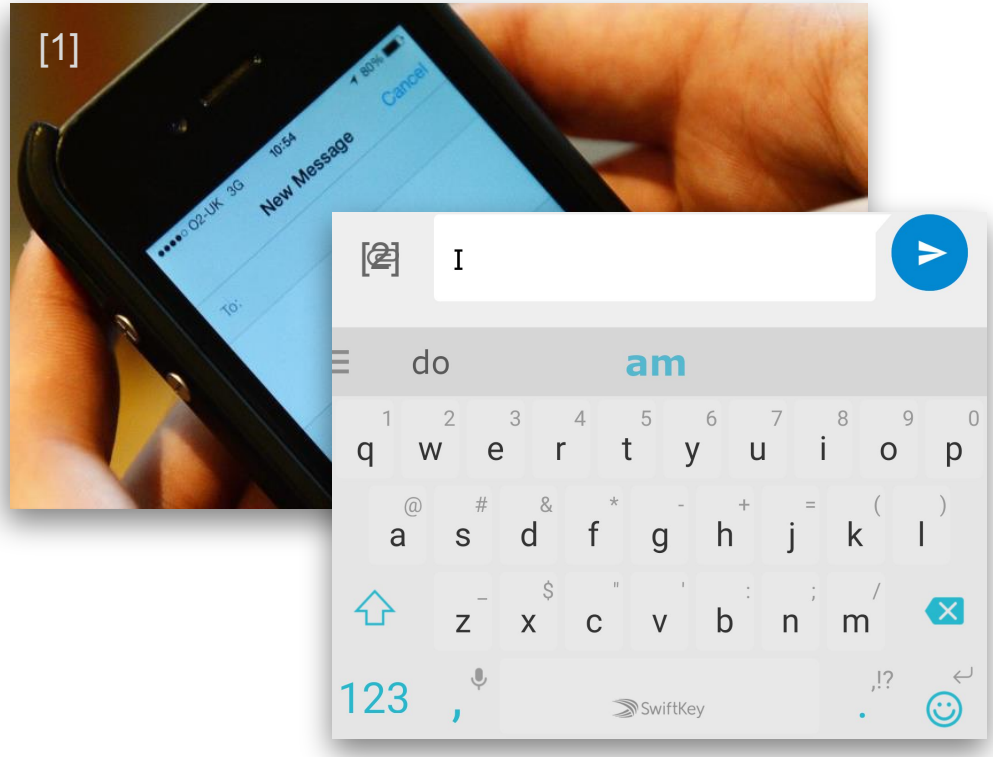
Jakapat Kannika
Forschungszentrum Jülich



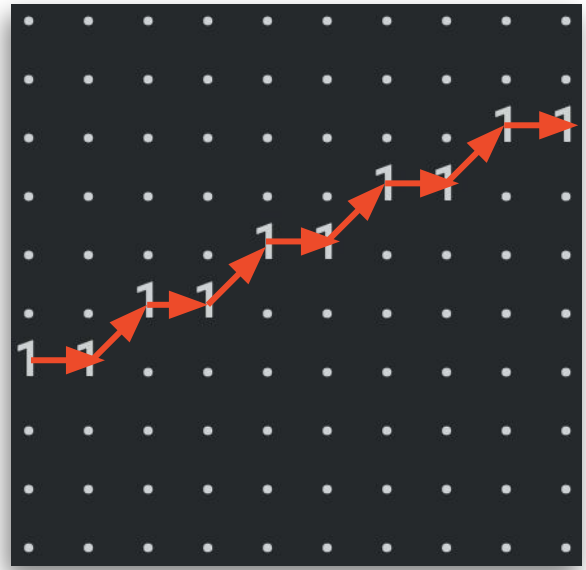
# Reducing noise in reconstructed tracks

# Language model: Next word prediction

“I am Sam”
“Sam I am”
“I do not like green eggs and ham”



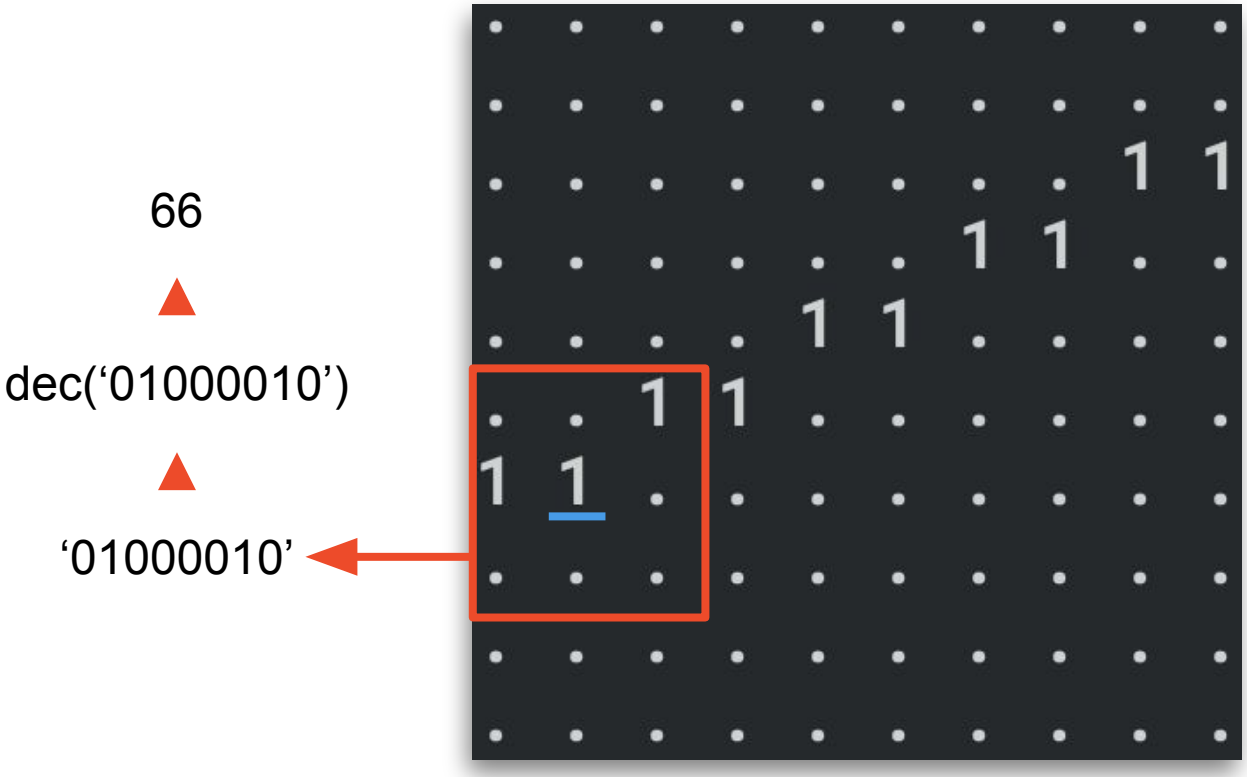

Sources:
[1] http://www.androidpolice.com
[2] http://qocall.com



Moving direction tokens:

90 45 90 45 90 45 90 45 90

Learning features: moving directions.




- 66 is a *neighbor pattern id*.
- There are 255 patterns (excluding one that has zero neighbor point).

# Track finding models

Trained language models:

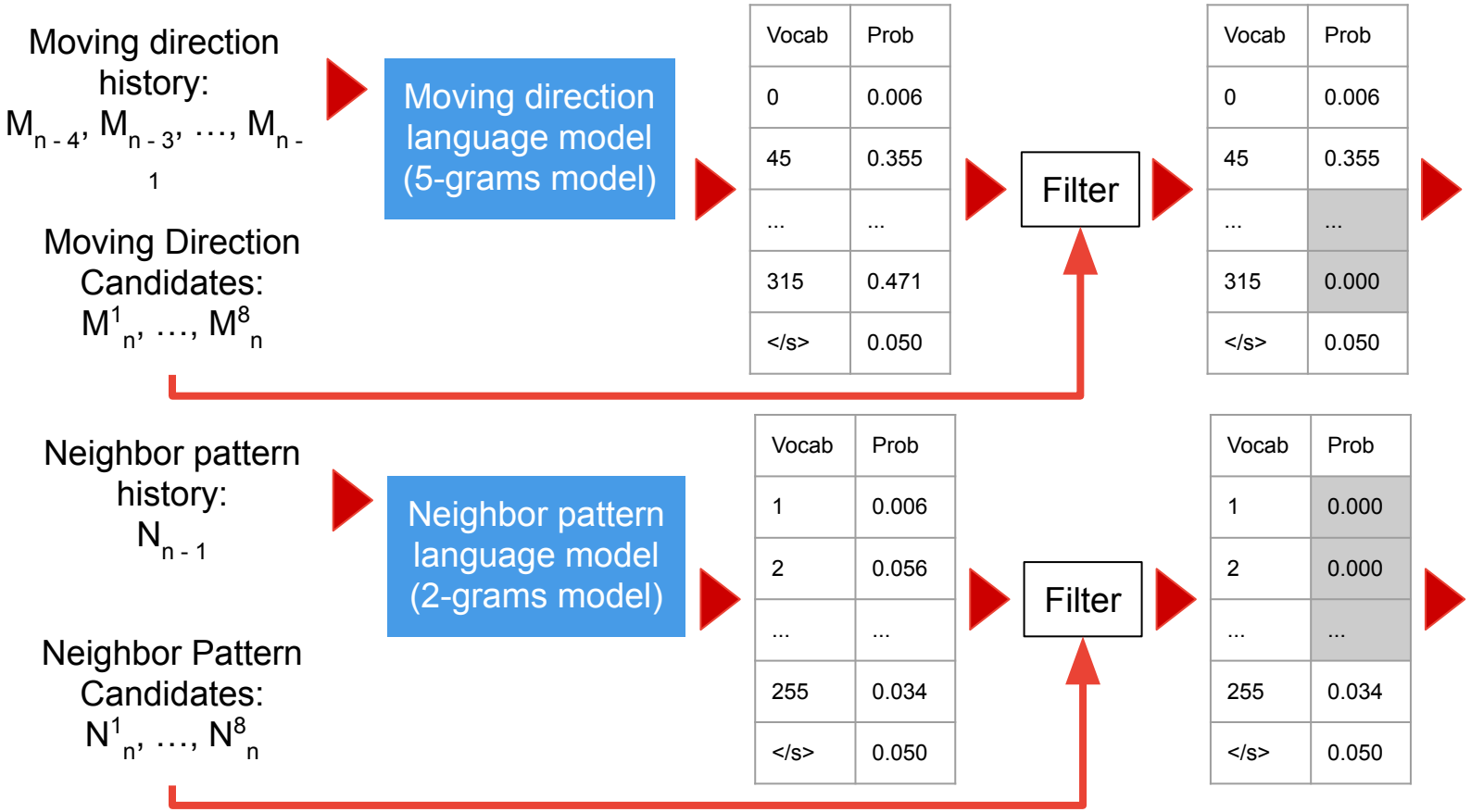
✓ *2-grams neighbor pattern model,
✓ 5-grams, 10-grams, 15-grams moving direction models.

Target MLP trainings:

- 5-grams moving direction model + 2-grams neighbor pattern model,
- 10-grams moving direction model + 2-grams neighbor pattern model,
- 15-grams moving direction model + 2-grams neighbor pattern model.

*The type of the 2-grams model is 2-skip-2-grams.

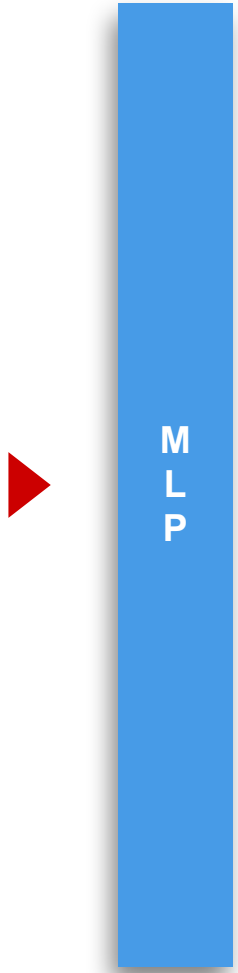
# Track finding using Multi-Layer Perceptrons (MLPs)




MLP training: data preprocessing.

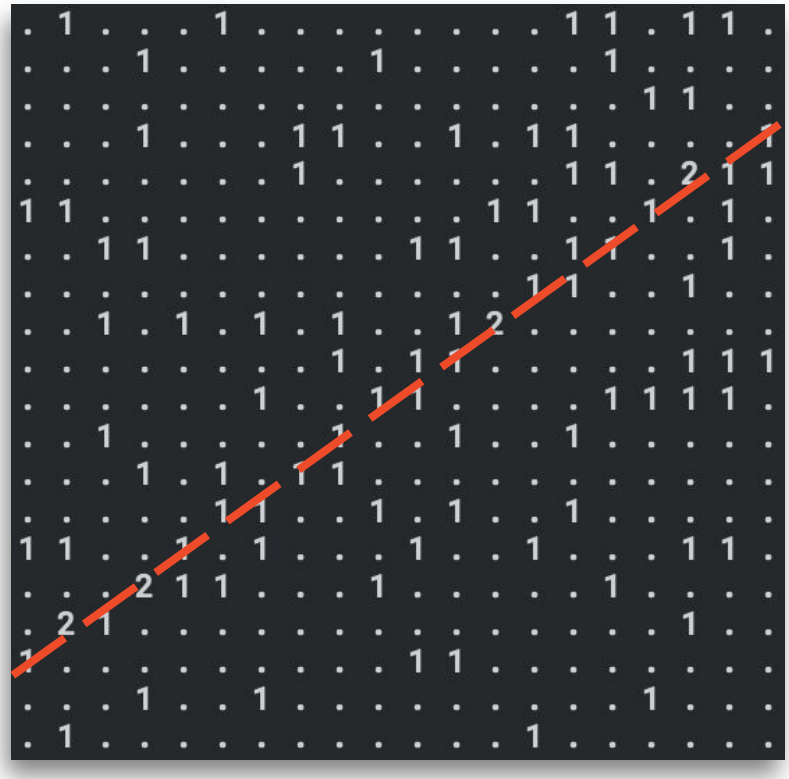
```
neighpat_outputs = [
        0.0, 0.0, …,
        ..., 0.015356
]  # Length : 256
```

```
movdir_outputs = [
        0.0, 0.097816, 0.0,
        0.092013, 0.0, 0.0,
        0.0, 0.0, 0.028828
]  # Length: 9
```



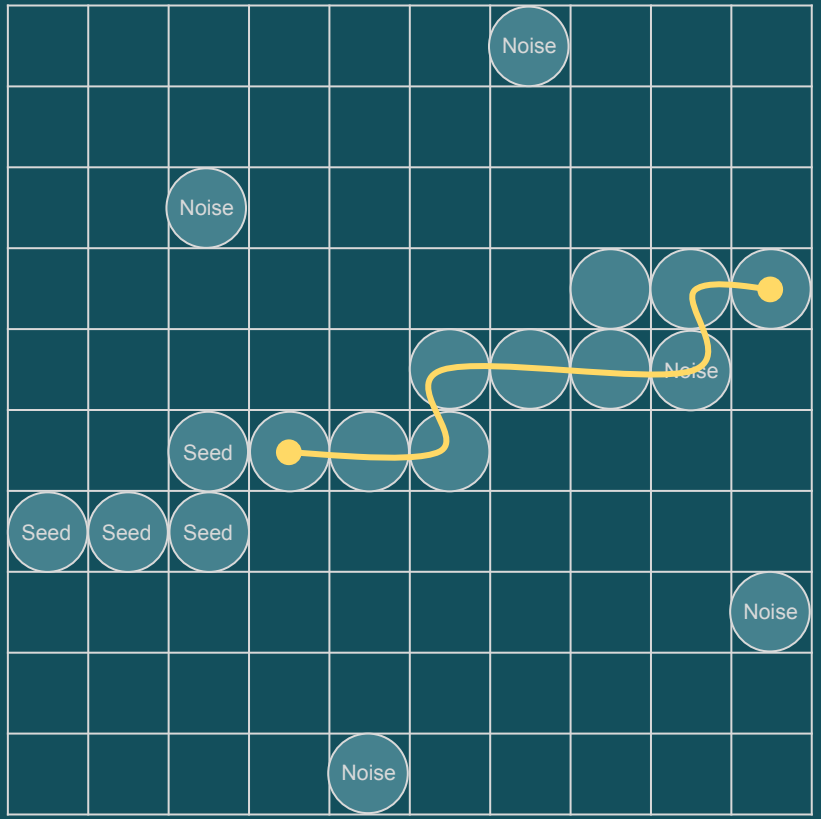

```
mlp_outputs = [
        0, 0, 0,
        1, 0, 0,
        0, 0, 0
]  # Length: 9
```

MLP training: inputs and outputs.



Track: straight line,
Number of track: 1,
Sizes of the simulation frames: 15 x 15, 20 x 20, 25 x 25,
Noise: 0 - 40%,

MLP training: sample generation.

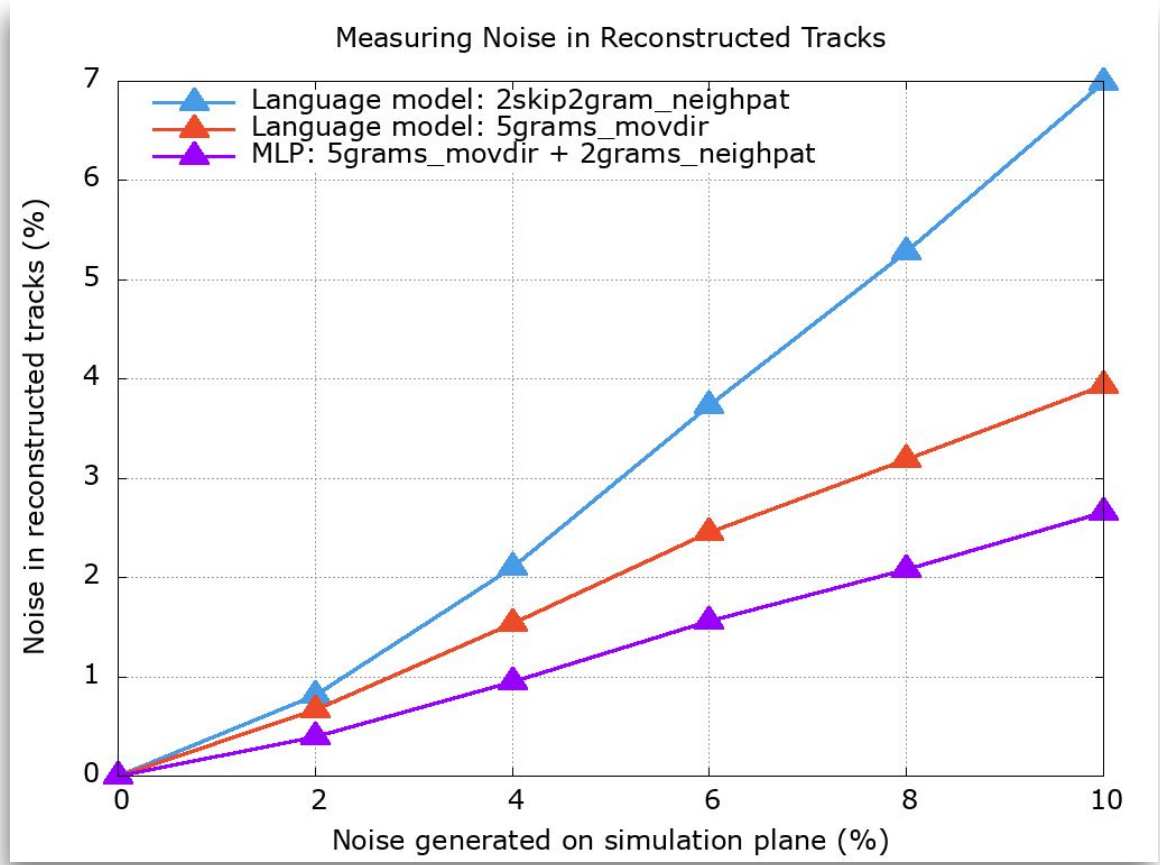


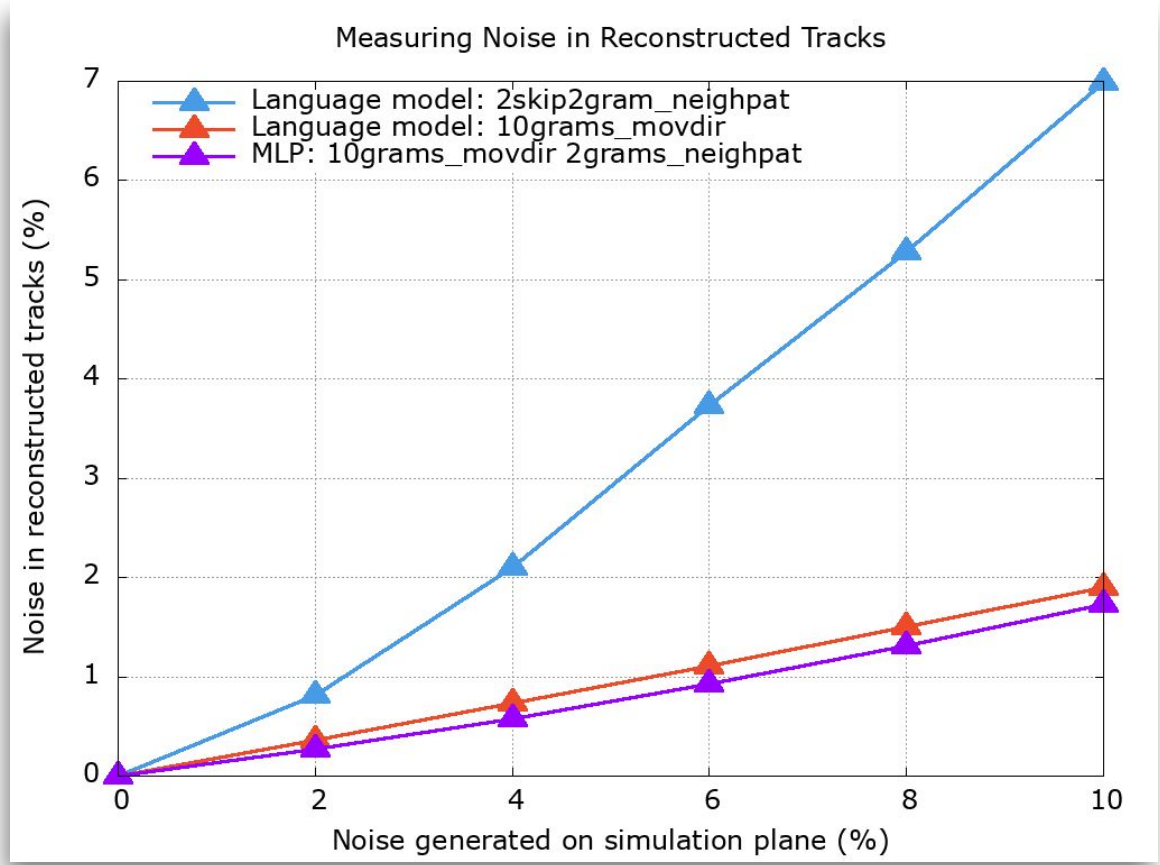

Generated noise: 5%
Track length: 9
Noise in track: 1

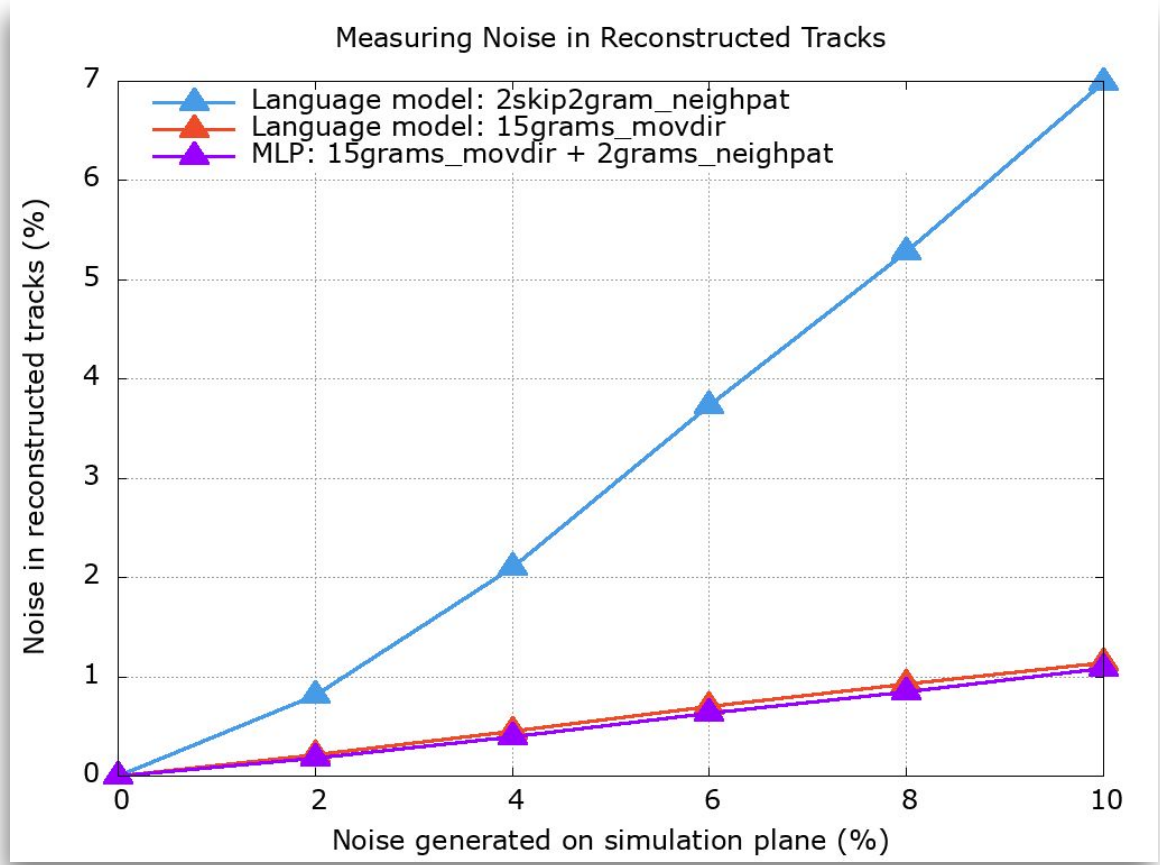
Noise in track (%):
100/9 = 11.11%

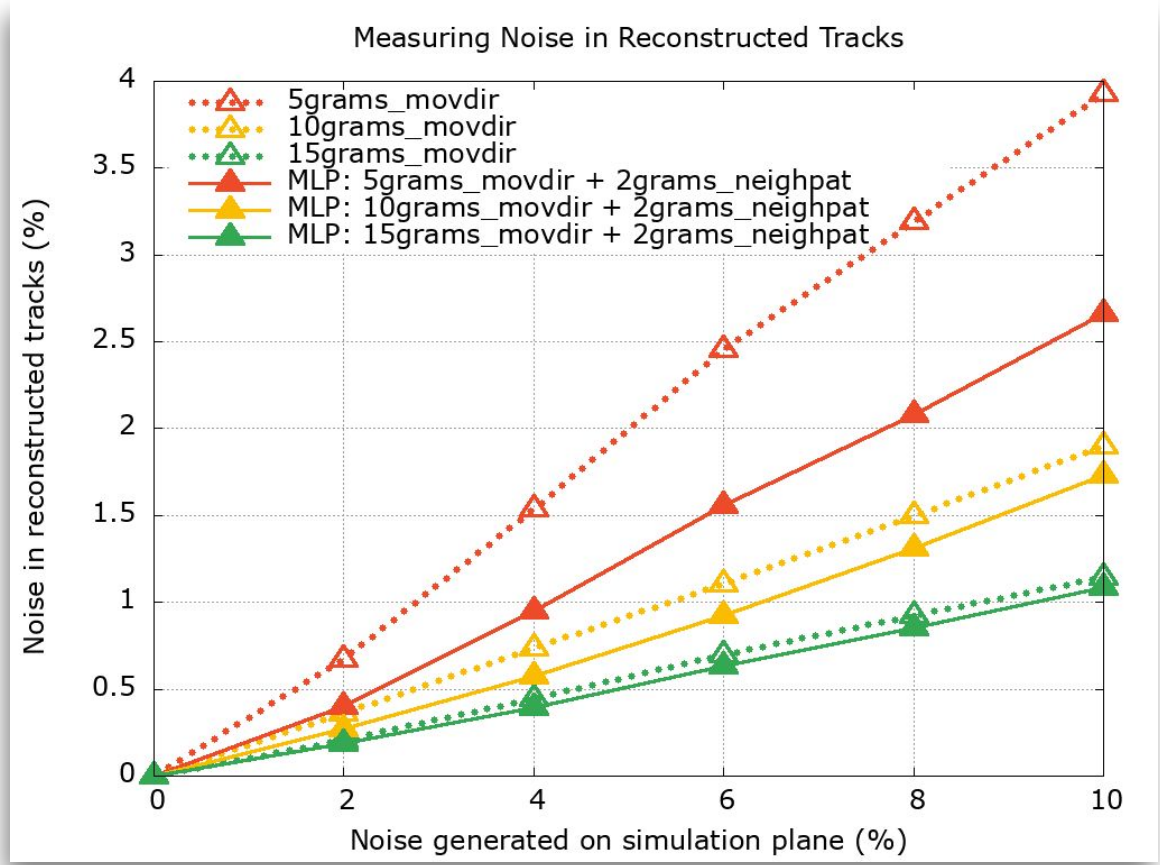
Model evaluation: calculating average noise in the reconstructed tracks.

# Results and discussions

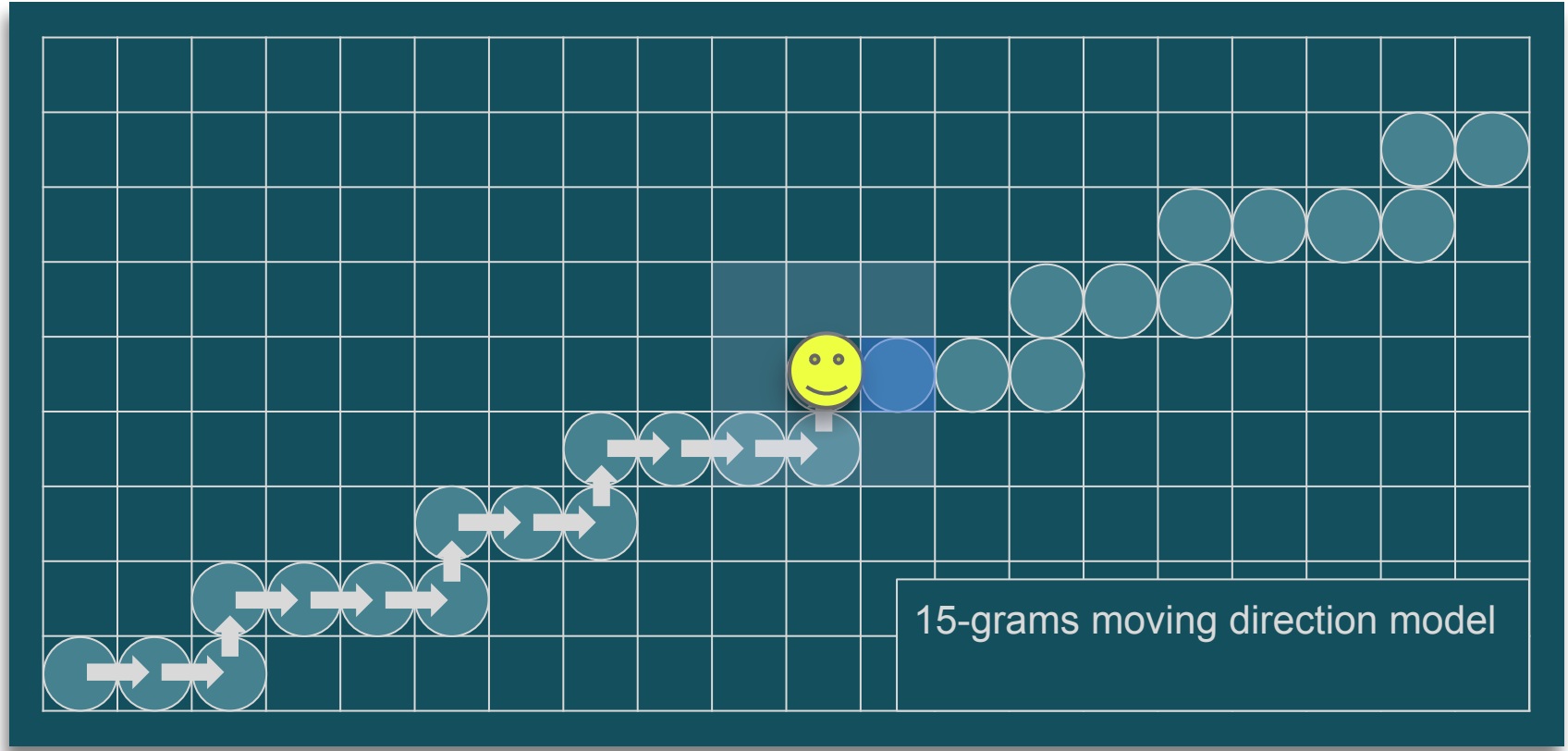

Measuring Noise in Reconstructed Tracks
Language model: 2skip2gram_neighpat
Language model: 5grams_movdir
MLP: 5grams_movdir + 2grams_neighpat
Noise in reconstructed tracks (%)
Noise generated on simulation plane (%)
0
1
2
3
4
5
6
7
0
2
4
6
8
10

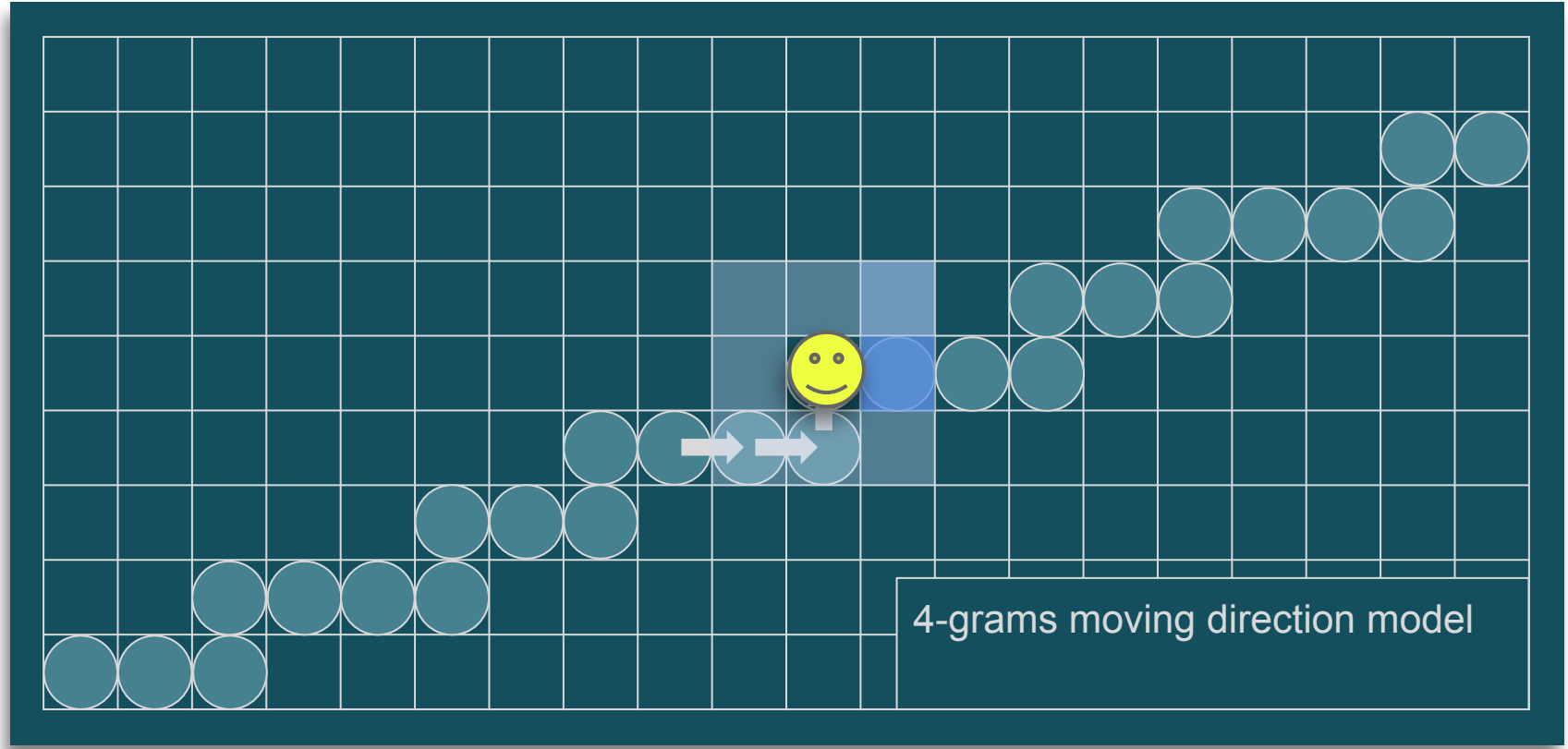

Measuring Noise in Reconstructed Tracks
Language model: 2skip2gram_neighpat
Language model: 10grams_movdir
MLP: 10grams_movdir 2grams_neighpat
Noise in reconstructed tracks (%)
Noise generated on simulation plane (%)
0
1
2
3
4
5
6
7
0
2
4
6
8
10

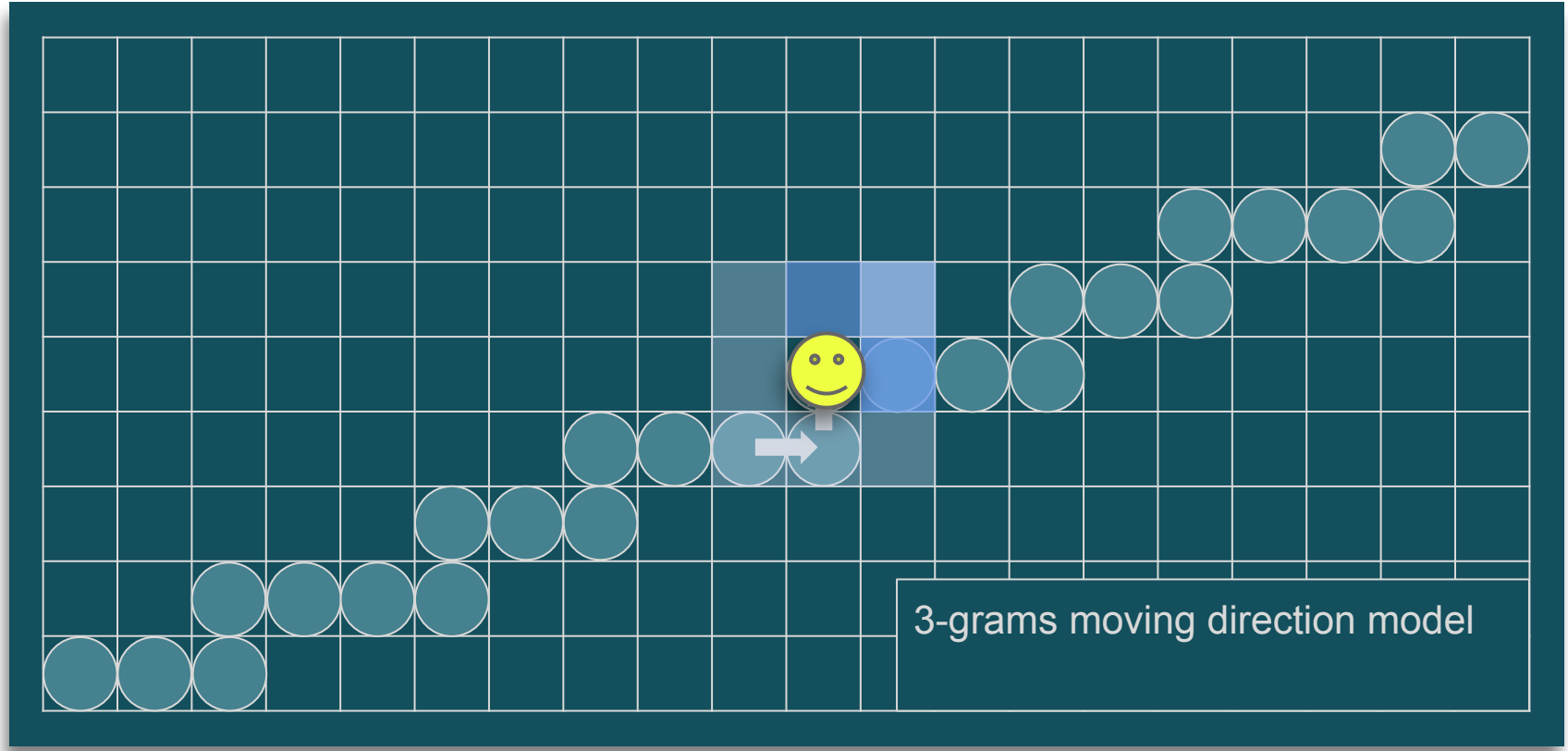

Measuring Noise in Reconstructed Tracks
Language model: 2skip2gram_neighpat
Language model: 15grams_movdir
MLP: 15grams_movdir + 2grams_neighpat
Noise in reconstructed tracks (%)
Noise generated on simulation plane (%)
0
1
2
3
4
5
6
7
0
2
4
6
8
10


Measuring Noise in Reconstructed Tracks
5grams_movdir
10grams_movdir
15grams_movdir
MLP: 5grams_movdir + 2grams_neighpat
MLP: 10grams_movdir + 2grams_neighpat
MLP: 15grams_movdir + 2grams_neighpat
Noise in reconstructed tracks (%)
Noise generated on simulation plane (%)
0
0.5
1
1.5
2
2.5
3
3.5
4
0
2
4
6
8
10

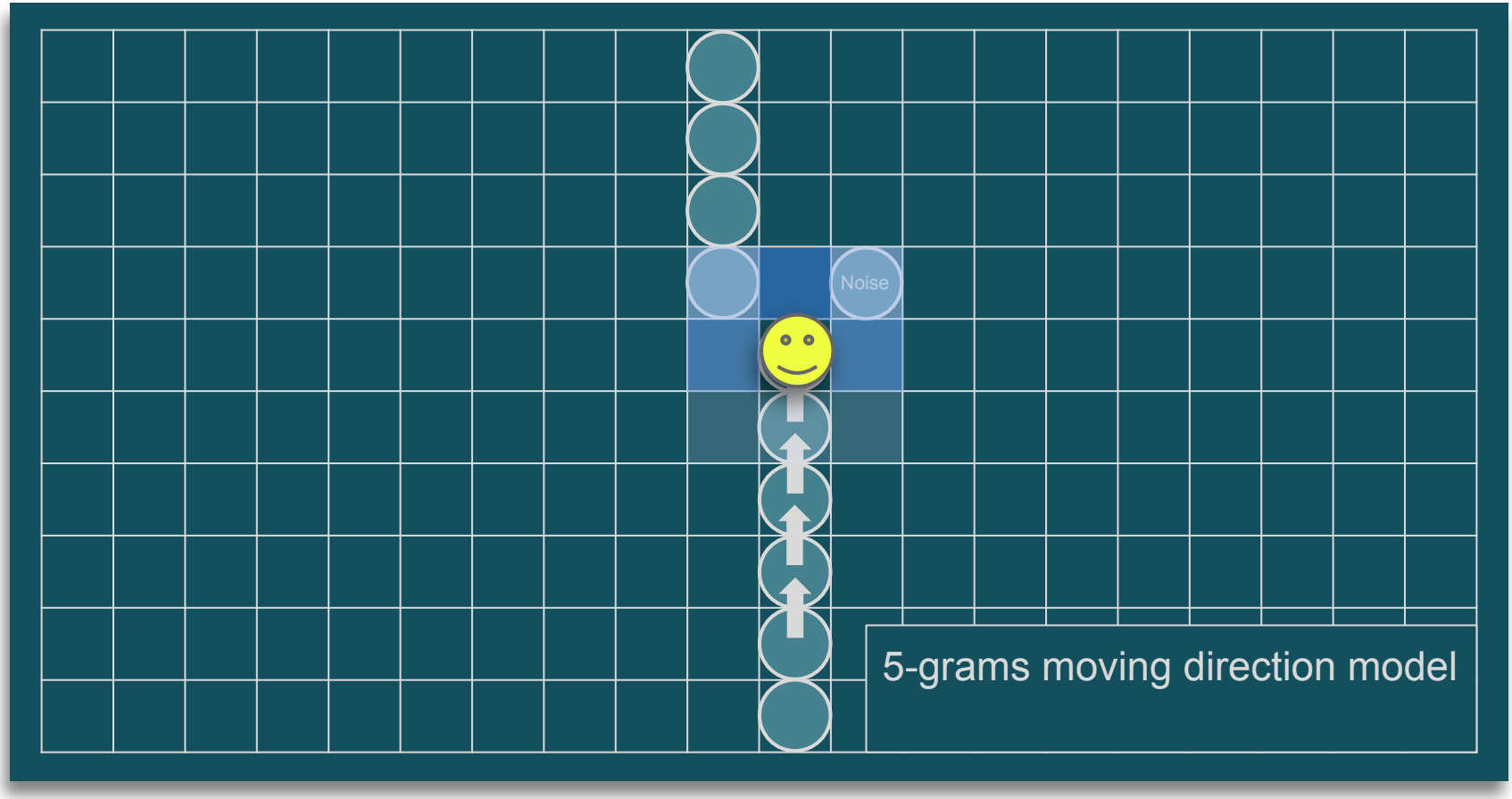
# The problem of tracking with a short history

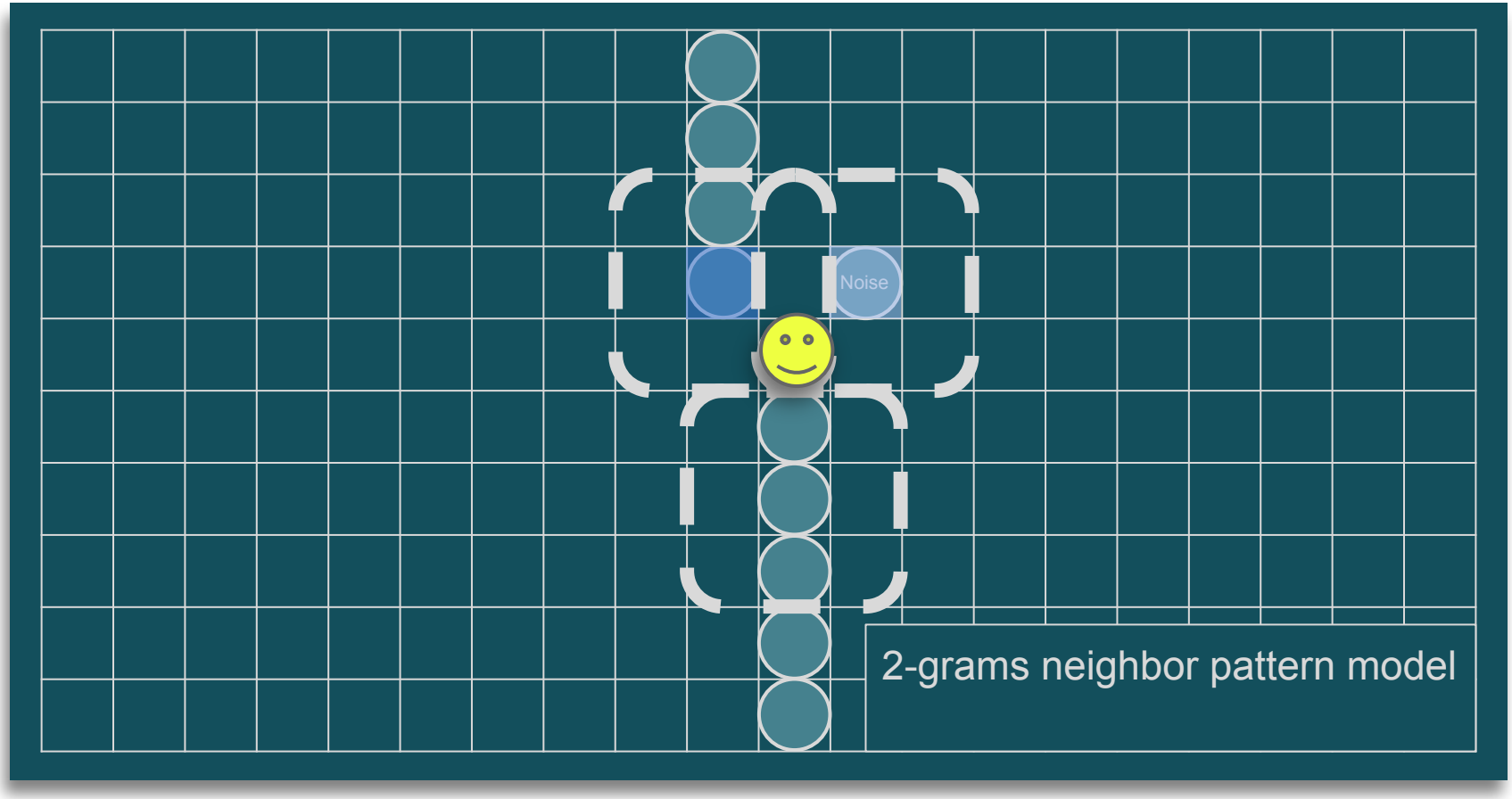

15-grams moving direction model

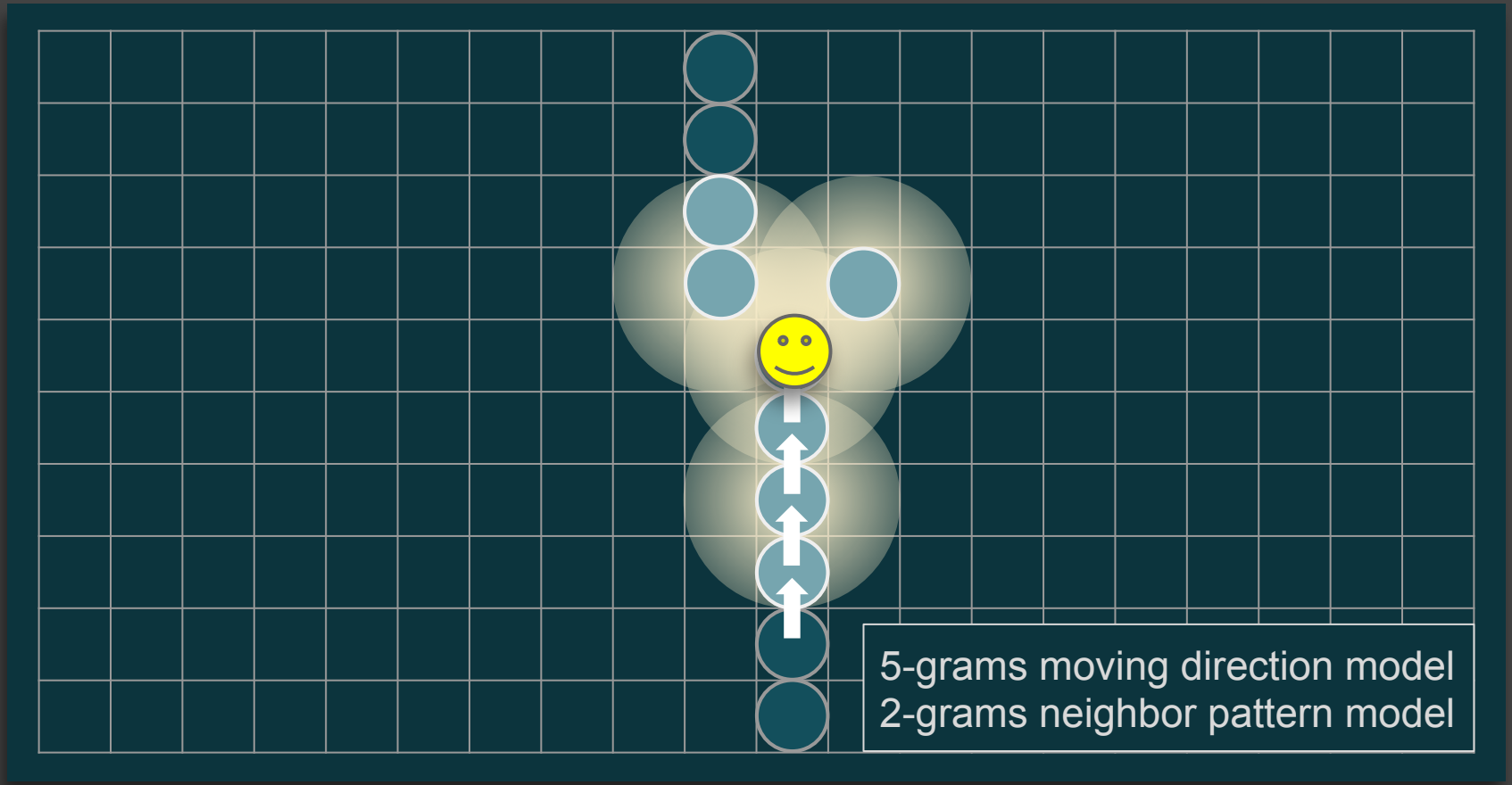

4-grams moving direction model


3-grams moving direction model

“Including neighbor pattern information can help the algorithm to eliminate more ambiguities.”


Noise
5-grams moving direction model


Noise
2-grams neighbor pattern model


5-grams moving direction model
2-grams neighbor pattern model

## Conclusions:

- MLP can correctly learn correlations between the output probabilities of the moving direction, and the neighbor pattern language models,
- The correlations between the two language models can improve noise toleration property in the track finding algorithm,
- The MLP-based track finding can effectively improve the accuracy of track findings in short-length history cases.

## Outlooks:

- Implement the same algorithms to support the data from PANDA Root,
- Include isochrone radius as one of the tracking features.

Thank you.