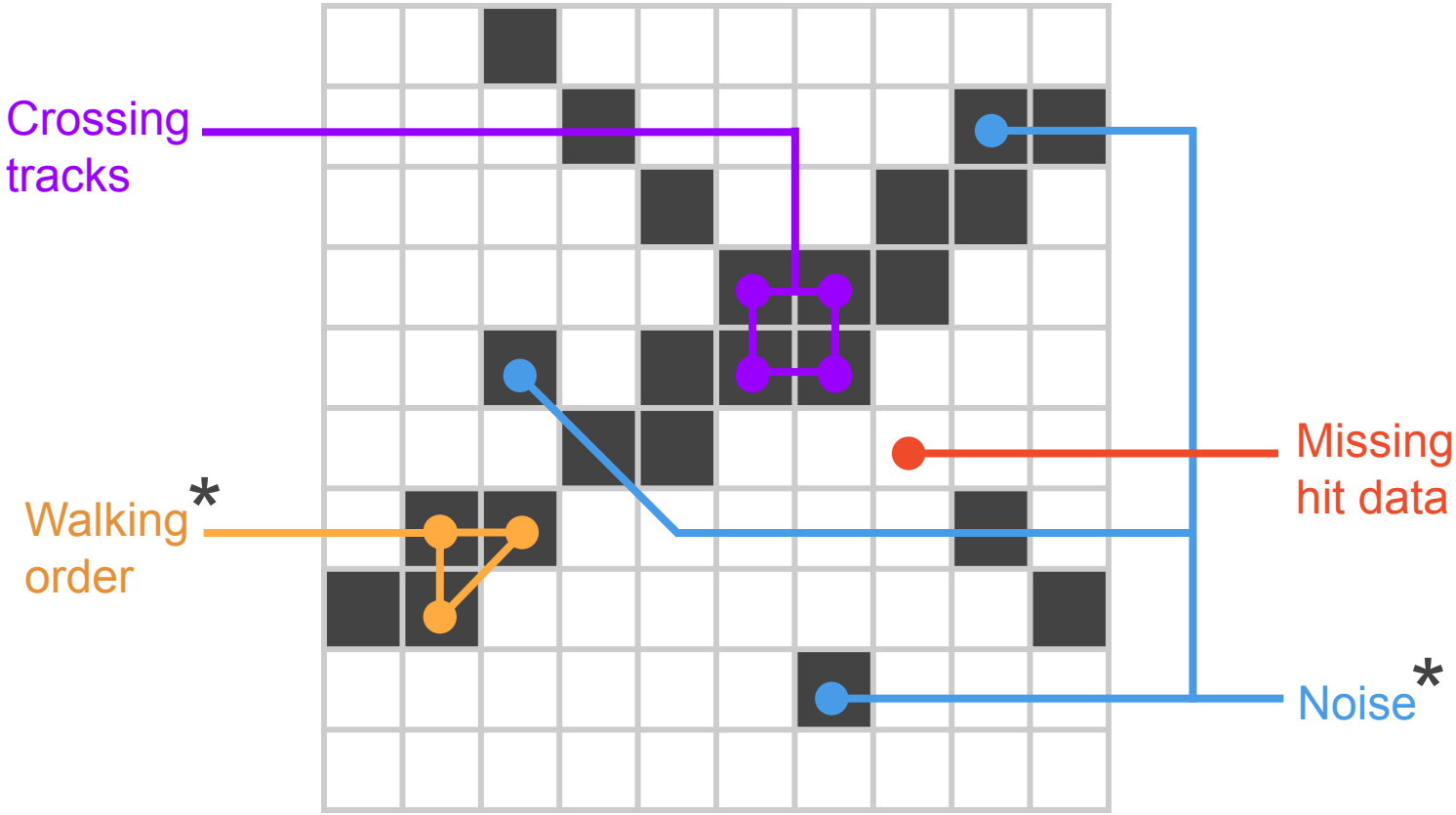
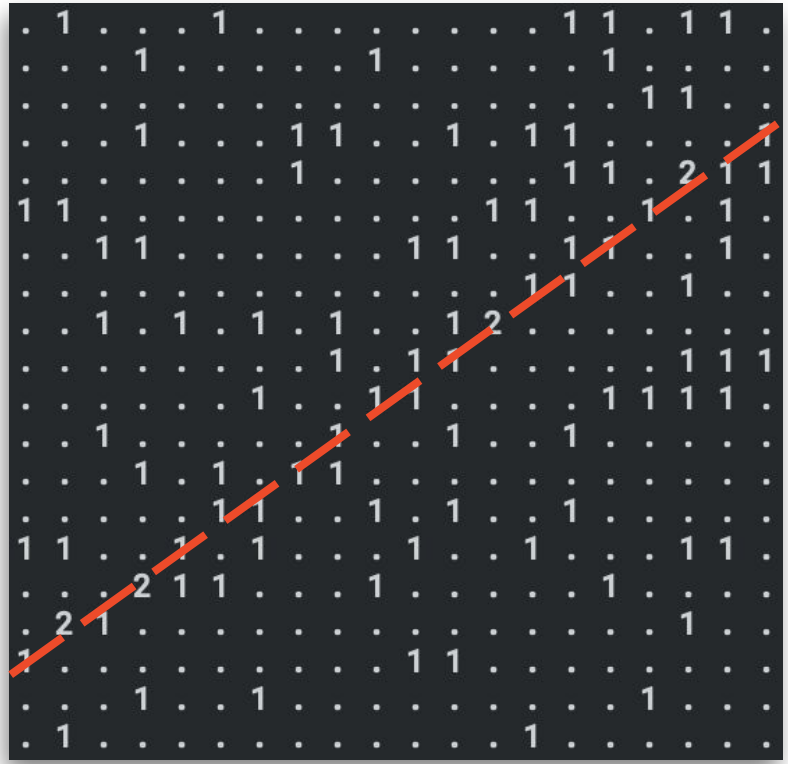
# Track Finding Using a Language Model

Jakapat Kannika
Forschungszentrum Jülich




*Our current work is focusing on these problems on the STT.

What makes track finding difficult?



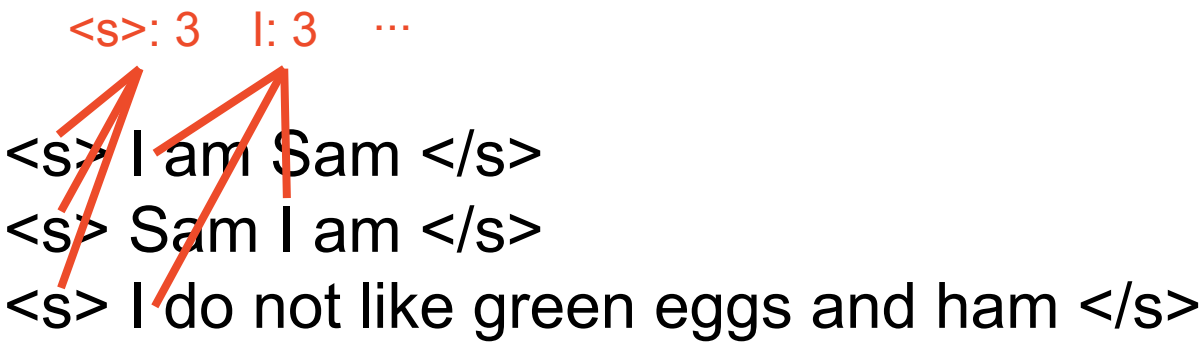
How to find a track from continuous hits
in the presence of noise?

# Language Models

# Example: 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

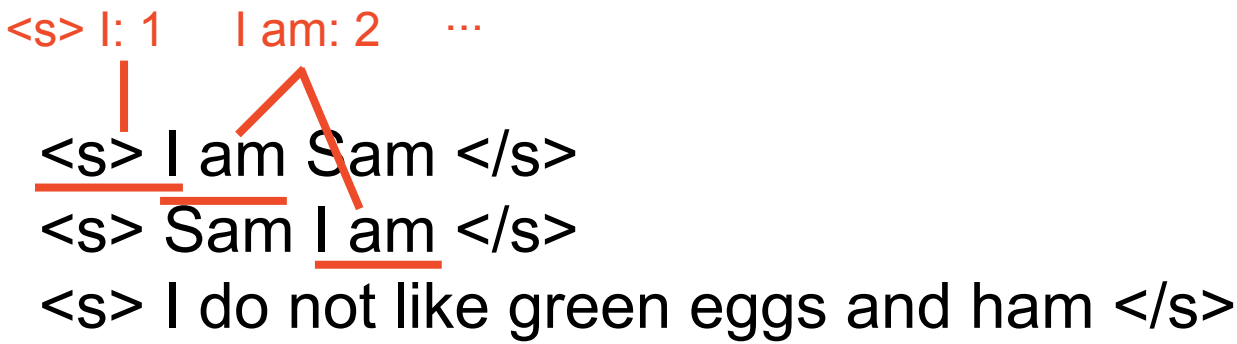

<s>: 3
I: 3
···
<s> I am Sam </s>
<s> Sam I am </s>
<s> I do not like green eggs and ham </s>

# Unigram Model

| Word | Count |
|---|---|
| <s> | 3 |
| I | 3 |
| am | 2 |
| </s> | 3 |
| Sam | 2 |
| do | 1 |

| Word | Count |
|---|---|
| not | 1 |
| like | 1 |
| green | 1 |
| eggs | 1 |
| and | 1 |
| ham | 1 |

Frequency distribution for a unigram model


<s> I: 1
I am: 2
···
<s> I am Sam </s>
<s> Sam I am </s>
<s> I do not like green eggs and ham </s>

# Bigram Model

| Word | Count |
|---|---|
| <s> I | 2 |
| I am | 2 |
| am Sam | 1 |
| Sam </s> | 1 |
| <s> Sam | 1 |
| Sam I | 1 |
| am </s> | 1 |
| <s> I | 1 |

| Word | Count |
|---|---|
| I do | 1 |
| do not | 1 |
| not like | 1 |
| like green | 1 |
| green eggs | 1 |
| eggs and | 1 |
| and ham | 1 |
| Ham </s> | 1 |

Frequency distribution for a bigram model

# Finding a probability distribution of the bigram model

$$P(w_n | w_{n-1}) = C(w_{n-1}w_n) / C(w_{n-1})$$

Probability of which $w_{n-1}$ could be followed by $w_n$

Bigram count for the word $w_{n-1}w_n$

Unigram count for the word $w_{n-1}$

| Word | Prob. |
|---|---|
| P(<s> \| I) | 2/3 = 0.67 |
| P(am \| I) | 2/3 = 0.67 |
| P(Sam \| am) | 1/2 = 0.5 |
| P(</s> \| Sam) | 1/2 = 0.5 |
| P(Sam \| <s>) | 1/3 = 0.33 |
| P(I \| Sam) | 1/2 = 0.5 |
| P(</s> \| am) | 1/2 = 0.5 |
| P(I \| <s>) | 1/3 = 0.33 |

| Word | Prob. |
|---|---|
| P(do\| I) | 1/3 = 0.33 |
| P(not \| do) | 1/1 = 1 |
| P(like \| not) | 1/1 = 1 |
| P(green \| like) | 1/1 = 1 |
| P(eggs \| green) | 1/1 = 1 |
| P(and \| eggs) | 1/1 = 1 |
| P(ham \| and) | 1/1 = 1 |
| P(</s> \| Ham) | 1/1 = 1 |

Probability distribution for a bigram model

# What is the next word after ‘I’?

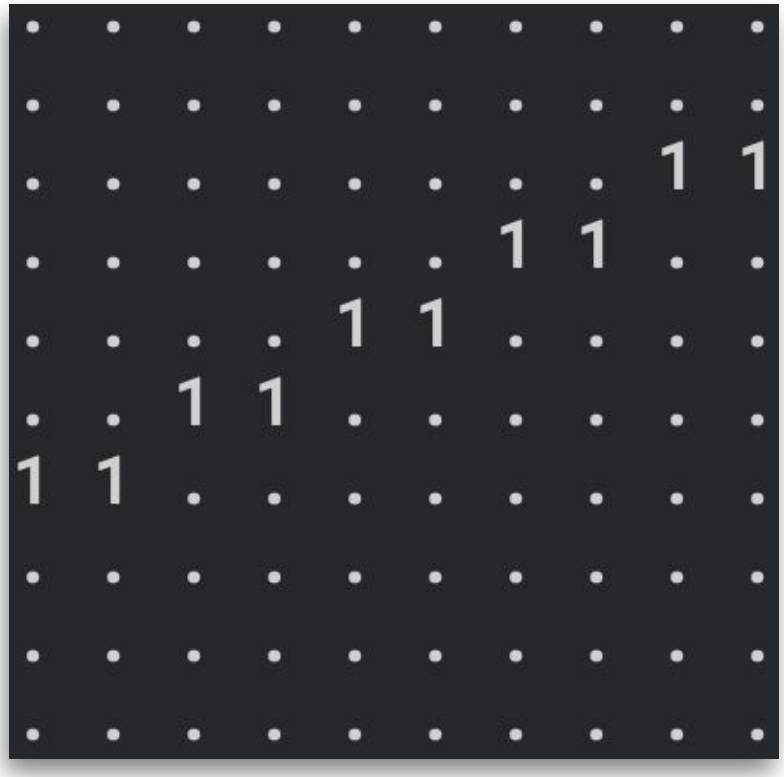
<s> I am Sam </s>
<s> Sam I am </s>
<s> I do not like green eggs and ham </s>

| **Word** | **Prob.** |
|---|---|
| P(am \| I) | 0.67 |
| P(do \| I) | 0.33 |

# How can we apply the language models to the track finding task?

(*x*, *y*) information:

[0, 3], [1, 3], [2, 4], [3, 4], [4, 5], [5, 5], [6, 6], [7, 6], [8, 7], [9, 7]



Features:

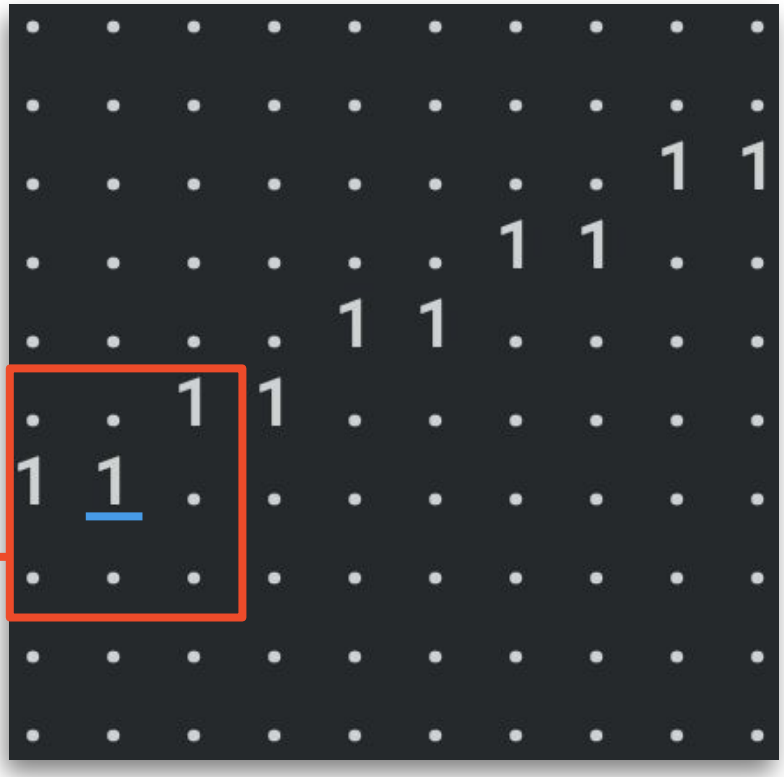
- neighbor pattern,
- moving direction.

# Neighbor Pattern Feature

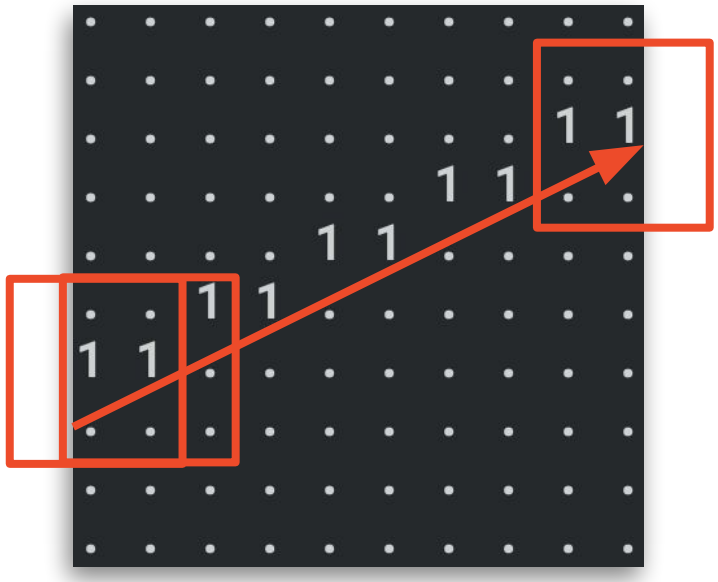
66

dec(‘01000010’)

‘01000010’



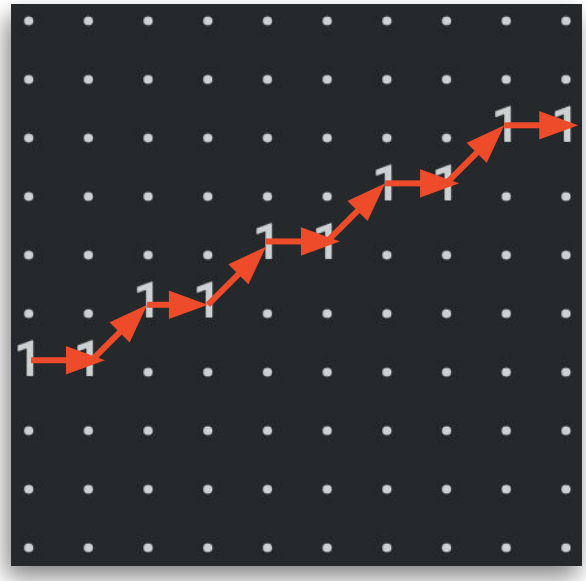
- 66 is a *neighbor pattern id*.
- There are 255 patterns (excluding one that has zero neighbor point).




Neighbor pattern tokens:

32 66 36 66 36 66 36 66 36 2

# Moving Direction Feature



Moving direction tokens:

90 45 90 45 90 45 90 45 90

# Training Language Models and Tracking Results

# Training language models

Neighbor pattern feature:

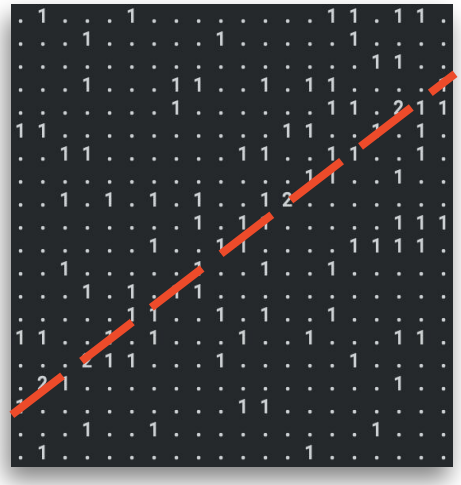
- Bigram,
- 1-skip-bigram,
- 2-skip-bigram.

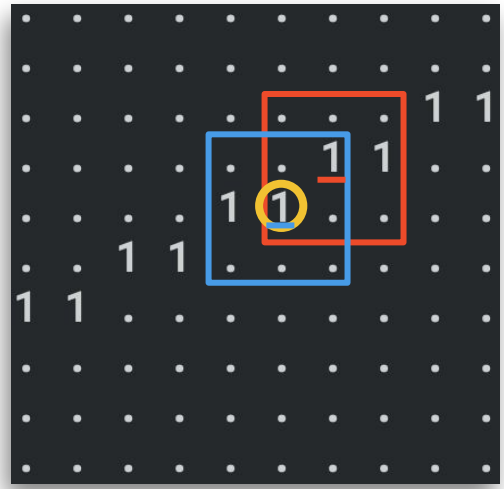
Moving direction feature:

- 5-gram,
- 10-gram,
- 15-gram.

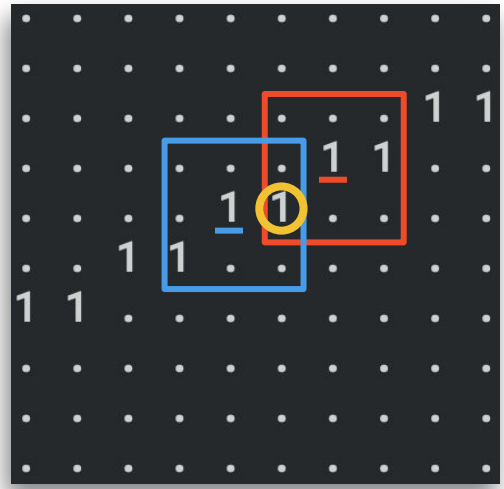
Toy data generators:

- Straight line generator,
- Noise generator.

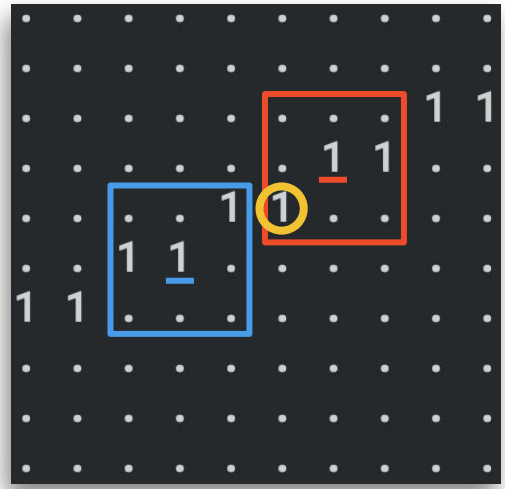




Bigram



1-skip-bigram

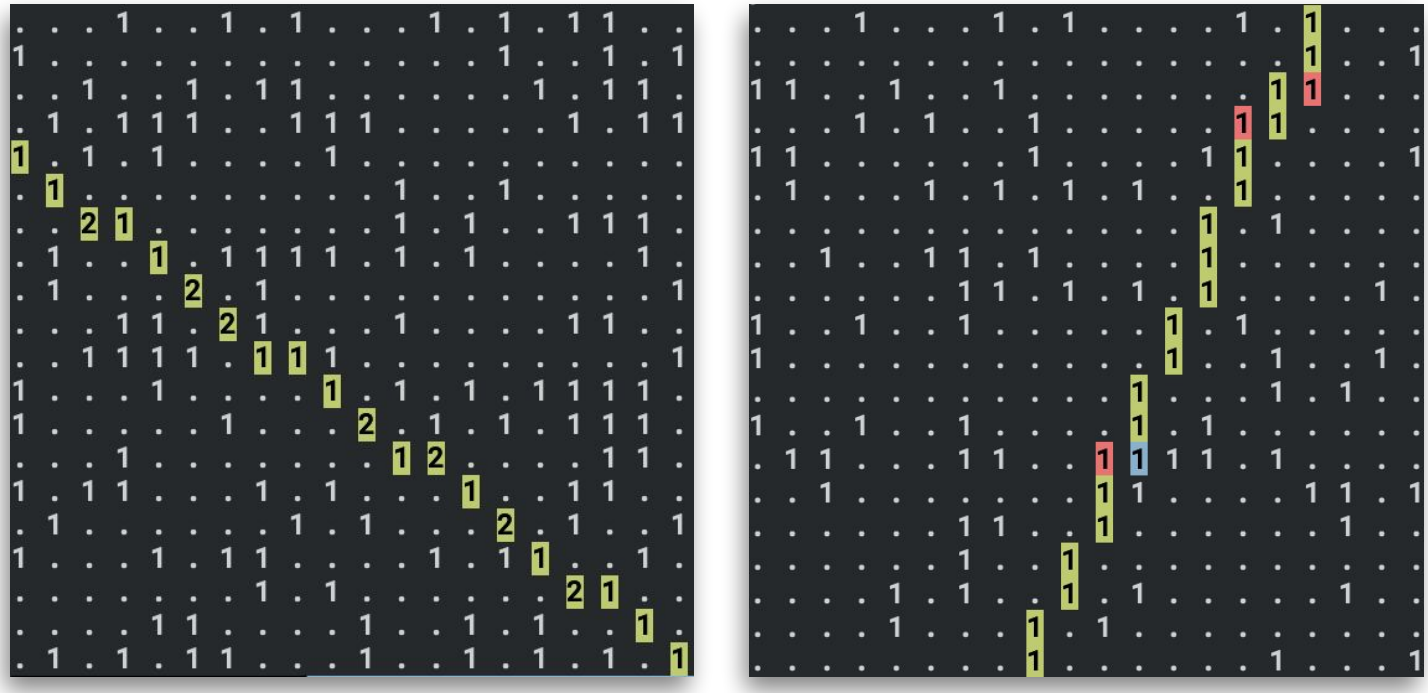


2-skip-bigram

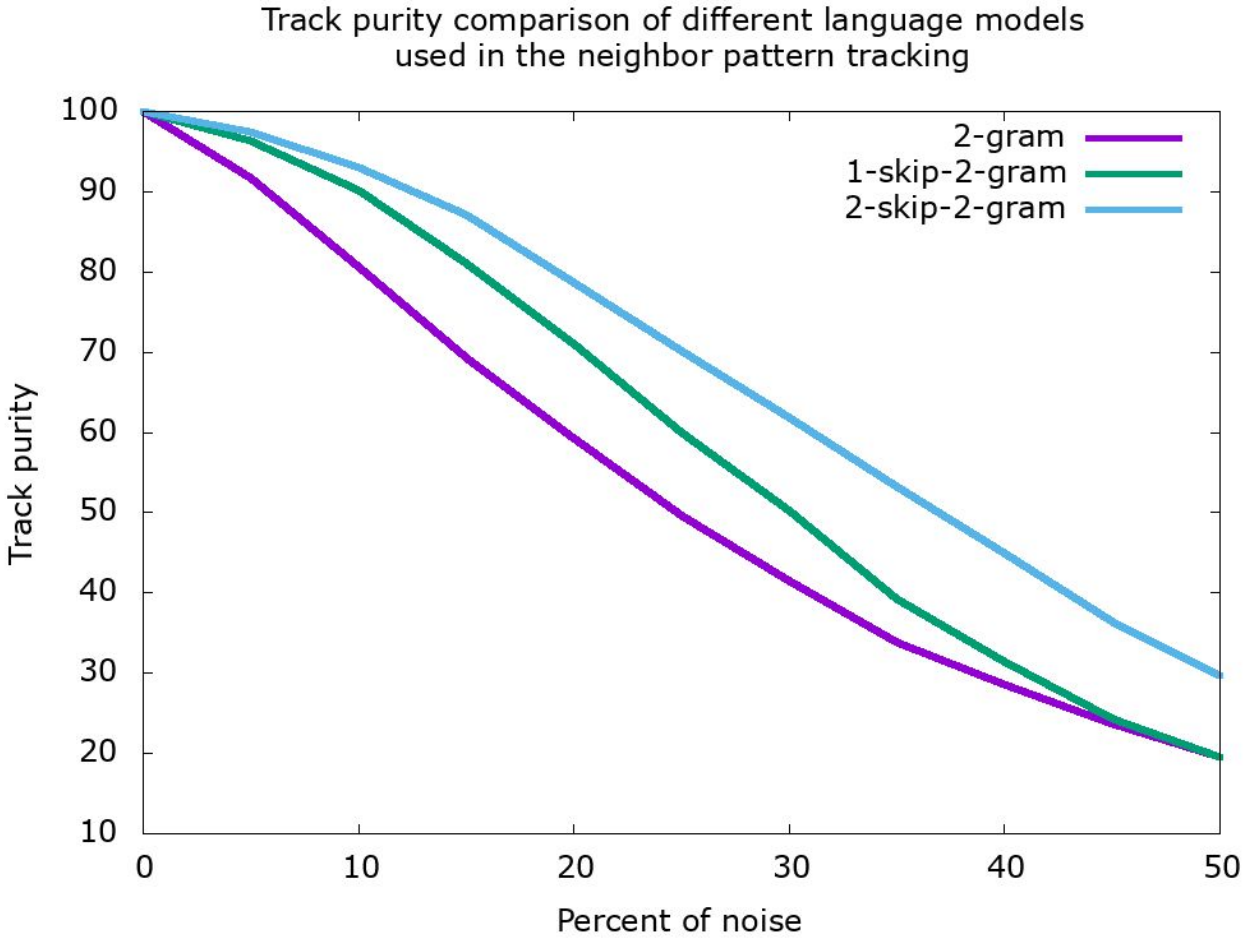
Skip-gram models for neighbor pattern feature

# Results: Neighbor Pattern Feature

# Track finding using the neighbor pattern feature

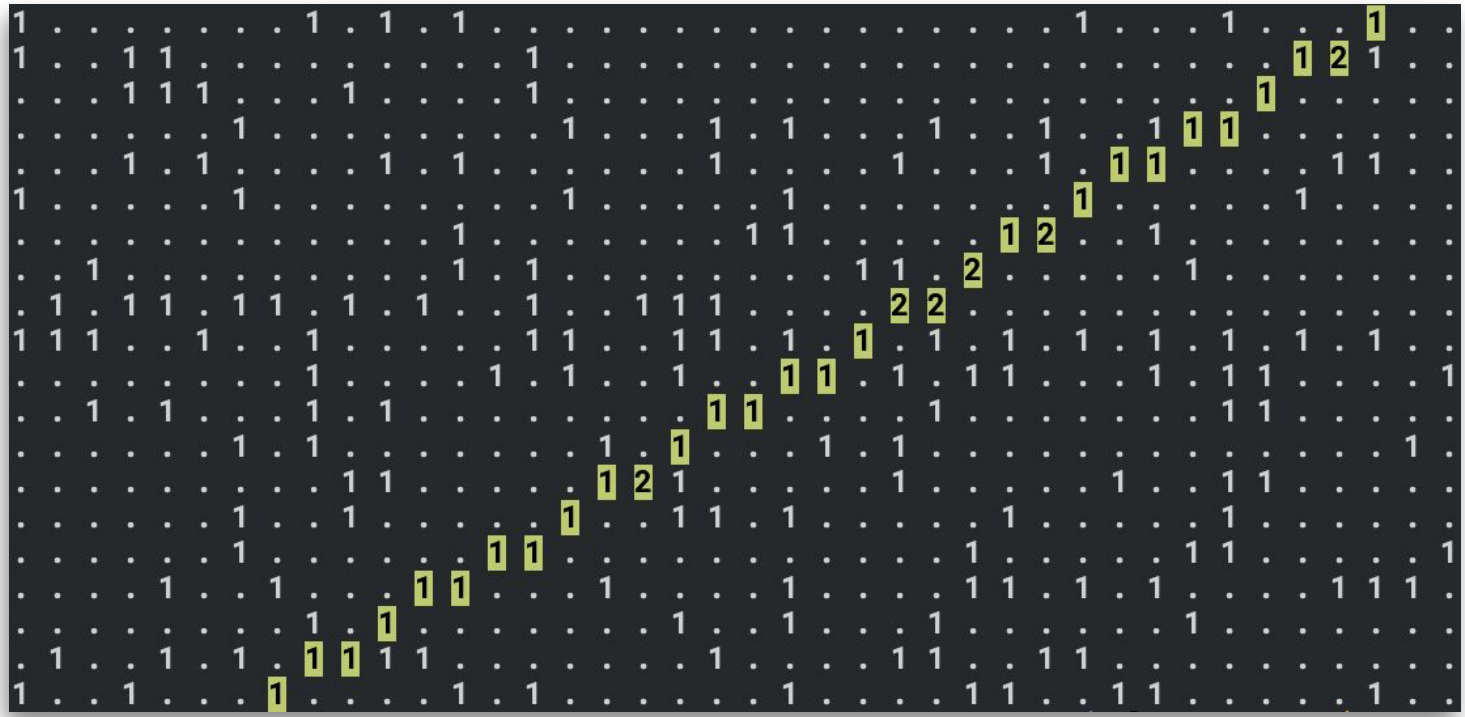


GREEN: correct predicted hit, RED: incorrect predicted hit, BLUE: missed correct hit


Track purity comparison of different language models
used in the neighbor pattern tracking
2-gram
1-skip-2-gram
2-skip-2-gram
Track purity
100
90
80
70
60
50
40
30
20
10
0
10
20
30
40
50
Percent of noise

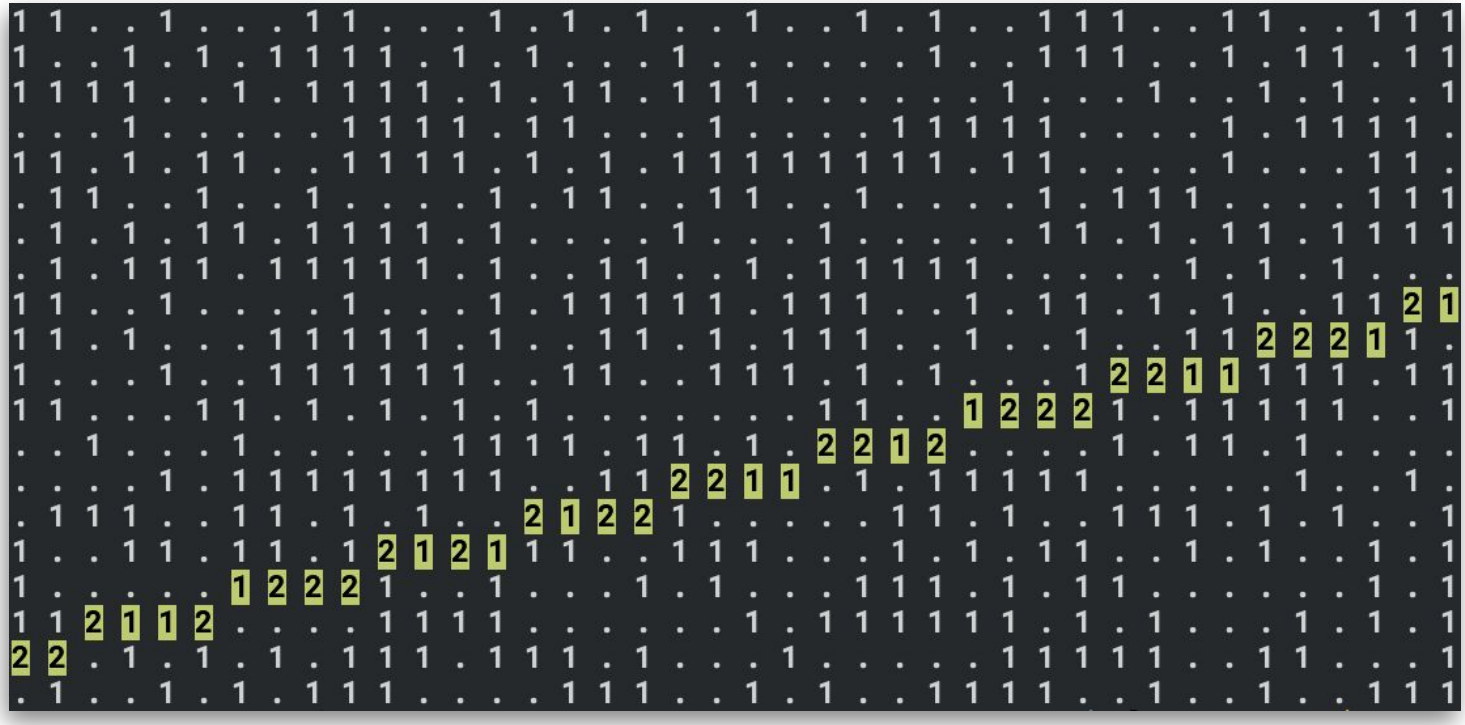
# Results: Moving Direction Feature

# Track finding using the moving direction feature (1)

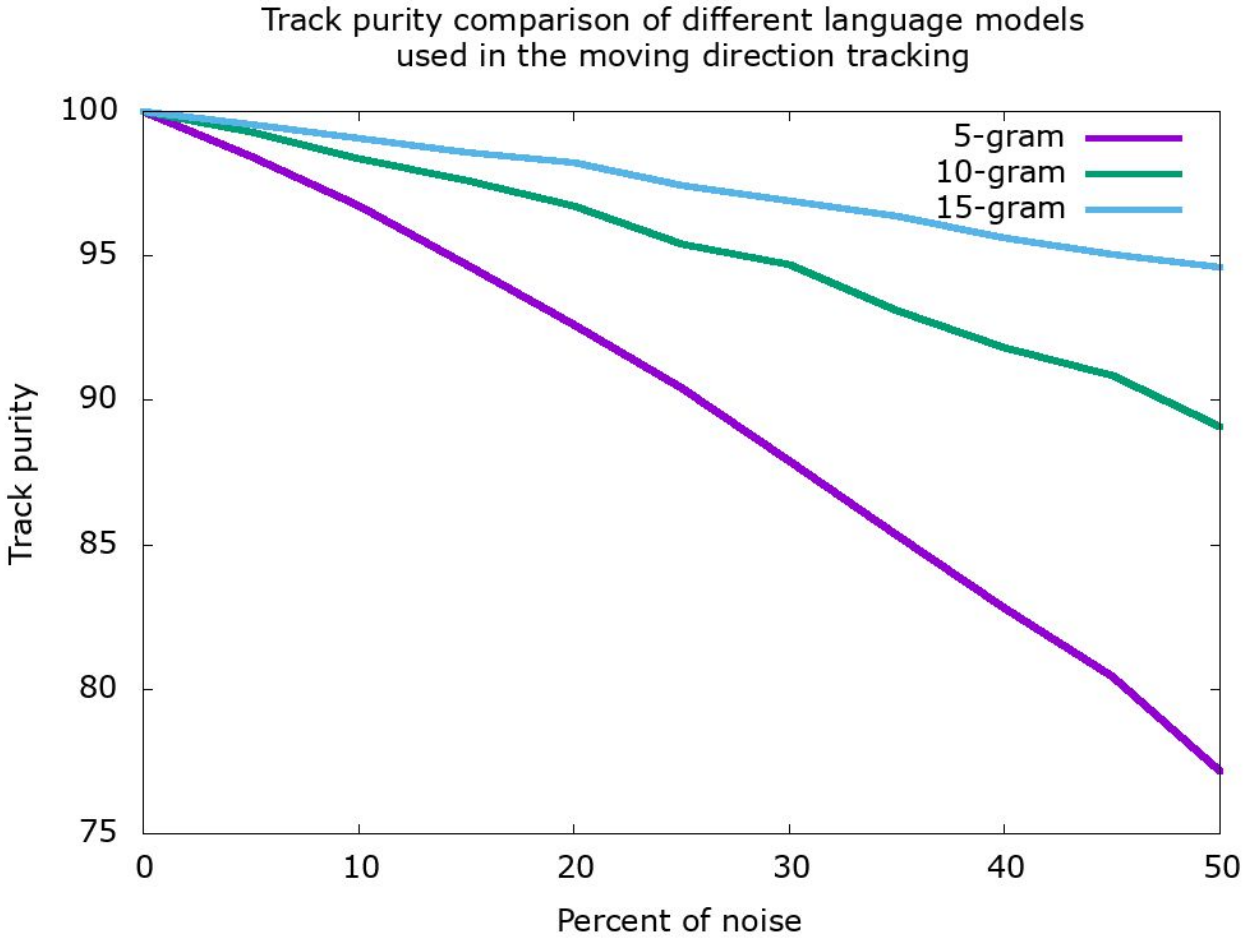


Feature: moving direction, lang model: 15-gram, noise: 20%

# Track finding using the moving direction feature (2)



Feature: moving direction, lang model: 15-gram, noise: 50%


Track purity comparison of different language models
used in the moving direction tracking
Track purity
Percent of noise
5-gram
10-gram
15-gram
100
95
90
85
80
75
0
10
20
30
40
50

# Conclusions

- Neighbor pattern and moving direction are potential features for the track finding task,
- 2-skip-n-gram is the most suitable model for the neighbor pattern tracking,
- Lack of direction information in the neighbor pattern feature causes many cases of incorrect prediction.
- Moving direction feature requires a high order ngram model for accurate hit predictions,
- Some long dependency pattern recognition issues cannot be solved by the neighbor pattern or moving direction features.

# Outlook

- Implement language models using artificial neural network,
- Study correlation between neighbor pattern and moving direction features,
- Test the models with a curved line,
- Include the isochrone radius information as a feature of the tracking model.

$P(\textit{You} \mid \textit{Thank})$