

# Графы в машинном обучении

Романов Владимир БПМИ192

Национальный исследовательский университет  
«Высшая школа экономики» (Москва)

23 ноября 2021 г.

# Какие задачи на графы нам интересны?

## Задачи на вершины

- link-prediction
- node-classification

## Задачи на графы

- graph-classification

# Vertex embeddings

Рассмотрим embedding, которые описывают локальные свойства

Мы хотим, чтобы соседи имели близкий вектор, а далекие разные

# Vertex embeddings

Рассмотрим embedding, которые описывают локальные свойства

Мы хотим, чтобы соседи имели близкий вектор, а далекие разные

Формулировка напоминает аналогичную задачу для текста

Идея: попытаемся применить word2vec для этой задачи

# DeepWalk (Постановка)

## Обозначения

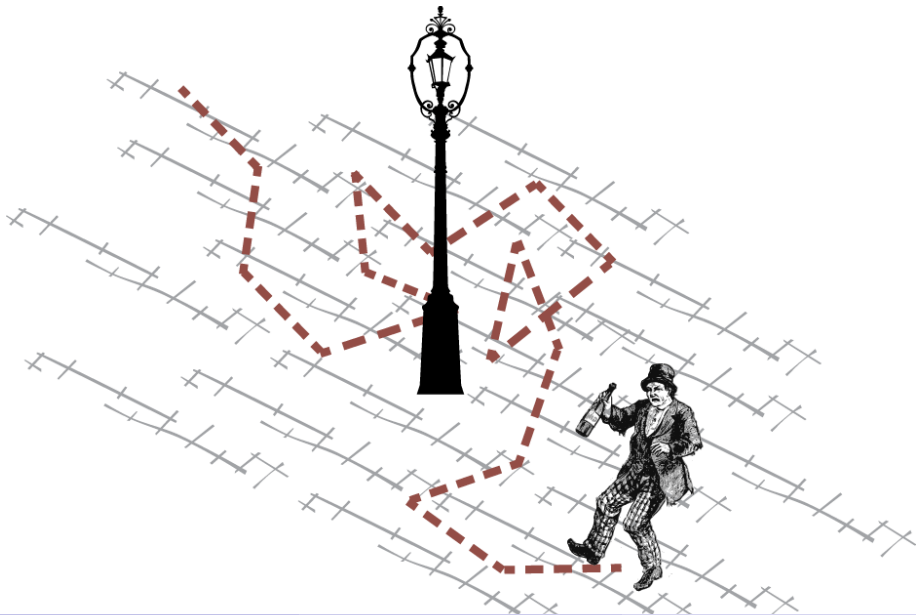
Embedding:  $\Phi : V \rightarrow \mathbb{R}^d$

Loss:  $\sum_{u \in N(v)} -\ln \Pr(u \mid \Phi(v))$

## Алгоритм

- 1 Переберем вершину  $s \in V$
- 2 Рассмотрим случайный путь  $\mathcal{W}_s$  размера  $t$
- 3 Прорелаксируем SkipGram пройдясь по  $\mathcal{W}_s$  окном  $w$
- 4 Повторим процесс  $\gamma$  раз

# DeepWalk (in a nutshell)



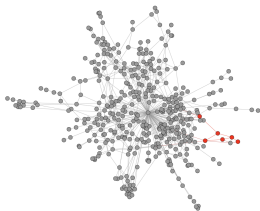
# DeepWalk (Hierarchical Softmax)

Проблема: делать Softmax долго

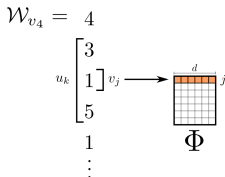
Решение: дерево отрезков (или negative sampling)

$$Pr(u | \Phi(v)) = \prod_{b_i \in B} Pr(b_i | \Phi(v))$$

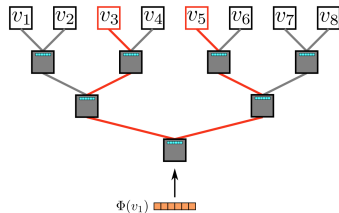
Итоговый алгоритм:



(a) Random walk generation.



(b) Representation mapping.



(c) Hierarchical Softmax.

# Node2Vec (search strategy)

## Обновление

Embedding:  $f : V \rightarrow \mathbb{R}^d$

$$\text{Loss: } \sum_{u \in N_S(v)} -\ln \text{Pr}(u \mid f(v))$$

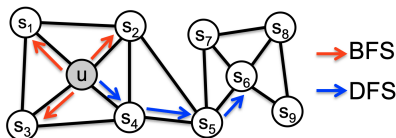


Figure 1: BFS and DFS search strategies from node  $u$  ( $k = 3$ ).

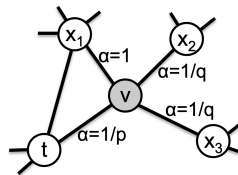


Figure 2: Illustration of the random walk procedure in *node2vec*. The walk just transitioned from  $t$  to  $v$  and is now evaluating its next step out of node  $v$ . Edge labels indicate search biases  $\alpha$ .



# Список литературы

- DeepWalk: <https://arxiv.org/pdf/1403.6652.pdf>
- Node2Vec: <https://arxiv.org/pdf/1607.00653.pdf>