

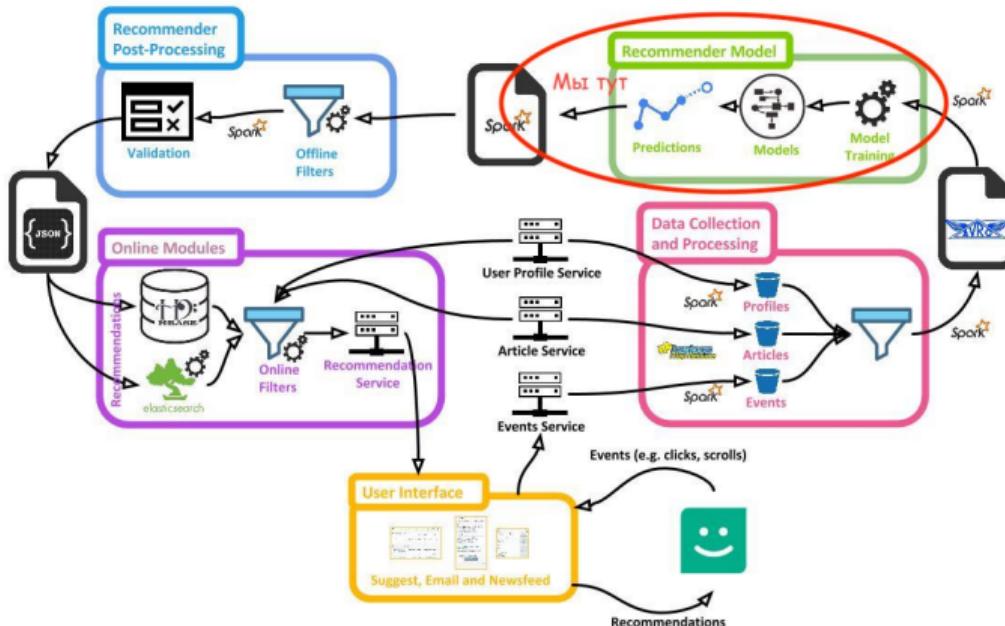
Рекомендеры на графах

Николай Анохин

31 октября 2023 г.



Контекст



Графы и рекомендации
●○○

Классика
○○○

Молниеносное введение в GNN
○○○○○○○○○○

GNN в рекомендациях
○○○○○○○○

Итоги
○○○○

Графы и рекомендации



Графы и рекомендации
ooo

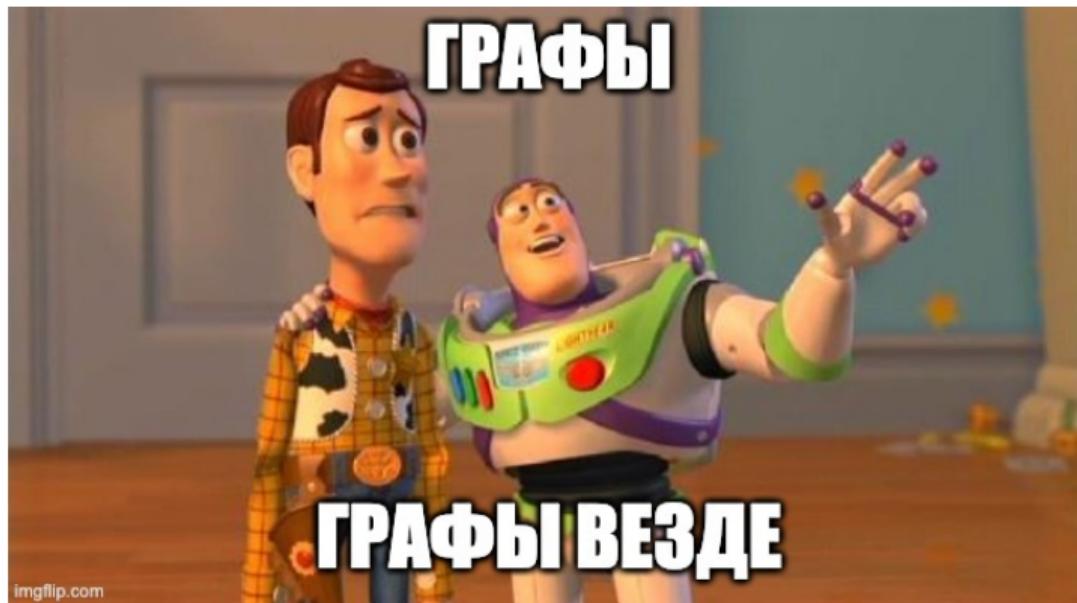
Классика
ooo

Молниеносное введение в GNN
oooooooo

GNN в рекомендациях
ooooooo

Итоги
oooo

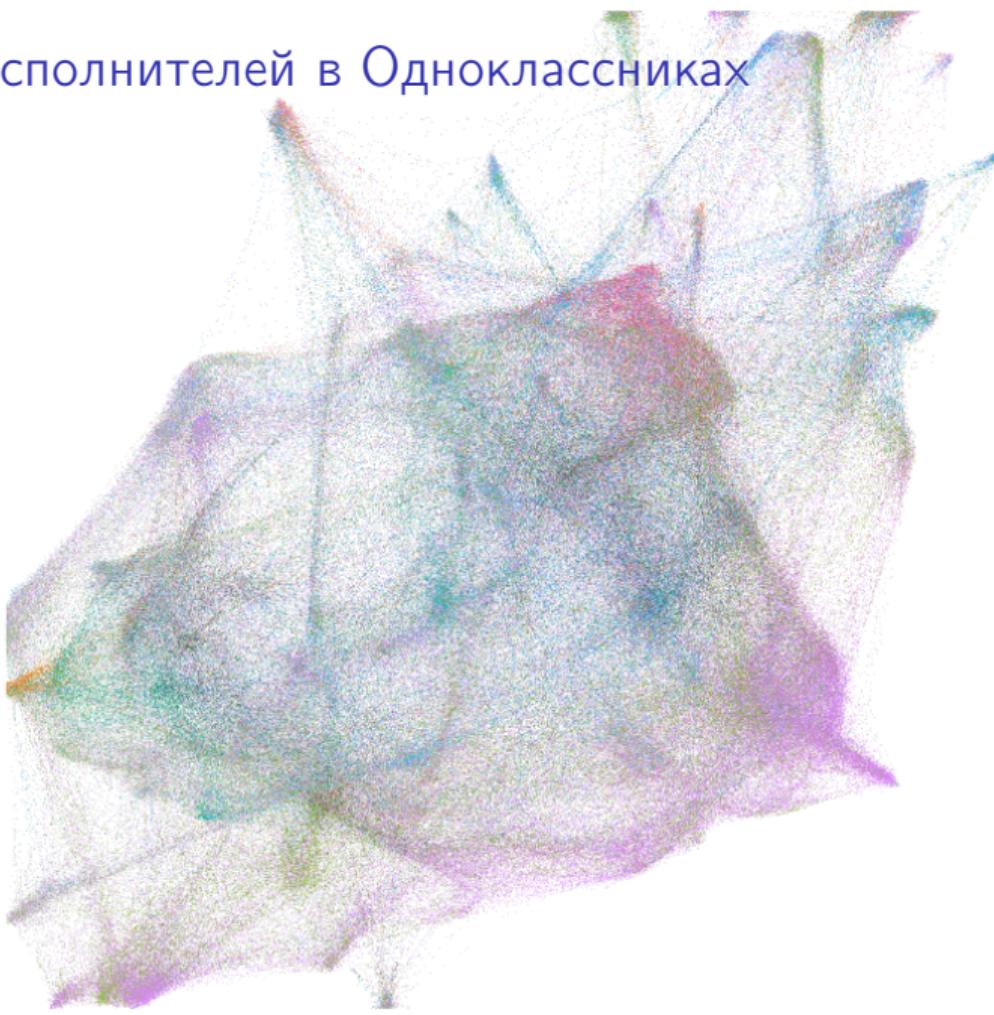
Почему графы



Граф исполнителей в Одноклассниках

20000 вершин
750000 ребер

unknown	(34.09%)
Neukazan	(14.32%)
Other	(11.68%)
Pop	(6.45%)
Blues	(3.7%)
Chanson	(2.86%)
Rock	(2.75%)
Rap	(1.31%)
Dance	(1.04%)



Графы и рекомендации
ооо

Классика
●оо

Молниеносное введение в GNN
oooooooooo

GNN в рекомендациях
ooooooooo

Итоги
oooo

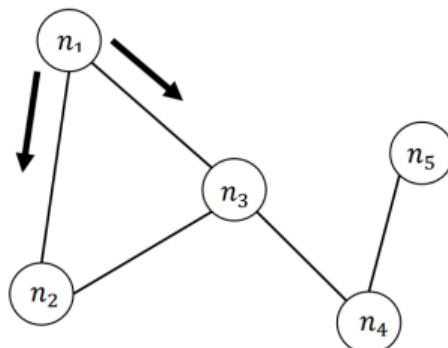
Классика



Personalized Page Rank (PPR) [PLCL19]

Начальное состояние

$$\vec{s}_0 = (1 \ 0 \ 0 \ 0 \ 0)$$



Вероятность перехода

$$\vec{p}_{n+1} = c \vec{s}_n \cdot P + (1 - c) \vec{s}_0$$

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 0 & 1/3 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$(1 - c)$ – вероятность рестарта

Где мы окажемся через бесконечное число шагов?



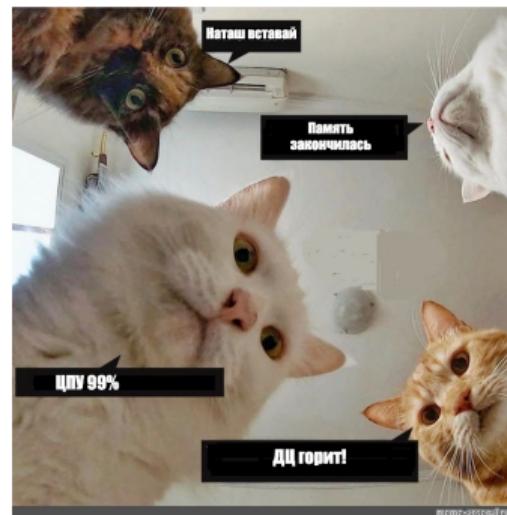
Вычисление рекомендаций

- Точное решение

$$\vec{r} = c \vec{r}P + (1 - c)\vec{s} \quad \Rightarrow$$

$$\vec{r} = (1 - c)\vec{s}(I - cP)^{-1}$$

- Сэмплирование Монте-Карло
- Оптимизированное прямое решение
- Оптимизированный power iteration



Графы и рекомендации
ooo

Классика
ooo

Молниеносное введение в GNN
●oooooooo

GNN в рекомендациях
ooooooo

Итоги
oooo

Молниеносное введение в GNN



Разные виды данных в нейросетях [Ma21]

i.i.d. sample



independent feats

architecture class

feed forward

inductive biases

interactions btwn feats



images

grid
convolution

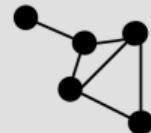
spatial correlations.



sequences

recurrent/
autoregression

temporal/sequential
dependencies.

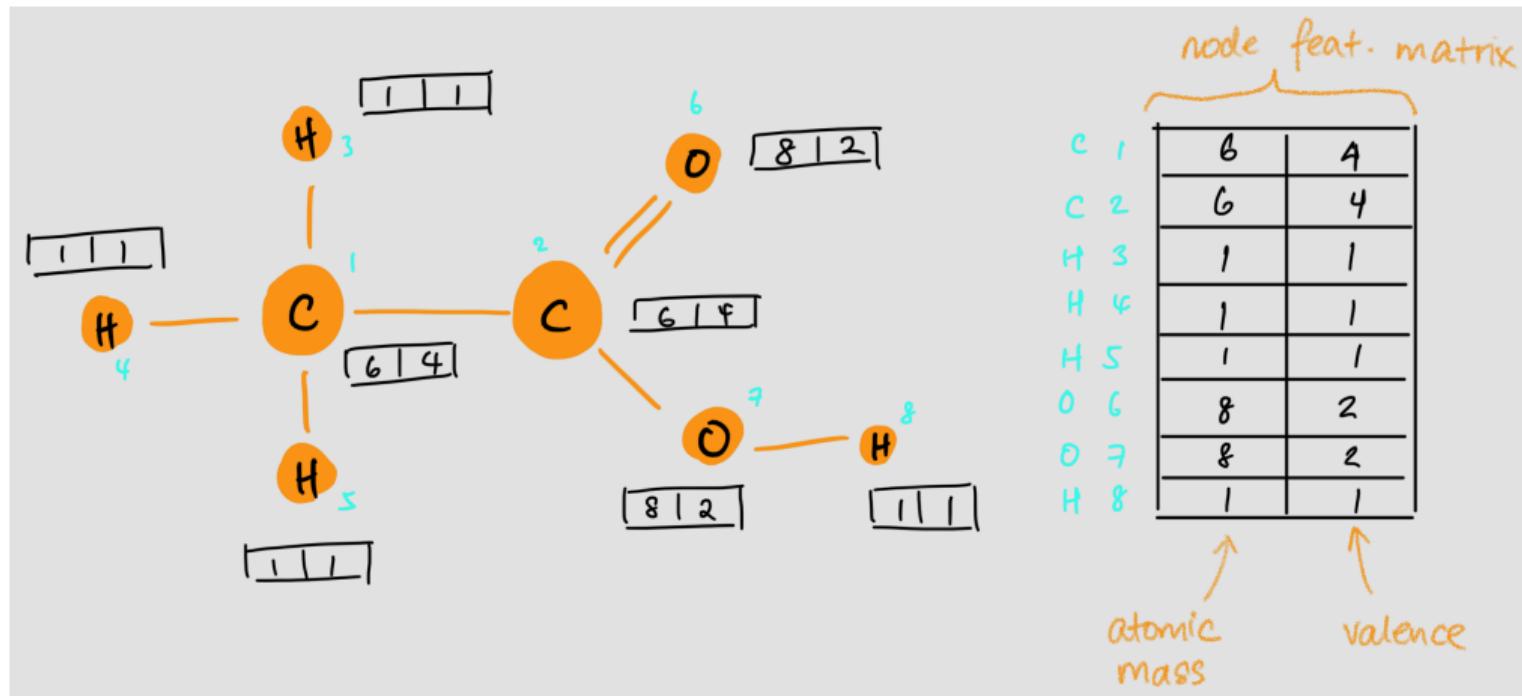


message
passing

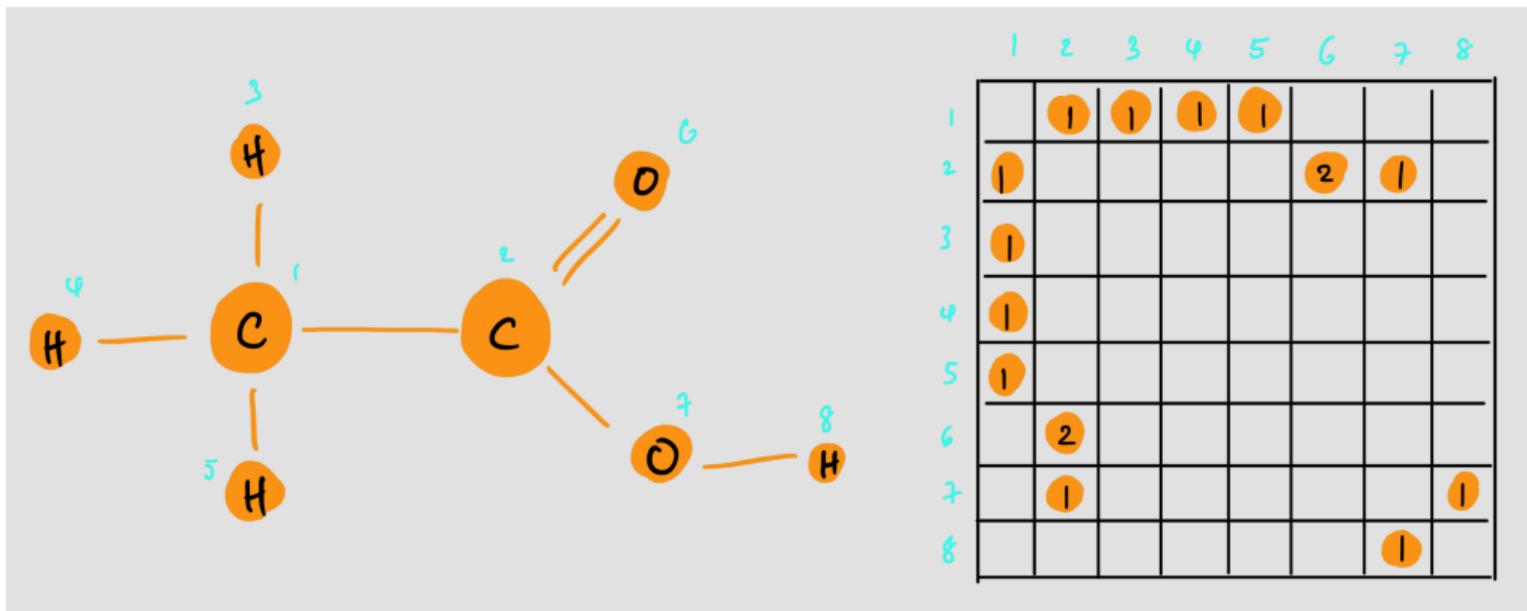
relational
(general)



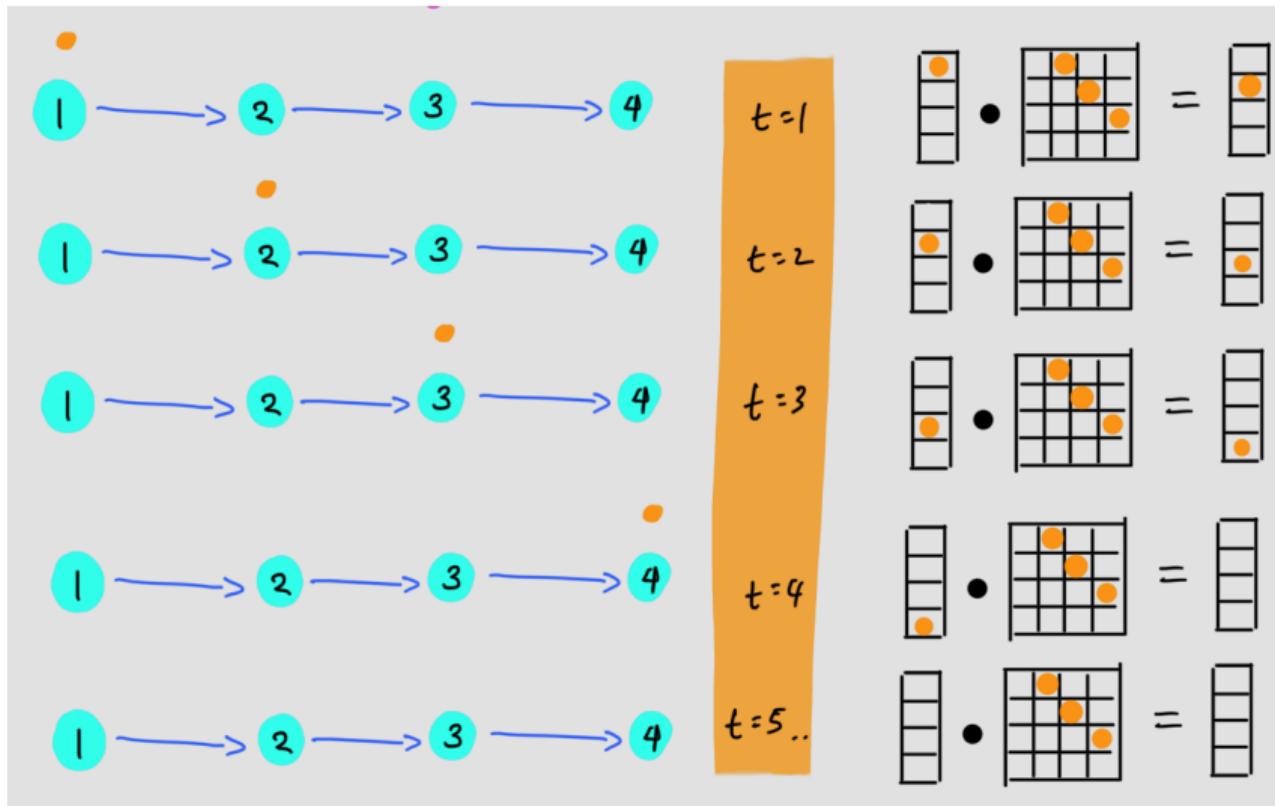
Признаки узлов



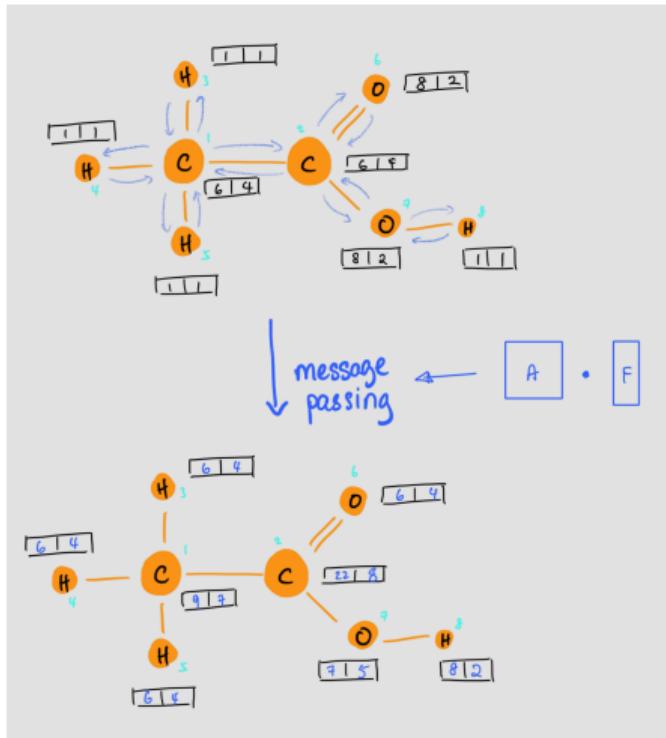
Структура графа



Интуиция message passing



Message passing



А где же нейросети?

$$\text{MP}(\mathbf{A}, \mathbf{F}) \rightarrow \mathbf{Z} \quad (\text{message passing})$$

$$\boxed{\mathbf{A}} \cdot \boxed{\mathbf{F}} \rightarrow \boxed{\mathbf{Z}}$$

$$\text{FF}(\mathbf{z}, \theta) \rightarrow \mathbf{F}' \quad (\text{feed-forward operation on node feats})$$

learnable params

$$\text{relu} \left(\boxed{\mathbf{z}} \cdot \boxed{\mathbf{w}} + \boxed{\mathbf{b}} \right) \rightarrow \boxed{\mathbf{F}'}$$

Общий фреймворк GNN [WSZ⁺22]

A. Message passing в узле $n_v^{(l)}$ на l слое:

1. Агрегация представлений соседей $h_u^{(l)}$ (Aggregation)

$$n_v^{(l)} = \text{Aggregator}_l(\{h_u^{(l)}, u \in N(v)\}),$$

2. Апдейт представления

$$h_v^{(l+1)} = \text{Updater}_l(n_v^{(l)}, h_v^{(l)})$$

B. Использование представлений узлов для решения задачи

- Классификация узлов $f(h_v)$
- Предсказание / классификация связей $g(h_v \cdot h_u)$
- Классификация графов $h(Pooling(h_v))$



Примеры GNN

GraphSage

$$\text{Aggregation: } \mathbf{n}_v^{(l)} = \text{Aggregator}_l \left(\left\{ \mathbf{h}_u^l, \forall u \in \mathcal{N}_v \right\} \right),$$

$$\text{Update: } \mathbf{h}_v^{(l+1)} = \delta \left(\mathbf{W}^{(l)} \cdot [\mathbf{h}_v^{(l)} \oplus \mathbf{n}_v^{(l)}] \right),$$

GAT

$$\text{Aggregation: } \mathbf{n}_v^{(l)} = \sum_{j \in \mathcal{N}_v} \alpha_{vj} \mathbf{h}_j^{(l)}, \alpha_{vj} = \frac{\exp \left(\text{Att}(h_v^{(l)}, \mathbf{h}_j^{(l)}) \right)}{\sum_{k \in \mathcal{N}_v} \exp \left(\text{Att}(h_v^{(l)}, \mathbf{h}_k^{(l)}) \right)},$$

$$\text{Update: } \mathbf{h}_v^{(l+1)} = \delta \left(\mathbf{W}^{(l)} \mathbf{n}_v^{(l)} \right),$$



Графы и рекомендации
ooo

Классика
ooo

Молниеносное введение в GNN
oooooooo

GNN в рекомендациях
●ooooooo

Итоги
oooo

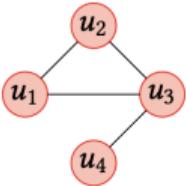
GNN в рекомендациях



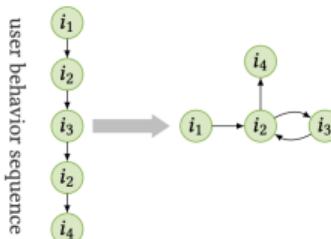
GNN в рекомендациях [WSZ⁺22]

	i_1	i_2	i_3	i_4	i_5
u_1	0	1	1	0	0
u_2	0	0	1	0	1
u_3	1	0	0	1	0
u_4	1	0	1	1	1

(a) User-item bipartite graph.



(c) Social relationship between users.

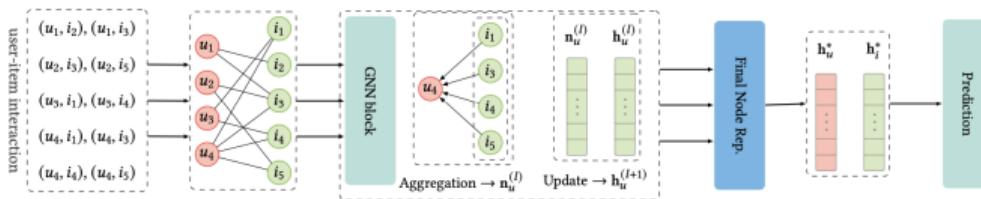


(b) Sequence graph.



(d) Knowledge graph

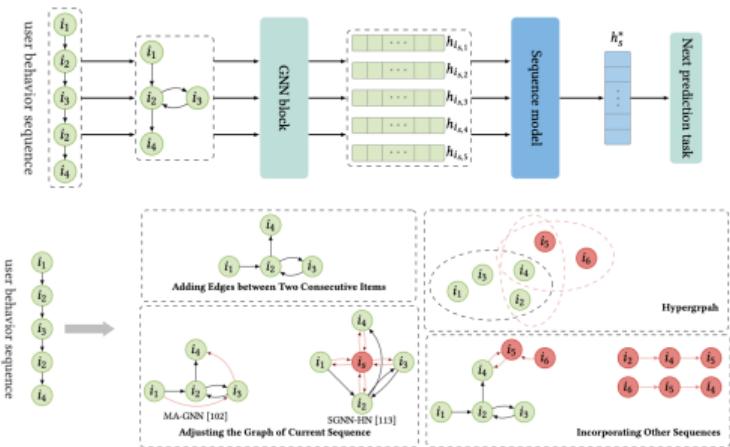
GNN для коллаборативной фильтрации



Вопросы

- Как построить граф?
- Как агрегировать информацию от соседей?
- Как сформировать финальное представление пользователей и айтемов?

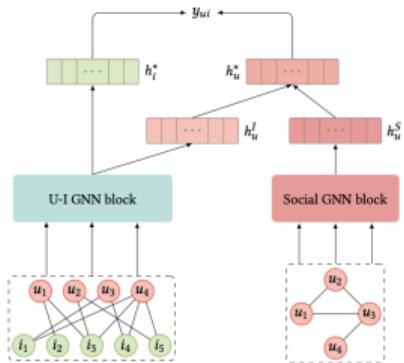
GNN на последовательностях айтемов



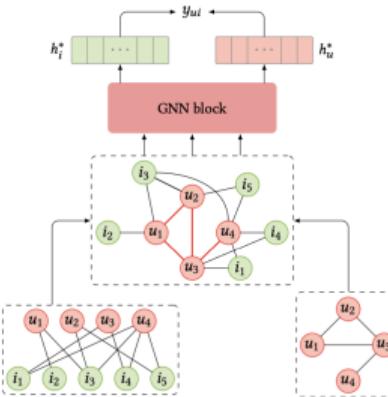
Вопросы

- Как построить граф?
- Как сагрегировать информацию всей последовательности?

GNN в задачах с соцсетью



(a) The framework of GNN on the bipartite graph and social network graph separately.



(b) The framework of GNN on the unified graph of user-item interactions and social network.

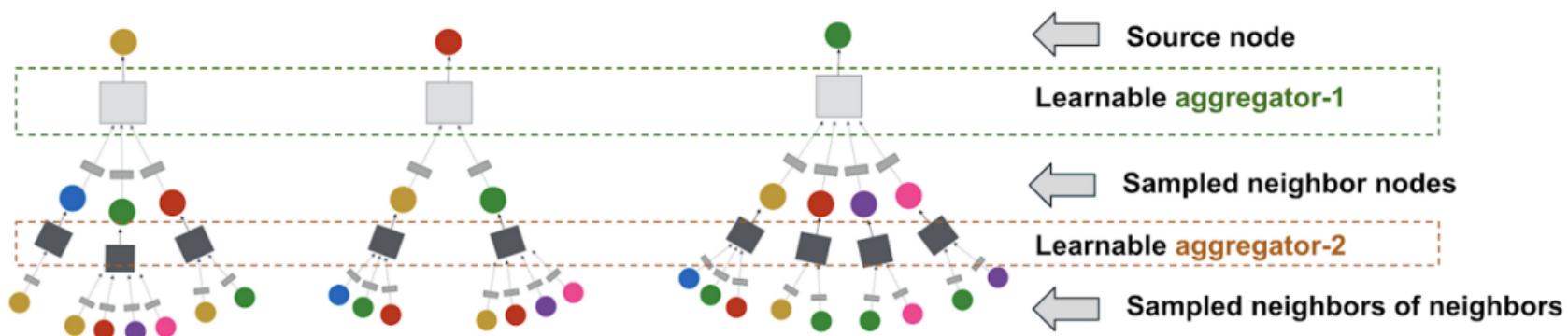
Вопросы

- Комбинировать ли граф CF с графом друзей?
- Разный ли вес у друзей?



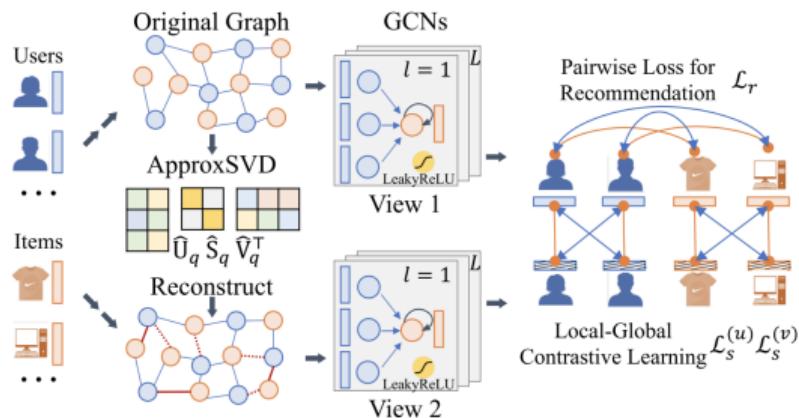
GNN в задачах с графом знаний

PinSage: A new graph convolutional neural network for web-scale recommender systems
[YHC⁺18]



А вообще работает?

LightGCL: Simple Yet Effective Graph Contrastive Learning for Recommendation
[CHXR23]



Rank	Model	Recall@20↑	nDCG@20	Paper	Code	Result	Year	Tags
1	BSPM-EM	0.1920	0.1597	Blurring-Sharpening Process Models for Collaborative Filtering			2022	
2	BSPM-LM	0.1901	0.1548	Blurring-Sharpening Process Models for Collaborative Filtering			2022	
3	LT-OCF	0.1875	0.1574	LT-OCF: Learnable-Time ODE-based Collaborative Filtering			2021	
4	SimpleX	0.1872	0.1557	SimpleX: A Simple and Strong Baseline for Collaborative Filtering			2021	
5	UltraGCN	0.1862	0.1580	UltraGCN: Ultra Simplification of Graph Convolutional Networks for Recommendation			2021	
6	LightGCN	0.1830	0.1554	LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation			2020	
7	SSCF	0.1775	0.1390	Sapling Similarity: a performing and interpretable memory-based tool for recommendation			2022	
8	NGCF	0.1570	0.1327	Neural Graph Collaborative Filtering			2019	

<https://paperswithcode.com/sota/recommendation-systems-on-gowalla>



Data	Metric	DGCF	HyRec	LightGCN	MHCN	SGL	SimGRACE	GCA	HCCF	SHT	SimGCL	LightGCL	p-val.	impr.
Yelp	R@20	0.0466	0.0472	0.0482	0.0503	0.0526	0.0603	0.0621	0.0626	0.0651	0.0718	0.0793	7e-9	10%
	N@20	0.0395	0.0395	0.0409	0.0424	0.0444	0.0435	0.0530	0.0527	0.0546	0.0615	0.0668	8e-9	8%
	R@40	0.0774	0.0791	0.0803	0.0826	0.0869	0.0989	0.1021	0.1040	0.1091	0.1166	0.1292	2e-9	10%
	N@40	0.0511	0.0522	0.0527	0.0544	0.0571	0.0656	0.0677	0.0681	0.0709	0.0778	0.0852	2e-9	9%
Gowalla	R@20	0.0944	0.0901	0.0985	0.0955	0.1030	0.0869	0.0896	0.1070	0.1232	0.1357	0.1578	1e-6	16%
	N@20	0.0522	0.0498	0.0593	0.0574	0.0623	0.0528	0.0537	0.0644	0.0731	0.0818	0.0935	2e-6	14%
	R@40	0.1401	0.1356	0.1431	0.1393	0.1500	0.1276	0.1322	0.1535	0.1804	0.1956	0.2245	3e-6	14%
	N@40	0.0671	0.0660	0.0710	0.0689	0.0746	0.0637	0.0651	0.0767	0.0881	0.0975	0.1108	3e-6	13%
ML-10M	R@20	0.1763	0.1801	0.1789	0.1497	0.1833	0.2254	0.2145	0.2219	0.2173	0.2265	0.2613	1e-9	15%
	N@20	0.2101	0.2178	0.2128	0.1814	0.2205	0.2686	0.2613	0.2629	0.2573	0.2613	0.3106	3e-9	18%
	R@40	0.2681	0.2685	0.2650	0.2250	0.2768	0.3295	0.3231	0.3265	0.3211	0.3345	0.3799	7e-10	13%
	N@40	0.2340	0.2340	0.2322	0.1962	0.2426	0.2939	0.2871	0.2880	0.3318	0.2880	0.3387	1e-9	17%
Amazon	R@20	0.0211	0.0302	0.0319	0.0296	0.0327	0.0381	0.0309	0.0322	0.0441	0.0474	0.0585	2e-7	23%
	N@20	0.0154	0.0225	0.0236	0.0219	0.0249	0.0291	0.0238	0.0247	0.0328	0.0360	0.0436	2e-6	21%
	R@40	0.0351	0.0432	0.0499	0.0489	0.0531	0.0621	0.0498	0.0525	0.0719	0.0750	0.0933	1e-7	24%
	N@40	0.0201	0.0246	0.0290	0.0284	0.0312	0.0371	0.0301	0.0314	0.0420	0.0451	0.0551	9e-7	22%
Tmall	R@20	0.0235	0.0233	0.0225	0.0203	0.0268	0.0222	0.0373	0.0314	0.0387	0.0473	0.0528	3e-5	11%
	N@20	0.0163	0.0160	0.0154	0.0139	0.0183	0.0152	0.0252	0.0213	0.0262	0.0328	0.0361	1e-4	10%
	R@40	0.0394	0.0350	0.0378	0.0340	0.0446	0.0367	0.0616	0.0519	0.0645	0.0766	0.0852	1e-5	11%
	N@40	0.0218	0.0199	0.0208	0.0188	0.0246	0.0203	0.0337	0.0284	0.0352	0.0429	0.0473	7e-5	10%



Графы и рекомендации
ooo

Классика
ooo

Молниеносное введение в GNN
oooooooo

GNN в рекомендациях
ooooooo

Итоги
●ooo

Итоги



Итоги

Архитектура GNN хорошо ложится на формулировку рекомендательной задачи.

Хотя это популярное направление в исследовательских статьях, польза от применения на продакшнене пока что под вопросом.



До следующего раза



<https://t.me/mlvok>



Литература I

-  Xuheng Cai, Chao Huang, Lianghao Xia, and Xubin Ren, *Lightgcl: Simple yet effective graph contrastive learning for recommendation*, 2023.
-  Eric Ma, *An attempt at demystifying graph deep learning*, 2021.
-  Sungchan Park, Wonseok Lee, Byeongseo Choe, and Sang-goo Lee, *A survey on personalized pagerank computation algorithms*, IEEE Access PP (2019), 1–1.
-  Shiwen Wu, Fei Sun, Wentao Zhang, Xu Xie, and Bin Cui, *Graph neural networks in recommender systems: A survey*, 2022.
-  Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L. Hamilton, and Jure Leskovec, *Graph convolutional neural networks for web-scale recommender systems*, Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (New York, NY, USA), KDD '18, Association for Computing Machinery, 2018, p. 974–983.

