



Санкт-Петербургский государственный университет  
Кафедра системного программирования

## Теория графов. Презентации 1-2

Команда 7: Елкин Леонид, Шишин Кирилл

Санкт-Петербург  
2025

# Single-source parent BFS, постановка задачи

Дан невзвешенный неориентированный граф  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , в котором выделена исходная вершина  $s \in \mathcal{V}$ . Для каждой вершины  $v \in \mathcal{V}$  определить вершину  $u \in \mathcal{V}$ , являющуюся предшественником вершины  $v$  на одном из кратчайших путей из  $s$ .

## Определение (Путь)

Путь  $\mathcal{P}$  из вершины  $s$  в вершину  $d$  — это такая последовательность вершин  $v_1, v_2, \dots, v_n$ , что  $v_1 = s$  и  $v_n = d$  и  $(v_i, v_{i+1}) \in \mathcal{E}$ .

## Определение (Кратчайший путь)

$\mathcal{P}^*$  — кратчайший путь из вершины  $s$  в вершину  $d$ , если  $|\mathcal{P}^*| = \min\{|\mathcal{P}| : \mathcal{P} \text{ — путь из } s \text{ в } d\}$

## Multi-source parent BFS, постановка задачи

Дан невзвешенный неориентированный граф  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , в котором выделено множество исходных вершин  $S \in \mathcal{V}$ . Для каждой исходной вершины  $s \in S$  и каждой вершины  $v \in \mathcal{V}$  определить вершину  $u \in \mathcal{V}$ , являющуюся предшественником вершины  $v$  на одном из кратчайших путей из  $s$ .

# PageRank, постановка задачи

Дан невзвешенный ориентированный граф  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ . Необходимо найти "важность" каждой вершины в графе.

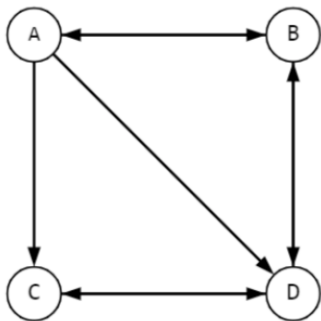
В классической версии алгоритма изначально "важность" каждой вершины равна  $\frac{1}{|\mathcal{V}|}$ . На каждой новой итерации алгоритма "важность" каждой пересчитывается в соответствии с формулой

$$PR(u) = \frac{1-d}{|\mathcal{V}|} + d \sum_{v \in B(u)} \frac{PR(v)}{L(v)}, \text{ где}$$

- $B(u)$  — множество вершин, из которых есть путь в вершину  $u$
- $L(v)$  — количество путей, исходящих из вершины  $v$
- $d$  — коэффициент затухания (обычно 0.85)

Пересчет "важности" вершин производится до сходимости.

$$L = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 1 & 0 \end{pmatrix}$$



$$M = (dL + (1 - d)E)$$

Где  $E$  – матрица из  $\frac{1}{N}$

$$PR^{(t+1)} = M * PR^{(t)}$$

# Сходимость Pagerank

- Матрица переходов  $M$  стохастическая (сумма элементов в столбцах равна 1)
- Элементы этой матрицы строго положительные

Из Т. Перрона-Фробениуса следует<sup>1</sup>

- $M$  имеет уникальный собственный вектор, который соответствует максимальному собственному числу  $\lambda=1$ . (Этот же вектор является стационарным)
- $PR^{(t+1)} = M * PR^{(t)}$  сходится к этому собственному вектору.

---

<sup>1</sup>Согласно материалам статьи

## Определение (SuiteSparse:GraphBLAS)

*SuiteSparse:GraphBLAS — полная реализация стандарта GraphBLAS.*

- Основные объекты: матрица, вектор. Операции над ними
- Несколько форматов представления матриц (напр., CSC, CSR)
- Алгоритм выполнения конкретной операции выбирается в зависимости от контекста задачи

Графы для экспериментов были взяты из Stanford Large Network Dataset Collection. Такой выбор был сделан в силу:

- Представлен широкий выбор графов, которые подходят в том числе для задач поиска кратчайшего пути
- Реалистичность и практическая значимость
- Разнообразие размеров
- Графы SNAP часто используются в публикациях



# Набор данных

Name	Description	Nodes	Edges
LiveJournal1	Social network LiveJournal	4,847,571	68,993,773
Orkut	Social network Orkut	3,072,441	117,185,083
Youtube	Social network YouTube	1,134,890	2,987,624
Pokec	Social network Pokec	1,632,803	30,622,564
DBLP	Co-authorship network from DBLP	317,080	1,049,866
Stanford	Web graph of Stanford University pages	281,903	2,312,497
Gowalla	Location-based social network Gowalla	196,591	950,327
Enron	Email communication network from Enron	36,692	367,662

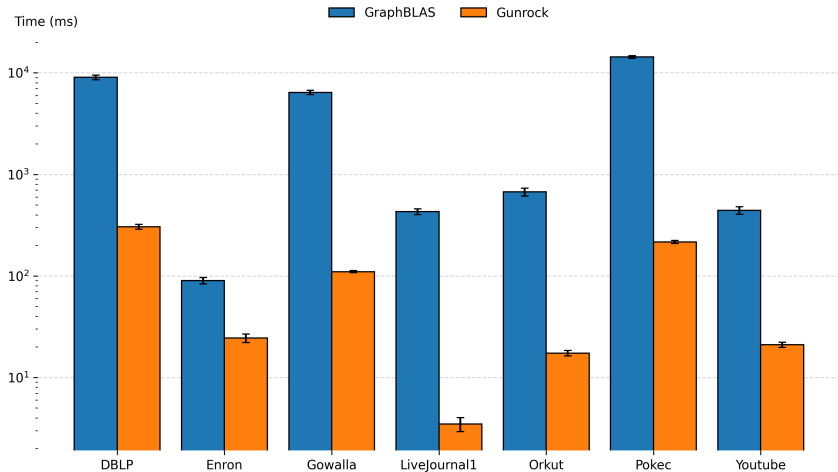
## Цели:

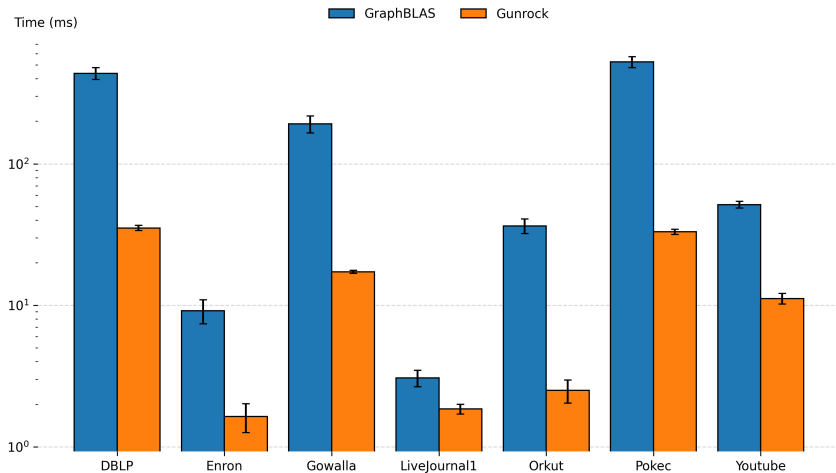
- Выяснить какой прирост дает GPU многопоточность в сравнении с CPU многопоточностью.
- Проанализировать на сколько растет время MSBFS в пределах одной библиотеки с разным количеством начальных вершин относительно SSBFS.

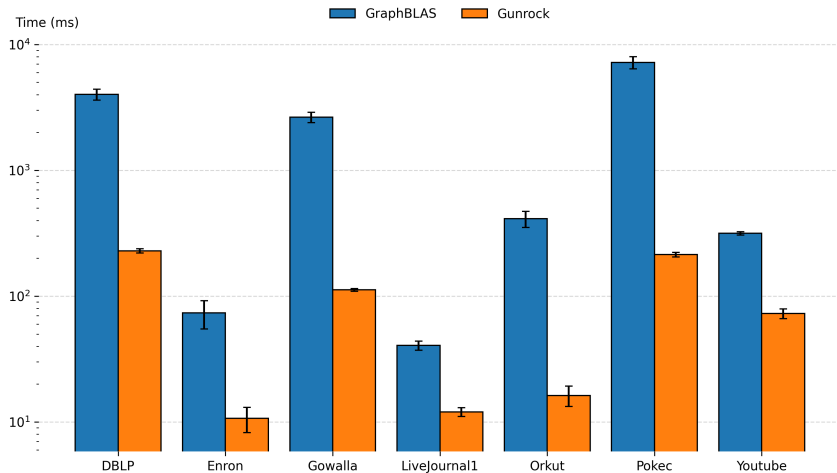
## Характеристики вычислительной машины:

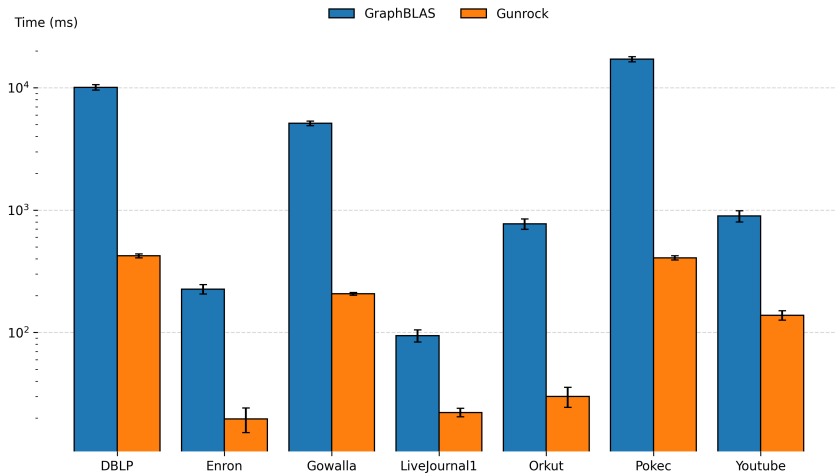
- 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz 2.30 GHz
  - ▶ 8 ядер, 16 логических процессоров
  - ▶ Кэш L1 — 640 КБ
  - ▶ Кэш L2 — 10 МБ
  - ▶ Кэш L3 — 24 МБ
- RAM: 32 GB
- GPU: NVIDIA GeForce RTX 3070 laptop
- OS: Windows 11 24H2

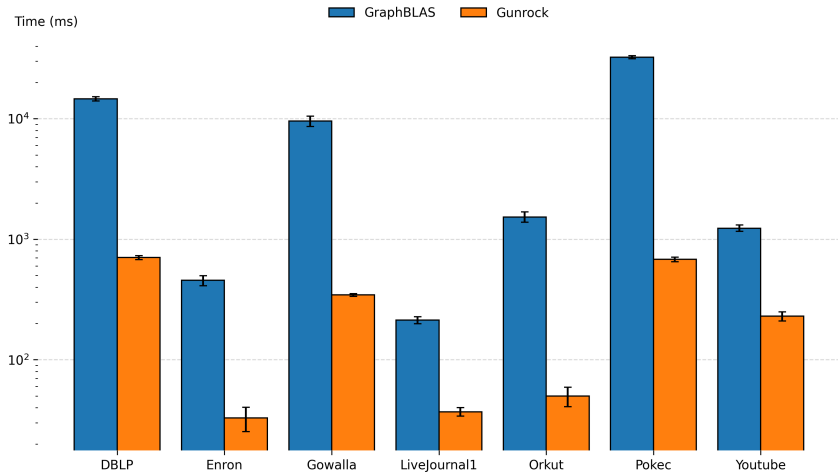
# Pagerank





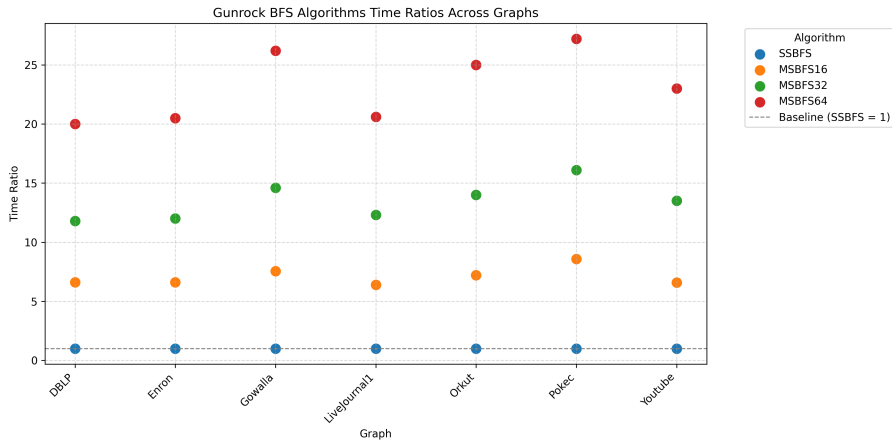








# Gunrock — сравнение разных parent bfs



# GraphBLAS — сравнение разных parent bfs

