

Исследование влияния спарсификации на скорость и точность работы алгоритмов на графах

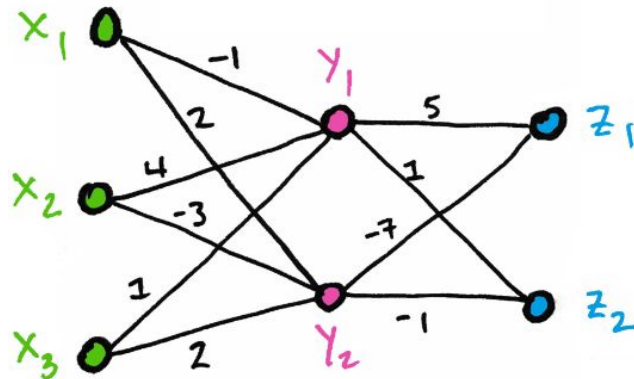
Грустный кефир

Илья Быков, Дмитрий Долбня, Анна Чистякова, Константин Щепин

Зачем?

- 1) для любых алгоритмов на графах (Поиск кратчайших путей, кластеризация графов, задачи машинного обучения на графах)
- 2) уменьшения затрат на хранение и обработку данных (социальные сети, географические сети, интернет)
- 3) визуализация графов, представление иерархических структур.
- 4) ускорение матричных операций поддерживающие графовое представление

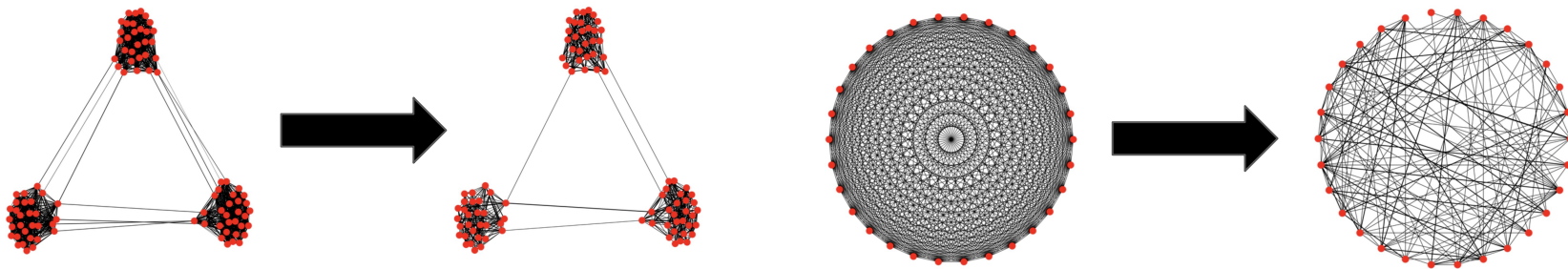
$$\begin{matrix} & M & & N \\ \begin{bmatrix} -1 & 2 \\ 4 & -3 \\ 1 & 2 \end{bmatrix} & & \begin{bmatrix} 5 & 1 \\ -7 & -1 \end{bmatrix} \end{matrix}$$



ОСНОВНЫЕ ВИДЫ:

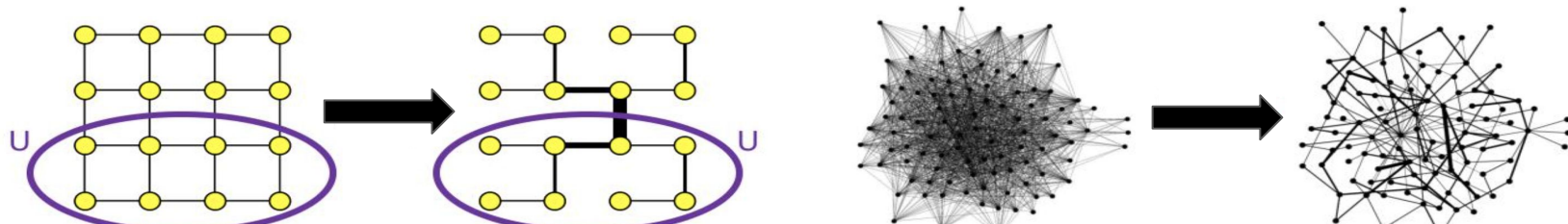
1) Spectral-Preserving Sparsification (алгоритм Спилмана-Шриваставы):

лапласиан графа $L \rightarrow$ для каждой пары вершин производит случайное блуждание от первой до второй \rightarrow используя L получает “Effective Resistance” для этой пары, интуитивно = разница в напряжений узлов, когда единичный ток подается в первый узел и извлекается во втором узле (что то около среднего времени блуждания) \rightarrow выбирается подмножество ребер, минимизирующие суммарное “сопротивление”



2) non-uniform cut sparsifier (Benczúr-Karger)

алгоритм стремится сохранить попарные расстояния (геодезические пути), путем перевзвешивания ребер в зависимости от вероятности, с которой ребро было выбрано



Постановка задач

Цель: исследовать влияния спарсификации плотных графов на работу алгоритмов.

Задачи:

- Реализация PageRank
- Проведение экспериментов с плотными и спарсифицированными графами

PageRank

$$r_k(v) = (1 - d)/n + d \sum_{u \neq v} (r_{k-1}(u)/d_{out}(u))$$

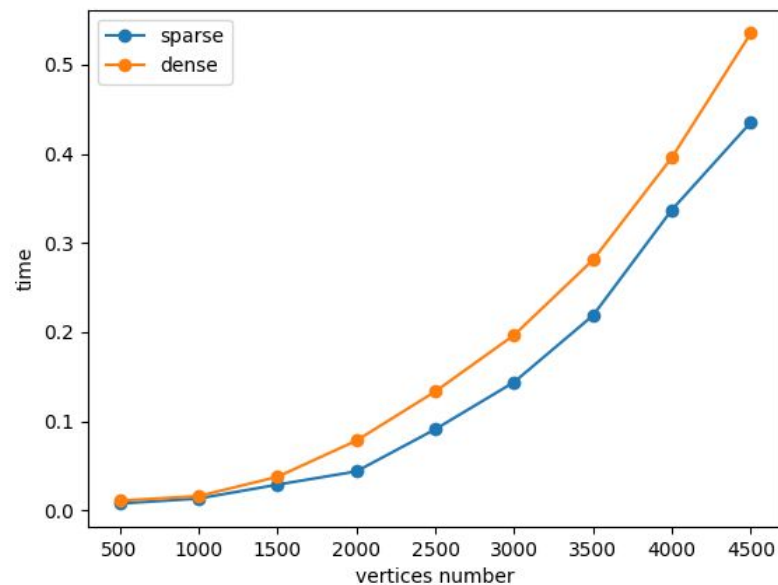
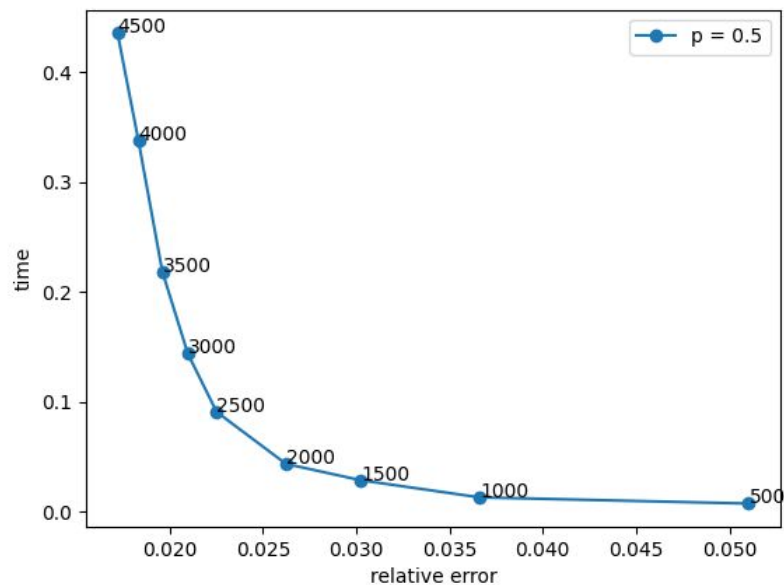
v — вершина

r_k — ранг на k -й итерации

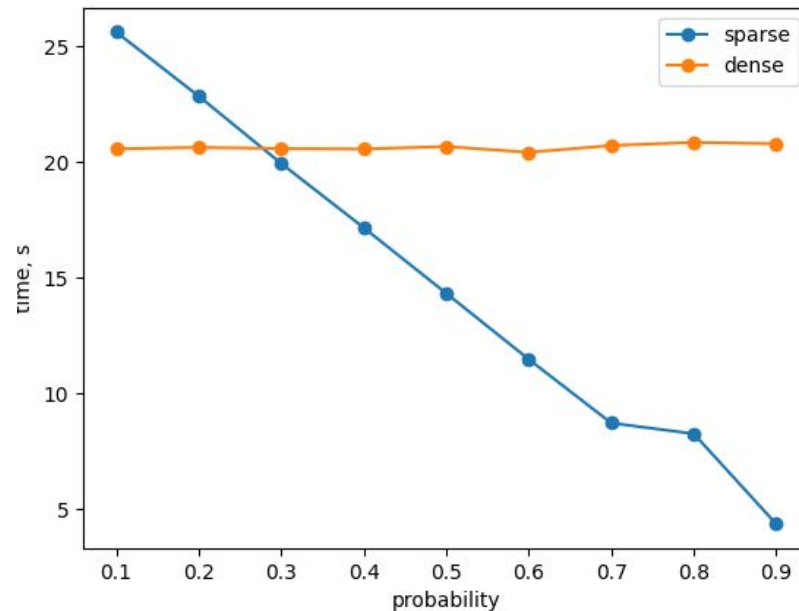
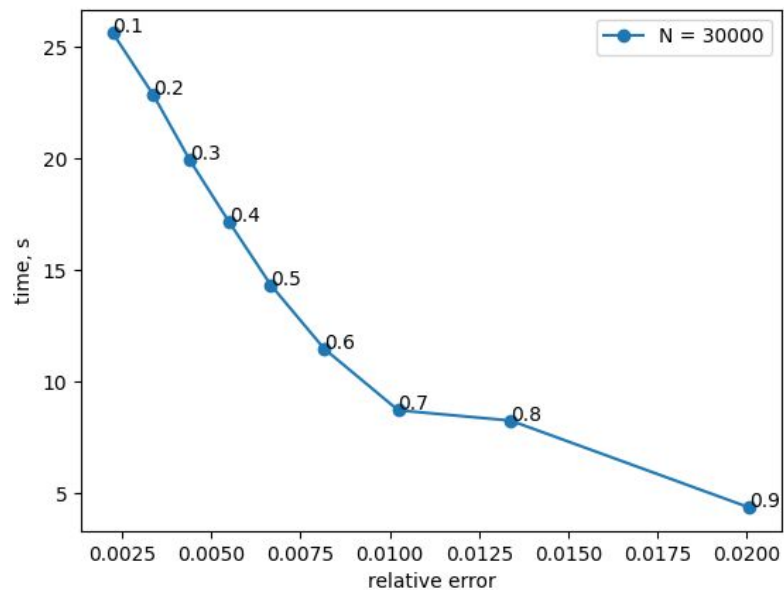
$d_{out}(v)$ — количество исходящих из v ребер

d - вероятность перехода в случайную соседнюю вершину

Эксперименты



Эксперименты



ОСНОВНЫЕ ССЫЛКИ

- <https://github.com/Any497/GraphSparsificationResearch> - код
- https://www.researchgate.net/publication/356707900_The_GraphBLAS_in_Julia_and_Python_the_PageRank_and_Triangle_Centralities - PageRank with linear algebra
- <https://www.slideserve.com/merlin/graph-sparsifiers-a-survey-powerpoint-ppt-presentation> - хорошее сравнение методов спарсификации