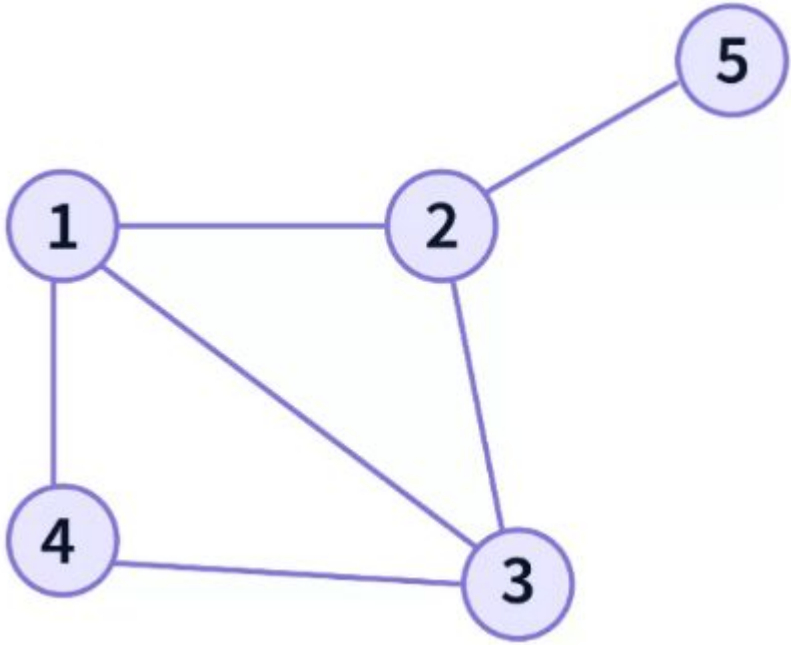
Теория графов. Презентация 1

Команда 1: Перевалов Ефим, Рожков Александр, Нурмухаметов Рафик

Треугольники



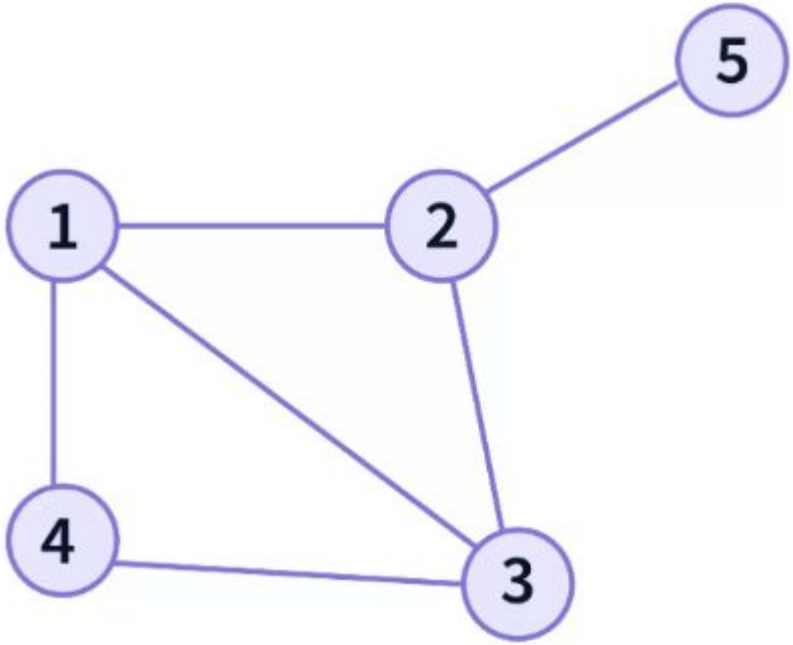
Задача

- Найти все уникальные треугольники в графе

Где может использоваться?

- Социальные сети
- Оптимизации

Треугольники



Burkhardt:

$$\frac{1}{6}\sum_j\sum_i (A^2 * A)$$

Sandia:

$$\sum_j\sum_i (U^2 * U)$$

Где

A - матрица смежности

U - треугольная матрица

* - поэлементное умножение

Треугольники Graph BLAS

A - матрица смежности исходного графа

U - треугольное представление матрицы смежности

* - фильтрация значений из матрицы

Burkhardt:

$$\frac{1}{6}\sum_j\sum_i\left(A^2 * A\right)$$

Sandia:

$$\sum_j\sum_i\left(U^2 * U\right)$$

Где

A - матрица смежности

U - треугольная матрица

* - поэлементное умножение

Треугольники SPLA

Функциональность SPLA для операций над матрицами.

mxm - умножение матриц

emult - поэлементное умножение

Burkhardt:

$$\frac{1}{6}\sum_j\sum_i \left(A^2 * A\right)$$

Sandia:

$$\sum_j\sum_i \left(U^2 * U\right)$$

Где

A - матрица смежности

U - треугольная матрица

* - поэлементное умножение

Набор данных

Были выбраны следующие графы[1]:

- roadNet-CA (1 965 206 вершин, 2 776 607 ребер, кк 0.0464)
- loc-Gowalla (196 591 вершина, 950 327 ребер, кк 0.2367)
- amazon0505 (410 236 вершин, 3 356 824 ребра, кк 0.4064)
- com-Amazon (334 863 вершины, 925 872 ребра, кк 0.3967)
- soc-Pokec (1 632 803 вершины, 30 622 463 ребра, кк 0.1094)

Их выбор обусловлен:

- Прикладным значением
- Разнообразностью
- Использовались для анализа задачи[2]

[1]Stanford Large Network Dataset Collection

[2]Pogozhelskaya-report.pdf

Кратчайшие пути

Задача:

- Для ориентированного (если без отрицательных весов, то можно и не ориентированного) графа без петель необходимо найти кратчайшие расстояния от заданной вершины до всех остальных (достижимых) вершин графа.

Где может использоваться?

- Картография
- Сети дорог

Кратчайшие пути

Задача:

- Для ориентированного (если без отрицательных весов, то можно и не ориентированного) графа без петель необходимо найти кратчайшие расстояния от заданной вершины до всех остальных (достижимых) вершин графа.

Алгоритм Беллмана-Форда:

- Граф $G = \langle V, E, \varphi \rangle$
- Количество ребер в минимальном пути будет не больше чем $|V|$
- Если ребро $e = (u,v)$ существует, то $d[v] = \min(d[v], d[u] + \varphi(e))$, где d - массив расстояний

Кратчайшие пути

Алгоритм Беллмана-Форда (псевдокод без отрицательных циклов):

```
for i in 0..(|V|-1)
  d[i] = +∞
d[s] = 0
for i in 1..(|V|-1)
  for (u,v) in E
      if d[v] > d[u] + w(u,v)
      then d[v] = d[u] + w(u,v)
return d
```

Кратчайшие пути Graph BLAS

Алгоритм Беллмана-Форда:

- Граф представлен в виде матрицы смежности A с весами
- Используем полукольцо min-plus:
  - $a \oplus b = \min(a, b)$
  - $a \otimes b = a + b$

Кратчайшие пути Graph BLAS

Алгоритм Беллмана-Форда (псевдокод без отрицательных циклов):

```
for i in 0..(|V|-1)
  d[i] = +∞
d[s] = 0

for i in 1..(|V|-1)
  d = min.+(d, d x A)

return d
```

Кратчайшие пути Google Pregel

Алгоритм Беллмана-Форда:

- Идея заключается в использовании итеративных супершагов, где вершины обмениваются сообщениями и обновляют свои расстояния до исходной вершины
- Инициализация как в обычном алгоритме
- На первом супершаге активна только исходная вершина
- На каждом супершаге вершины:
  - Получают сообщения с расстояниями
  - Обновляют своё расстояние, если полученное значение меньше текущего
  - Рассылают обновленные расстояния соседям

Набор данных

Были выбраны следующие графы[1]:

- cit-HepTh (27 770 вершин, 352 807 ребер)
- cit-HepPh (34 546 вершин, 421 578 ребер)
- roadNet-PA (1 088 92 вершины, 1 541 898 ребер)
- roadNet-TX (1 379 917 вершин, 1 921 660 ребер)
- roadNet-CA (1 965 206 вершин, 2 776 607 ребер)
- cit-Patents (3 774 768 вершин, 16 518 948 ребер)

Их выбор обусловлен:

- Прикладным значением
- Разнообразностью

1]Stanford Large Network Dataset Collection

Эксперимент

Ход эксперимента:

- Сравнить реализации на разных библиотеках
- Провести несколько запусков
- Собрать информацию о средних значениях
- Провести анализ полученных результатов

Вычислительная система:

- Процессор: 11th Gen Intel(R) Core(TM) i3-1115G4
  - Количество ядер: 2
  - Количество логических ядер: 4
- RAM: 8 GB
- ОС: Ubuntu 20.04.4