

**CAS CS 562: Advanced Database**  
**Programming Assignment#2**  
Yanshan Song U93089867  
Juntao Wang U18801732

**Hadoop naive algorithm**

**Job1:**

**Map:** Input  $v_1 \ v_2$ , emit  $\langle v_1, v_2 \rangle$

**Reduce:** Input  $\langle v, \Gamma(v) \rangle$

For each  $(u, w) (u\_id < w\_id)$  in  $\Gamma(v)$ , produce  $\langle (u, w), v \rangle$

**The result of Job1:**

1	2,43	1
2	2,60	1
3	2,128	1
4	2,161	1
5	2,185	1
6	2,199	1
7	2,200	1
8	2,4954	1
9	2,5673	1
10	2,6020	1

**Job2:**

**Map:** Input  $\langle (u, w), v \rangle$

Produce three items:  $\langle (u, w), v \rangle$ ,  $\langle (u, v), \$ \rangle$ ,  $\langle (w, v), \$ \rangle$

**Reduce:** Input  $\langle key, values \rangle$

If  $\$$  in *values*, For each  $v$  in *values*, *triangleCount* += 1/3

Finally, output *triangleCount*

**The result of Job2:**

---

1	TriangleCount	1597979
---	---------------	---------

### Hadoop improved algorithm

#### Job1:

**Map:** Input  $v_1 \ v_2$ , emit  $\langle v_1, v_2 \rangle$

**Reduce:** Input  $\langle v, \Gamma(v) \rangle$

For each  $u$  in  $\Gamma(v)$ ,

if  $v\_id < u\_id$ , produce  $\langle (v, u), ("smallerIdDegree", |\Gamma(v)|) \rangle$

else produce  $\langle (u, v), ("largerIdDegree", |\Gamma(v)|) \rangle$

#### The result of Job1:

1	1,2	smallerIdDegree,51
2	1,43	smallerIdDegree,51
3	1,60	smallerIdDegree,51
4	1,128	smallerIdDegree,51
5	1,161	smallerIdDegree,51
6	1,185	smallerIdDegree,51
7	1,199	smallerIdDegree,51
8	1,200	smallerIdDegree,51
9	1,4954	smallerIdDegree,51
10	1,5673	smallerIdDegree,51

#### Job2:

**Map:** Input  $\langle key, value \rangle$ , produce  $\langle key, value \rangle$

**Reduce:** Input  $\langle (v, u), [("smallerIdDegree", smallerIdDegree), ("largerIdDegree", largerIdDegree)] \rangle$

If  $smallerIdDegree \leq largerIdDegree$ , produce  $\langle v, u \rangle$ , else produce  $\langle u, v \rangle$

#### The result of Job2:

1	1	10127
2	103827	1
3	103917	1
4	104012	1
5	104013	1
6	104452	1
7	104453	1
8	104454	1
9	105008	1
10	105058	1

#### Job3:

**Map:** Input  $\langle v, u \rangle$ , produce  $\langle v, u \rangle$

**Reduce:** Input  $\langle v, \Gamma(v) \rangle$

For each  $u$  in  $\Gamma(v)$ , if  $v\_id \leq u\_id$ , produce  $\langle v, (u, null) \rangle$  else produce  $\langle u, (v, null) \rangle$

For each  $(u, w)(u\_id < w\_id)$  in  $\Gamma(v)$ , produce  $(v, (u, w))$

**The result of Job3:**

1	1	10127,null
2	1	43,null
3	1	43,10127
4	1	2,null
5	1	2,10127
6	1	2,43
7	9	10,null
8	100	201,null
9	62	100,null
10	100	62,201

**Job4:**

**Map:** Input  $\langle key, values \rangle$

if  $values[1] = null$ , produce  $\langle values, key \rangle$  else produce  $\langle (key, values[0]), \$ \rangle$

**Reduce:** Input  $\langle key, values \rangle$

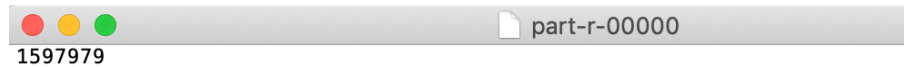
If  $\$$  in  $values$ , For each  $v$  in  $values$ ,  $triangleCount += 1$

Finally, output  $triangleCount$

**The result of Job4:**

1	TriangleCount	1597979
---	---------------	---------

**Pig**



1597979