Now, I’m going to introduce our proposed work.

We designed 5 datasets of a serious of number of person nodes.

Start from 1000 to 5000

And for each dataset, we designed 6 kinds of data structures.

Totally we have 30 databases.

Then each database has 37 number of queries to test the query performance

There are three kind of nodes in the database: Person node, gender node, age node.

Person node will have a label, Name, age and gender attributes.

Age nodes has an age attribute to identify age, range from 1 to 99

Gender nodes has gender attributes to identify its gender.

Now this is six data structures.

The first one only has separate person nodes.

The second add two gender nodes and connect them with person nodes of that gender

D3 add 100 age nodes and connect them with person nodes of that age.

D4 add 100 age nodes and 2 gender nodes. Then connect them with person nodes respectively

D5 first add 100 age nodes. Then each age nodes has two gender nodes. Then each gender nodes connect with person nodes of that gender and age.

D6 first added 2 gender nodes. Then each gender nodes has 100 age nodes. Then each age nodes connect with person nodes of that age and gender.

Now I’m going to introduce our queries.

First query only query by age in a range.

By changing the age range, we can obtain selection factor from 10% to 100%

2ed query only by gender.

By male or by male and female.

Query 3 query by male and a age range.

By changing the age range we can obtain selection factor from 5% to 50%

Query 4 query by mix gender and each of the gender has a range of age. By optimize work load, it's a upper triangular matrix, because the lower triangular will be a mirror of the upper side.

Two unique contribution of our work.

读一下