Classwork

Roll: 1605107

4.1. T0 tuple will go into N0 node, T1 into N1 and so on. After T39 node again the Tuples will cycle from T0 to T39 nodes.

4.2. We have to create a hash function in such a way that we ~~get~~ as ~~evenly~~ distribute tuples as evenly as possible. For 'a' we can create a hash function using NID %40. Again for 'b' hash(street, city, district)% such that we get even distribution.

4.3 we have 400 tuples and 20 nodes.

So there will be 19 partitioning attribute.

a) so, partition vector can be,

[021, 041, ..., 381] where each attribute represents last 3 digits of ID

b) P0 → v < 021

P1 → 021 ≤ v < 041

P19 → v ≥ 381

1605107

4.4

As same age and of-age are in same node we don't need to search other than the partition we need for the particular query. So all node can be ~~run~~ ~~ex~~ used paralally for execution.

As we can utilize all nodes speed-up will be very close to linear n and also scale-up will be very close to linear 1.

5.1

(a) As nodes are evenly distributed scanning entire relation will be performant.

(b) for point query we ~~need to~~ cannot tell where the point is actually located so need to search who all nodes.

(c) For range query same problem, so not very v performant.

5.2

(a) Good if the hash function can destribute evenly.

(b) good for point query if the hash parameter is included.

(c) Not very efficient for range query as tuples are not destributed based on ranges.

5.3

(a) Good as tuples are destributed evenly.

(b) good if query parameter is same as the attribute used for creating ranges.

(c) good for range queries for same as b.

6.1 (a) If we have too many tuples with same hash value. (attribute value skew)

(b) If we have too many tuples with same range and execution of some range of values are much higher than others.

6.2

(a) Equi-width histogram

(b) Total Tuples = 5x(45+35+25+50+15)

= 850

so as we have 4 nodes there will 3 partioning attribute and each node will ~~consit~~ consist nearly 213 tuple.

partition veetor = [6, 13, 19]

(c) New approximate histogram