Tri-variate Optimization Strategies of Semi-Join Technique on Distributed Databases

Volume 66 - Number 6

Year of Publication: 2013

Authors:
Sunita M. Mahajan
Vaishali P. Jadhav

10.5120/11091-6043

Abstract

The problem of finding an optimal strategy to minimize the data transmission cost in distributed database systems, even with the one join attribute is a NP-Hard problem. Determining the optimal sequence of join operations in query optimization leads to exponential complexity. To deal with such a problem, there is need to develop a heuristic approach to solve the problem in polynomial time. This paper mentioned the use of semi-join operation. Beneficial Semi-join operation reduces the amount of data transmission required to perform the join sequences. This paper addresses the optimization of queries with one and more than one join attributes.

References

- Subir Bandyopadhyay, Qiuling Fu, Joan Momssey, and A. Sengupta. A multiattribute semijoin operation for query optimization in distributed databases 1996.
- Pauray S. M. Tsai and Arbee L. P. ChenOptimizing Entity Join Queries by Extended Semijoins in a Wide Area Multidatabase Environment,IEEE international conference ob parallel and distributed databases 1994
- Lin Zhou et al., The semi-joinoptimization in distributed databases, National Conference
Tri-variate Optimization Strategies of Semi-Join Technique on Distributed Databases

- Chen and Yu, Interleaving a join sequence with semi-joins in distributed query processing.
- Stoker et al., Integrating semi-join reducers in the state of the art query processors, German Research Concil (DFG).
- Zemanek et al., Optimizing SPARQL Queries over Disparate RDF Data Sources through Distributed Semi-joins.

Index Terms

Computer Science
Information Sciences
Keywords
Distributed Database  Query Optimization  Gainful Semi-Join  Beneficial Semi-Join