List of Figures

List of Tables

List of Algorithms

Glossary of Terms

ACA	Ant Colony Algorithm
AMOSA	Archived Multi-objective Simulated Annealing
AMOSA-MVS	Archived Multi-objective Simulated Annealing for Materialized View Selection
AST	Abstract syntax tree
CPU	Central Processing Unit
CSA	Clonal Selection Algorithm
DAG	Directed Acyclic Graph
DE	Differential Evolution
DEMO^{NS-II}	Differential Evolution based variants of NSGA-II
DFS	Distributed File System
DIMMQ	Discardable In-Memory Materialized Query
DTLZ	K. Deb, L. Thiele, M. Laumanns and E. Zitzler defined test problems for multi-objective optimization techniques.
EA	Evolutionary Algorithm
GA	Genetic Algorithm
GDE3	Generalized Differential Evolution - 3
HDFS	Hadoop Distributed File System
HDP	Hortonworks Data Platform
HRU	Venky Harinarayan, Anand Rajaraman and Jeffrey D. Ullman, Stanford University
MA	Memetic Algorithm
MB	Mega Byte
MODE-BE	Multi-Objective Differential Evolution for Binary Encoded data
MOEA	Multi-Objective Evolutionary Algorithm
MOGA	Multi-Objective Genetic Algorithm
MOO	Multi-Objective Optimization
MOSA	Multi-objective Simulated Annealing
MVPP	Multiple View Processing Plan
NPGA	Niched Pareto Genetic Algorithm
NP-hard	Non-deterministic polynomial-time hard
NSGA-II	Non-dominated Sorting Genetic Algorithm-II
OLAP	On-Line Analytical Processing
OPTICS	Ordering Points to Identify Clustering

	Structure
PB	Petabyte
PGA	Polynomial Greedy Algorithm
PSA	Parallel Simulated Annealing
PSO	Particle Swarm Optimization
RDBMS	Relational Database Management System
RDD	Resilient Distributed Data-set (Spark's)
SA	Simulated Annealing
SMC	Simple Matching Coefficient
SQL	Structured Query Language
SSD	Solid-state drives
TB	Terabyte
TPC	Transaction processing Performance Council
TPC-H	Transaction processing Performance Council (Benchmark version H)
ZDT	Zitzler, Deb and Thiele's test problem for evaluating Evolutionary Multi-objective optimization techniques

Symbols and Notations

$\mathbb{B}$	Set of all truth values
$\mathbb{E}(S)$	The set of all equivalence relations over a given set S
Δdom	Amount of domination between solutions
$\Delta dom_{average}$	Average domination between solutions
Δdom_{min}	Minimum amount of domination between solutions
δE	Energy difference between solutions generated in Simulated Annealing algorithm
$\equiv$	Equivalent
γ	Convergence metric of multi-objective optimization technique
Γ	A real valued constant
μMax	Mean of maximum distances between each solution vectors w.r.t all other solution vectors obtained
$\not\prec$	Does not dominate
$\prec$	Dominates
σMax	Standard deviation of maximum distances between each solution vectors w.r.t all other solution vectors obtained
$\sim$	A relation
$\triangleq$	Defined to be equal to
Ξ	The set of *formae* or equivalence classes
T	System temperature in Simulated Annealing process
A	Availability of memory space for materializing views
$A_{V'}$	Space required for materializing the set of sub-queries V'
A_M	Space required for materializing the set of views M
$B(v, S)$	Total benefit of materializing view v if S is the set of view selected at an iteration of HRU-greedy algorithm
$C(v)$	Cost of processing view v
$C_a^q(v)$	Cost of accessing query q when view or vertex v of an MVPP graph is materialized
C_M^Q	The total cost of responding queries Q when the set of views M is materialized
$C_m^r(v)$	Cost of maintaining a view or vertex v due to change in the base relation r
$C_{total}(v)$	Total processing cost of view v in an MVPP graph.
$C_{V'}^Q$	The total MapReduce cost of processing set of

	queries Q when a set of sub-queries or sub-expressions as views V' is materialized
CR	Real valued cross-over probability constant in the range[0,1]
d	Number of dimensions of a data cube
D_Ψ	The set of different parameters between equivalence relations using basis Ψ .
$DE/x/y/z$	Version of Differential Evolution with x vector to be mutated, y specifying whether *random* or *best* and z specifies cross-over scheme
F	Real valued amplification vector in the range [0,1] in DE
$\tilde{\mathscr{F}}$	The current estimate of Pareto front
$\mathscr{F}$	Pareto front
f_q	Frequency of a query q
f_u	Updating frequency of base relations of an MVPP graph
f_v	Frequency of queries on view v of an AND-OR view graph
g	Generation number in an evolutionary process
g_{max}	Maximum number of generations in an evolutionary process
g_v	Updating frequency of a view v of an AND-OR view graph
L_i	Number of levels of hierarchies associated with dimension i of data warehouse
$\mathbf{M}$	Objectives of a multi-objective optimization problem
M_{de}	Formae basis based DE mutation operator template
$\mathbb{P}$	Probability (function value)
$\mathscr{P}$	Set of Pareto optimal solutions
$Q(v, M)$	Cost of answering sub-query v of an AND-OR view graph when set of views M is materialized
$Q_G(M)$	Total query processing cost of MVPP graph G, if the set of views M is materialized
$R(v)$	Resultant relation corresponding to vertex v in an MVPP graph
R_v	Cost incurred in reading the materialized view v of an AND-OR view graph
$randI()$	Random integer index generator
$randR()$	Evaluation of a uniform random number generator in the range [0,1]
$\mathbf{S}$	Set of vectors
T_q	Query processing tree for query q
$\tau(G, M)$	Processing cost of AND-OR view graph G when the set of views M is materialized
$U(M)$	The total updating cost of set of materialized views M
$U_G(M)$	Updating cost of MVPP graph G when set of views M is materialized
$UC(v, M)$	Maintenance cost of a sub-query or view v when the set of views M is materialized