Chapter 5

Random Walk Grey Wolf Optimizer Algorithm for Materialized View Selection (RWGWOMVS)

Anjana Gosain
Guru Gobind Singh Indraprastha University, Delhi, India

Kavita Sachdeva
Shree Guru Gobind Singh Tricentenary University, Haryana, India

ABSTRACT

Optimal selection of materialized views is crucial for enhancing the performance and efficiency of data warehouse to render decisions effectively. Numerous evolutionary optimization algorithms like particle swarm optimization (PSO), genetic algorithm (GA), bee colony optimization (BCO), backtracking search optimization algorithm (BSA), etc. have been used by researchers for the selection of views optimally. Various frameworks like multiple view processing plan (MVPP), lattice, and AND-OR view graphs have been used for representing the problem space of MVS problem. In this chapter, the authors have implemented random walk grey wolf optimizer (RWGWO) algorithm for materialized view selection (i.e., RWGWOMVS) on lattice framework to find an optimal set of views within the space constraint. RWGWOMVS gives superior results in terms of minimum total query processing cost when compared with GA, BSA, and PSO algorithm. The proposed method scales well on increasing the lattice dimensions and on increasing the number of queries triggered by users.

DOI: 10.4018/978-1-7998-2975-1.ch005
INTRODUCTION

OLAP (Online Analytical Processing) queries triggered by the database users, demand aggregated and compiled results instantly. Therefore, for effective decision making, information systems are required, which disseminates aggregated and instant results. Data warehouse is a kind of information system used for decision making. A data warehouse (DW) (Han & Kamber, 2001; Morse & Issac, 1998) is a data repository integrated from numerous varied sources of data, used for decision support querying and analysis. The biggest concern is to handle such datastore in a cost effective way. ETL process extracts the data from source systems, then transform the data and finally loads the data into the data warehouse. Therefore, data is stored in the warehouse as native data sources along with derived views. Such derived views are referred as materialized views (Jain & Gosain, 2012).

Materialized views raises the processing efficiency of triggered queries, by avoiding the access to native data sources so as to reduce the total processing cost of queries. Such approach minimizes the query processing cost, but it increases the maintenance cost of chosen views. Thus, the most important task to design a data warehouse is to minimize the total query processing cost which includes both the query evaluation cost and view maintenance cost, while finding an appropriate number of views, within the storage space constraint. This problem is classified as materialized view selection (MVS) problem.

For solving the problem of MVS, various frameworks and solutions exist in literature. Frameworks like Data cube and lattice (Harinarayan et al., 1996), Multiple View Processing Plan (MVPP) (Yang et al., 1997), AND-OR view graphs (Gupta & Mumick, 2005; Mami et al., 2011; Tamiozzo et al., 2014; Horng et al., 2003) have been used for representation of views. Lattice framework (Harinarayan et al., 1996) has the ability to capture dependency among aggregate views and thus creates data cubes at multiple dimensions. To answer a query in minimum possible time, lattice framework access the smallest data cube available. MVPP framework (Yang et al., 1997), a global query processing plan for complete set of queries, exploits the existence of common sub-expressions for most of the queries. AND-OR view graph (Gupta, 1997) is used to express all the possible execution plans for evaluating a query in the query set. To represent all the data cubes at distinct aggregation levels, lattice framework (Harinarayan et al., 1996) portrays user queries easily. Therefore, numerous studies have chosen lattice framework for representing the problem space of MVS problem. Further, to minimize the total query processing cost, various researchers have implemented numerous evolutionary optimization algorithms (EA’s) like Particle Swarm Optimization (PSO) (Gosain & Heena, 2016; Sun & Ziqiang, 2009), Bee Colony Optimization (BCO) (Kumar & Arun, 2015), Adaptive Genetic algorithm (AGA) (Yu et al., 2015) etc., in selecting the materialized views. This
Related Content

A Methodology for Improving Business Process Performance through Positive Deviance
[www.igi-global.com/article/methodology-improving-business-process-performance/80242?camid=4v1a](www.igi-global.com/article/methodology-improving-business-process-performance/80242?camid=4v1a)

Redesigning a SAD Course to Promote Problem-Based Learning
[www.igi-global.com/chapter/redesigning-sad-course-promote-problem/21092?camid=4v1a](www.igi-global.com/chapter/redesigning-sad-course-promote-problem/21092?camid=4v1a)
Conformance Analysis of Organizational Models: A New Enterprise Modeling Framework using Algebraic Graph Transformation
www.igi-global.com/article/conformance-analysis-organizational-models/75464?camid=4v1a

The SOA Frontier: Experiences with Three Migration Approaches
www.igi-global.com/chapter/soa-frontier-experiences-three-migration/72215?camid=4v1a