Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0409U005036

Особливі позначки: відкрита

Дата реєстрації: 16-11-2009

Статус: Захищена

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Петров Олександр Валерійович
2. Petrov Olexandr Valeriyovich

Кваліфікація:

Ідентифікатор ORCID ID: Не застосовується

Вид дисертації: кандидат наук

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 05.13.06

Назва наукової спеціальності: Інформаційні технології

Галузь / галузі знань: Не застосовується

Освітньо-наукова програма зі спеціальності: Не застосовується

Дата захисту: 16-10-2009

Спеціальність за освітою: 8.080401

Місце роботи здобувача: Державний вищий навчальний заклад "Донецький національний технічний університет"

Код за ЄДРПОУ: 02070826

Місцезнаходження: пл. Шибанкова, 2, м. Покровськ, Донецька обл., 85300

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): К11.051.08

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

IV. Відомості про підприємство, установу, організацію, в якій було виконано дисертацію

Повне найменування юридичної особи: Державний вищий навчальний заклад "Донецький національний технічний університет"

Код за ЄДРПОУ: 02070826

Місцезнаходження: пл. Шибанкова, 2, м. Покровськ, Донецька обл., 85300

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

V. Відомості про дисертацію

Мова дисертації:

Коди тематичних рубрик: 50.41.21

Тема дисертації:

1. Моделі, методи та інструментальні засоби оптимізації розподілених сховищ даних

2. Models, methods and tools for distributed data warehouses optimization

Реферат/анотація:

1. Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 - Інформаційні технології, Донецький національний технічний університет, 2009. У дисертації розглядається важлива науково-практична задача оптимізації складу матеріалізованих уявлень, а також розміщення даних у розподіленому сховищі даних (РСД). Дисертаційна робота присвячена розробці нових моделей та методів для оптимізації розподілених сховищ даних. Проведено аналіз розподіленого сховища даних, визначені основні його компоненти. Використовуючи об'єктно-орієнтований підхід, було розроблено об'єктні моделі типових компонентів РСД. Загальну модель розподіленого сховища даних побудовано як систему об'єктних моделей типових компонентів РСД, що взаємодіють. Запропоновано новий підхід до розв'язання задачі оптимізації розподілених сховищ даних, який базується на сумісному використанні розробленої об'єктної моделі РСД та модифікованого генетичного алгоритму, що дозволяє підвищити ефективність роботи розподіленого сховища даних за рахунок зменшення середнього часу виконання користувацьких запитів.

Розроблено нову модифікацію генетичного алгоритму для задачі оптимізації РСД. Для кодування схем розміщення даних у розподіленому сховищі даних вперше використані подвоєні мультихромосоми. Для виконання операції з подвоєними мультихромосомами розроблено нові версії генетичних операторів рекомбінації, схрещування, мутації та відбору. На основі об'єктної моделі та модифікованого генетичного алгоритму створено інструментальний засіб для моделювання та оптимізації розподілених сховищ даних, що складається з блоків моделювання, оптимізації та аналізу результатів моделювання. Проведено серію машинних експериментів з моделювання та оптимізації РСД. На основі проведених експериментів визначені раціональні параметри модифікованого генетичного алгоритму, а також було розроблено рекомендації щодо підвищення ефективності розподіленого сховища даних. Ключові слова: розподілене сховище даних, таблиця фактів, таблиця виміру, матеріалізоване уявлення, об'єктно-орієнтоване моделювання, генетичний алгоритм, схрещування, мутація, рекомбінація.

2. Candidate of technical sciences' thesis on specialty 05.13.06 - Information technologies. Donetsk national technical university, 2009. In this dissertation a new scientific and practical problem of optimization of a set of materialized view, and also of optimization of data allocation in a distributed data warehouse is being observed. The thesis is devoted to development of new models and methods of distributed data warehouses (DDW) optimization. Structural analysis of distributed data warehouse has been conducted and on the basis of this analysis the main components of its physical and logical architecture have been revealed. Then using object-oriented approach object models of the distributed data warehouse's typical components were developed. For these models' development UML language was used. A new approach to the problem of optimization of distributed data warehouses has been proposed. This approach is based on a simultaneous use the object-oriented model of distributed data warehouse that had been developed before and a modified genetic algorithm. The modified genetic algorithm is used to obtain new schemas of data allocation in a distributed data warehouse (by manipulations with schemas of data allocation that are represented by doubled multichromosomes with genetic operators of recombination, crossover and mutation). In the same time object model of distributed data warehouse receives schemas of data allocation in DDW and then it is being used to calculate values of criterion of efficiency of distributed data warehouse. These values then are sent to the genetic algorithm which uses them as values of fitness-function. The approach proposed allows increasing distributed data warehouse's efficiency by decreasing of the average time of user select queries processing. A new modification of genetic algorithm for the problem of distributed data warehose optimization was developed. For encoding schemas of data distribution in distributed data warehouse for the first time doubled multichromosomes were used. New versions of genetic operators of recombination, crossover and mutation were developed. This modification of genetic algorithm allows obtaining suboptimal solutions of the problem of distributed data warehouse optimization. Also in this modification not only data allocation between the nodes of distributed data warehouse is being optimized but the set of materialized views is being optimized. Basing on the developed object model and the modified genetic algorithm a new tool for distributed data warehouse modeling and optimization was developed. This tool consists of three subsystems: subsystem of modeling, analysis and optimization of distributed data warehouse. The tool that has been developed can be used at the stage of distributed data warehouse development as well as it can be used when a distributed data warehouse is already implemented and functioning. As an object of experiments the distributed data warehouse of Donetsk regional office of Prominvestbank was selected. Real data of functioning of this DDW were obtained and used to test model adequacy. Then a series of computational experiments of modeling and optimization of distributed data warehouse was made. At the beginning an influence of data allocation and technical characteristics of DDW to its efficiency was researches. These experiments showed that doubling of channels' capacity leads to a slight increasing of DDW efficiency. Also it has been researched that creation of some new materialized views and allocation them on the nodes of most use leads to increasing of DDW efficiency as well. But this increase isn't enough to solve the distributed data warehouse optimization problem since the solutions obtained are far from optimal anyhow. Basing on the experiments' results rational parameters of the modified genetic algorithm were determined (such as probabilities genetic operators, number of generation and number of individuals in the population). Then these parameters were used for genetic algorithm that was used for

data warehouse optimization. The suboptimal solution obtained with use of genetic algorithm was compared with optimal solution that was obtained used brute force method and the difference between these solutions was just 4,37%. This means that the genetic algorithm developed allows to receive suboptimal solutions that are very close to the global optimum. Also some recommendations for distributed data warehouse efficiency increasing were developed. Keywords: distributed data warehouse, fact table, dimension table, materialized view, fragmentation, replication, object-oriented modeling, genetic algorithm, crossover, mutation, and recombination.

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки:

Стратегічний пріоритетний напрям інноваційної діяльності:

Підсумки дослідження:

Публікації:

Наукова (науково-технічна) продукція:

Соціально-економічна спрямованість:

Охоронні документи на ОПІВ:

Впровадження результатів дисертації:

Зв'язок з науковими темами:

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Лаздинь Сергій Володимирович
2. Lazdyn Sergiy Volodymirovich

Кваліфікація: к.т.н., 05.13.06

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Левикін Віктор Макарович

2. Левикін Віктор Макарович

Кваліфікація: д.т.н., 05.13.06

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Власне Прізвище Ім'я По-батькові:

1. Ульшин Віталій Олександрович

2. Ульшин Віталій Олександрович

Кваліфікація: д.т.н., 05.13.07

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Рецензенти

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові голови ради Каргин Анатолий Алексеевич

Власне Прізвище Ім'я По-батькові головуючого на засіданні Каргин Анатолий Алексеевич

Відповідальний за підготовку облікових документів

Реєстратор

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності



Юрченко Т.А.