International Journal of Engineering & Technology, 7 (2.13) (2018) 133-135

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

Optimization structure database of complex systems

A. M. Korneev *, A. K. Pogodaev, L. S. Abdullakh, M. V. Salfetnikov

Lipetsk State Technical University
**Corresponding author E-mail: d.48@rambler.ru*

Abstract

In the article such ways of optimization of the structure of a database of complex structured objects that allow simplifying the scheme of an information system. As a result, the response time of the information system to the user's request is reduced, and it becomes possible to more effectively manage and monitor the progress of ongoing processes. A formalized model of the relational expression is developed, a model of the corresponding equivalent transformation of the chosen relational expression is created, a formal description of the developed algorithm for optimizing the queries is presented.

Keywords: *Optimization Structure; Database; Information; System; Modeling Processes.*

1. Introduction

On large complex structured objects that contain a large number of units, the information system consists of a number of subsystems participating in the management of the units. This leads to complex algorithms for obtaining, transmitting and processing data of processed materials, for data loss, for delays in solving operational problems, making it difficult to track the progress of individual stages and production as a whole.

The development of complex structured systems takes into account the availability of modern automated tracking systems, the collection of information about processes and the management of these processes [1-4].

However, during the operation of the information system, unused attributes that take up additional memory and reduce the performance of the database, complicated connections between the elements and their attributes are identified.. This is due to the redundancy of the information system.

Information support of the automated system for modeling processes with complex structure takes into account a set of different data reflecting the economic and production activities of the research object, regulatory documentation, language tools and organizational methods of creating and maintaining information arrays.

To store data, a relational dbms is used, which makes it possible to provide the basic principles of the organization of information support, which are necessary for such a class of tasks:

- Minimization of redundancy of stored data;;
- Organization of maximum use of video terminals to provide direct user access to information;
- Availability of information presented to the user in the form of tables, and the ability for the user to manipulate this data using the pl / sql query language;
- Ability to use high-level programming languages to implement data access;;
- Ensuring reliability and safety.

2. Research method

In the first step, a group of final attributes is identified, associated with one common attribute located above the level, and denoted by a relational expression, where the common attribute will be the key attribute of the normalized relationship.

For a given set of attributes, a set of relations is created.

$A=\{A_1,\dots,A_n\}$ - set of attributes. Any element is described by the set of values accepted by the attribute.

$X(B), B\subseteq A$ - cartesian product of elements in the set B, which is a subset A.

$\beta(X(B)), B\subseteq A$ - set of subsets of set $X(B), B\subseteq A$, which is the set of all relations defined on the attributes $B\subseteq A$.

Since each attribute is used once in a relationship, you must specify each attribute individually to establish a one-to-one correspondence between the attribute and its semantics.

$$M=\bigcup_{\forall B\subseteq A}\beta(X(B))$$

- Set of all relations defined on all subsets a. Each relationship $m\in M$ is a set of tuples representing an enumeration of values each relation has type $B\subseteq A$.

a relational expression is a set of arguments that assign values from a given «set of relations M, logical expressions, subset of the set of attributes, operations of relational algebra».

A relational expression contains a set of elements.

B - set of elements of a relational expression.

$O\subseteq B$ - a set of relational operations participating in the expression.

$G\subseteq B^2$ - set of relations of expression

$$(a,b)\in G, a\in O, b\in B$$

As a result, a statement is formed:

$$\forall a \forall b[(a,b)\in G \to a\in B \wedge b\in B] \wedge \forall a \exists b[a\in B \to (a,b)\in G \vee (b,a)\in G]$$

Any element of a relational expression must be indirectly or directly linked to other elements of the relational expression. Violation of this principle entails a loss of correctness of the relational expression. Any operation can be an argument of only one other operation.

$$\forall a \forall b \begin{bmatrix} a\in O \wedge b\in O \wedge (a,b)\in G \\ \to \overline{\exists c[c\in O \wedge (c,b)\in G]} \end{bmatrix}$$

For any operation b, that is an operation argument, there is no other operation, b.

To optimize the conversion of a relational expression, it is formalized. Any selection is presented in the form of a cascade of selections. Each selection moves to the side of greater nesting. In this case, the projections can fade and split. Each cascade of selections and projections is combined into single selection, single projection or selection with subsequent projection [5, 6].

The operation of the projection cascade has the domain of definition

$$f(C_2,\theta,o) = o\in\Pi \wedge \arg_\theta o \in \Pi$$

Operation of the cascade of projections:

$$C_2(\theta,o) = C_2'(\theta,o,\arg_\theta o)$$

The operation of transforming an expression according to the law of the projection cascade is of the type $C_2 : \Theta\times\Pi\times\Pi \to \Theta$ and is written in a functional form $\theta' = C_2(\theta,a,b)$. The operation is performed on projection operations a и b, for which the assertion holds that b argument in favor a, and attributes a are among the attributes b.

$$\begin{aligned} & a\in\Pi \wedge b\in\Pi \wedge (a,b)\in G \wedge \forall p[p\in\Pi \wedge (a,p)\in G \to (b,p)\in G] \to \\ & \forall d\forall e[(d,e)\in G \wedge (d,e)\neq(a,b) \wedge d\neq b \to (d,e)\in G'] \wedge \\ & \exists g[(b,g)\in G \wedge g\in H \wedge (a,g)\in G'] \wedge \\ & \forall d\forall e[(d,e)\in G' \to (d,e)\in G \vee \exists g[g\in H \wedge (d,e)=(a,g) \wedge (b,g)\in G]] \wedge \\ & \forall h[h\in B \wedge \exists k[(k,h)\in G'] \to h\in B'] \wedge \\ & \forall h[h\in B' \to h\in B] \end{aligned}$$

The operation of the selection cascade has the domain of definition

$$f(C_3,\theta,o) = o\in\Sigma \wedge \arg_\theta o\in\Sigma \cdot$$

Operation of the cascade of selections:

$$C_3(\theta,o) = C_3'(\theta,o,\arg_\theta o)$$

The operation of transforming an expression according to the law of a cascade of selections is of the type $C_3 : \Theta\times\Sigma\times\Sigma \to \Theta$ and is written in functional notation $\theta' = C_3(\theta,a,b)$. The operation is defined on such two selection operations a и b, for which it is true that b is an argument a.

$$\begin{aligned} & a\in\Sigma \wedge b\in\Sigma \wedge (a,b)\in G \to \\ & \forall d\forall e[(d,e)\in G \wedge (d,e)\neq(a,b) \wedge d\neq b \to (d,e)\in G'] \wedge \\ & \forall g[(b,g)\in G \to (a,g)\in G'] \wedge \\ & \forall d\forall e[(d,e)\in G' \to (d,e)\in G \vee [d=a \wedge e\in S\cup H]] \wedge \\ & \forall h[h\in B \wedge h\neq b \to h\in B'] \wedge \\ & \forall h[h\in B' \to h\in B] \end{aligned}$$

All attributes can be divided into participating and not participating in queries. The decision on the relation of a specific attribute to a group is taken by the designer, relying on a denormalized or normalized graph of functional dependencies. In the graphs, all the attributes specified before the stage of construction are marked.

A group of attributes that do not participate in queries can be divided into those on which the logical integrity of the information system depends, and those in the storage of which there is no need [7-11]. It is not necessary to store the final attributes of the functional dependency graph that are not involved in the queries, and the key attributes of the junior relationships that were denormalized, since this attribute becomes final. When deleting the final attributes, you must also exclude all functional dependencies combined with these attributes.

3. Results and analysis

Application of the query optimization algorithm when creating an information system for complex structured objects. The number of queries that can be performed in an information system of complex structured objects is quite large. And in the process of operating the system, the number of requests increases with the growing need for research on incoming data.

The algorithm for query optimization is presented on the example of the query: "determine the dates for obtaining the types of products sold when performing operations that started earlier 1.01.2018"». The relationships required to implement this query are shown in fig.3.

Table 1: Example Relationship

relationship r			
Product			A
Role in the operation			B
Operation			C
relationship q			
Product		A	
Date creation		E	
Date of useing		F	
relationship w			
Operation	C		
Date of the beginning	L		
Date of completion	J		

The original query is written as the following relational expression:

$$\pi_F(\sigma_{I<'1.01.2018'}(\pi_{R.A,B,R.C,E,F,I,J}\,\sigma_{R.A=Q.A\wedge R.C=W.C}(R\times Q\times W)))$$

The disadvantage of this relational expression is the large expenditure of computational resources for its implementation. The graphical representation of the described relational expression is shown in figure 1.

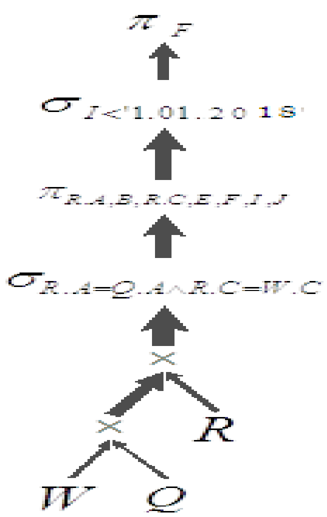

Fig. 2: Initial Relational Expression.

Applying the formalized algorithm of query optimization to the original relational expression, the optimal relational expression is formed:

$$\pi_F(\sigma_{R.A=Q.A}(\pi_{Q.A,Q.F}(Q)\times\pi_{R.A}(\sigma_{R.C=W.C}$$
$$(\pi_{R.A,R.C}(R)\times\pi_{W.C}(\sigma_{I<'1.01.2002'}(W))))))$$

It is structurally more complex than the original relational expression, but its implementation is more efficient graphical representation of the developed optimal relational expression is shown in figure 2.

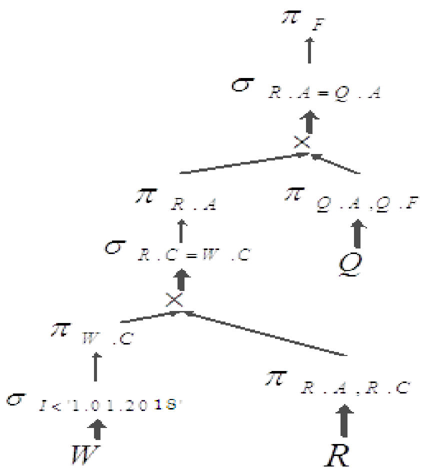

Fig. 3: The Optimal Relational Expression.

In the figures, the thickness of the lines shows the relative amount of information processed at each level of the relational expression being examined. In the optimal relational expression, at each level, a lower amount of information is realized relative to the original relational expression. The result justified the higher efficiency of the resulting query.

4. Conclusion

The method of query optimization when creating an information system for complex structured objects is considered. The original relational expression and the resulting optimal relational expression are given.. From the initial formal model, a lot of basic functional dependencies are obtained; the optimal structure of the implemented database was developed on its basis, the algorithm of query optimization. As a result, we developed a formal model of each relational expression, established a formal model of the corresponding equivalent transformation of the selected relational expression, a formal description of the algorithm to optimize queries investigated.

References

[1] A.m. Korneev, l.s.abdullakh, t.a. Smetannikova, international journal of applied engineering research. 2015. T. 10. № 11. C. 27961-27980.

[2] Korneev, a. M. Methods for the identification of cross-cutting technology production of steel products: monograph / lipetsk state pedagogical university. - lipetsk: lgu, 2009. - 286 p.

[3] Korneev, a. M. Cellular structural-hierarchical modeling of complex prostranstvenno-distributed systems. - news of higher educational institutions of the central chernozem region, 2011, no. 1, pp. 62-66.

[4] Pogodaev a.k. A method for optimizing a query-based alternative soedi- neniya database tables for an information systems governmental enterprises [text] / ak pogodaev, a.y muraveyko, d.v dyatchin, // proceedings of the tsu. Series: knowing machines, systems of drives and a detail of machines. Spets.vyp. - tula: tula state university publishing house, 2006 - pp 331- 337.

[5] kuznetsov, l. A., a. Korneev, m. Automated system of selecting the optimal technology for the production of steel. - izvestiya vuzov. Ferrous metallurgy, 1994, no. 5, pp. 45-48.

[6] S.l. Blyumin, a.m. Shmyrin, o.a. Shmyrina. Bilinear neighborhood systems. Monograph. Lipetsk, lipetsk state technical university publishers, 2006. 131 p, (in russian).

[7] A.l. Genkin, i.v. Nikulin, the use of simulation in the control sheet-rolling complex, informacionnye tehnologii i vychislitel'nye sistemy, 2011, № 2, p. 75-79, (in russian).

[8] J.k. Alekseev, a.p. Sukhorukov, introduction to the theory of catastrophes, librokom, 2013, 176 pp.

[9] Pogodaev a.k., muraveyko a.u, dyatchin d.v., alternative connection for query optimization in relational systems // vesti universities chernozem 2006, №1 (3). S.13-15.

[10] A.m. Korneev, l.s.abdullakh, t.a. Smetannikova,.simulation of difficult industrial systems in the form of the probable automatic machine, metallurgical and mining industry volume 7, issue 2, 2015, pages 257-261.

[11] Abdullah l.s, korneev a.m, smetannikova t.a. Structural cell-hierarchical identification of complex spatially distributed production systems // proceedings of the 3rd international academic conference. 2013, st. Louis, missouri, usa. C.75-79.