Reduce the cost of transporting the parts with Genetic Algorithm

[1]Seyed Masoum Hoseini , [2*]Mina Mahdavi

[1]Department Of Computer Engineering , Amol Branch , Islamic Azad University , Amol , Iran

[*2]Department Of Computer Engineering , Amol Branch , Islamic Azad University , Amol , Iran

_____

Abstract

The main objective of the allocation algorithm to determine the proportion of the various parts of the site is to reduce the cost of shipping the piece. This study presents an algorithm to reduce the cost of data transfer when allocating parts is, in such a manner that First of all, parts of the of the site, it should be, second reduce the cost of transmission components sites should Third update costs for all components that are present at the site should be optimized, However, all the conditions for the transfer of part of the site is selected, so that the reliability and availability of the site increases. In this study, the assignment of parts to a site in a distributed database system is done using genetic algorithms, Efforts to reduce the cost of data transmission. To review and consider the advantages of the previous method, the initial population, we have tried to better optimality of previous approaches is the initial population And a genetic algorithm to select the best chromosome instead of two parameters, the three parameters used. Analytical study shows that the proposed algorithm to other algorithms, the benefits and availability of data and reduce costs in the allocation of parts.

Key words: Distributed database, segmentation data, genetic algorithms, communication costs, allocation of parts

_____

1 – Introduction

Some of the benefits are distributed database system reliability and availability. System reliability as the probability of being active at a particular time, to be defined. Definition of availability, the probability that the system is continuously available for a period of time. In a common centralized database system, failure of a site as a failure of the system. In a distributed database of a site failure affects only the access to the site. Clearly, this problem would be to improve the reliability and availability. Another advantage of better performance in the distributed database and distributed processing is observed.

A possible drawback of added complexity and overhead of a distributed database that is designed and implemented there. Strategies for distributed queries and transactions must be expanded. The design process involves dividing the distribution and allocation of these pieces. The purpose of segmentation to better distribution. Optimal allocation of parts of the existing site is very important. The placement of redundant components, increased consistency and efficiency of the available redundancy allocated segment is created, but in other parts of this problem is resolved. Finally, the distributed database must be able to fully repair the damage then good.

2- Background And the Literature Research

- Literature Research

Communication costs

In a distributed database transfer from one site to another site with a piece of transfer fees or the cost of communication. Communication costs, the most important parameter for evaluating the allocation algorithms are given parts. Hence all parts allocation algorithms are proposed to reduce the parameter structure. The most important factor for the transfer fee or the cost of communication in distributed database is a component allocation algorithms.

Assign parts

Distributed database performance increases when the parts are stored in the nodes, which are gaining more space. The problem is finding the specific nodes for each piece. Counting the number of accesses per node to a piece offers a practical solution for this purpose. With the maximum amount available for a particular piece, a node can be a primary candidate for the storage device. In order to allocate parts processing algorithms is proposed based on the cost function is assigned questions and to accurately determine whether a piece of a site devoted to or deleted from the site.

- Literature
- ✓ Mir and Handel (2002), have proposed a design method and the clustering of sites used Communication costs between sites in the allocation of plots in a distributed database system, To a minimum, resulting in increased performance in the network

environment. In this way, the allocation of the components so that a series of interrelated components are assigned to sites near. In this way, a simple and comprehensive model, to find a near optimal allocation, so that the reduced communication costs have developed.

✓ Mr. Chen and his colleagues in 2002, and the allocation method for the design of components in distributed database provided. It uses a clustering algorithm based on genetic search for partitioning data Database to retrieve high-performance, Have proposed a method in which formulate the problem of clustering data partitioning as a traveling salesman problem is presented.

3 - Research Methodology

3-1- Function mutation

```
for (int i = NP + NC; i < NP + NM + NC; i++)//☻ Operation MUTATION ☻
{
    int Fragment = 0, CHromosome = 0;
    Fragment = Rand.Next(0, NumberFragment);
    CHromosome = Rand.Next(0, NP);
    mutation(CHromosome, Fragment, i);
}
private void mutation(int CHromosome, int Fragment, int NumberHomeArray)
{
    for (int i = 0; i < NumberFragment; i++)
    {
        if (i == Fragment)
        {
            ArrCHromosome[NumberHomeArray, i] = "1";
        }
        else
        {
            ArrCHromosome[NumberHomeArray, i] = ArrCHromosome[CHromosome, i];
        }

    }
}
```

3-2- Fitness function

```csharp
private void OptimalAllocation(int SelFra, int NumNode),
{

    string s = string.Empty;
    string[] x = new string[10];
    s = string.Empty;
    s = ArrayHNode[SelFra];
    x = s.Split(',');
    int Cost = 0, Row = 0, Column = NumNode, MainCost = 100;
    if (x.Length > 1)
    {
        for (int i = 0; i < x.Length; i++)
        {
            Row = Convert.ToInt32(x[i]);
            Cost = 0;
            if (Row > Column)
            {
                Column = Convert.ToInt32(x[i]);
                Row = NumNode;
            }
            Cost = CostTransmission[Row, Column];
            if (MainCost > Cost)
            {
                MainCost = Cost;
            }
        }
        CostAll += MainCost;
    }
}
```

Difference between repetitive and non-repetitive in that the iterative method Because there are several versions of a piece of reliability and high availability So that if a site has a certain piece, fail, there are other sites that contain the fragment is But this will be no additional overhead Because it requires more memory and a lot of time spent in the practice update The unique method to reduce overhead and time update. As seen in the figure above, the proposed algorithm in the range of the number of sites and the number of pieces, the optimality is greater than the threshold.

4- Results

In this study, an allocation algorithm called genetic algorithms were tested components. This algorithm is based on a threshold algorithm, but different strategies for transferring data to other sites using the parts. In this study, we have unique algorithms in distributed database systems studied. Based on this model and transaction data, there are three parameters to find the optimal allocation of the overall communication cost which reduces our main problem. These parameters include shipping costs, reliability or availability of the update. In time transfer experiments, reliability or the ability to access and update of the characteristics associated components that were studied. Allocation algorithm components were implemented using C # .NET application. To measure the performance of the proposed algorithm is compared with a threshold algorithm.

References

1- 1 Hababeh I. O. 2005. A Method for Fragment Allocation Design in the Distributed Database Systems, The Sixth Annual U.A.E. University Research Conference.

2- Baseda R, Tasharofi S, Rahgozar M. 2006. Near Neighborhood Allocation: A Novel Dynamic Data Allocation Algorithm in DDB.

3- Huang Y. F. , Chen J. H. 2001. Fragment Allocation in Distributed Database Design

4- Journal of Information Science and Engineering 17, 491-506.

5- Baseda R, Tasharofi S, Rahgozar M. 2006. Near Neighborhood Allocation: A Novel Dynamic Data Allocation Algorithm in DDB.

6- Ulus T and Uysal M. 2003. Heuristic Approach to Dynamic Data Allocation in Distributed Database Systems, Pakistan Journal of Information and Technology 2 (3): 231-239, ISSN 1682-6027.

7- Huang Y. F. and Chen J. H. 2001. Fragment Allocation in Distributed Database Design , Journal of Information Science and Engineering 17, 491-506.

8- Ahmad I, Karlapalem K. 2002. Evolutionary Algorithms for Allocating Data in Distributed Database Systems, International Journal of Distributed and Parallel Databases, 11: 5-32, The Netherlands.

9- Hababeh I. 2005. A Method for Fragment Allocation Design in the Distributed Database Systems, The Sixth Annual U.A.E. University Research Conference.

10- Baseda R, Tasharofi S, Rahgozar M. 2006. Near Neighborhood Allocation: A Novel Dynamic Data Allocation Algorithm in DDB.

11- Ozsu M, Valduriez T. 2011. Principles of Distributed Database Systems, 3rd edn, pp. 245–293.

12- Tosun U. Cosar A. 2012. Parallel mutation operator for the quadratic assignment problem. In: Proceedings of WIVACE, Italian Workshop on Artificial Life and Evolutionary Computation, Parma.

13- Sevinc E, Cosar A. 2011. An evolutionary genetic algorithm for optimization of distributed database queries. Comput. J. 54(5), 717–725.

14- Jing L, Ng MK, Huang JZ. 2010. Knowledge-based vector space model for text clustering. J Knowl Inf Syst (Springer) 25(1):35–55.

15- Kranen P, Assent I, Baldauf C, Seidl T. 2010. The ClusTree: indexing micro-clusters for anytime stream mining. J Knowl Inf Syst (Springer).

16- Ni X, Quan X, Lu X, Wenyin L, Hua B. 2010. Short text clustering by finding core terms. J Knowl Inf Syst (Springer).

17- Sakuma J, Kobayashi S. 2009. Large-scale k-means clustering with user-centric privacy-preservation. J Knowl Inf Syst (Springer).