

Center of Information Technology and Scientific Computing

Department of Software Engineering

Database Administration (DBA)

*Assignment seven*

Data Allocation Model: literature

*Submitted by Henok Addis*

*Id: ATR/6271/09*

Submitted to Mr. Natnael A

Submitted Date: 07/10/2019

Abstract

Allocation of data is one of the key design issues of distributed database. A major cost of query execution in a distributed database system is the data transfer cost from one site to another site. The allocation of fragments among the different sites over the network plays an important role in performance of the distributed database system. The main objective of a data allocation in distributed database is to place the data fragments at different sites in such a way, so that the total data transfer cost can be minimized while executing a set of queries.

Keywords: distributed database systems, data allocation, database design, database maintenance

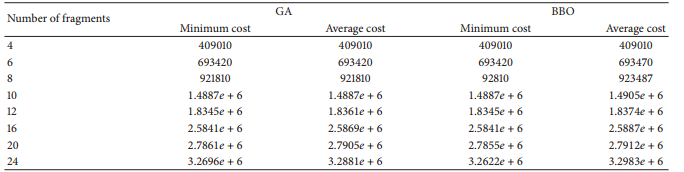
# 1. Introduction

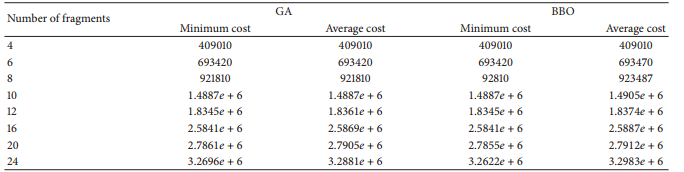
# The Data Allocation Models

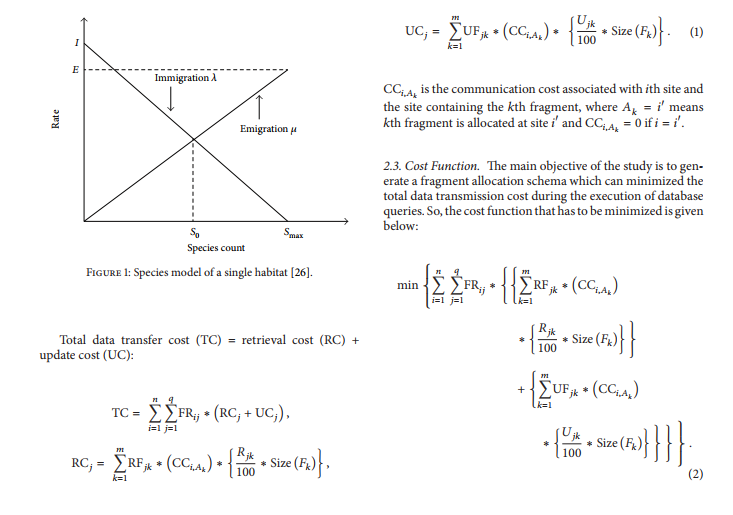
*1.* ***Fragment Allocation Problem.***Assume a distributed database system consisting of sites 𝑆 = {𝑆1, 𝑆2, . . . , 𝑆𝑛} on which a set of queries 𝑄 = {𝑞1, 𝑞2, . . . , 𝑞𝑞} is running. Each site has its own processing power, memory, and local database system and all the sites are connected by a communication link network. Let 𝐹 = {𝐹1, 𝐹2, . . . , 𝐹𝑚} be the set of fragments after partitioning all global relations during fragmentation phase of distributed database design. The allocation problem involves finding the optimal placement of the fragments (𝐹) to the sites (𝑆). The optimality can be defined with respect to two measures, minimal cost and performance.

*2.* ***The Cost Model.*** Table 1 gives the description of various notations used to draw the cost model of data allocation. There are primarily two types of costs associated with execution of a query. The first type is the cost of retrieval of fragments to process a query and the second type is the cost to update fragments to process that query. The formula to calculate total cost of data transfer is given as follows.

*3.* ***Cost Function.***The main objective of the study is to generate a fragment allocation schema which can minimized the total data transmission cost during the execution of database queries. So, the cost function that has to be minimized is given below:







# 5. Conclusion

This paper presents a new biogeography-based optimization technique for nonreplicated static allocation of data fragments during the design of distributed database. To evaluate the performance of proposed algorithm, results are compared with GA. From the results, it is clearly evident that the proposed technique for data fragment allocation is providing quality solutions in quick time.

The proposed algorithm significantly minimizes the data transfer cost during the execution of a set of queries. However, in some cases the average cost of allocation for BBO is more than GA, but for fast running time and quality solution, BBO can be introduced as a capable algorithm for fragment allocation during distributed database design.

# References

1. M. Ozsu and P. Valduriez, *Principles of Distributed Database Systems*, Prentice Hall, 2nd edition, 1999.