Solving Job Shop Scheduling Problem with Genetic Algorithm using Cloud Computing Approach

Saeed teimoori

Master Graduate of Shahid Bahonar University, Kerman

Abstract
In this research, we introduced a map-reduce framework from cloud computing to execute the genetic algorithms with large population sizes and overcome limited resource problem of common single computers. In comparison with related works such as MRPGA and Huang's method. Huang and Park, using the model presented in, an algorithm for job shop scheduling problem solving is presented in the context of cloud computing. In this algorithm, the encryption method used based on start times of jobs and another field as the generation of genetic algorithm is defined in display individuals of the population. The algorithm MRPGA in which the claim is made that the current map-reduce model due to a fall in local optimization does not support as well as a genetic algorithm. So this method has tried to do changes in map-reduce architecture, So that the possibility of falling into the local optimum reduced. By comparing the proposed method with the MRPGA method and the standard Map-Reduce method, we concluded that the proposed algorithm performs better for different tests in terms of finding the exact results and thus produces shorter schedules.

Keywords:
Job shop scheduling problem, MRPGA method, Huang and park method

1. Introduction

Some of decisions are known in the industry as planning decisions. Resource planning process for production determines set of activities necessary to scheduling. Scheduling is the allocation of limited resources to activities during the time, to optimize one or several the objective function [1].
Suppose there are number of different job that must be processed on the number of different machines. Everything is processed by every machine once and only once. Processing a job on a machine called an operation. Everything must be processed with particular succession on different machines. This constraint is called Technological constraint. Anything is independent of other jobs and arrangement perform of acts does not depend on order acts of other jobs. Finding scheduling all of the jobs with compliance with all provisions, also be implemented and has the fewest completion time, is called Job shop scheduling problem [2].

Manuscript received September 5, 2025
Manuscript revised September 20, 2025
https://doi.org/10.22937/IJCSNS.2025.25.9.22

To implement a genetic algorithm in the context of map-reduce framework, studies have been done. From the first jobs is done in this area by way of MRPGA [3]. In this way, the initial population was created by a coordinator node and this node has the responsibility for implementing the general structure of genetic algorithm rings. Chromosomal analysis at the stage of mapping and Parents' Choice is performed in two reduction stages. Except for the evaluation and selection of parent's steps, other operations including the production of initial population, combined, mutations and review of check out conditions happens in the single-node coordinator algorithm. In the Method based on Map-reduce framework the standard [4], to solve the problem parents choose the appropriate in method MRPGA, The output of the random distribution stage of mapping, reduction step is used internode and thus need to reduce the second stage has been eliminated in MRPGA. In this algorithm, as well as other operations are carried out in a coordinator node. Huang [5] used from this implementation Standard, to solve job shop scheduling problem by genetic algorithm in the context of mapping-reduction.

Although previous algorithms are able to implement genetic algorithm in the context map-reduce framework, but is has three main problems which makes it difficult to use them on the big problems. These three problems can be listed as follows:

1-Selection mechanism in this algorithm has not been implemented to form effectively. In other words, a large collection of good people population may be sent to a reducer and the result of acts selection operator, many of these good people have not the possibility of transmission to the next generation and therefore reduced the efficiency and speed of convergence of genetic algorithm.
2-In all these ways, the cutting and mutation operation that are costly operations, apply in a node and therefore if population size is selected large, reduces speed enforcement of Genetic algorithms rings.
3-In these algorithms, production the initial population is performed for series and by a node, thus the production of large initial populations was time consuming and reduces from degree of parallelism of the algorithm.

In this research, a method introduced based on map-reduce framework so that genetic algorithms with very large

populations has produced to solve job shop scheduling problem and overcome on difficulties of previous methods.

2-2-1. Describes of job shop scheduling problem

Job shop scheduling problem can be described as follows [1,5,6]:

Suppose there are number n of different job $\{J_1, J_2, \cdots, J_n\}$ that should on m different machines $\{M_1, M_2, \cdots, M_m\}$ be processed. Everything is processed by every machine once and only once. Processing a job on a machine called an operation. Job action i-th on machine j-th shows with O_{ij}. Everything must be processed with particular succession on different machines. This constraint is called to technical constraints. Anything is independent of the other jobs and arrange perform of acts not related to the other jobs actions. Every action O_{ij} has a special processing time is t_{ij}.

The purpose of on job shop scheduling problem can be stated as follows:

$$C = \min\{C_1, C_2, \cdots, C_n\}$$

Where C_i is represents the completion time$J_i, i = 1,2, \cdots, n$. The purpose of job shop scheduling problem can be expressed in terms of disjunctive graph:

If we could specified direction of edges of non-oriented at any subset and create path on all nodes in each subset, then we can identify arrange operations that are processed on each machine. In this case found an answer to job shop scheduling problem.

So the choice scheduling is equivalent to choose a graph like $\boldsymbol{G(N, A, E')}$ where $\boldsymbol{E' = E'_1 \cup E'_2 \cup \cdots \cup E'_m}$ and $\boldsymbol{E'_j}$ path connected is oriented that passes through all the vertices assigned to $\boldsymbol{E_j}$. The number of edges found in $\boldsymbol{E'_j}$ is equal to $\boldsymbol{d_j - 1}$. If the graph is selected, non-circular, in this case shows a possible scheduling.

Each node is corresponds O_{ij} with a weight t_{ij} that indicates processing time j-th operation from i-th job. Nodes 0 and 1 are zero weight. The longest path from node 0 to node 1 is called the critical path. So during the critical path represents the maximum completion time is scheduling. Find at least the critical path of all oriented-graphs $\boldsymbol{G(N, A, E')}$ is equal to find the optimal solution for job shop scheduling problem.

3. Genetic Algorithm

Genetic algorithm, kind of heuristic algorithm is that has been from theory "survival of the fittest" idea of Darwin's [7]. Genetic algorithms can be categorized as a kind of evolutionary algorithm that potential solutions to a particular problem in the form of data structures similar to chromosomes, encoding and recombination operations on these structures acts to protect of vital information. Genetic algorithm compared with traditional algorithms has two important advantages:

1-This algorithm has parallel structure and behavior. This algorithm population of solutions possible kept at any moment and tried to push this crowd towards better spaces for searching.

2-This algorithm has a random structure, So if it's getting caught in local optimum can using a random walk, to release themselves from this local optimum.

Therefore, from this algorithm in many areas those need to find optimal or near to be optimal, especially on issues that have large search space, has been abundant use. Genetic algorithm can be generally expressed as following [7,8,9]:

First the selected initial population and the quality of this initial population based on the evaluation function is determined. Then, in each iteration are selected parents by applying the combination operators (cutting), to produced children. These children with very little probability (close to zero) found of mutated and are added to the population. Then some of the people based on the fitness value, removed from the population to remains stable population size. Each step of the algorithm is called a generation.

3.1 Cloud computing and map-reduce model

The evolution of computing in such a way that it can be assumed to be an important element. In such a case users try based on their needs and no matter how it where it is a service and or how is delivered, giving it access. Examples a variety of computing systems presented that attempt to provide such services to users. Some of these computational systems including cluster computing, grid computing and cloud computing recently.

Grid computing concepts in this chapter is the prelude to understand cloud computing, and then consider on cloud computing and we examined how to use map-reduction models for solving optimization.

3-2. Grid computing

The first and most widely used definition for Grid computing definition used by Foster and Keselman [10] provided: "A grid computing, hardware and software structure that access to a reliable, consistent, pervasive and inexpensive to provide high levels of computing capabilities". The real problem related to grid consists of resource sharing and coordinated problem solving in virtual organizations that are composed of several sub-organizations. Shared that we're looking for, just is not simple file exchange a direct but access to computers, software, data and other resources will also be included [11].

The potential benefits of Grid computing in the IT management can be summarized as follows [12]:

- Grids gathering heterogeneous systems together to simulate a large computer and thus can provide computing power higher for a job. In other words, using Grid computing, existing resources are being used more appropriately.
- Grid computing, caused saving in companies IT sector, because it reduces Total Cost of Ownership (TCO).
- Grid computing due to parallel processing build capacity and access to additional resources and also load balancing, improve the efficiency in computing resources and data storage.
- Grid computing to improve efficiency and also increase reliability. Because the resources that in grid are ruined, quickly replaced by other sources in a grid (them duty as the balance is distributed among other sources). For this reason, grid users can continue to its activities regardless of probable damages.
- Grid computing allows companies to provide distributed resource management. It does this through virtualization of physical resources a heterogeneous in different geographical locations is being carried out. It provides possibility to manage resources.

4. Cloud computing

Foster and colleagues [10] defines cloud computing as follows:
"Cloud computing is model distributed to Large-scale that set of scalable and virtualized from managed computing power, storage space and clients puts at the disposal of customers via the Internet".the basic characteristics of cloud computing can be summarized as follows

- Cloud computing is a new computing model.
- In this model, resources (hardware, storage space and system apps) would be presented for service.
- The basic characteristic of cloud, virtualization and dynamic scalable is based on user demand.
- Cloud services via the web or via API offered to users.

4.1 Map-reduce:

Map - reduce a programming model that was created by Google Inc. in 2004 [13]. This model to simplify parallel data processing on large clusters has been designed [13]. Map – reduce model is a model of efficient and scalable and therefore can be used for processing sets of massive data. When mapping - reduction framework was introduced, Google indexing system your search was completely rewritten to be able to use of these new programming model. Indexing system, which produces the data structures used by Google web search engines. In general average hundreds of thousands of action map-reduce are implemented on Google's clusters, to enable them twenty-peta bytes of data to process daily [14].
Google first library map-reduce wrote in 2004 to language C ++.

Phoenix [15], the implementation of map-reduce for systems with shared memory, multi-core and multi-processor is designed. This means that a computer with multiple processor cores. Mars [16], is a map-reduce framework that are designed for GPUs. A graphics processor unlike central processors have Processors high power and as well as high memory bandwidth. But programming is hard for them because they are designed for graphics application.

From the beginning map-reduce framework, extensive research regarding the use of this framework has been made in solving various problems [3,17,18,19,20,21]. Many calculations can easily be done by programming model mapping-reduction.

4-2. The implementation of the Genetic algorithm in the context of cloud computing

Of the first things that were observed in this area, the algorithm MRPGA [3] in which the claim is made that the current map-reduce model due to a fall in local optimization does not support as well as a genetic algorithm. So this method has tried to do changes in map-reduce architecture, So that the possibility of falling into the local optimum reduced.

The overall cloud architecture that runs MRPGA algorithm is presented in Figure 1. It is logical that this algorithm has operations evaluation and selection of the parents, the operation was time consuming and should therefore be implemented by the framework of map-reduce. Evaluation is done by a mapping stage and also selection by two reducing stages. This system is consists of a coordinator node, a master node and worker multi-node. Master nodes are responsible for scheduling the execution of parallel activities [3].

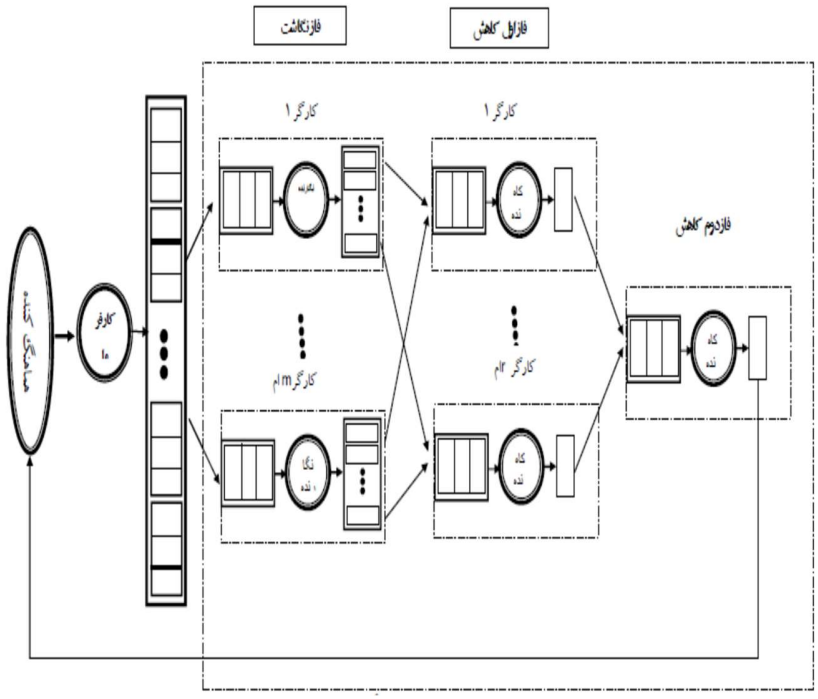

Figure 1: general architecture of cloud algorithm MRPGA [3]

.

4-2-2. Implementation based on the standard model of map-reduce

Each of an iteration of genetic algorithms in the form of a full stage map-reduce is implemented separate. In this method is performed evaluation of people in mapping stage and them selection in the reducing stage. Figure 2 overall structure of this algorithm can be displayed.

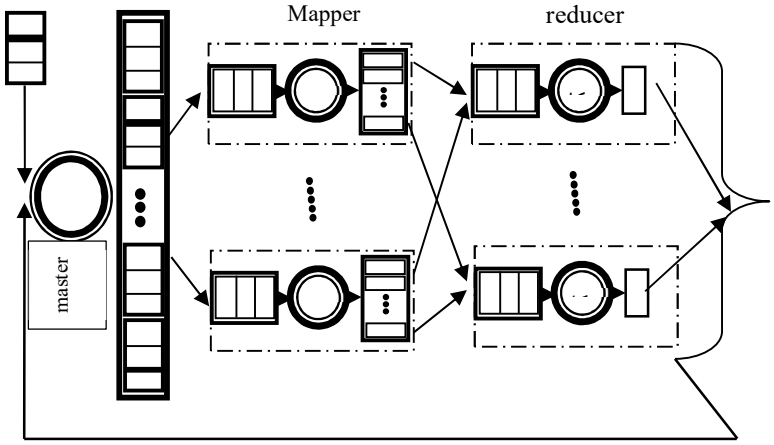

Figure 2: The overall structure of the proposed algorithm in [4]

This algorithm has tried to solve the problems associated with parental choice in MRPGA method, but many of disadvantages and shortcomings of this algorithm are inherited. One of these weaknesses as follows:

1-In this algorithm similar to MRPGA, from a single master node to apply the combination and mutation operators are used. So the load serial implementation of this operator, from degree of parallelism of the algorithm is reduced.

2-In this algorithm, it is assumed that the initial population is generated by your receiver service; as a result, the production of initial population, in addition to form serial is done, dependent to intensity computing power and receiver service speed. If the receiver service is a node with low computing power (eg a mobile handset), in this case the production of very large populations, component points will be very critical of this algorithm.

3-In this algorithm may also in distribution of people stage among reducers a large part of good people, to a corresponding reducer and therefore will be there possibility of losing good people of population.

4-3. job shop scheduling problem solving by genetic algorithm in map-reduce framework

Huang and Park [5], using the model presented in [4], an algorithm for job shop scheduling problem solving is presented in the context of cloud computing. In this algorithm, the encryption method used based on start times of jobs and another field as the generation of genetic algorithm is defined in display individuals of the population. This field helps to algorithm could after a certain number of generation to end his jobs (Unlike the method proposed in [4], a genetic algorithm until runs that to reach a certain level of fitness).

In the previous sections the structure of the three algorithms were presented in previous studies [3,4,5], was investigated. Compare this algorithms described in this section. Compare these methods are shown in Table 1.

Table 1: Comparison of investigated methods

	MRPGA[11]	Standard Model [13]	Huang method [14]
Goal	Implementation of genetic algorithms in the context of map-reduce	Implementation of genetic algorithms in the context of map-reduce	Implementation of job shop scheduling problem by genetic algorithm in the context of map-reduce
The number of mapping steps for each the genetic algorithm loop	1	1	1
the number of Reducing steps for each Genetic Algorithm loop	2	1	1
The task of master nodes	Distribution of individuals among mapping	Distribution of peoples among mapping, the distribution people among Reducer, the implementation of the general process of genetic algorithms, conducted combined operations, mutation and update, review conditions	Distribution of peoples among mapping, the distribution people among Reducer, the implementation of the general process of genetic algorithms, conducted combined operations, mutation and update, review conditions

		Completion of algorithm	Completion of algorithm
The task of mapping nodes	Assessment of individuals of the population	Assessment of individuals of the population	Assessment of individuals of the population
The task of Reducer nodes	Parental choice	Parental choice	Parental choice
Production of the initial population	To form Random and by the coordinator node	By receiver service	By receiver service
condition Completion of algorithm	On the basis of individual fitness values	On the basis of individual fitness values	Based on individual fitness and the number of generations
Time of combined and mutation operation	Much	much	Much
Production time initial population	Much	It depends on receiver service	It depends on receiver service
The possibility of losing good people on selection stage	Much	Exist	Exist

5- Proposed research algorithm

Although previous algorithms [3,4,5], are able to implement a genetic algorithm in the context of map-reduce framework, But has three main problem that the use of them in big problems is difficult. These three problems can be listed as follows:

Selection mechanism in this algorithm has not been implemented to form effectively. In other words, a large collection of good people population may be sent to a reducer and the result of what selection operator, many of these good people not have the possibility of transmission to the next generation and thus decreases efficiency and speed of convergence of genetic algorithm.

In all these methods, cutting and mutants operations that are costly operations, applied in a node (coordinator node in method of MRPGA and master node in two other methods) and therefore if the size populations is select to large, genetic algorithms reduces speed enforcement circles.

In this algorithm, production the initial population is done for series by a node (Coordinator Node in MRPGA and receiver service in two other methods), thus producing a large initial populations was time consuming and reduces from degree of parallelism of the algorithm.

In this research to solve these problems has been introduced a new method. In this method, used in addition to different operators for genetic algorithms, map-reduce structure also changed in such a way to be able combine and mutate operation performed also in worker nodes and thus lose a large part of the time these nodes and improved runtime. In addition, used from a specific reducer node in reducing stage to process always the best population solutions and thus the possibility of the loss of good solutions to be resolved at the selection stage. In the proposed algorithm, from an initial mapping -reducing stage and specific is used to generate the initial population and therefore the proposed method, do not have production problems the previous methods of initial population.

The proposed algorithm is composed of a master node and the number worker nodes. In this algorithm map-reduce a separate step is used to generate the initial population. At this stage is responsible mapper nodes duty the production of raw chromosomes and reducer nodes duty the production of structures 5-1 for each of the raw

chromosomes. As stated, for this master nodes are not involved the distribution of mappers outputs to reducers, each reducer node your inputs received from outputs of local mappers. After generating the initial population is beginning genetic algorithms rings. Each of these rings, by a map-reduce stage is done complete. For this purpose, the master node is distributed the individuals of the population among mappers nodes. These people are values of evaluation -1. Mappers nodes in is responsible the task of evaluating these people and they identified vale of length scheduling based on decoding of section chromosome. These nodes in addition to the evaluation of people, identified the best people and them with unique identifier NULL, to their takes output. In addition, other people (except the best people), is taken with random ID to the output. Then master node is responsible task assignment these output between reducer nodes. This allocation is done in a random so that the same number of people assigned to each of reducers. In addition, all those who are NULL ID to a specific node (node 0) are assigned. Reducer nodes are responsible the task of applying the selection operators, combination and mutations and to number your inputs produces offspring. This children has the values random ID, evaluation -1 and generation of a larger unit than their parents. This children as the new generation is divided by master node between mapper. Figure 3, the overall structure of the proposed algorithm shows schematically .

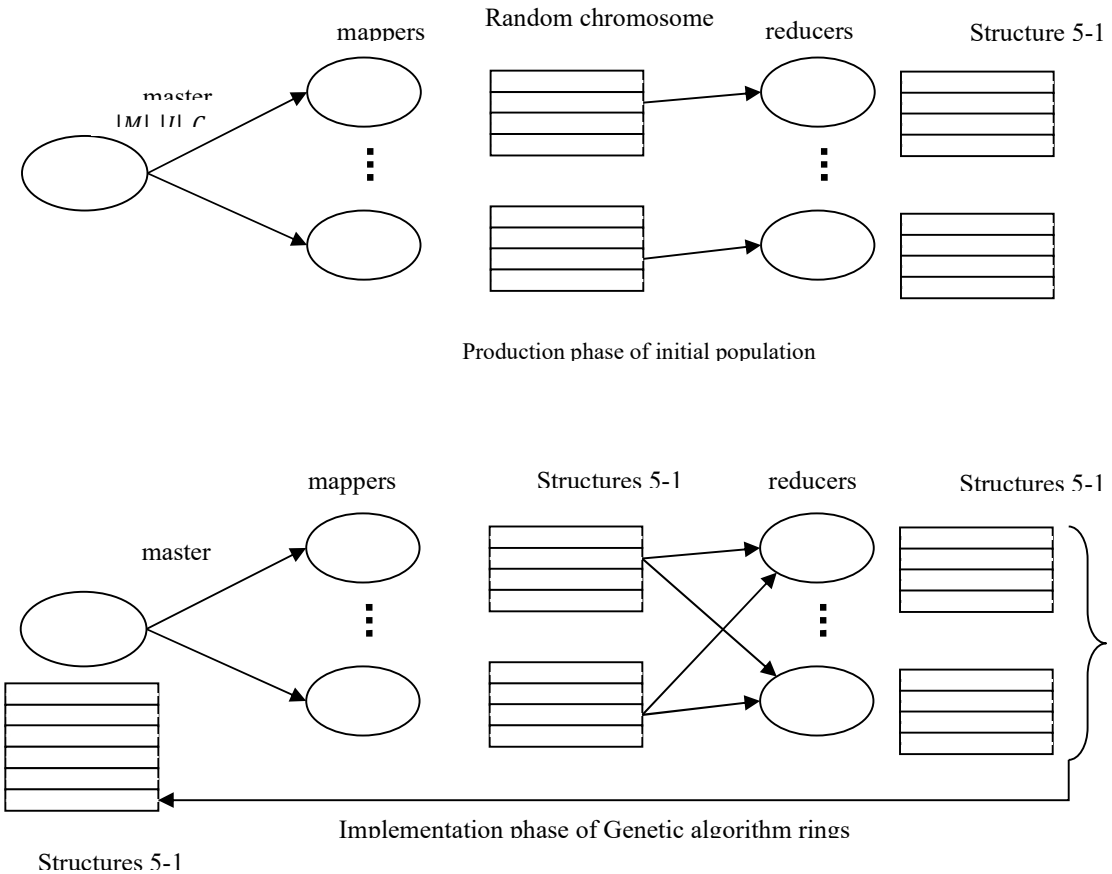

Figure 3: The overall structure of the proposed algorithm

7. Conclusions

.

Older algorithms due to deterministic nature and their serial, not able to manage space nonlinear of job scheduling problem and if stuck in a local optimum, no way to get rid of it. To solve this problem can use evolutionary algorithms and in particular genetic algorithm that random and parallel nature. For the genetic algorithm can give us a proper solution, should have chance produce large sections of solution space in the population, so for a very difficult problem like job scheduling, this population needs to be as large as possible. But large population is means bigger time costs because conventional computers because the structures of its resources, Genetic algorithms are implemented in serial form and therefore cannot control their large populations. To solve this problem also needs to be larger hardware that have the ability to manage big data, can be used. Cloud computing is a new model that tries to solve the problem of limited resources of computer (computational and storage) and puts almost unlimited resources at our disposal. Map-reduce introduced framework from cloud computing by Google Inc. and provides ability to write algorithms for execution in cloud computing environments. This framework hides the all the intricacies associated with cloud computing infrastructure and enables to programmer which can use to write some simple function of the possibilities of cloud computing. In this book, we have introduced a method based on mapping-reduction framework so has produced genetic algorithms to solve job scheduling problem with very large populations and overcome to the problem limited resources of conventional computers. By comparing the proposed method with previous methods were introduced for this purpose (MRPGA methods and Huang methods), reached the conclusion that used the proposed algorithm for four benchmarks, perform better and therefore produces shorter schedules.

8. Suggestions for future studies

Considering that the proposed algorithm in early generations did not do better in comparison with other methods, in this case, can used other innovative methods or providing operators new genetic and or applied to restructuring of map-reduce these improvements. Future studies can be classified in three categories:
1-Study about innovative and meta-heuristic methods (other than genetic algorithms) to solve job scheduling problem: can be used the newer Meta-heuristic algorithms instead of genetic algorithms such as algorithms based on the immune system, genetic algorithms that uses from local searches on their structure and so which may have higher performance, can be used to solve job scheduling problem.
2-Study about Operators of genetic algorithm: If we assume on use of genetic algorithm, can be this algorithm about the structure different operators, and in particular the new coding techniques for job scheduling problem that has a dramatic effect on the performance of the algorithm, are conducted further studies.
3-Study about the structure of map-reduce framework: It can be about other methods to arrangement map and reduce nodes and also master nodes and even more studies are done the structure of mapping -reducing functions.

References

[1] K. R. Baker, D. Trietsch, Principles of Sequencing and cheduling,Wiley,2009
[2] P. Deepu ,Scheduling Reentrant Flexible Job Shops With Sequence Dependent Setup Times, Montana State University, vol. 165, pp. 1–19, 2005.
[3] C.Jin, C. Vecchiola, R. Buyya, Mrpga: An extension of mapreduce for parallelizing genetic algorithms, . IEEE Fourth International Conference on eScience, Indianapolis, IN, pp. 214-221,2008.
[4] A. Verma, X. Llor,D. E. Goldberg, Scaling Genetic Algorithms Using MapReduce, Proceedings of the Ninth International Conference on Intelligent Systems Design and Applications, Washington, DC, USA : IEEE Computer Society, pp. 13-18,2009.

[5] D. W. Huang, J. Lin**,** Scaling Populations of a Genetic Algorithm for Job Shop Scheduling Problems Using MapReduce*,* IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom), Indianapolis, IN, pp. 780 – 785, 2010.

[6] P. Larranaga, Genetic Algorithms for the Travelling Salesman Problem: A Review of Representations and Operators*.* Artificial Intelligence Review, Vol. 13, pp. 129-170, 1999.

[7] A. E. Eiben, J. E. Smith, Introduction to Evolutionary , Springer, vol. 165, pp. 1–19,2007.

[8] S. N. Sivanandam, S. N. Deepa, Introduction to Genetic Algorithms*,* New York ,Springer, 2008.

[9] P. Larranaga, C. M. H. Kuijpers, R. H. Murga, I. Inza, S. Dizdarevic, Genetic Algorithms for the Travelling Salesman Problem: A Review of Representations and Operators. Artificial Intelligence Review, Vol. 13, pp. 129–170, 1999.

[10] I. Foster, Y. Zhao, I. Raicu, S. Lu, Cloud Computing and Grid Computing 360-Degree Compared*,* Grid Computing Environments Workshop, Austin ,TX , pp. 1-10, 2008

[11] K. Stanoevska, T. Wozniak, S. Ristol, Grid and Cloud Computing: A Business Perspective on Technology and Applications*,* Berlin, Springer-Verlag, 2010.

[12] M.Armbrust, Above the Clouds – A Berkeley View of Cloud, California , Technical Report UCB/EECS-2009-28.

[13] T. Yamada, R. A. Nakano, Genetic Algorithm Applicable to Large-Scale JobShop Problems, Parallel Problem Solving from Nature, Vol. 2, pp. 281-290, 1992.

[14] J. Dean, S. Ghemawat, MapReduce: simplified data processing on large clusters, Communications of the ACM - 50th anniversary issue: 1958 - 2008, Vol. 51, pp. 107-113, 2008.

[15] C. Ranger, R. Raghuraman, A. Penmetsa, G. Bradski, C. Kozyrakis, Evaluating mapreduce for multi-core and multi processor systems, High Performance Computer Architecture, Scottsdale, AZ : IEEE, pp. 13-24, 2007.

[16] B. He, W. Fang, Q. Luo, N. K. Govindaraju, T. Wang, Mars: a mapreduce framework on graphics processors. Proceedings of the 17th international conference on Parallel architectures and compilation techniques, New York, NY, USA : ACM, pp. 260–269, 2008.

[17] J. Feldman, Using many machines to handle an enorenormous error-correcting code, Information Theory Workshop, Punta del Este : IEEE, pp. 180-182, 2006.

[18] J. Ekanayake, S. Pallickara, G. Fox, Mapreduce for data intensive scientific analyses, IEEE Fourth International Conference on eScience, Indianapolis, IN : IEEE, pp. 277–284, 2008.

[19] A. McNabb, C. Monson, K. Seppi, Parallel pso using mapreduce, IEEE Congress on Evolutionary Computation, Singapore : IEEE, pp. 7–14, 2007.

[20] O. J. Haggarty, W. Knottenbelt, J. T. Bradl, Distributed response time analysis of gspn models with mapreduce, International Symposium on Performance Evaluation of Computer and Telecommunication Systems, pp. 82–90, 2008.

[21] T. Xia, Large-scale sms messages mining based on map-reduce, International Symposium on Computational Intelligence and Design. Vol. 1, pp. 7-12, 2008.