Asynchronous Parallel Distributed Genetic Algorithm with Elite Migration

Kazunori Kojima, Masaaki Ishigame, Goutam Chakraborty, Hiroshi Hatsuo, and Shozo Makino

Abstract—In most of the popular implementation of Parallel GAs the whole population is divided into a set of subpopulations, each subpopulation executes GA independently and some individuals are migrated at fixed intervals on a ring topology. In these studies, the migrations usually occur ‘synchronously’ among subpopulations. Therefore, CPUs are not used efficiently and the communication do not occur efficiently either. A few studies tried asynchronous migration but it is hard to implement and setting proper parameter values is difficult.

The aim of our research is to develop a migration method which is easy to implement, which is easy to set parameter values, and which reduces communication traffic. In this paper, we propose a traffic reduction method for the Asynchronous Parallel Distributed GA by migration of elites only. This is a Server-Client model. Every client executes GA on a subpopulation and sends an elite information to the server. The server manages the elite information of each client and the migrations occur according to the evolution of sub-population in a client. This facilitates the reduction in communication traffic.

To evaluate our proposed model, we apply it to many function optimization problems. We confirm that our proposed method performs as well as current methods, the communication traffic is less, and setting of the parameters are much easier.

Keywords—Parallel Distributed Genetic Algorithm (PDGA), asynchronous PDGA, Server-Client configuration, Elite Migration

I. INTRODUCTION

GENETIC Algorithm (GA) [1] is a multi-point search technique imitating the survival of the fittest rule of nature. The approximate solutions are obtained by repeating the selection, the crossover and the mutation operations called genetic operations, on a set of approximate solutions. An individual in the population has the chromosome information which is an encoded solution of the problem.

Since J. H. Holland et al. proposed GA in the early 1970s, GA had been successfully applied to various function optimization problems, combinational optimization problems, control problems, machine learning and so on [2]. This is because GA is easy to implement, it is robust, and it could find global optimum solution.

In recent years, it has been required to get the better approximate solutions faster and more efficiently for larger and more complicated problems. However, it is hard to obtain the approximate solution for these problems by using GA running on a single machine, because it is computationally heavy.

There are two basic approaches to make GA faster. In approach one, the algorithm itself is modified from its generic model to a more specific algorithm to solve the problem at hand, or hybridize GA with artificial neural network, simulated annealing or some other heuristics.

The second approach is to speed up the computation by parallel distributed implementation (PDGA). Many researches in the PDGA are coarse-grained parallel GA which is implemented as follows. The population is divided into some subpopulations and each subpopulation is assigned to a different processor. GA is executed in parallel on each processor. To prevent premature convergence of the subpopulation, some individuals from each processor are exchanged (migrated) among different subpopulations.

There are two main methods to migrate individuals. One is the synchronous method where migration occurs at fixed generation intervals. In this method, a fast CPU needs to wait for the response from a slow CPU for synchronization of migration. Therefore, it is not efficient for CPU utilization especially in a heterogeneous environment. The other is the asynchronous method where individuals are migrated with fixed probability. But it is hard to implement the asynchronous event in general. Furthermore, the migrations are not driven by the searching need in both the methods.

To solve these problems, Munetomo et al. [10] proposed a migration method that controls the migration timings by the difference of the standard deviation in a particular subpopulation. However, it is very hard to set the parameter to control the migration timing.

In this paper, we propose a migration method, we named Elite Migration. This is an Asynchronous Parallel Distributed Genetic Algorithm implementation on a Server-Client topology. In this method, a client station executes GA independently on a subpopulation and send an elite chromosome information to the server when an elite is updated. The server manages elite information received from each subpopulation. When an elite of a client is not updated for several generations, the server sends some manipulated chromosome information to the client and the client receives the chromosome information as migrants. Therefore, migration is less and controlled by the server.

To evaluate the proposed method, we applied our method to find the global optimum for some uni-modal and multimodal functions, and confirmed that the results are as good as obtained by other methods with much less message passing.

II. PARALLEL DISTRIBUTED GA

Researches on PDGA can be classified roughly into 4 approaches [3].
migrations are not motivated by necessity.

To solve the above problems, Munetomo et al. proposed Sigma-Exchange model [10]. In this model, the standard deviation of fitness in each subpopulation is observed. When fitness of members in a subpopulation converges, migration occurs. However, it is hard to set the parameter that controls migration because we need to set the parameter according to the characteristics of the application and/or the population size.

Recently Erick et al. [3] reported various researches and experiments with PDGA. However, they didn’t discuss about the communication traffic.

Adachi et al. proposed Parameter-free GA [16], [17] and Hiroyasu et al. proposed Dual Individual Distributed Genetic Algorithm [18]. They implemented their algorithm with Master-Slave configuration. However, the migrants pass between the master and the slave synchronously, creating heavy communication load. They did not discuss about the communication traffic either.

Thus two main problems of PDGA are reducing unnecessary migration which involves communication cost, and setting different parameters for maximum efficiency. To overcome these problems, we propose Asynchronous elite migration PDGA as explained in section III.

III. ASYNCHRONOUS ELITE MIGRATION PDGA

In PDGA using subpopulation based methods, the island model is usually used. In the island model, the population is divided into some subpopulations and each subpopulation executes GA independently. To prevent convergence of each subpopulation, individuals need to be exchanged between subpopulations at fixed intervals or with fixed probability.

In most of the island models, ring topology is used as shown in Figure 1. Individuals are sent to next island and are received from the previous island synchronously at fixed intervals. Hereafter, we call this model Synchronous model. However, in general, it is hard to process such asynchronous events. Furthermore, from the viewpoint of genetic search, it is not efficient if migration occurs either randomly or at fixed intervals, because

![Subpopulation Diagram](image)

**Fig. 1.** Island model: migrants are sent synchronously in fixed intervals on a ring topology

1) Global Parallelization [4], [5], where the whole population is assigned to a single processor and the calculations of fitness or the genetic operations are assigned to different processors.

2) Coarse-grained Parallelization [6], [7], [8], [9], [10], where the population is divided into some subpopulations and the individuals are exchanged between these subpopulations.

3) Fine-grained Parallelization [11], [12], [13], where individuals are assigned on a grid structure and an individual can crossover with its neighbouring individuals.

4) Works that combine the above [14], [15].

Among these researches, the subpopulation based methods [6], [7], [8], [9], [10] are the most natural parallelization method and can be implemented easily.

In PDGA using subpopulation based methods, the island model is usually used. In the island model, the population is divided into some subpopulations and each subpopulation executes GA independently. To prevent convergence of each subpopulation, individuals need to be exchanged between subpopulations at fixed intervals or with fixed probability.

In most of the island models, ring topology is used as shown in Figure 1. Individuals are sent to next island and are received from the previous island synchronously at fixed intervals. Hereafter, we call this model Synchronous model. In the Synchronous model, a fast processor has to wait for a slow processor for the migration to be synchronized. Therefore, in an heterogeneous environment, where there are some slow processors, the available computation power can not be used efficiently.

On the other hand, there is a model where individuals are exchanged with fixed probability. In this model, each subpopulation requests migration asynchronously. Hereafter, we call this model Random-Exchange model. However, in general, it is hard to process such asynchronous events. Furthermore, from the viewpoint of genetic search, it is not efficient if migration occurs either randomly or at fixed intervals, because
A Subpopulation Client does not send an elite chromosome information at every elite updating, but once in every $\nu$ times of updates. By increasing $\nu$, our system is able to reduce the communication frequency from a Subpopulation Client to Elite Server. We need to optimize this interval $\nu$.

2) Longevity Parameter: The Longevity parameter ($\lambda$) is to control the timing of the request for chromosome import from Elite Server. To simplify the implementation, $\lambda$ is decremented by 1 when, in a generation, an elite is not updated. And the migration occurs when $\lambda$ is decremented to 0. When an elite is updated, $\lambda$ is reset to the initial value. Thus, when GA searches well within a Subpopulation Client, it does not need to send migration request to Elite Server and the communication frequency is low.

3) The Number of Migrants: As described before, the migration occurs when Longevity Parameter in a Subpopulation Client is decremented to 0. The Subpopulation Client requests to Elite Server for migrants and Elite Server sends $\mu$ chromosome information to the Subpopulation Client. Individuals selected randomly in the Subpopulation Client are replaced by newly received individuals. The ensuing traffic could be kept low when the number of migrants ($\mu$) is small.

B. Implementation

We implemented the simulation using C and MPI, and executed on SGI Origin2000. Origin2000 is a parallel processing machine with 32 CPUs. MPI (Message Passing Interface) [19] is the standard library for message communication. MPI is used on parallel computers but there are similar packages for UNIX or WindowsNT, so that one can use MPI on WSs or PCs on a TCP/IP network.

When we implement Parallel GA on WSs or PCs on a TCP/IP network, we can also use PVM (Parallel Virtual Machine), Socket library, Java language and so on. It is hard to implement Random-Exchange and Sigma-Exchange using the above techniques. However, it is easy to implement our algorithm using them.

Figure 4 shows Problem Analysis Diagram (PAD) of Elite Server program that we used, and Figure 5 shows the PAD of the Subpopulation Client program.

Elite Server waits for messages from Subpopulation Clients until they exist. When a message arrives, Elite Server processes it according to its content and then waits for further messages. In our programs, messages are identified using TAG of MPI. Messages can be identified by including TAG in the message on Socket programming or Java language.

The Subpopulation Client executes GA and sends an elite information once every $\nu$ times the elite in that island is updated. When an elite is not updated, Longevity parameter ($\lambda$) is decremented. If Longevity parameter is 0, the Subpopulation Client sends a request to Elite Server for migration and receives migrants from Elite Server.

A message is composed of a TAG and the message itself. The TAG is used to identify the message. The main body of the message mainly is the chromosome itself. Thus, a message that a Subpopulation Client sends to the Elite Server is the chromosome information of the elite in that Subpopulation Client. A message that a Subpopulation Client receives from the Elite Server is some chromosome information that the Elite Server manages.

Figure 4 and Figure 5 omit the details. But it is evident from these figures that the model is easy to implement and that the communication cost is low.
IV. EXPERIMENTS AND RESULTS

In this section, we describe the experiments and the results. For comparison, we implemented Synchronous model, Random-Exchange, Sigma-Exchange and APDGA-EM on Origin2000.

A. Experimental conditions

We experimented the above mentioned algorithms on 8 function optimization problems. Table I shows the list of functions Func1 to Func8, we experimented with.

Func1 and Func2 are a $n$-dimensional uni-modal function and a $n$-dimensional multi-modal function respectively. We created these two functions. Each variable is encoded to 10-bit gray code in the chromosome. Func3 and Func4 are De Jong’s Function F1 and F5. Func5, Func6, Func7 and Func8 are the functions that were used in the first International Contest on Evolutionary Optimization (ICEO) in 1996 and the second ICEO in 1997. From Func3 to Func8, the precision of a solution is set to 6 decimal places and each variable is encoded to 22–31 bits gray code. The optimization criterion is to get maximum value for all the functions.

The parameters for GA are as follows. Total population size is 1024. Selection is roulette selection. Crossover is single point crossover for each variable. Crossover rate is 0.6. Mutation is applied at bit level with 0.03 probability. An elite is preserved at every generation and it is terminated at the 500th generation.

The parameters for PDGA are follows. Migration rate is 0.2. Migration interval is 100 generations in Synchronous model. Migration probability is 0.01 in Random-Exchange. From preliminary experiments of Sigma-Exchange a suitable value for the parameter to control migration timing is selected. And ring topology is used.

In the proposed APDGA-EM, interval $\nu$ to send an elite is set at 10, initial Longevity $\lambda$ is 100 and the number of migrants is 2.

B. Comparison of fitness and traffic

In this set of experiments, the number of islands is 16 and each algorithm was executed for 25 times.

Table II shows the mean of best fitness till termination. In this table, Sync, Rand-ex and Sig-ex indicate Synchronous model, Random-Exchange model and Sigma-Exchange model respectively.

From this table, we can confirm that the fitnesses obtained by APDGA-EM are as good as obtained by other methods.
Therefore, in spite of lower communication cost, which we will show later, APDGA-EM could deliver at least equally good results.

Table III shows the mean of the traffic that one Subpopulation Client has received in the simulations. Table IV shows the mean of the traffic that one Subpopulation Client has sent to Elite Server in the simulations.

| TABLE II |
|------------|-------------------|-------------------|------------------|
|            | Sync              | Rand-ex           | Sig-ex           |
| De Jong F1 | -1.9696E-08       | -2.2352E-08       | -2.5914E-08      |
| Sphere     | -1.1312E-05       | -9.8701E-06       | -1.6630E-05      |
| Griewank   | -3.1375E+02       | -3.0876E+02       | -3.1379E+02      |
| Michalewicz| 4.6875E+00        | 4.6875E+00        | 4.6875E+00       |

| TABLE III |
|------------|-------------------|-------------------|------------------|
|            | Sync              | Rand-ex           | Sig-ex           |
| Uni-modal  | 1200.00           | 2317.60           | 410.24           |
| Multi-modal| 1200.00           | 2279.04           | 324.16           |
| De Jong F1 | 720.00            | 1357.60           | 208.80           |
| De Jong F5 | 480.00            | 837.12            | 0.00             |
| Langerman  | 1200.00           | 2269.92           | 139.52           |
| Sphere     | 1200.00           | 2441.76           | 362.40           |
| Griewank   | 1200.00           | 2201.76           | 348.96           |
| Michalewicz| 1200.00           | 2480.96           | 54.24            |

C. Parameters of APDGA-EM

In APDGA-EM, there are the following three parameters.
- Initial value of Longevity Parameter ($\lambda$)
- The number of elite for migrants ($\mu$)
- Interval to send an elite information to Elite Server ($\nu$)

Among these parameters, initial value of Longevity Parameter ($\lambda$) and the number of elite for migrants ($\mu$) affect the communication cost for received data, and interval ($\nu$) to send an elite information affects the communication cost for sent data.

Figure 6 shows the dependence of final fitness on initial Longevity Parameter. Figure 7 shows the dependence of final fitness on the number of migrants. Figure 8 shows the variation of final fitness on the interval to send an elite information.

From Figure 6 and Figure 7, we can confirm a tendency of negative and positive gradient respectively. For example, when the initial value of Longevity Parameter is small or when the number of migrants is large, the fitness is better. In summary, this means that the fitness is better when more messages are passed between Elite Server and Subpopulation Client, and it is worse when the traffic is less. However, the fitness fluctuates within a narrow range. Therefore, we can also confirm that these parameters do not affect the fitness strongly.

From Figure 8, it is seen that the resulting fitness does not really depend on the interval at which elite is passed from Subpopulation Client to Elite Server, and the fitness varies within a very narrow range. Similar behaviour is confirmed for all other functions too. This is an important difference from the parameter of Sigma-Exchange that has to be changed according to population size, GA operators and the application of received data is the same as the one of sent data. In Synchronous model because the number of message passing is same. The traffic in Random-Exchange is almost twice as that of Synchronous model, and here too the traffic of received data is almost same as the sent data. In Sigma-Exchange, both the received data traffic and sent data traffic is less than the ones in Synchronous model and Random-Exchange. Convergence of population in De John’s Function F5 (Func4) is difficult, because the difference between higher fitness value and lower fitness value is large. Therefore, migration takes place only occasionally and the traffic is less. Also in Michalewicz Function (Func8), migration happens seldom and the traffic is less. In these two functions, the traffic of the received data is almost same as the sent data.
problem. We thus conclude that the proposed algorithm is more robust to the selection of the parameters for various applications.

From all the experiments, we recommend the general optimum values for the parameters as:

- Initial Longevity parameter ($\lambda$) to be 100
- The number of elite information to send for migrants ($\mu$) to be 2
- Interval to send an elite information to Elite Server ($\nu$) to be 10

The above setting could deliver near optimum values for all the experiments we performed.

V. CONCLUSION

In this paper, we proposed Elite Migration for Asynchronous Parallel Distributed Genetic Algorithm which uses Server-Client topology. Our proposed method mainly sets some rules to communicate between Elite Server and Subpopulation Client which run genetic operations on a small population of chromosomes. A Subpopulation Client sends an elite chromosome information to Elite Server when an elite is updated. To control migration timing, Longevity Parameter ($\lambda$) is introduced on each Subpopulation Client. And some chromosome information that the Elite Server manages are used for migrants. Thus the proposed algorithm reduces the message passing traffic in PDGA.

We applied our proposed method and other PDGA methods to various function optimization problems and confirmed the effectiveness of our proposal. The traffic in proposed method is less than other current PDGA methods, and the results by our proposed method is better or at least equally good as those obtained by other methods.

Three parameters of the proposed method could affect the fitness values and the generated traffic. However, we have shown that the effect of the parameter values are only marginal and do not need to be tuned for different problems. Thus the proposed algorithm is robust to variations in parameter values, and the same set of values work well for every problem.

Therefore, the proposed algorithm is robust as well as its communication cost is lower compared to its competitive algorithms.

ACKNOWLEDGMENTS

The authors would like to thank Iwate Prefectural University Media Center Administrative staffs, who made possible the access to the computation environment needed to pursue this work.

REFERENCES

Kazunori Kojima received the B.Eng. and M.Eng. degree in mining college from Akita University in 1993 and 1995, respectively. He is currently a research associate of Faculty of Software & Information Science of Iwate Prefectural University. He has been engaged in research on genetic algorithms.

Masaaki Ishigame received the Ph.D. degree in Tohoku University in 1974. He was a research associate of Tohoku University, an employee of Matsushita Denso, and an associate professor of Akita University. He is currently Dean of the graduate school of Iwate Prefectural University. He has been engaged in research on signal processing, image processing, and knowledge engineering.

Goutam Chakraborty received his Ph.D. in 1993 from Tohoku University, Japan. Presently he is Professor and head of the Intelligent Informatics lab., Department of the Software and Information Science, Iwate Prefectural University, Japan. His main research interests are Soft Computing algorithms and their applications to solve pattern recognition, prediction, scheduling and optimization problems including applications in wired and wireless Networking problems.

Hirosi Matsuo received the Ph.D. degree in Tohoku University in 1990. He was an employee of Canon, and a research associate and an assistant professor of Akita University. He is currently an associate professor of Tohoku Seikatsu Bunka College.

Shozo Makino received the Ph.D. degree in Tohoku University in 1974. He was a research associate, and an associate professor of Tohoku University. He is currently a professor of Tohoku University. He has been engaged in research in voice recognition and understanding, voice database, voice signal processing, voice CALL system, image processing, character recognition.