computing@tanet.edu.te.ua
www.computingonline.net

ISSN 1727-6209
International Journal
of Computing

ISLAND MIGRATION MODEL WITH PARALLEL MUTATION STRATEGIES FOR COMPUTING THE TRAVELING SALESMAN PROBLEM ON MULTICOMPUTER PLATFORM

Plamenka I. Borovska [1], **Subhi A. Bahudaila** [2], **Milena K. Lazarova** [3]

[1] Technical University of Sofia, 8 boul. "Kliment Ohridsky", 1756 Sofia, Bulgaria, pborovska@tu-sofia.bg,
http://www.tu-sofia.bg/Bul/faculties/fksu/ks/leader_eng.htm
[2] Technical University of Sofia, 8 boul. "Kliment Ohridsky", 1756 Sofia, Bulgaria, subhudail@yahoo.com
[3] Technical University of Sofia, 8 boul. "Kliment Ohridsky", 1756 Sofia, Bulgaria, milaz@tu-sofia.bg

Abstract: *This paper investigates the efficiency of a model of parallel genetic computation of the traveling salesman problem with circular periodic chromosomes migration. The parallel model is verified by MPI-based program implementation on a multicomputer platform. The correlation of the application and architectural spaces has been investigated by exploring the impact of the scalability of the application and the parallel machine size over the efficiency of the parallel system. Performance profiling, evaluation and analysis have been made for different numbers of cities and different sizes of the multicomputer platform. The paper also investigates the impact of the mutation strategy on the solution quality of coarse-grained parallel genetic algorithm with circular periodic migration for the traveling salesman problem. We propose an approach to improve the quality of solution by applying parallel variable mutation rates for the local evolutions in the concurrent processes. A series of experiments has been carried out with parallel fixed and variable mutation rates in order to estimate the efficiency of the suggested approach. The best quality solutions have been obtained for the strategy with parallel fixed mutation rates.*

Keywords: *Euclidean TSP, parallel genetic algorithm, island model, chromosomes migration, mutation rate, multicomputer, parallel programming with MPI, parallel profiling, parallel performance.*

1. THE PROBLEM AREA

The traveling salesman problem (TSP) is a well-known NP-hard combinatorial optimization problem [1]. The goal is to find the minimum cost for the entire tour. A tour is a path that starts from a city, visits each city exactly once, and goes back to the starting city. Local optimization algorithms give an improved solution iteratively by making local changes. Genetic algorithms [2] are local optimization algorithms. They require an enormous amount of computations because of the genetic operations of selection, crossover, mutation and evaluation of the fitness of each chromosome (presenting a tour) in the populations. The computational time is improved when utilizing parallel systems. There are several parallel genetic approaches that are used to reduce the large amount of computation time associated with the serial genetic algorithms. Parallel genetic algorithms (PGA) with migration can speedup the slowly evolving subpopulations by introducing chromosomes that are better than the locally best ones. One method of building a PGA is to apply the master/workers paradigm [4]. Island models imply migration approach [5]. A hybrid approach which combines the segmentation and migration approaches is proposed in [7].

The genetic mutation is a random change that occurs in the characteristics of a gene. Mutation tends to bring about major unpredictable changes in the fitness of an individual. Increasing the number of mutations results in increasing the algorithm's freedom to search outside the current region of variable space. Mutation rate (μ) is a limitation ratio for the number of mutation events on each generation. Different mutation strategies have been applied to facilitate the convergence of PGA. As shown in [6] self-adaptation of mutation fails to increase the performance of genetic computations over best fixed mutation GA in both dynamic and static environment. Fuzzy adaptive search method for estimation of genetic parameters is suggested in [7] providing improved search efficiency but the overall speedup of the parallel computations is expected to be poor due to the overhead time of calculation of the mutation and migration parameters for each generation. The study of the effects of migration rates on the distribution of fitness suggests that increasing the migration rate results in greater selection intensity and thus delays the convergence of the algorithm and allows evolution of partial solutions independently in different demes [8].

The purpose of this paper is to propose an efficient parallel genetic model with periodic bidirectional migration of circular topology with parallel mutation rates for solving TSP on a multicomputer platform as well as to explore the impact of the parallel mutation rates on the quality of solutions.

2. ISLAND MODEL WITH CIRCULAR PERIODIC MIGRATION AND PARALLEL MUTATION RATES

The suggested circular migration model of parallel genetic computation of TSP is shown in Fig.1. The migration topology constitutes a logical ring of processes. Each processor evolves its local population and performs the genetic operations independently on an isolated subpopulation of the individuals, periodically sharing its best individuals with the other processors through migration. Best local chromosomes migration occurs between fixed pairs of processors periodically related to the number of the iterations (generations) of the PGA in order to facilitate the convergence of the PGA. The topology of the interconnection between demes is an important factor in the performance of the PGA because it determines how fast or how slow a good solution disseminates to other demes. The choice of the ring topology with bidirectional migration paths is motivated by the fact that the ring has a long diameter and allows a better differentiation of individuals i.e. a great diversity of genetic material spread over the islands.

Once a process receives the chromosomes-migrants it compares the fitness of the migrants with the fitness of the best local chromosomes. The local individuals of the worst local chromosomes are replaced by the migrants in case their chromosomes are better. The migration policy is as follows: process P_i (i denotes the rank of the process) sends its best chromosomes obtained so far to process $P_{(i+1)mod(N)}$ and receives the best chromosomes from process $P_{((i-1)+N)mod(N)}$, where N stands for the number of processes in the parallel system. Migration is implemented by means of the *MPI_Sendrecv()* communication operation for each migration period. Each process is assigned various mutation rates within the specified range $[\,0.01 \le \mu \le 0.2\,]$. Migration of the best individuals (chromosomes) occurs periodically in every migration period (per 20 generations). The number of migrants is 10 i.e. the best chromosomes of 10 individuals are transferred to another process following the circular migration topology and intended to replace the local worst chromosomes i.e. elitism is enforced. In PGA, for instance, the migration approach is an evaluation process between the local and global optimizations, (i.e.), the worst local optimizations are replaced with migrants of the best global optimizations.

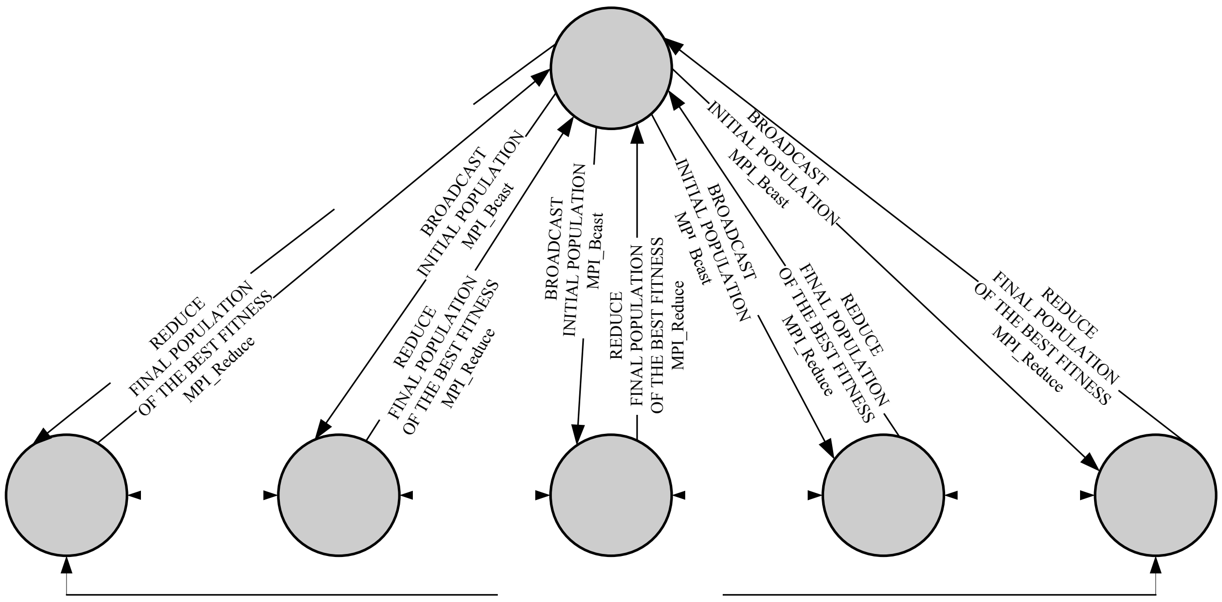

Fig. 1 - Parallel computational model of island model with periodic bidirectional chromosome migration

In the migration approach, the parameters for chromosome migration, such as migration topology, migration size, migration frequency, migrant composition, and replacement policy, must be carefully chosen to maintain a balance between subpopulation homogeneity (caused by chromosome migrations) and subpopulation diversity (caused by independent evolutions). If migrations occur too frequently or too many chromosomes are migrated, each processor might end up with the same subpopulation. If migrations occur too infrequently, or too few chromosomes are migrated, the migration approach degenerates to the independent approach.

Each processor generates its own initial subpopulation by random shuffles of the received chromosomes.

Next each processor starts performing the genetic operations of selection, recombination and mutation for creating new generation. The offspring chromosomes are generated by random single-point crossover e.g. the child copies the left part of its mother genome to the point of cross-over, and than copies the right part of its father genome from the point of cross-over on. Mutation occurs at the gene level within a chromosome by exchanging two randomly selected genes.

3. THE EXPERIMENTAL FRAMEWORK

The machine size of the multicomputer varies from 1 to 5 workstations (Intel Pentium 4 2,8 GHz) in order to explore the scalability of the parallel computer platform in respect to the scalability of the parallel application under investigation. The computers communicate via Fast Ethernet switch (100 Mbps). The parallel programming environment is MPICH Version 1.2.4, the language compiler is Microsoft Visual C++ 2005, and parallelism profiling is made utilizing Jumpshot Version 3.0.

Table 1. Experimental variable genetic parameters

#Processors	1	2	3	4	5	**#Migrants**
#Cities	**Population size**	**Subpopulation size**				
100	120	60	40	30	24	4
200	240	120	80	60	48	9
300	360	180	120	90	72	14
400	480	240	160	120	96	19
500	600	300	200	150	120	24
600	720	360	240	180	144	28

Table 2. Experimental parallel mutation strategies

Group	Strategy	Abbreviation	Description	Mutation rate μ
fmr	Fixed mutation rate	Fmr0.01	Fixed mutation rate $\mu = 0.01$	$\mu = 0.01$
		Fmr0.05	Fixed mutation rate $\mu = 0.05$	$\mu = 0.05$
		Fmr0.1	Fixed mutation rate $\mu = 0.1$	$\mu = 0.1$
		Frm0.15	Fixed mutation rate $\mu = 0.15$	$\mu = 0.15$
		Fmr0.2	Fixed mutation rate $\mu = 0.2$	$\mu = 0.2$
vmrg	Variable mutation rate for each generation	vmrginc0.001	Variable mutation rate for each generation increasing $\mu + = 0.001$; $0.01 \leq \mu \leq 0.2$	$\mu + = 0.001$; $0.01 \leq \mu \leq 0.2$
		vmrgdec0.001	Variable mutation rate for each generation decreasing $\mu - = 0.001$; $0.01 \leq \mu \leq 0.2$	$\mu - = 0.001$; $0.01 \leq \mu \leq 0.2$
pvmr	Parallel variable mutation rate different for each process	pvmr0.05	Parallel variable mutation rate - each process has different fixed mutation rate: $\mu = 0.2 \rightarrow$ P0, $\mu = 0.15 \rightarrow$ P1, $\mu = 0.1 \rightarrow$ P2, $\mu = 0.05 \rightarrow$ P3 and $\mu = 0.01 \rightarrow$ P4	$0.01 \leq \mu \leq 0.2$
		pvmrg±0.001	Parallel variable mutation rate for each generation; for even processes → incrementing $\mu + = 0.001$; $0.01 \leq \mu \leq 0.2$ and for odd processes → decrementing $\mu - = 0.001$; $0.01 \leq \mu \leq 0.2$	$0.01 \leq \mu \leq 0.2$

The experimental constant genetic parameters are as follows: 200 generations, random single-point cross-over, the migration topology is bidirectional ring, the migration period is 20 generations, the number of migrants is 4% of the population size, subpopulation size is determined by the ratio of the population size to the number of processors.

The experimental variable genetic parameters are summarized in Table 1.

We have made numerous experiments implying various types of mutation strategies (Table 2).

The experimental mutation strategies can be divided into 3 major groups. The first group, denoted by fmr, comprises identical fixed mutation rates for all local evolutions. The second group, denoted by vmrg implies variable mutation rate for each generation and has two sub-groups – increasing and decreasing mutation rates. The third group, denoted by pvmr, has two sub-groups – each local evolution applies different fixed mutation rate (pvmr0.05) and each generation evolves with different variable mutation rate (pvmrg±0.001).

4. PARALLELISM PROFILING AND SCALABILITY ANALYSIS

The parallelism profiling is shown for the instance of 600 cities. The connected states of the communication transactions during the parallel computational process are shown in Fig.2. Gantt's chart for the parallel genetic computation on a multicomputer of five workstations, showing the behavior of the circular migration model implementation is given in Fig.3. For one migration period there are 28 migrants of the best chromosomes migrating from a process to the neighboring process.

The scalability of the parallel system is shown in Fig.4. Obviously, the genetic implementation of the parallel application of TSP with circular periodic chromosomes migration scales well and almost proportionally with the size of the multicomputer platform.

Comparison of the speedup obtained with different communication topologies of the migration is shown in Fig.5. As can be seen bidirectional circular migration outperforms both global broadcast migration PGA model and independent island PGA without migration providing 5 to 10% faster convergence for different sizes of the multicomputer.

The execution on 10 processors show 10% better speedup for circular migration topology compared to independent island model without migration and 13% better speedup for circular migration topology compared to global broadcast migration.

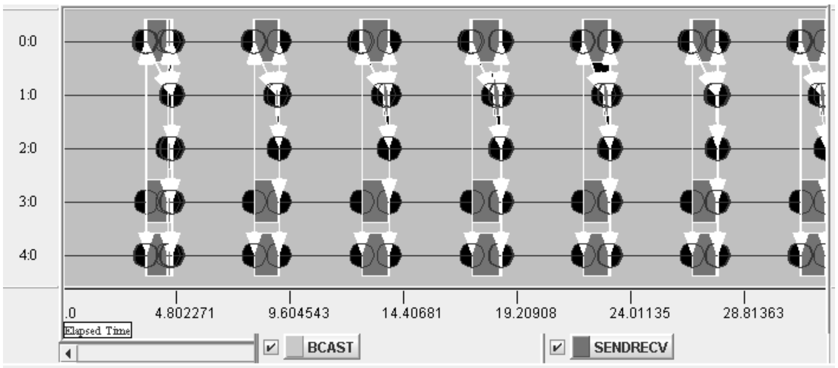

Fig. 2 - Connected states of communication transaction

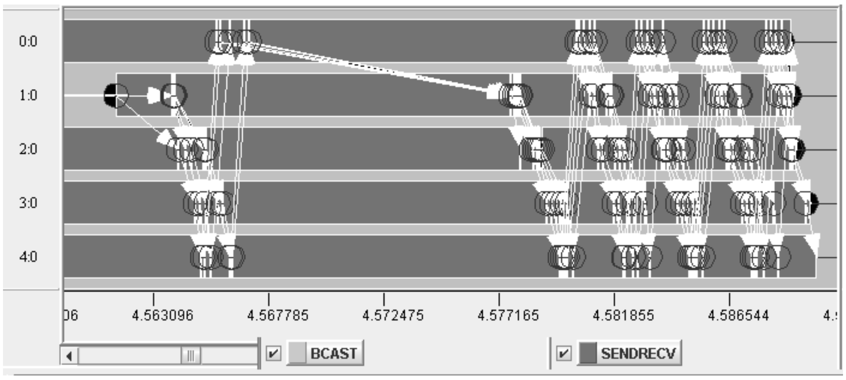

Fig. 3 - Gantt's chart for one migration period (28 migrants) for TSP of 600 cities

The experimental results draw to a conclusion that is only theoretically evaluated in [9] – PGA can integrate large numbers of processors and reduce significantly the execution time of many practical applications.

5. QUALITY OF SOLUTION ANALYSIS

Our experimental study of the impact of the mutation strategy on the solution quality of the parallel genetic algorithm with circular migration topology and variable types of mutation rates is based on 50 runs on the multicomputer platform. On each island (computer) the local evolution starts with different initial population. The statistics of the best fitness values obtained for solving the TSP for 200 cities implementing genetic approach with circular migration topology and parallel fixed mutation rates for local evolutions is presented in Fig.6.

The percentage quality difference D [6] of the best fitness obtained in one execution of PGA on the multicomputer platform compared to the perfect solution of TSP is calculated as:

$$D = \frac{best\_fitness - perfect\_fitness}{perfect\_fitness} \times 100\% \quad (1)$$

where, *best_fitness* – length of the shortest tour found by the parallel genetic algorithm, *perfect_fitness* – length of the shortest tour (optimal).

The quality difference for 50 runs for the strategy with parallel fixed mutation rates for local evolutions are shown in Fig.7. The deviation of the quality difference of 50 runs is shown in Fig.8. The comparison of the average quality difference of the different experimental mutation strategies is shown in Fig.9.

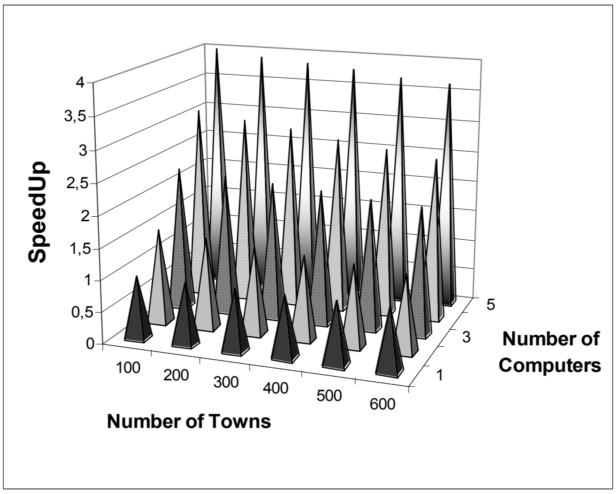

Fig. 4 - Scalability of the parallel system

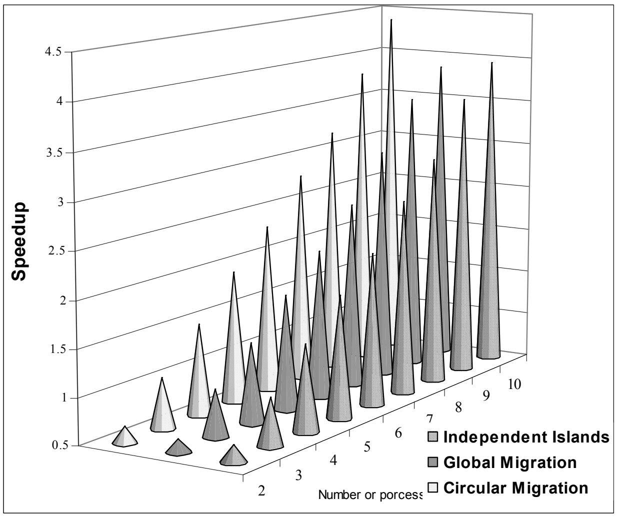

Fig. 5 - Speedup comparison of migration topologies

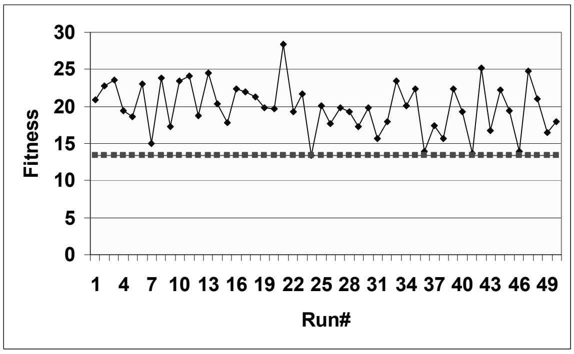

Fig. 6 - Fitness statistics for parallel mutation strategy pvmr0.05

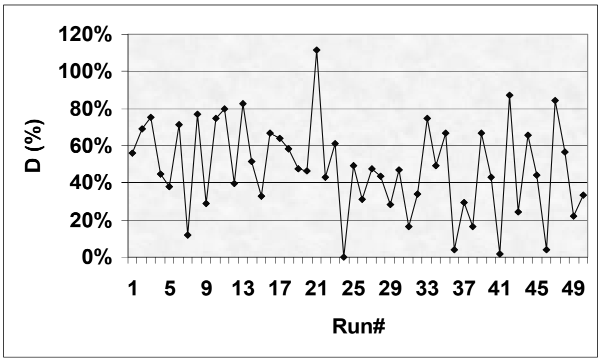

Fig. 7 - Quality difference for parallel mutation strategy pvmr0.05

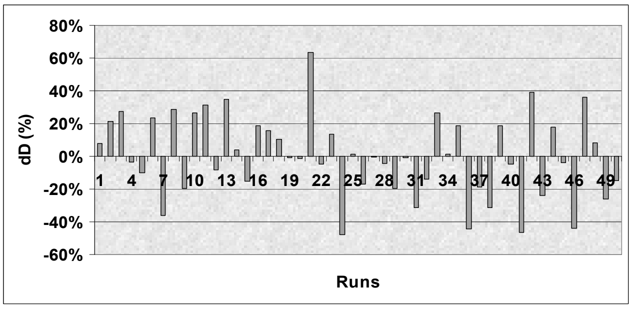

Fig. 8 - Deviation of the quality difference for parallel mutation strategy pvmr0.05

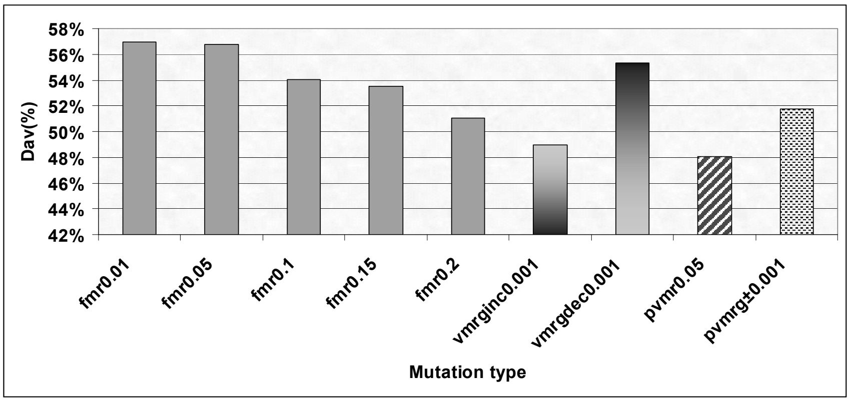

Fig.9 - Deviation of the quality difference for the experimental parallel mutation strategies

Obviously, the best results are obtained for the mutation strategy with parallel fixed mutation rates *pvmro.05* i.e. different fixed mutation rate is applied for each local evolution. Experimental results for the mutation strategy *pvmro.05* on a network of 10 workstations show that the average quality difference improves significantly.

The mutation strategy with incrementing parallel mutation rates for generations *vmrginc0.001* and the mutation strategy with parallel mutation rates for generations *pvmrg* ± *0.001* also lead to satisfactory quality solutions. Out of the strategies with identical fixed mutation rates for local evolutions the strategy *fmr0.2* gives solutions of the best quality. The worst quality of solutions is obtained for the strategies with identical fixed mutation rates for local evolutions *fmr0.01*, *fmr0.05* and the strategy with decrementing parallel mutation rates for generations *vmrgdec0.001*.

The results for the solution quality obtained with bidirectional circular migration utilizing parallel fixed mutation rates *pvmrg±0.5* are compared with global broadcast migration PGA model and independent island PGA model without migration (Fig.10).

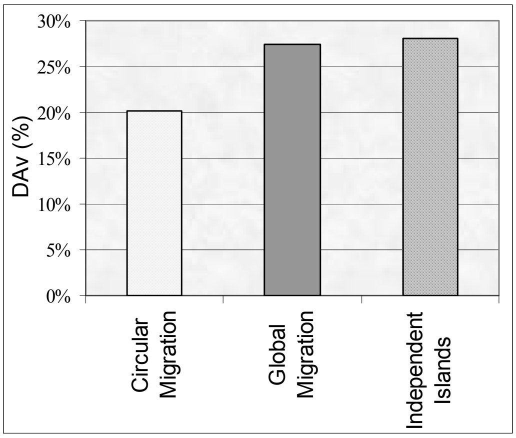

Fig. 10 - Comparison of the deviation of the quality difference for the experimental communication topologies of the migration

Obviously the circular migration gives not only fastest convergence but provides highest quality of the solution obtained – deviation of the quality difference calculated on the base of 50 runs is more than 15% better for the circular topology of the migration compared to the other migration topologies.

6. CONCLUSIONS

In this paper we have suggested parallel genetic model with circular bidirectional migration for computing the TSP with parallel mutation rates. The proposed parallel computational model has been verified by means of MPI-based parallel implementation on multicomputer platform. Parallelism profiling and performance estimation and analysis have been made on the basis of the MPI parallel program implementation for a scalable multicomputer platform for various numbers of cities. The experimental results show that the parallel coarse-granule implementation scales almost proportionally in respect to the machine size. The speedup slightly slows down increasing the problem size from 100 to 600 cities. Speedup comparison of different migration topologies show that bidirectional circular migration topology outperforms both global broadcast migration PGA model and independent island PGA without migration.

The study of the impact of the mutation strategy on the solution quality of the parallel genetic algorithm with circular migration topology and variable types of mutation rates draw to a conclusion that the variation of mutation rate for each generation and for each process can improve the performance and get the best fitness of PGA. The best fitness achieved experimentally is for the case of parallel fixed mutation rates for each local evolution. This result is explained by the fact that in this case we get a great diversity of populations and a greater chance to evolve near optimal population.

The aim of the paper was to explore the influence of variation of mutation parameters and dynamic migration topology on the speedup and solution quality provided by the PGA. Compared to the results achieved by other researchers [5, 6, 7] the suggested parallel fixed mutation rates and bidirectional circular migration topology require less computations than fuzzy or adaptive parameters tuning thus providing both better convergence computation speed of the PGA. Moreover, our investigation uses an optimization problem of TSP as a case study and not particular function as in [6] or dynamic environment problem as in [7]. Future work will include further research to apply our experience with the TSP to other scheduling and networking problems and to explore the implementation of the suggested migration and mutation strategies in solving other optimization problems by island PGA.

7. REFERENCES

[1] Traveling Salesman Problem www.tsp.gatech.edu

[2] R. Haupt, S. Haupt, *Practical Genetic Algorithms*, John Wiley & Sons, 2004.

[3] Genetic algorithms for TSP http://www.tsp.umu.se/~top/travel.html

[4] P. Borovska, “Solving the TSP in Parallel by Genetic Algorithm on Multicomputer Cluster”, *International Conference on Computer Systems and Technologies (CompSysTech’06)*, June 2006, Veliko Turnovo, Bulgaria.

[5] L. Wang, A. Maciejewski, H. Siegel, V. Roychowdhury, and B. Eldridge, “A study of five parallel approaches to a genetic algorithm for the traveling salesman problem”, *Intelligent Automation and Soft Computing*, Vol.11, No.4, 2005, pp. 217÷234.

[6] W. Rand, R. Riolo, “The Problem with a Self Adaptive Mutation Rate in Some Environments”, *Proc. of the 2005 Conference on Genetic and evolutionary computation GECCO'05*, 2005, pp.1493-1500.

[7] Y. Maeda, Q. Li, “Parallel Genetic Algorithm with Adaptive Genetic Parameters Tuned by Fuzzy Reasoning”, *International Journal of Innovative Computing, Information and Control*, Vol.1, No.1, ICIC International, March 2005, pp. 95-107.

[8] E. Cantu-Paz, “On the Effects of Migration on the Fitness Distribution of Parallel Evolutionary Algorithms”, *Workshop on Parallel Processing and Evolutionary Computation*, Las Vegas, July 8th, 2000.

[9] E. Cantu-Paz, D. Goldberg, “On the scalability of Parallel Genetic Algorithms”, *Evolutionary Computation*, Vol.7, No.2, 1999, pp. 429-449.

Assoc. Prof. PhD Plamenka Ivanova Borovska *is currently head of Computer System Department, Technical University of Sofia, Bulgaria. She has graduated from Technical University of Sofia, specialty “Computer Systems and Technologies” in 1978. She defended PhD thesis in the area of parallel computing at the same university. Research areas: parallel computing, parallel algorithms, parallel programming, parallel computer architectures.*

Subhi Abdulrahim Bahudajla *(Republic of Yemen) is a PhD student at Computer Systems Department, Technical University of Sofia, Bulgaria. He has graduated from Technical University of Varna, specialty “Computer Science and Technologies” in 1988г. Research areas: parallel computing, parallel algorithms, parallel programming, parallel computer architectures, multithreading.*

Milena Kirilova Lazarova, PhD, *is an Assist.Prof. in the Computer System Department, Technical University of Sofia, Bulgaria. She has graduated from Technical University of Sofia, specialty “Computer Systems and Technologies” in 1995. Research areas: parallel computing, parallel algorithms, parallel programming, parallel computer architectures.*