A Hybrid Genetic - Differential Evolution Algorithm (HybGADE) for a Constrained Sequencing Problem

Didem Abidin
Computer Engineering Department
Manisa Celal Bayar University
Manisa, Turkey
didem.abidin@cbu.edu.tr

Abstract— **For researchers, evolutionary algorithms are mostly preferable because of their effectiveness in finding the optimum solutions to many problems. Among these problems, sequencing is one of the most popular. In daily life, it is a must to find the best solution to a sequencing problem in order to save time, money and labour. Education is also one of the application areas of optimization where sequencing matters. In this paper, a hybrid genetic - differential algorithm is introduced, which finds better solutions to sequencing problem in education. The correct order of educational data is crucial because it directly affects the students' performance. In this study, educational material of Database course in Ege University Tire Kutsan Vocational School Computer Programming Department is used as the data set with two different evolutionary algorithms (EA). In these data sets, there are some constraints which should be considered while sequencing. We called them "prerequisites" that tells us the rules about the order of the modules of a course. That is why, the study can be considered as an application of Precedence-Constrained Sequencing. The sequencing performances of pure genetic algorithm (GA) and hybridized differential evolution (DE) with genetic algorithms (HybGADE) are compared with a computer program implemented. It is observed that, HybGADE can be used with 99.54% of reliability where pure GA has an effectiveness of 98.53%. The percentage of the students passing the class has been observed for four years. The ratio of students passing the class has increased from 39% to 65%, which can be considered as a remarkable increase.**

Index Terms—**Differential Evolution Algorithm, Hybrid algorithms, Genetic Algorithm, Precedence-Constrained Sequencing.**

I. Introduction *(Heading 1)*

The idea behind the precedence-constrained sequencing problems (PCSP) is to obtain the best sequence of data given with less cost especially from the time consumption point of view. Since there are various application areas of PCSPs like scheduling [1], routing, networking, travelling salesman problem (TSP) [2] or project management [3], there emerged different solution techniques to these problems. In some of them, heuristic solution mechanisms are offered [4][5], while for others artificial intelligence algorithms are used.

Genetic Algorithms (GA) [6][7] are one of the most common solution techniques to PCSPs. In previous studies, researchers also presented some alternative and effective solutions to these types of problems. GA is used to solve precedence-constrained sequencing and scheduling in production [8] and in process planning [9]. Some studies showed the significant differences between the applications of different versions of GA operators. Route planning problems can be considered as sequencing problems and GA was used successfully in [10] with a standard one – point crossover operator. Order – acceptance problems can be classified as another branch of sequencing problems, in which a two – point crossover technique can be applied [11]. Curriculum sequencing is also handled in [12] for different courses having different features from each other, which can be considered as a rule based approach.

Differential Evolution (DE) is a parallel direct search method which works on vector populations [13]. Unlike GA, the most important operator of DE is mutation. This means that randomness plays a greater role in DE than it plays in GA.

Traditionally, DE is applied to continuous optimization problems [14]. On the other hand, GA can be applied to global optimization problems. There are some studies which compared the performances of the two algorithms [15] on very popular NP-Complete combinatorial optimization problems like TSP and N-Queen [16]. Another study has combined the two in their study to solve unit commitment scheduling problem [17], which is a mixed-integer optimization problem. Trivedi et al [17] obtained a remarkable cost saving with the stated structure. In previous studies, GA had been introduced as an effective solution for sequencing precedence constrained data [18][19]. DE is also used in solving multi-objective optimization problems with GA [20].

There are some very strong hybrids of GAs with other evolutionary algorithms [21]. GA can be merged with particle swarm optimization (PSO) technique [22] and ant colony optimization [23] for continuous domains successfully. In the same way, DE can also be combined with other evolutionary algorithms to obtain powerful hybrid solutions. PSO is an appropriate couple to DE to be hybridized [24]. Similarly, covariance matrix adaptation algorithm has been successfully hybridized with DE for parameter optimization in [25].

In this study, GA and DE are combined in a different way to solve a combinatorial optimization problem. Both GA and DE operators are used together in one algorithm to obtain the

most reliable sequencing results on educational data. This data consists of the parts of a database class in Computer Programming department of Ege University Tire Kutsan Vocational School. The performance of students has been observed for four fall semesters and the results are given in this paper. Sequencing the education material with a pure genetic algorithm has also increased students' performance; but with the application of HybGADE explained here, more reliable results were obtained. The performance of the students increased from 39% to 65%, which can be considered remarkable.

The layout of the paper is as follows: In Section 2, the educational data and the structure of the hybrid algorithm are introduced. In Section 3 results of the study are given and in Section 4, the paper is concluded with discussions and future work.

II. FORMULATION OF PROBLEM

A. Data

This paper seeks for the performance analysis results for sequencing educational data. Sequencing the topics of a course in correct order can be considered a constraint satisfaction problem [26] [20]. Among the topics, there exist some which should be introduced before or after others. The constraints among the topics influence the order lectures will be taught to students. The symbolic representation of the educational data is saved in an XML file. Each module tag in the XML represents one of the topics in the course. The topics are enumerated from 0 to 19 (attribute *mno*). For each module tag, there are several *prereq* tags, indicating the prerequisite topics for the module. This XML file is converted to a sparse matrix and the sparse matrix is used in HybGADE. In the matrix, first line belongs to first module *mno*=0. For example, the *prereq* tags for one of the modules are transferred to a line of the sparse matrix as: "0-0-1-0-0-1-0-1-0-0-0-0-1-0-0-0-0-1-0".

B. HybGADE Structure

In HybGADE, mutation mechanism of GA is replaced with the mutation of DE. In DE, since the mutation operation is considered the main operation, the mutation rate variable for classical GA is also omitted. The pseudocode for HybGADE is given in Figure 1.

Although the general structure of the algorithm looks like that of the GA's, there are some remarkable differences. In [17], either GA or DE is applied in the algorithm, when it is time to apply the genetic operators. But in this study, the most powerful weapons of GA and DE are used together as one algorithm. The possible output of the program will look like Figure 2.
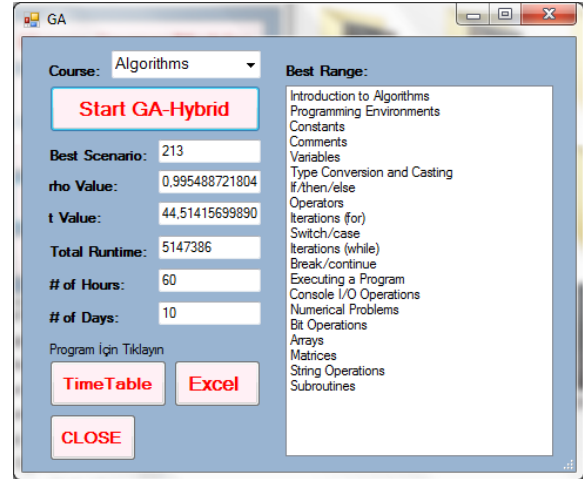

Fig. 1. Pseudocode for HybGADE

Fig. 2. A possible output in the program

As shown in the figure, the genes of the solution consist of the modules of the course. The order given in the list box on the right is the order offered for the modules to be taught. The reliability percentage of the order can be read from the rho value calculated with Spearman Rank Correlation.

1) Generating initial population

Each individual in the population is supposed to be a possible solution for HybGADE. In educational data, the individuals are formed according to the permutation encoding principle [27]. Each gene in the chromosome represents a part of the sequence; therefore each gene should have a unique integer value. The initial population is generated randomly, with randomly sequenced genes. Once the first randomly sequenced chromosome is generated, the rest of the population can be derived by taking different combinations of the first individual randomly again. The number of individuals in the population of educational data is used as NP = 100, 120, 140, 160, 180, 200.

2) Fitness function

For the educational data, the genes in a chromosome should be sequenced according to precedence rules. Each gene can have prerequisite genes for itself. When one or some of the prerequisite genes for a certain gene (say p) take place after p, in this case, this individual cannot be an appropriate sequence solution. The aim of the fitness function is to minimize the precedence violations. The precedence rules of the data to be

sequenced are given in a prerequisite matrix and the fitness function is responsible for checking the precedence violations for each gene in each individual. The equation for the fitness function is given in Equation (1).

$$fitness\,[i] = \frac{1}{\sum_{k=1}^{m} \frac{pcount_k}{prcount_k}} \qquad (1)$$

where m is the number of modules in a chromosome, *pcount* is the number of prerequisite modules existing in the chromosome for module k and *prcount* is the total number of prerequisite modules for module k. When the fitness values for all individuals are calculated, they are ranked in the population matrix from the least possible to the most possible solution according to their fitness values. By doing so, it is more likely to choose better chromosomes to survive to the next generation.

3) Selection and crossover operations

Selection and crossover operations are used exactly as they are in a pure GA, which means that two individuals are selected from the population according to linear-ranking selection [28]. As the crossover operation, order crossover (OX1) [29], 2-point order crossover (OX2) [30] and partially - mapped crossover (PMX) [31][32] are used.

4) Mutation operation

The mutation operator of GA is replaced with the mutation operator of DE. HybGADE is improved to work without a mutation rate, which means that the new mutation operator is applied in all iterations. In DE, mutation operator is usually followed by the selection operation of DE, which decides the mutant individual's survival according to the fitness value. In HybGADE, the mutant individual is replaced with the chosen individual if its fitness value is better. In this case, mutation operator of DE is moved to GA with its decision step. This step helps HybGADE to come to a decision about the mutant individual quickly. In this study, four different mutation operators DE/rand/1 [25], DE/best/1 [13], Simplex1 [33] and Simplex2 [33] are used. Parameters used in HybGADE are given in Table 1. F value is called the differential weighting factor and it generally takes values in [0.4, 1] interval [34].

TABLE I. PARAMETERS FOR HYBGADE

Parameter	Value
Population Size	100, 120, 140, 160, 180, 200
# of Generations (G)	500, 750, 1000
Crossover Rate (crate)	0.8
F value	0.5

III. EXPERIMENTAL RESULTS

The solutions obtained with GA and HybGADE are applied on students in Ege University Tire Kutsan Vocational School for four fall semesters. The data is tested on the computer program developed in Java. In 2011-12 Fall semester, none of the sequencing techniques were applied to the educational material of Database course and the passing percentage was 39%. In 2012-2013 Fall, 33 students were registered on the course. The sequence of topics, which was offered as the results of pure GA was applied during semester. At the end of the semester, 17 students were successful with a percentage of 51%. In 2014-2015 Fall semester, 44 students were registered in the same course. This time, the sequence of topics as offered by the HybGADE was applied. Among these, 23 students were successful, which shows an increase in the overall performance of the class with 52%. In 2015-16 Fall semester, HybGADE was again applied to educational material of Database course. Thirty-one students out of 47 have passed the class with a 65% success ratio, which shows the effect of the sequencing of course material. The increase in the success of the class is given in Figure 3.
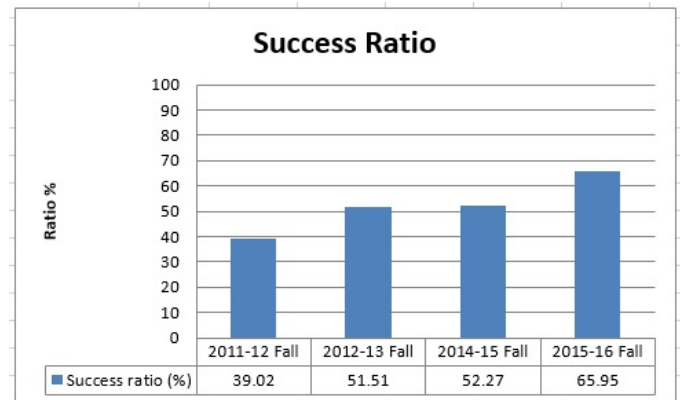

Fig. 3. Success ratios for database classes

When pure GA results are compared to that of the HybGADE's, results of HybGADE are almost always better than pure GA's. The reliability percentages of crossover operators OX1, OX2 and PMX combined with four DE mutation operators are given in Figure 4, 5 and 6 respectively.
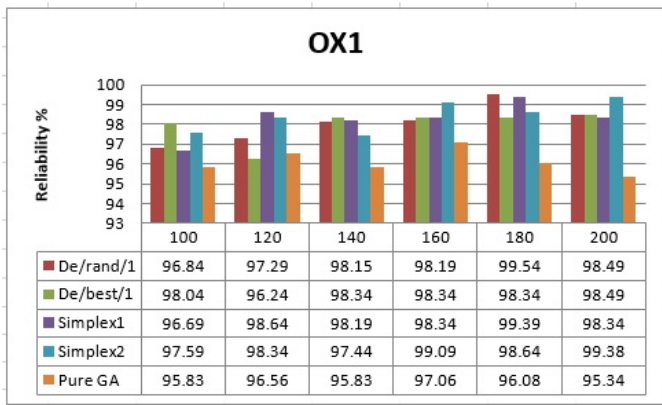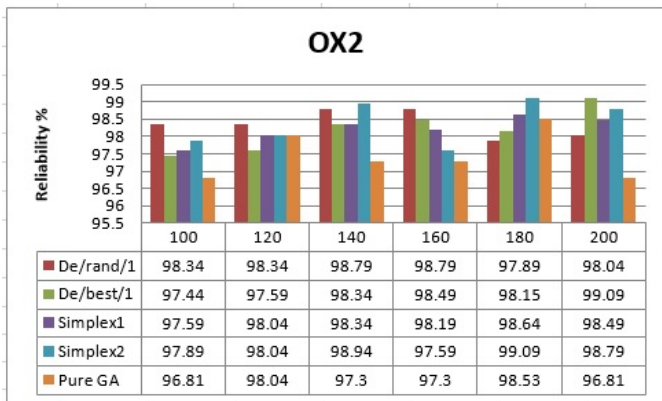
Fig. 4. Reliability percentages for OX1

	100	120	140	160	180	200
De/rand/1	96.84	97.29	98.15	98.19	99.54	98.49
De/best/1	98.04	96.24	98.34	98.34	98.34	98.49
Simplex1	96.69	98.64	98.19	98.34	99.39	98.34
Simplex2	97.59	98.34	97.44	99.09	98.64	99.38
Pure GA	95.83	96.56	95.83	97.06	96.08	95.34

Fig. 5. Reliability percentages for OX2

	100	120	140	160	180	200
De/rand/1	98.34	98.34	98.79	98.79	97.89	98.04
De/best/1	97.44	97.59	98.34	98.49	98.15	99.09
Simplex1	97.59	98.04	98.34	98.19	98.64	98.49
Simplex2	97.89	98.04	98.94	97.59	99.09	98.79
Pure GA	96.81	98.04	97.3	97.3	98.53	96.81

Fig. 6. Reliability percentages for PMX

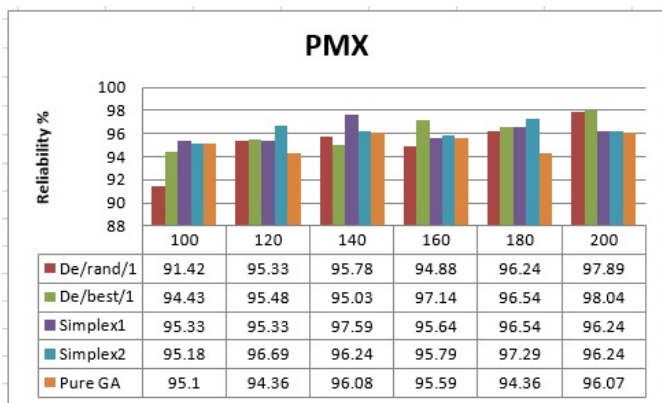|	100	120	140	160	180	200
De/rand/1	91.42	95.33	95.78	94.88	96.24	97.89
De/best/1	94.43	95.48	95.03	97.14	96.54	98.04
Simplex1	95.33	95.33	97.59	95.64	96.54	96.24
Simplex2	95.18	96.69	96.24	95.79	97.29	96.24
Pure GA	95.1	94.36	96.08	95.59	94.36	96.07

The results show that HybGADE performs better than applying pure GA operators for this application. The best result is obtained for NP=180 for OX1 crossover and DE/rand/1 operator combination. Pure GA had obtained the best reliability for NP=180 in previous studies [19]. But the results for NP=180 in hybrid versions even got better. Also when results are observed for the same number of individuals in the population, hybrids of GA crossover and DE mutation have always given better results. On the other hand, when runtimes are compared for hybrid results, OX1, having the best performance, is not the fastest. Out of 24 crossover-mutation combinations, PMX is the fastest in the 19 combinations.

When fitness values obtained from the results are examined, maximum fitness values in each combination is obtained as 1, where minimum fitness values change with the number of individuals in the population. Figure 7 shows the minimum fitness values, which tend to increase with an increasing number of individuals in the population.
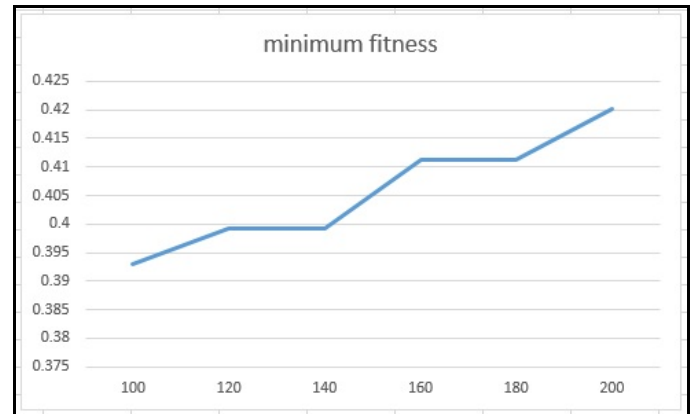

Fig. 7. Minimum fitness values for six different population values (NP)

IV. 4. DISCUSSION AND FUTURE WORK

In this study, the main purpose was to highlight the power of hybridized algorithms in education domain. In earlier studies, pure GA has also obtained reliable results. But with the combination of mutation of DE to GA, HybGADE came forward as more powerful mechanism to be used in combinatorial optimization problems than pure GA. Here, four different mutation mechanisms for DE and three different crossover mechanisms for GA are used. Since the mutation mechanism of GA was replaced with that of DE, no mutation rate value was used in the implemented algorithm of HybGADE.

The students of Ege University Tire Kutsan Vocational School had very low success rates in courses. Optimizing the sequence of topics in a course directly affected the understanding capability of the students. With this study, the order of the topics was arranged according to the theoretical results of pure GA first. There occurred a remarkable increase in their success ratios. Next year, the order of the topics was arranged according to the theoretical results of HybGADE. The increase in the success ratios of the students showed that the theory of HybGADE is better than pure GA and applicable in practice.

Next move about HybGADE will be applying the hybrid sequencing algorithm to a different - but well known - problem to provide different application areas to the algorithm.

REFERENCES

[1] Y. Yun, M. Gen, C. Moon, "Hybrid genetic algorithm with adaptive local search for precedence – constrained sequencing

problems". 40th IEEE Computers and Industrial Engineering Conference, 2010, pp. 1 – 6.

[2] A. Otman, A. Jaafar, "A comparative study of adaptive crossover operators for genetic algorithms to resolve the travelling salesman problem". International Journal of Computer Applications 31(11), 2011, pp. 49-57.

[3] Y. Yun, H. Chung, C. Moon, "Hybrid genetic algorithm approach for precedence - constrained sequencing problem". Computers & Industrial Engineering 65, 2013, pp. 137-147.

[4] E. Duman, I. Or, "Precedence constrained TSP arising printed circuit board assembly". International Journal of Production Research 42(1), 2004, pp. 67-78.

[5] J. Renauld, F.F. Boctor, J. Ouenniche, "A heuristic for the pickup and delivery travelling salesman problem". Computers and Operations Research 27, 2000, pp. 905-916.

[6] H.J. Holland, Adaptation in Natural and Artificial Systems. Cambridge: MIT Press, 1975.

[7] D.E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning. Reading: Addison Wesley, 1989.

[8] S.D. Dao, R. Marian, "Modeling and optimization of precedence - constrained production sequencing and scheduling using multi-objective genetic algorithms". Proceedings of the World Congress on Engineering London, UK. Vol II, 2011, pp. 1027-1032.

[9] Y. Su, X. Chu, D. Chen, D. Chu, "A modified genetic algorithm for precedence-constrained operation sequencing problem in process planning". IEEE International Conference on Industrial Engineering and Engineering Management, Bandar Sunway, 2014, pp. 84-88.

[10] M.C. Wu, C.F. Shih, C.F. Chen, "An efficient approach to cross - tab route planning for water manufacturing". International Journal of Expert Systems with Applications 36, 2009, pp. 11962 – 11968.

[11] W.O. Rom, S.A. Slotnick, "Order acceptance using genetic algorithms". Computers & Operations Research Elsevier Ltd. 36, 2009, pp. 1758 – 1767.

[12] L. De Marcos, R. Barchino, J-J. Martínez, J-A. Gutiérrez, J-R. Hilera, "Competency-based curriculum sequencing: Comparing two evolutionary approaches". IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Sydney, Australia, 9 – 12 December 2008, pp. 339-342.

[13] R. Storn, K. Price, "Differential evolution: A simple and efficient adaptive scheme for global optimization over continuous spaces". Journal of Global Optimization 11, 1997, pp. 341 – 359.

[14] J. Zhang, V. Avasarala, A.C. Sanderson, T. Mullen, "Differential optimization for discrete optimization: An experimental study on combinatorial auction problems". IEEE Congress on Evolutionary Computation, Hong Kong, 2008.

[15] B. Hegerty, C.C. Hung, K. Kasprak, "A comparative study on differential evolution and genetic algorithms for some combinatorial problems". 1st Workshop on Intelligent Methods in Search and Optimization, 2009.

[16] A.S. Farhan, W.Z. Tareq, F.H. Awad, "Solving N queen problem using genetic algorithm". International Journal of Computer Applications 122(12), 2015, pp. 11-14.

[17] A. Trivedi, D. Srinivasan, S. Biswas, T. Reindl, "Hybridizing genetic algorithm with differential evolution for solving the unit commitment scheduling problem". Swarm and Evolutionary Computation 23, 2015, pp. 50-64.

[18] D. Abidin, Ş. Çakır, "Rule – based genetic algorithm for in – service training curriculum plan". International Conference on Computers, Digital Communications and Computing, Barcelona, Spain, 2011, pp. 160 – 166.

[19] D. Abidin, Ş. Çakır, "Analysis of a rule based curriculum plan optimization system with spearman rank correlation". Turkish Journal of Electrical Engineering and Computer Sciences 22, 2014, pp. 176-190.

[20] P. Xiaoying, Z. Jing, C. Hao, C. Xuejing, H. Kaikai, "A differential evolution - based hybrid NSGA-II for multi-objective optimization". IEEE 7th International Conference on Cybernetics and Intelligent Systems and Robotics, Automation and Mechatronics; 2015, pp. 81-86.

[21] M.A. Rigi, S. Mohammadi, "Finding a hybrid genetic algorithm constraint satisfaction problem based solution for resource constrained project scheduling". IEEE International Conference on Emerging Technologies, Islamabad, Pakistan, 2009, pp. 52-56.

[22] C.F. Juang, "A hybrid of genetic algorithm and particle swarm optimization for recurrent network design". IEEE Transactions on Systems, Man and Cybernetics Part B: Cybernetics 34(2), 2004, pp. 997-1006.

[23] S. Fidanova, M. Paprzycki, O. Roeva, "Hybrid GA-ACO algorithm for a model parameters identification problem". Federated Conference on Computer Science and Information Systems, Warsaw, Poland, 7-10 September 2014, pp. 413-420.

[24] M. Pant, R. Thangaraj, C. Grosan, A. Abraham, "Hybrid differential evolution - particle swarm optimization algorithm for solving global optimization problems". 3rd International Conference on Digital Information Management, London, UK, 13-16 November 2008, pp. 18-24.

[25] S. Ghosh, S. Das, S. Roy, S.K. Minhazul Islam, P.N. Suganthan, "A differential covariance matrix adaptation evolutionary algorithm for real parameter optimization". Information Sciences 182, 2012, pp. 199-219.

[26] J.H. Moreno-Scott, J.C. Ortiz-Bayliss, H. Terashima-Marín, S.E. Conant-Pablos, "Challenging heuristics: Evolving binary constraint satisfaction problems". Genetic and Evolutionary Computing Conference (GECCO), Philadelphia, Pennsylvania, USA, 2012, pp. 409-416.

[27] B.A. Julstrom, "String and permutation coded genetic algorithms for the static weapon-target assignment problem". 11th Annual Conference on Genetic and Evolutionary Computation Conference, Montreal, Quebec, Canada, 2009, pp. 2553-2558.

[28] N.M. Razali, J. Geraghty, "Genetic algorithm performance with different selection strategies in solving TSP". Proceedings of the World Congress on Engineering WCE, London, UK. Vol II, 6-8 July 2011, pp. 1134-1139.

[29] L. Davis, "Applying adaptive algorithms to epistatic domains". Proceedings of the International Joint Conference on Artificial Intelligence, 1985, pp. 162-164.

[30] G. Syswerda, Schedule Optimization Using Genetic Algorithms. Editors: Davis L., Van Nostrand Reinhold. Handbook of Genetic Algorithms, New York, NY, USA, 1991, pp. 332-349.

[31] D.E. Goldberg, R. Lingle, "Alleles, loci, and the traveling salesman problem". Proceedings of the First International Conference on Genetic Algorithms and Their Application, 1985, pp. 154-159.

[32] G. Üçoluk, "Genetic algorithm solution of the TSP avoiding special crossover and mutation". Journal of Intelligent Automation & Soft Computing 8(3), 2013, pp. 265-272.

[33] D. Kamiyama, K. Tamura, K. Yasuda, "Down-Hill simplex method based differential evolution". SICE Annual Conference, Taipei, Taiwan, 2010, pp. 1641 – 1646.

[34] D. Swagatam, P.N. Suganthan, "Differential evolution: A survey of the state-of-the-art". IEEE Transactions on Evolutionary Computation 15(1), 2011, pp. 4-31.