Assessing the Performance of the Non-dominated Sorting Genetic Algorithm in Optimizing Construction Site Planning

A. M. Khalafallah[1]

[1]Department of Architectural & Manufacturing Sciences, Western Kentucky University, Bowling Green, KY 42101; PH (270) 745-5949; FAX (270) 745-5946; email: ahmed.khalafallah@wku.edu

ABSTRACT

Genetic algorithms are techniques inspired by the mechanics of natural selection and genetics. They adopt survival of the fittest and structured exchange of information as basic mechanisms for optimization, and they are deemed suitable for construction site planning. Among the commonly-used multi-objective genetic algorithms, the Non-dominated Sorting Genetic Algorithm stands out as a robust optimization algorithm for such type of construction optimization problems. The main objective of this study is to investigate the performance of the algorithm under the alteration of optimization parameters. Specifically, the study investigates the performance of the algorithm under the variation of the population size, number of generations, type and value of cross-over probability, and mutation likelihood. This study should provide a better understanding of the behavior of the algorithm in solving site planning problems, and identify strategies that could help accelerate convergence toward optimal solutions.

INTRODUCTION

Designing a construction site layout is a process that entails identifying a location for every temporary facility with the objectives of (i) cutting on the cost of material handling, (ii) reducing travel times for equipment and personnel, and (iii) promoting safety and security on site (Khalafallah and El-Rayes 2008, Anumba and Bishop 1997, and Tommelein et al. 1992). Temporary facilities may include field offices, storage facilities, staging areas, fabrication yards, toilets, disposal areas, batch plants, among others (Khalafallah and El-Rayes 2008, Hegazy and Elbeltagi 1999, Yeh 1995). The process of designing an efficient layout for the construction site is crucial to the success of a construction project and could result in significant savings, if properly performed, or excessive risk exposure, if recklessly overlooked.

Several studies investigated optimizing construction site layouts, using a variety of methodologies, including genetic algorithms (Khalafallah and El-Rayes 2008, 2006a, 2006b; Mawdesley et al. 2002; Tawfik and Fernando 2001a, 2001b; and Harmanani et al. 2000), neural networks (Yeh 1995), simulation (Dawood and Marasini 2001), among others. Genetic algorithms have been proven to effectively facilitate the optimization of construction site layouts and improve the search process, especially for problems characterized by a large search space (Khalafallah 2006). These algorithms (GAs) are inspired by the mechanics of natural selection and genetics. They adopt the survival of the fittest and the structured exchange of genetic materials among population members over successive generations as a basic mechanism for the search process (Goldberg 1989). GAs are considered effective tools for solving site layout optimization problems for several reasons, including (1) the usual non-continuity and large size of the search space; (2) the ease of representing solutions in the form of strings; and (3) the sufficiency of near-optimal solutions for this type of optimization problems (Khalafallah 2006). Multi-objective genetic algorithms are special types of GAs that are devised to optimize a number of objectives simultaneously. The Non-dominated Sorting Genetic Algorithm (NSGA, Deb et al. 2000) and its derivative (NSGA II) are among the most widely used multi-objective genetic algorithms, and have been proven effective in optimizing site layouts that involve multiple optimization objectives, if these objective functions are not combined together using utility functions.

Despite the robustness of the Non-dominated Sorting Genetic Algorithm in optimizing construction site layouts, the process of utilizing these algorithms is intricate and involves specifying several optimization parameters such as the population size, the number of generations, the type and value of cross-over probability, and the mutation probability. Improper specification of these parameters hinders the performance of the algorithm, and in many cases leads to improper convergence toward local maxima/minima. As such, the study of these parameters and their effects on optimization is deemed critical for an efficient and successful execution of these algorithms.

The main objective of this study is to investigate the performance of the Non-dominated Sorting Genetic Algorithm II (NSGA II) under the alteration of optimization parameters. Specifically, the study aims at evaluating the performance of the algorithm under the variation of the population size, number of generations, type and value of cross-over probability, and mutation likelihood.

RESEARCH METHODOLOGY

In order to investigate the performance of NSGA-II under the alteration of the optimization parameters, a benchmark optimization problem is selected. The problem first appeared in the literature in 2005 and involves multi-objective optimization of a construction site layout for a multistory garage building (El-Rayes and Khalafallah 2005). The objectives of optimization were to maximize construction safety and minimize the travel costs of resources on site, while satisfying all the practical layout constraints. NSGA II is selected for multi-objective optimization as it can readily depict the tradeoff between these two conflicting objectives.

The benchmark problem is studied under a set of scenarios that involve doubling the value for each optimization parameter, while fixing the values for all other parameters, in order to evaluate the effect of the parameter on optimization. For example, the population size was initially set to approximately the number of decision variables (20) and then doubled from one scenario to the next, while fixing the values of other parameters. The results for each set of scenarios are then plotted in order to graphically compare them and identify the effect of altering the parameter's value on the performance the algorithm.

OPTIMIZATION MODEL DEVELOPMENT

A model for the aforementioned benchmark problem was developed using an MS Visual C++ compilation of the Non-dominated Sorting Genetic Algorithm II (NSGA II, Deb et al. 2000). In that compilation, NSGA II is the main driving engine for optimizing the two main objectives; maximizing construction safety and minimizing travel cost of resources, simultaneously. It is designed to search for and identify optimal solution(s) for all the decision variables; which in this case are the location coordinates of the temporary facilities. The input and output data is stored in a relational database, similar to the one described by Khalafallah (2006), in order to facilitate handling the results and controlling the optimization parameters.

POPULATION SIZE EFFECT

A first set of scenarios were designed in order to assess the behavior of the model under the alteration of the population size, as shown in Figure 1. In these scenarios, all the optimization parameters were set to fixed values except for the population size which was doubled from one scenario relative to the next. The parameters that were fixed are the number of generations (100 generations), the cross-over probability (0.5), the type of cross-over (multiple point), and the mutation probability (0.001). In the initial run, the population size was set to a value of 20 individuals (which is the approximate number of decision variables in this optimization problem) and then doubled with each run (50, 100, 200), as shown in Figure 1.

As shown in Figure 1, doubling the population size helped significantly in improving the performance of the algorithm. The improvements were depicted by better-generated Pareto-optimal fronts (solution curves) that dominate the old fronts that represent lower number of generations. However, doubling the population size from 100 to 200 individuals did not result in a significant improvement in the results. It is also worth mentioning that an uninformed specification of optimization parameters could result in a very poor performance for the algorithm due to generating an initial population that is not in the ballpark of good solutions.

The above observation indicates that selecting the population size in the range between approximately 5 – 10 times the number of decision variables might be the best specification for the population size parameter. Setting that parameter outside the aforementioned range caused an inefficient algorithm performance.

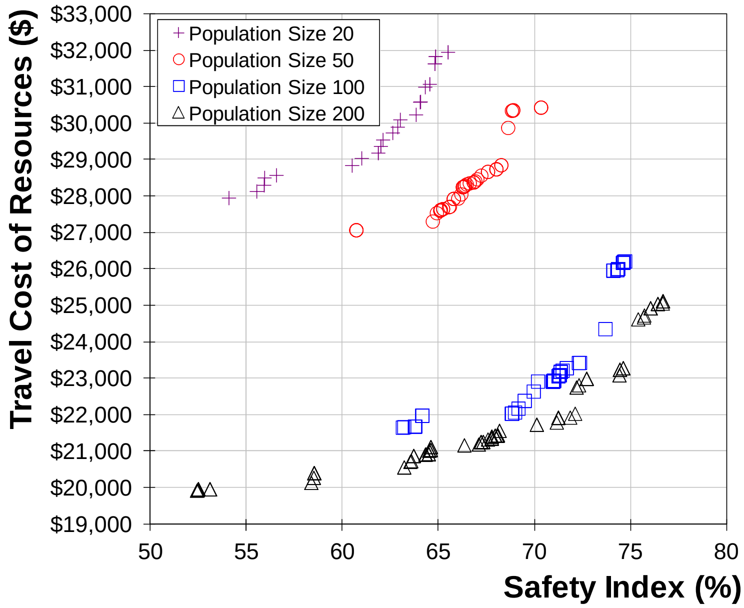

Figure 1. Effect of Doubling the Population Size on the Pareto-optimal Front

NUMBER OF GENERATIONS EFFECT

Several runs were performed, in a second set of scenarios, in order to assess the behavior of the model with the alteration of the number of generations. As shown in Figure 2, the number of generations was doubled in a series of successive runs, starting with 50 generations. The results showed that 200 generations (10 times the number of decision variables) were sufficient to achieve a reasonable Pareto-optimal front. This indicates that setting the number of generations to approximately 10 times the number of decision variables should improve the efficiency of optimization.

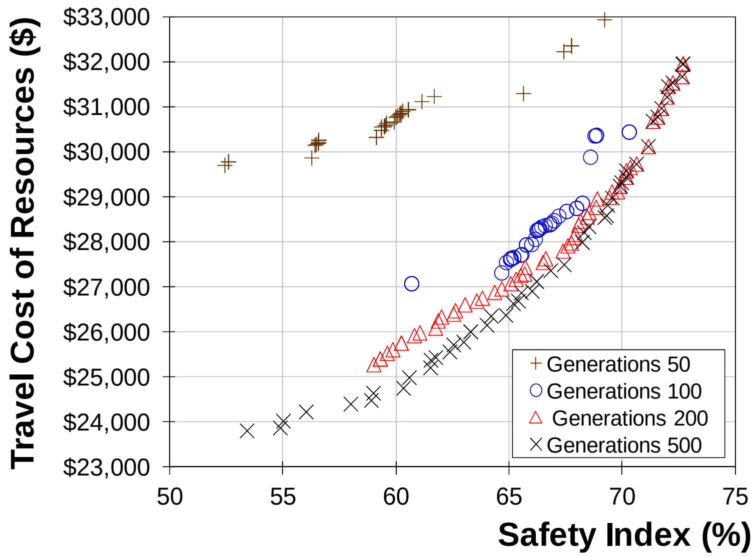

Figure 2. Effect of Doubling the Number of Generations

It is also worth mentioning that doubling the number of generations from 200 to 500 resulted in a better Pareto-optimal front and a better distribution range for the solutions. However, on the other hand, the computational cost increased linearly with increasing the number of generations. This additional computational cost is compensated for by achieving better solutions that provide less travel cost of resources. Thus, the additional computational cost is justified.

EFFECT OF CROSS-OVER PROBABILITY

In a third set of scenarios, the model was tested under various cross-over probabilities while fixing the population size to 50 individual solutions, the number of generations to 200 generations, the mutation probability to 0.001 and using multiple cross-over points as recommended in literature. Generally, increasing the cross-over probability would help to improve the final solutions and the final Pareto-optimal front as more individuals are permitted to mate and reproduce, which improves the chances of reproducing individual solutions of better qualities. As shown in Figure 3, increasing the cross-over probability over 0.5 did not offer significant improvement in the generated Pareto-optimal fronts.

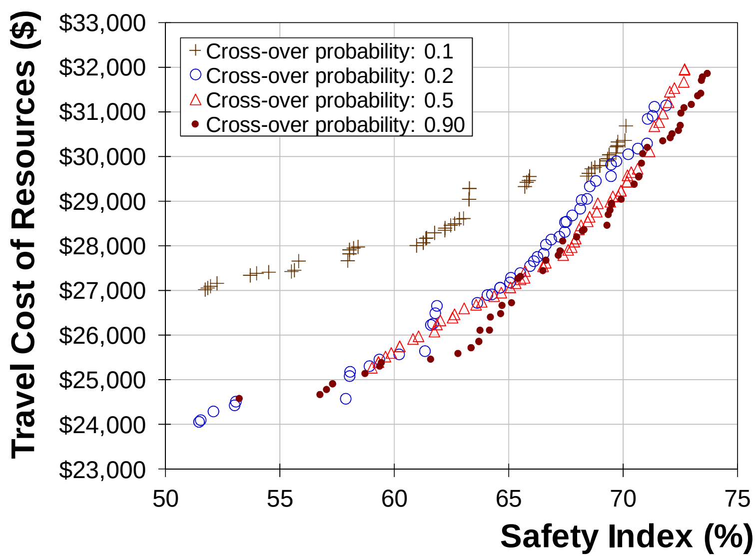

Figure 3. Effect of Altering the Cross-over Probability on the Pareto-optimal Front

CROSS-OVER TYPE & MODEL PERFORMANCE

The model was also tested under two types of cross-over; single point and multiple points, as shown in Figure 4. With population size fixed to 50 individual solutions, 200 generations, 0.5 cross-over probability and 0.001 mutation probability, multiple point cross-over provided a better Pareto-optimal front and a better distribution range for the non-dominated optimal solutions.

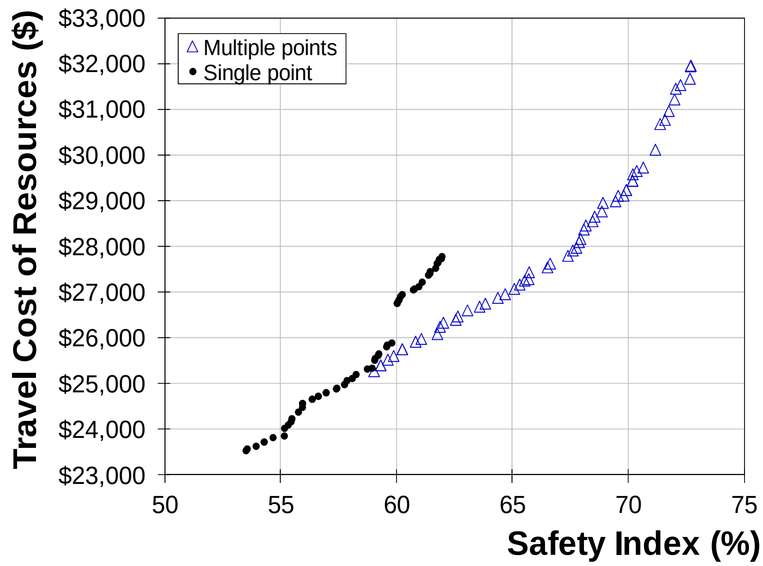

Figure 4. Effect of Single and Multiple Cross-over Operators

BEHAVIOR WITH VARIOUS MUTATION PROBABILITIES

Lastly, the model was tested under various mutation probabilities, while fixing the population size to 50 individuals, the number of generations to 200, and using multiple points cross-over with a probability of 0.5. As shown in Figure 5, doubling the probability did not result in a significant improvement in the Pareto-optimal front. Thus selecting a mutation probability in the range between 0.001 and 0.01 would be appropriate, keeping in mind that it would not affect the solution significantly, as shown in Figure 5.

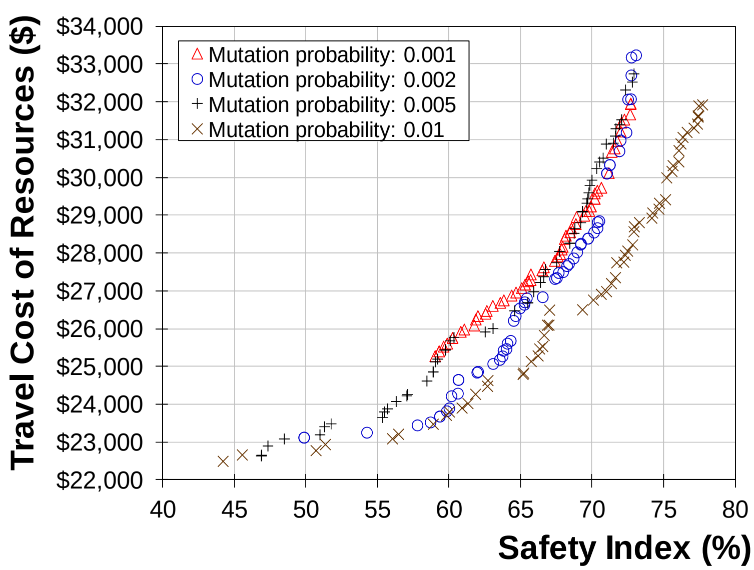

Figure 5. Behavior with the Alteration of the Mutation Probability

Mutation is used in general, to overcome possible convergence of solution to a local minimum. The experiments showed that increasing the mutation rate does not always improve the results, e.g. increasing the mutation rate from 0.002 to 0.005 as illustrated in Figure 5.

SUMMARY AND CONCLUSION

This study investigated the performance of NSGA II under the alteration of optimization parameters, which include the population size, number of generations, type and value of cross-over probability, and mutation likelihood. An uninformed specification of such optimization parameters could result in a very poor performance for the algorithm. The main goal of the study is to provide guidelines on specifying such parameters in order to boost the chances of running an effective algorithm.

A benchmark site layout optimization problem was selected from the literature to study the effects of such alterations of the optimization parameters on the quality of the Pareto-optimal front and distribution range of the non-dominated optimal solutions. The benchmark problem is studied under a set of scenarios that involve doubling the value for each optimization parameter, while fixing the values for all other parameters

The results indicate that selecting (i) a population size in the range between 5 – 10 times the number of decision variables, (ii) a number of generations in the ballpark of 10 times the number of decision variables, (iii) a multiple points crossover with a 0.5 probability, and (iv) a mutation probability in the range between 0.001 – 0.1 might help accelerate the convergence, improve the Pareto-optimal front, and expand the distribution range of the non-dominated optimal solutions.

The study provided insight on the performance of the Non-dominated Sorting Genetic Algorithm (NSGA II) with the variation of optimization parameters. It identifies the expected ranges for these parameters that should help accelerate convergence toward the best Pareto-optimal front and solutions. This should prove useful to researchers and planners, alike, especially those who are interested in construction site layout optimization and closely–related problems.

REFERENCES

Anumba, C., and Bishop, G. (1997). “Importance of Safety Considerations in Site Layout and Organization.” *Canadian J. Civil Eng.*, 24(2), 229 – 236.

Dawood, N., and Marasini, R. (2001). “Stockyard Layout Management for Precast Concrete Products Using Simulation.” *Proceedings of the CIB-W78 International Conference, IT in Construction in Africa*, Mpumalanga, South Africa, 26/1-13.

Deb, K., Agrawal, S., Pratab, A., and Meyarivan, T. (2000). “A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II.” *KanGAL report 200001*, Indian Institute of Technology, Kanpur, India.

Goldberg, D. E. (1989). *“Genetic Algorithms in Search, Optimization and Machine Learning.”* Addison-Wesley, Reading, Massachusetts.

El-Rayes, K., and Khalafallah, A. (2005). "Tradeoff between Safety and Cost in Planning Construction Site Layouts," *Journal of Construction Engineering and Management*, ASCE, 131(11), 1186-1195.

Harmanani, H., Zouein, P., and Hajar, A. (2000). "An Evolutionary Algorithm for Solving the Geometrically Constrained Site Layout Problem." *Proceedings of the 8th ASCE Int. Conf. on Comput. Civ. and Build. Eng. (ICCCBE-VIII),* August 2000, Stanford University, USA, 1442-1449.

Hegazy, T., and Elbeltagi, E. (1999). "EvoSite: An Evolution-Based Model for Site Layout Planning." *Journal of Computing in Civil Engineering,* ASCE, 13 (3), 198-206.

Khalafallah, A. (2006). "*Optimal Construction Site Layout Planning for Airport Expansion Projects.*" PhD. Dissertation, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, USA, October 2006.

Khalafallah, A., and El-Rayes, K. (2006a). "Optimizing Airport Construction Site Layouts to Minimize Wildlife Hazards." *Journal of Management in Engineering*, ASCE, 22(4), 176-185.

Khalafallah, A., and El-Rayes, K. (2006b). "Minimizing Construction-Related Hazards in Airport Expansion Projects." *Journal of Construction Engineering and Management*, ASCE, 132(6), 562-572.

Khalafallah, A., and El-Rayes, K. (2008). "Maximizing Security of Operating Airports during Expansion Projects." *Journal of Construction Engineering and Management in Engineering*, ASCE, 134(1), 40-48.

Mawdesley, M. J., Al-jibouri, S. H., and Yang, H. (2002). "Genetic Algorithms for Construction Site Layout in Project Planning." *Journal of Construction Engineering and Management,* ASCE, 128 (5), 418-426.

Tawfik, H. M., and Fernando, T. (2001a). "A Simulation Environment for Construction Site Planning." *Proceedings of the 5th Int. Conf. Information Visualisation*, London, UK.

Tawfik, H. M. and Fernando, T. (2001b). "A Parallel Genetic Algorithm for Optimising Construction Site Layouts." *Proceedings of IASTED Int. Conf. Applied Simulation and Modelling*, Marbella, Spain.

Tommelein, I. D., Levitt, R. E., and Hayes-Roth, B. (1992). "SightPlan Model for Site Layout." *Journal of Construction Engineering and Management*, ASCE, 118 (4), 749-766.

Yeh (1995). "Construction-site Layout Using Annealed Neural Network." *Journal of Computing in Civil Engineering,* ASCE, 9(3), 201-208.