Selection Pressure and Performance in Spatially Distributed Evolutionary Algorithms

Jayshree Sarma Kenneth De Jong

Computer Science Department
George Mason University
Fairfax, VA 22030
E-mail: jsarma,kdejong@cs.gmu.edu

Abstract

Recent studies of spatially distributed EAs have formally characterized the selection pressure induced by various selection strategies applied to local neighborhoods of various sizes and shapes. These analyses provide us with the ability to predict the expected behavior of the local neighborhood EAs. In this paper we empirically validate these predictions using the domain of function optimization. We demonstrate the various ways selection pressure can be varied in a spatially distributed EA and show that, from a performance point of view, no optimal selection pressure can be defined since it also depends on the fitness landscape of the problem being solved. Our results suggest that it may be possible to adaptively tune selection pressure by varying a single parameter, the neighborhood radius.

1. Introduction

The increasing availability of parallel computing has led to the development of various forms of spatially distributed evolutionary algorithms (EAs). In these systems the population is distributed over a lattice of cells with each individual assigned to a cell. The evolutionary process of selection, recombination and reproduction is done in parallel at each cell by specifying a local neighborhood of interaction. These neighborhoods overlap providing implicit migration of information through the lattice.

Several design decisions are required for these spatially distributed EAs. The topology for the population is usually influenced by the architecture of the parallel machine on which the spatially distributed EA is implemented. The other major decisions involve choosing the size of the neighborhood and the local selection algorithm to be used. These choices were mainly based on the results of the empirical studies done in the past [2, 8, 9, 12].

However, recent studies have produced new analytical tools for spatially distributed EAs [10, 11]. In this paper we evaluate these tools to see whether or not they make useful predictions about the performance of spatially distributed EAs on difficult function optimization problems. The result is a better understanding of both the tools and how to design more effective EAs.

2. Background

A major part in the design of an effective EA is the maintenance of a proper balance between exploration and exploitation. The selection algorithm used is the primary form of exploitation of information. Traditionally, the selection pressure induced by selection algorithms has been formally studied by analyzing the growth curves of the best fitness class in an infinite population model in which no genetic operators are active. Such growth curves are approximated by a logistic equation of the form:

$$P_{b,t} = \frac{1}{1 + (\frac{1}{P_{b,0}} - 1)e^{-at}}$$

where a is the growth coefficient (which controls the growth rate), $P_{b,t}$ is the proportion of the best fitness class in the population at time t, and $P_{b,0}$ is the initial proportion of the best fitness class.

For centralized sequential EAs, the selection pressure induced by a selection algorithm is well understood theoretically [7, 1]. What has been recently shown is that for spatially distributed EAs the *emergent* global selection pressure induced by selection methods operating in local neighborhoods is qualitatively similar to but quantitatively weaker than their centralized sequential counterparts [5]. As a consequence, much of our understanding of the selection algorithms in sequential EAs can also be applied to spatially distributed EAs.

The emergent global selection pressure in a spatially distributed EA is quantitatively weaker than the centralized sequential EA mainly due to the time taken to propagate the information through the different neighborhoods in the lattice. Sarma and De Jong [10, 11] show that the propagation delays are directly related to the ratio of the neighborhood radius to the population topology radius when the radii are calculated using a distance measure adapted from spatial analysis techniques:

$$rad = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 + \sum_{i=1}^{n}(y_i - \bar{y})^2}{n}}$$

where $(\bar{x}, \bar{y})$ is the central cell of a neighborhood of size n.

Furthermore, for a given selection scheme, there is an inverse exponential relationship between this ratio and the growth coefficient a of the logistic equation describing the selection pressure. The result is an insightful analysis of the selection pressure resulting from choices of selection method and neighborhood topologies as illustrated in Figure 1.

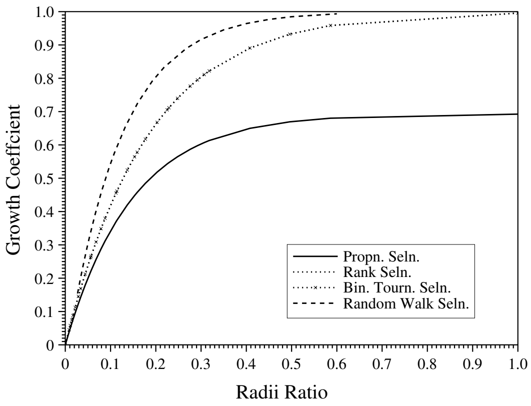

Figure 1. Growth curve coefficients for fitness proportional, linear ranking, binary tournament, and random walk (path elitism) selections as a function of the radii ratio.

These models, in turn, allow us to make quantitative predictions concerning the selection pressure produced by particular choices of selection method and neighborhood topology. For example, by reading horizontally across the graph in Figure 1, we are able to determine which combinations of selection methods and neighborhood topologies produce equivalent selection pressure. Similarly, by reading vertically one can determine the quantitative change in global selection pressure effected by changing the local selection method for a given neighborhood topology.

This ability, combined with the fact that the emergent selection pressures induced by the various local selection methods is qualitatively similar to the well-understood sequential versions, suggests we can also use these models to make predictions about the actual performance of real EAs on difficult function optimization problems. We explore these ideas in the remaining sections.

3. Experimental Setup

The first step is to choose a set of test functions to be used for evaluating the predicted and actual behavior of various spatially distributed EAs. Design of a suitable set of test problems is an extensive research issue in itself [3, 13]. We decided initially to select a fairly standard subset from the EA literature: the Rosenbrock, Schwefel, Rastrigin, and Griewangk functions, and boolean satisfiability problems like the N-Peak and Hamiltonian circuit problems from [6]. We are also looking into the possibility of using some of the recently developed test problem generators [4].

The spatial topology used throughout our experiments is a two-dimensional torus, the most common topology in practice. Each individual is assigned to a cell on this torus. So, for example, a 32×32 grid implies a population size of 1024. Each individual interacts locally with individuals in neighboring cells identified as those reachable from the central cell via a fixed number of steps. This, in turn, specifies the radius of the neighborhood and it is the ratio of this radius to that of the entire grid that affects selection pressure (as discussed earlier).

Unless explicitly stated otherwise, we used two-point crossover with a crossover rate of 1.0 and bit flipping mutation at a rate of 0.001. Evolution proceeds in parallel in each cell by selecting parents from the local neighborhood to produce an offspring to replace the current individual in each cell. In the experiments described the replacement strategies were not elitist, although this is a frequent addition in the literature. It should also be noted that the particular form of random walk selection used (path elitism) does implicitly result in an elitist replacement policy.

Since our goal is to predict and analyze performance, standard best-so-far curves are used to present the search effectiveness as a function of time (measured in generations). In addition, we found two other measures useful: allele loss and inbreeding coefficients.

Allele loss measures the number of gene positions that are completely fixed, i.e., all population members have the same allele values. The (lack of) diversity of the alleles is a strong indication of whether there is any possibility for improvement via the search process.

Inbreeding coefficients are a measure borrowed from population genetics and used by EA researchers to study the similarity of the two individuals used to produce an offspring. The main effect of inbreeding in natural populations is to increase the proportion of homozygous genotypes at the expense of the proportion of heterozygous genotypes. The effect of inbreeding is measured in terms of the reduction in the heterozygosity that occurs in the population. The inbreeding coefficient F is defined as:

$$F = \frac{H_0 - H}{H_0}$$

where H is the current proportion of heterozygous genotypes and H_0 is the initial heterozygosity and is given by $H_0 = 2p(1-p)$ and p is the frequency of the 0 allele. The frequency is measured for each location in the genotype. In the initial population, the H_0 frequency is around 0.5 for each location. As the generations increase and the diversity reduces, heterozygosity reduces, i.e., the individuals are fairly similar to each other and H will be closer to 0. Hence the higher the value of F, the higher the similarity of the two individuals that are being mated, and the higher the similarity of the offspring produced. Although similar in spirit to allele loss, the inbreeding coefficient takes into account the effects of selection.

The stopping criterion for all experiments is 100 generations and the results are averaged over 100 runs.

4. Experimental Results and Discussion

Because of the concern over communication costs in finely grained parallel architectures, neighborhood sizes are typically chosen to be as small as possible while still preserving enough diversity in the local neighborhoods. With this in mind, for the experiments reported here, we chose neighborhoods which produced a radii ratio of 0.15. Using Figure 1 we can use the specified growth coefficients to generate the predicted growth curves for four selection algorithms as shown in Figure 2. Notice how much stronger the selection pressure predicted for random walk is than the others, while ranking and binary tournament selection result in equal pressure, and fitness proportional selection the weakest. This implies a much more exploitative and focused search behavior with random walk selection, with fitness proportional selection the most explorative.

Figure 3 shows a summary of the best-so-far performance for the Schwefel test function, and which is typical of the behavior observed on the others. Random walk shows early aggressiveness, but (frequently) plateaus early at a suboptimal level. Ranking and binary tournament selection produce nearly identical behavior and perform the best on this function. Fitness proportional selection makes (exasperatingly) slow but steady progress, and in this case is terminated before the optimum is found.

As is the case with centralized EAs, the external search behavior is a direct consequence of the effects that selection is having on the internal gene pool. This is illustrated in Figures 4 and 5. The higher selection pressures result in higher allele loss rates and higher inbreeding coefficients, producing a much narrower search focus and, if the selection pressure is too high, premature convergence to a suboptimal point.

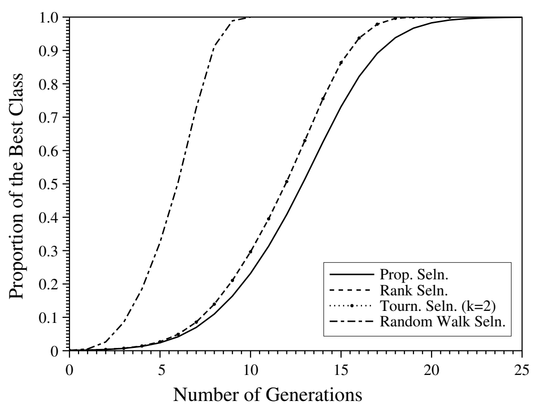

Figure 2. Predicted growth curves for fitness proportional, linear ranking, binary tournament, and random walk (path elitism) selections for radii ratio of 0.15.

One of the reasons that random walk selection produces higher selection pressure than the others is that it involves an implicit form of elitism. Since the central cell is always included in the path of the random walk, if the individual in the central cell is the best in the neighborhood, it always gets selected.

It is not difficult to modify the other selection methods to include elitism in the replacement strategy: an individual is replaced by the offspring only if the offspring has higher fitness. In this case adding elitism improves the performance in the case of fitness proportional selection considerably and can result in modest improvements for ranking and binary tournament selection [5].

The Hamiltonian Circuit and NP-peak problems, however, serve as reminders that there is no globally optimal level of selection pressure. Figure 6 illustrates this for Hamiltonian Circuit problems of size 10. In this case random walk selection is consistently outperformed by the weaker selection methods.

Table 1 shows clearly that the higher selection pressure of random walk results in (premature) convergence by 30 generations, while the weaker selection methods find the optimum much more reliably. Moreover, as the size of these problems increase, the need for weaker selection pressure also increases.

One of the nice features of the spatially distributed EAs is that selection pressure can be controlled by both the choice of selection method and the neighborhood radius. This gives us a finer degree of control over selection pressure than with centralized EAs. The models in Figure 1 provide us with the information necessary to make appropriate choices. For

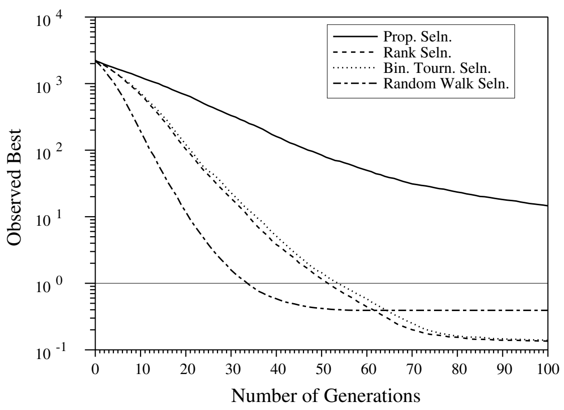

Figure 3. Best-so-far Performance measure for Schwefel function.

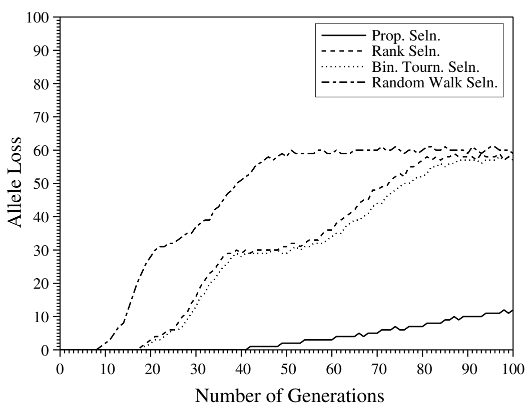

Figure 4. Allele loss statistics for Schwefel function.

Table 1. Number of runs in which the maximum was found for function HC10

No. of Gens.	Fit. Prop	Lin. Rank	Bin. Tourn.	Random Walk
30	0	45	43	23
100	63	63	69	23

example, in many finely grained architectures, the communication overhead of ranking selection is much higher than that of binary tournament selection. The models indicate that there is no reason to even consider ranking in such cases.

A second use of these models is to effect changes in selection pressure via changes in neighborhood size. As we saw in the previous example, the random walk selection with a length of 5 was too strong. So, for example, one can reduce the length of the random walk from 5 to 2 to reduce the neighborhood size and radius. This leads to a radii ratio of 0.096 which has a corresponding growth coefficient slightly closer to that of binary tournament (and rank) selection. Reducing the radii ratio in the case of random walk selection results in much better search behavior (as well as lower communication costs).

5. Conclusions and Future Work

The formal models of selection pressure for the commonly used selection algorithms in spatially distributed EAs have been evaluated for their effectiveness in making predictions about the behavior of real EAs on difficult function optimization problems. The models were shown to make accurate quantitative and qualitative predictions. In addition the models were shown to be useful in making important design decisions involving selection methods and neighborhood topologies.

As is the case with centralized EAs, it was shown that there is no globally optimal level of selection pressure, but varies as a function of the fitness landscape. It was noted, however, that the selection pressure in these spatially distributed models can be more finely tuned by adjusting the neighborhood radius. This suggests a simple mechanism for adapting the selection pressure during the search process in order to adjust the balance between exploration and exploitation during the search process. We are currently exploring this idea.

References

[1] T. Bäck. *Evolutionary Algorithms in Theory and Practice*. Oxford University Press, New York, 1995.

[2] R. J. Collins and D. R. Jefferson. Selection in massively parallel genetic algorithms. In *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 249–256, San Diego, CA, 1991. Morgan Kaufmann.

[3] L. Davis. Bit climbing, representational bias, and test suite desgin. In *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 18–23, San Diego, CA, 1991. Morgan Kaufmann.

[4] K. De Jong, Mitchell A. Potter, and W. M. Spears. Using problem generators to explore the effects of epistasis. In *Proceedings of the Seventh International Con-*

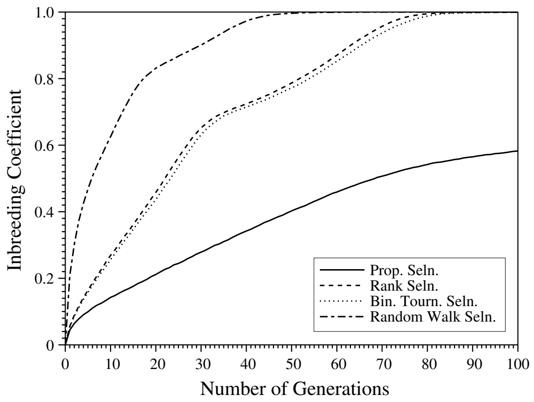

Figure 5. Inbreeding coefficient for Schwefel function.

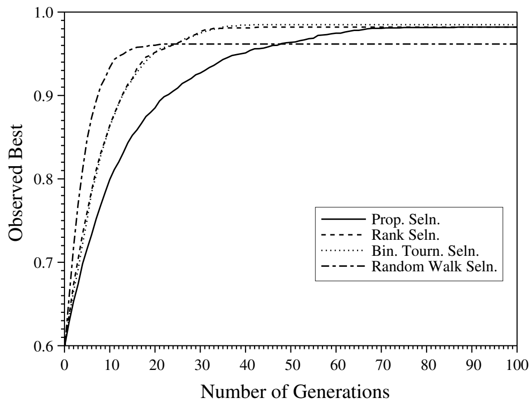

Figure 6. Best-so-far Performance measure for HC10 function.

ference on Genetic Algorithms, pages 338–345, East Lansing, MI, 1997. Morgan Kaufmann.

[5] K. De Jong and Jayshree Sarma. On decentralizing selection algorithms. In *Proceedings of the Sixth International Conference on Genetic Algorithms*, pages 17–23, Pittsburgh, PA, 1995. Morgan Kaufmann.

[6] K. A. De Jong and W. M. Spears. Using genetic algorithms to solve NP-Complete problems. In *Proceedings of the Third International Conference on Genetic Algorithms*, pages 124–132, Fairfax, VA, 1989. Morgan Kaufmann.

[7] D. E. Goldberg and K. Deb. A comparative analysis of selection schemes used in genetic algorithms. In *Foundations of Genetic Algorithms*, pages 69–93, San Mateo, CA, 1991. Morgan Kaufmann.

[8] M. Gorges-Schleuter. ASPARGOS an asynchronous parallel genetic optimization strategy. In *Proceedings of the Third International Conference on Genetic Algorithms*, pages 422–427, Fairfax, VA, 1989. Morgan Kaufman.

[9] J. McInerney. *Biologically Influenced Algorithms and Parallelism in Non–linear Optimization*. PhD thesis, University of California, San Diego, 1992.

[10] Jayshree Sarma and K. De Jong. An analysis of the effects of neighborhood size and shape on local selection algorithms. In *Proceedings Parallel Problem Solving from Nature - PPSN IV*, pages 236–244, Berlin, Germany, 1996. Springer-Verlag.

[11] Jayshree Sarma and K. De Jong. An analysis of local selection algorithms in a spatially structured evolutionary algorithm. In *Proceedings of the Seventh International Conference on Genetic Algorithms*, pages 181–186, East Lansing, MI, 1997. Morgan Kaufmann.

[12] P. Spiessens and B. Manderick. A massively parallel genetic algorithm: Implementation and first analysis. In *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 279–287, San Diego, CA, 1991. Morgan Kaufmann.

[13] D. Whitley, K. Mathias, S. Rana, and J. Dzubera. Building better test functions. In *Proceedings of the Sixth International Conference on Genetic Algorithms*, pages 239–246, Pittsburgh, PA, 1995. Morgan Kaufmann.