Growth Curves and Takeover Time in Distributed Evolutionary Algorithms

Enrique Alba and Gabriel Luque

Departamento de Lenguajes y Ciencias de la Computación
E.T.S.I. Informática, Campus Teatinos, 29071 Málaga (España)
`eat@lcc.uma.es, gabriel@lcc.uma.es`

Abstract. This paper presents a study of different models for the growth curves and takeover time in a distributed EA (dEA). The calculation of the takeover time and the dynamical growth curves is a common analytical approach to measure the selection pressure of an EA. This work is a first step to mathematically unify and describe the roles of the migration rate and the migration frequency in the selection pressure induced by the dynamics of dEAs. In order to achieve these goals we evaluate the appropriateness of the well-known logistic model and of a hypergraph model for dEAs. After that, we propose a corrected hypergraph model and two new models based in an extension of the logistic one. Our results show that accurate models for growth curves can be defined for dEAs, and explain analytically the migration rate and frequency effects.

1 Introduction

The increasing availability of clusters of machines has endorsed the fast development of parallel EAs (PEAs) [1]. Most popular PEAs split the whole population in separate subpopulations that are dealt with independently (islands). A sparse exchange of information among the component subalgorithms leads to a whole new class of algorithms that do not only perform faster (more steps by unit time), but that often lead to superior numerical performance [2, 3].

In the core of these parallel EAs we can find a spatially structured distributed EA (dEA) that has been implemented (usually) in parallel on a cluster of machines interconnected by a communication network. Many interesting parallel issues can be defined and studied in PEAs, but in this work we are interested in the distributed algorithm model using multiple populations, that is really the responsible of the search features. In this article we concentrate on the dynamics of the distributed EA, in particular in developing a mathematical description for the takeover time, i.e., the time for the best solution to completely fill up all the subpopulations of the dEA. We first will propose and analyze several models for the induced growth curves, and then address the calculation of the takeover time. In this work, only tournament selection is considered. Also, since we only focus on selection, we expect an easy extension of the results to many other EAs.

In order to design a dEA we must take several decisions. Among them, a chief decision is to determine the migration policy: topology, migration rate (number

of individuals that undergo migration in every exchange), migration frequency (number of steps in every subpopulation between two successive exchanges), and the selection/replacement of the migrants. In general, decisions on these choices are made by experimental studies. Therefore, it would be interesting if we could provide an analytical basis for such decisions.

Several works have studied the takeover time and growth curves for other classes of structured EAs [4–10]. In general, these works are oriented to study cellular EAs (with the important exception of the Sprave's one [10]) and it really exists a gap in the studies about dEAs from which something could be gained for other researchers or applications.

In the present work we focus on the influence of migration rate and migration frequency in the takeover time and in the growth curves. To achieve this goal we fix the topology to a simple and, at the same time, widely used one: a static directional ring. We also preset the policies of selection/replacement of the migrants. The emigrants are selected by binary tournament while the immigrants are included in the target population only if they are fitter than the worst-existing solution. In our analysis we will use the binary tournament mechanism with an elitist replacement (concretely, we use a $(\mu + \mu)$-dEA). We defer for a future work the theoretical analysis on other topologies and selection methods. Our contribution is to put to work the logistic model [5] and the hypergraph models [10], since they have never been tested and compared in practice (to the best of our knowledge). Then, we will propose three new mathematical models for the dynamics of selection: a corrected hypergraph plus two extended logistic models. Our aim is to improve on the accuracy (low error) of the initially tested models and consequently compute takeover times.

This paper is organized as follows. Section 2 is an introduction containing some preliminary background about previous works. Section 3 studies the effects of the migration frequency in the resulting growth curves; just after that, we extend the analysis by considering also the migration rate (Section 4). In Section 5, we analyze the predicted takeover times provided by the models. In the last section we summarize the conclusions and give some hints on the future work.

2 Performance of the Existing Theoretical Models

A common analytical approach to study the selection pressure of an EA is to characterize its takeover time [11], i.e., the number of generations it takes for the best individual in the initial population to fill the entire population under selection only. The growth curves are another important issue to analyze the dynamics of the dEAs. These growth curves are functions that associate the number of generations of the algorithm with the proportion of the best individual in the whole population. In this section we describe briefly the main models found in the literature defining the behavior of structured population EAs.

2.1 The Logistic Model

Let us begin by discussing the work of Sarma and De Jong (1997) for cellular EAs. In that work, they performed a detailed empirical analysis of the effects

of the neighborhood size and shape for several local selection algorithms. They proposed a simple quantitative model for cellular EAs based in the logistic family of curves already known to work for panmictic EAs [11]. In summary, the proposed equation is (1):

$$P(t) = \frac{1}{1 + \left(\frac{1}{P(0)} - 1\right)e^{-at}} \ .$$

(1)

where a is a growth coefficient and $P(t)$ is the proportion of the best individual in the population at time step t. This model threw accurate results for synchronous updates of square shaped cellular EAs. Recently, for the asynchronous case, improved models has been proposed in [9] not following a logistic growth. Anyway, using a logistic curve represents an interesting precedent that however should be validated for dEAs. In brief, we will do so in this article.

2.2 The Hypergraph Model

Sprave (1999) has proposed a unified description for any non-panmictic population structured EA, that could even end in an accurate model for panmictic populations (since they can be considered as fully connected structured populations). He modelled the population structure by means of *hypergraphs*. A hypergraph is an extension of a canonical graph. The basic idea of a hypergraph is the generalization of edges from pairs of vertices to arbitrary subsets of vertices.

He developed a method to estimate growth curves and takeover times. This method is based on the calculation of the diameter of the actual population structure and on the probability distribution induced by the selection operator. In fact, Chakraborty et al. (1997) previously calculated the success probabilities for the most common selection operators (p_{select}), what represents an interesting complement for putting hypergraphs to work in practice. A complete description of the hypergraph model can be found in [10].

2.3 Other Models

Although the logistic model is relatively well known, and hypergraphs could play an important role in the field, they are not the only existing models that can inspire or influence the present study. Gorges-Schleuter (1999) also accomplished a theoretical study about takeover times for a cellular ES algorithm. In her analysis, she studied the propagation of information over time through the entire population. She finally obtained a linear model for a ring population structure and a quadratic model for a torus population structure.

In a different work, Rudolph (2000) carried out a theoretical study on the takeover time in populations with array and ring topologies. He derived lower bounds for arbitrary connected neighborhood structures, lower and upper bounds for array-like structures, and an exact closed form expression for a ring topology.

Later, Cantú-Paz (2000) studied the takeover time in dEAs where the migration occurs in every iteration, which is the lower bound of the migration frequency value. He generalized the panmictic model presented in [11] by adding

a policy-dependent term. That term represents the effects of the policy used to select migrants and the individuals that they replace at the receiving island.

Giacobini et al. (2003) studied the takeover time in cellular EAs that use asynchronous cell update policies. The authors presented quantitative models for the takeover time in asynchronous cellular EAs with a ring topology.

In the present work, we focus on the described models: logistic and hypergraphs. The first one (logistic) is based on biological processes and it is well known in the case of cellular EAs, the other type of structured EAs. The second model (hypergraphs) posses a unique unified-like feature for all non-panmictic algorithms. We do not use the results of the other works directly since either they are linked to specialized algorithms or have a different focus (selection policy).

3   Effects of the Migration Frequency

In this section we analyze the effects of the migration frequency over the growth of the best individual in dEAs. In this aim we begin by performing an experimental set of tests for several migration frequencies. First, we describe the parameters used in these experiments, and later we analyze the obtained results.

3.1   Parameters

We have performed several experiments with different values of migration frequencies: 1, 2, 4, 8, 16, 32, and 64 generations. In general, researchers use frequencies in this range. Notice that a low frequency value (e.g., 1) means high coupling, while a high value (e.g., 64) means loose coupling (large gap). The rest of parameters are kept constant. In the experiments, we use a $(\mu + \mu)$-dEA with 8 islands (512 individuals per island), binary tournament selection, a static ring topology, and a preset moderate migration rate (8 individuals chosen by binary tournament are exchanged at each migration step). For all tests, we use randomly generated populations with individual fitness between 0 and 1023. Then we introduce a single best individual (fitness = 1024) in a randomly selected island. In hypergraphs we have used an expected level of accuracy of $\varepsilon = 2.5 \cdot 10^{-4}$. For the actual curves we have performed 100 independent runs.

In order to compare the accuracy of the models we proceeded to calculate the mean square error (2) between the actual values and the theoretically predicted ones (where k is the number of points of the predicted curve). The MSE gives the error for an experiment. But we also define a metric that summarizes the error for all experiments, thus allowing to perform a quantitative comparison between the different models easily. We studied several statistical values (mean, median, standard deviation, etc.) but finally we decided to use the $\|\cdot\|_1$ (3) that represents the area below the MSE curve (E is the number of experiments).

$$MSE(model) = \frac{1}{k} \sum_{i=1}^{k} (model_i - experimental_i)^2 \ . \tag{2}$$

$$\|model\|_1 = \sum_{i=1}^{E} |MSE(model)| \ . \tag{3}$$

3.2 Analysis of the Results

Now, we analyze the curves that have been obtained in the experiments. Figure 1 contains the lines of the actual takeover time for different migration frequencies. This figure shows that, for low frequency values, the dEA resembles the panmictic case [11]. This is common sense, since there exists high interaction among the subalgorithms. However, for higher frequency values (uncoupled search), the observed behavior is different: the subpopulations in the islands having the best solution converge quickly, and then the global convergence of the algorithm stops progressing (flat lines) until a migration of the best individual happens to take place. The observed effect is that of a *stairs*-like curve. The time span of each step in such a curve is governed by the migration frequency. The higher the migration frequency value, the largest the span of the step.
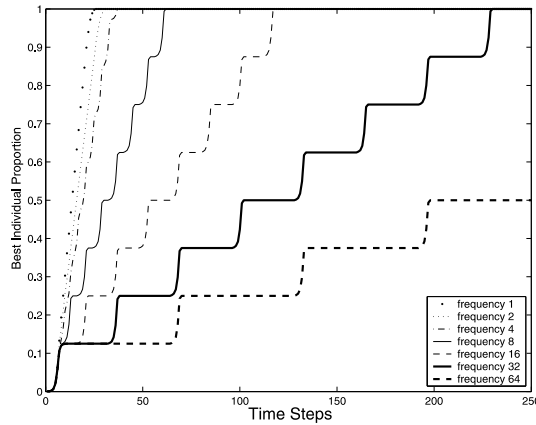

Fig. 1. Actual growth curves for several migration frequencies (100 independent runs)

Once we have understood the basic regularities behind, our goal is to find a mathematical model that allows an accurate fitting to all these curves. We begin this task by trying to use the mentioned logistic and hypergraph models. Let us first address the logistic case:

$$P(t) = \frac{1}{1 + a \cdot e^{-b \cdot t}} \ .$$

(4)

To strictly adhere to the original work of Sarma and De Jong for cellular EAs, the a parameter should be defined as a constant value ($a = \frac{1}{P(0)} - 1$). We call this model LOG1, and we plot its accuracy in Fig. 2a. We can quickly arrive to the conclusion that for low values of migration frequency (panmictic-like scenario), the error is small, what means good news for a logistic fitting. However, as the interaction among the islands decreases, it turns to be very inaccurate. One could think that trying a fitting with the a parameter fixed is provoking such an inaccuracy, and this is why we propose a new variant of the logistic model called LOG2 (see Fig. 2b). In this case, we consider a and b as free variables (in the previous model LOG1 a was a constant parameter). LOG2

allows a fitting with a smaller error than LOG1, but it still seems harmful since the actual steps are ignored both in LOG1 and LOG2.

Then, a clear conclusion is that the basic logistic model, even when enhanced, cannot be used for distributed EAs, as the existing literature also claim for most non canonical cellular EAs [6–9].
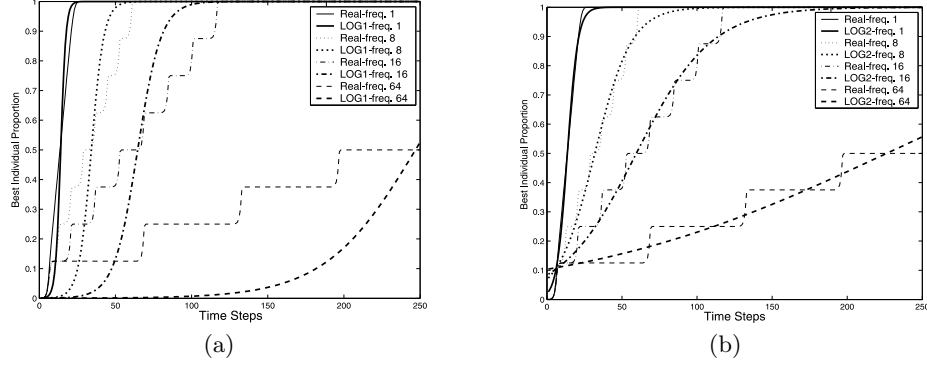

Fig. 2. Comparison between actual/predicted values with LOG1 (a) and LOG2 (b)

Therefore, we now turn to consider the hypergraph approach. In fact we present two variants of hypergraphs: the one in which p_{select} (5) accounts only for the probability of selection (HYP1), and the one where this probability (6) accounts both for selection and for replacement (HYP2). We introduce such distinction since in the seminal work [12] this second choice (combining selection and replacement within a probability) is said to be more exact.

$$p_{select1}(i, N) = 2 \cdot \frac{i}{N} - \left(\frac{i}{N}\right)^2 \ .$$

(5)

$$p_{select2}(i, N) = \frac{i}{N} + \left(1 - \frac{i}{N}\right) p_{select1}(i, N) \ .$$

(6)

where N is the population size and i denotes the total number of best individuals in the population.

When the hypergraph model is put to work, we can notice a clear improvement over the logistic models, obtaining an almost perfect curve fitting. As expected, HYP1 (Fig. 3a) generates a slightly worse fit than the HYP2 (Fig. 3b), because HYP1 is not accounting for the replacement effects.

To end this section we introduce a more accurate extension of the logistic model (N is the number of islands in the dEA):

$$P(t) = \sum_{i=1}^{i=N} \frac{1/N}{1 + a \cdot e^{-b \cdot (t - freq \cdot i)}} \ .$$

(7)

This expression is an extension of the logistic model based in the idea that each island converges according to a logistic model, and that the entire population grows up as a sum of the growth of each component island. If the subpopulations are in turn structured in some way this assumption could not hold, but
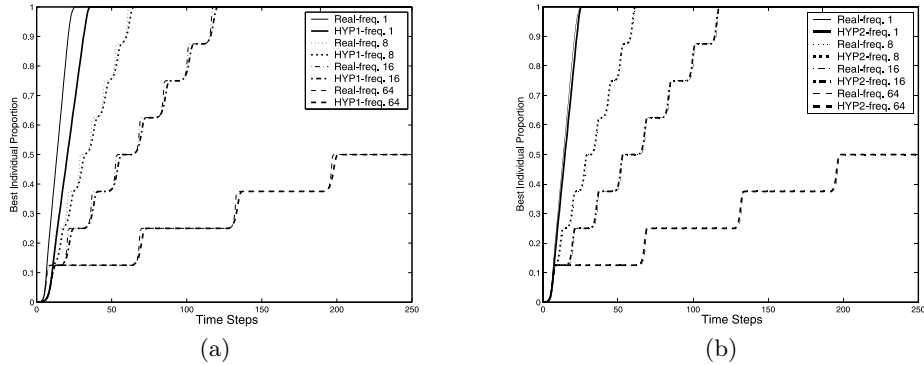
Fig. 3. Comparison between actual/predicted values with HYP1 (a) and HYP2 (b)

much must be said on this special subject (e.g., on distributed cellular EAs [3]), and thus it is left for a future work. To find the takeover time, we simply iterate the model until it reaches 1. We should notice that since it is an extension of the logistic approach, two variants could be also defined as we did before with LOG1 and LOG2. The first (SUM1) in which a is constant ($a = \frac{1}{P(0)/N} - 1$), and the second (SUM2) where a and b are adjustable parameters.
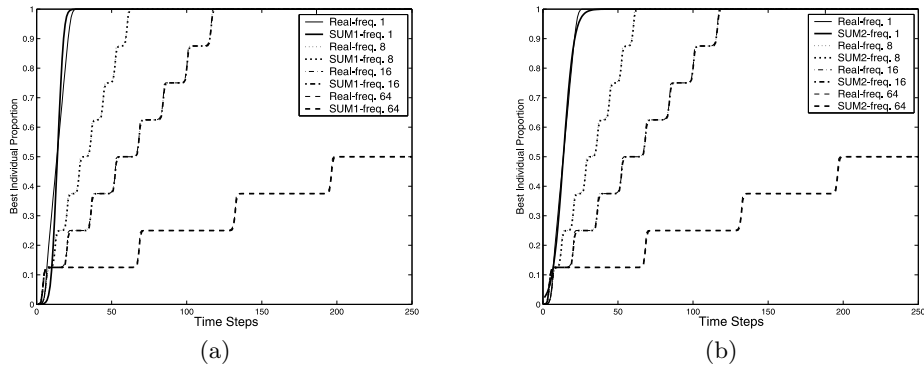

Fig. 4. Comparison between actual/predicted values with SUM1 (a) and SUM2 (b)

In figure 4 we plot the behavior of such a model. For low values of the migration frequency, SUM1 shows a less accurate behavior with respect to SUM2, but they two are equally or more precise than any other existing model, in particular with respect to the basic logistic and hypergraph variants. For the rest of migration frequencies, the SUMx models outperform the rest.

We conclude this section with a summary of the results. We have fitted the actual growth curves of a dEA with six different theoretical models. In Fig. 5 we graph together the mean square error and the $\| \cdot \|_1$ of that error for all the models under different migration frequencies. We can notice that in the case of low values of the migration frequency, most of models obtain a large error (with the exceptions of LOG2 and SUM2 models). In general, the second variant of each model is always better than the first one. The LOGx models show a stable

behavior for all frequencies, but, while LOG1 is always very inaccurate (it is the worst model), the LOG2 performs well, and it is only worse than SUM2 model. The SUMx and HYPx models reduce their errors as the migration frequency enlarges, although the SUMx models are always more accurate than the HYPx ones.
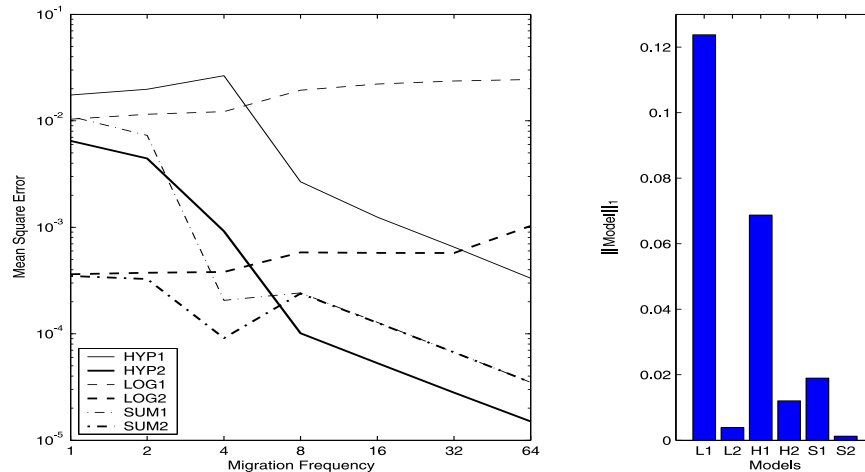

Fig. 5. Error (MSE) and $\| \cdot \|_1$ between actual and predicted values for all the models

4 Effects of the Migration Rate

In the previous section, we analyzed the effects of the migration frequency over the growth curves of a dEA. However, the number of individuals undergoing migration was fixed at a given value (rate = 8). Now, in this section we answer a common sense question: are the results somehow biased by the utilization of a concrete migration rate? To extend the previous study we now also analyze the effects of the migration rate. As we made for the migration policy, we begin by inspecting the induced behavior by different values of the migration rate.

As done before, we first compute the proportion of the best individual in a dEA when utilizing these values of the migration rate: 1, 2, 4, 8, 16, 32, and 64. The rest of the parameters are similar to the presented in Subsection 3.1. We proceed to perform a comparative study with the models discussed before.

In Fig. 6 we plot the way in which the migration rate influences the growth curve (frequency = 1). From this figure we can infer that the value of the migration rate determines the slope of the curve. The reason is that, when the migration rate value is high, the probability of migrating the best individual increases, and then the target island converges faster than if the migration rate were smaller.

Let us now proceed with the fitting of these curves with all the considered mathematical models. In Fig. 7 (left) we show the error for any combined value of the migration frequency and migration rate. To interpret the graph you must
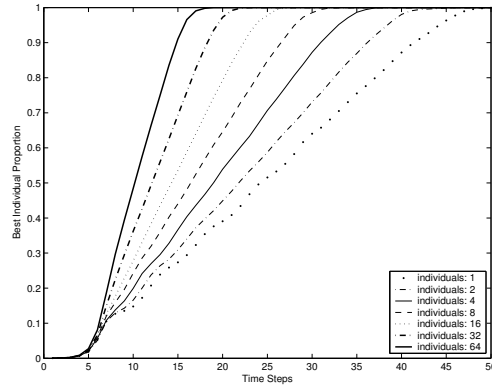
Fig. 6. Actual growth curves for several migration rates (100 independent runs)

notice that the first seven points of a line correspond to the error incurred by the associated predictive model for the seven different values of the migration rate at the same frequency, and that there exists seven groups of such points, one for each migration frequency from 1 to 64 (from left to right in the horizontal axis).
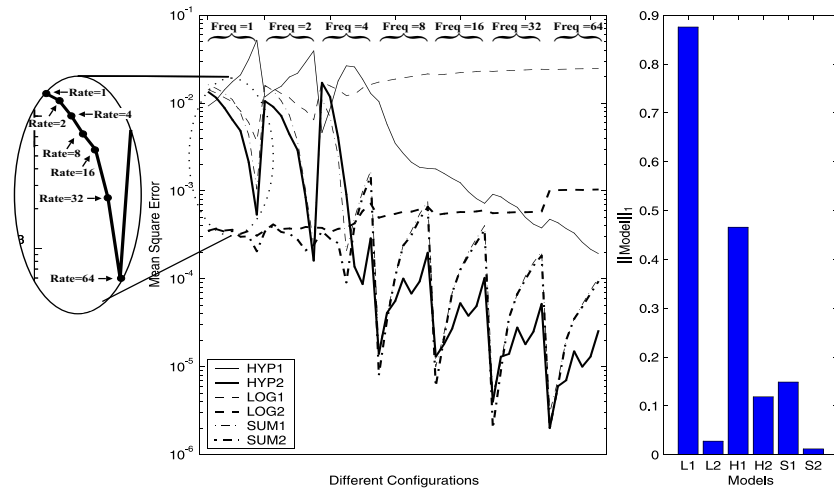

Fig. 7. Error (MSE) and $\| \cdot \|_1$ between actual and predicted growth curve for all the values of migration rate and frequency

We can see in Fig. 7 some behavioral patterns of the models with respect to the final MSE error they exhibit. First, we observe that the logistic behavior is very similar to the showed previously (Fig. 5), i.e., both variants are very stable; LOG1 obtains always larger errors while LOG2 is very accurate for all configurations. Second, the hypergraph model obtains low error for larger frequency values, while their inaccuracy is more evident for smaller values of migration frequency (high coupling). The proposed SUMx model also is somewhat sensitive to low frequency values, but it is quite stable and accurate for larger values of

migration frequency. Both, the HYPx and the SUMx models seem to perform a cycle: reduction/enlargement (respectively) of error as the migration rate enlarges (for any given frequency). The $\| \cdot \|_1$ summarizes quantitatively the MSE results in a single value per model. We observe that the LOG1 model is the most inaccurate one. Also, LOG2 obtains very accurate results and only is worse than the SUM2 model. Although the HYPx models are very accurate for lager frequency values, they show high errors for smaller values, thus making its $\| \cdot \|_1$ value larger than the rest. Clearly, SUM2 obtains the lowest overall error.

5 Takeover Time Analysis

In the previous sections we have studied the effects of the migration frequency and rate over the takeover growth curves. Now, we analyze the effect of these parameters over the takeover times themselves. Figure 8 contains the value of the actual takeover time for different migration frequencies and rates. We can notice that the takeover value increases for higher frequencies and smaller rates. However, the rate effect over the takeover time is smoother than the frequency one.
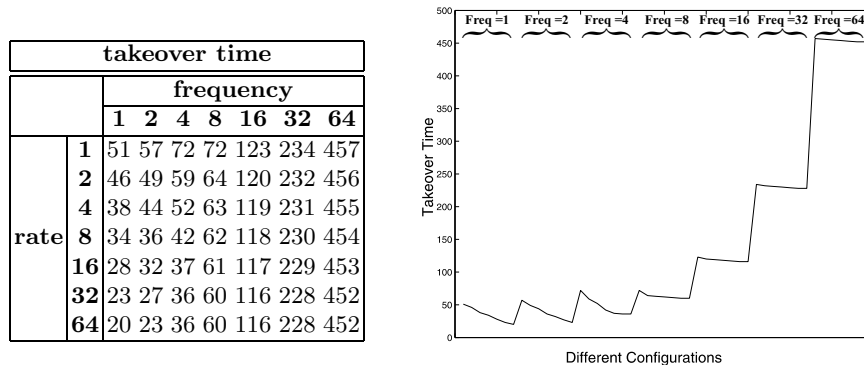
takeover time							
		frequency					
		1	2	4	8	16	32	64
	1	51	57	72	72	123	234	457
	2	46	49	59	64	120	232	456
	4	38	44	52	63	119	231	455
rate	**8**	34	36	42	62	118	230	454
	16	28	32	37	61	117	229	453
	32	23	27	36	60	116	228	452
	64	20	23	36	60	116	228	452

Fig. 8. Actual takeover time values for all configurations (100 independent runs)

Once we have observed the effect of the migration rate and frequency over the takeover time, we analyze the predicted values provided by the models. Figure 9 shows the error of the predicted takeover time with all the models. We can notice that the predictions of the models are very sensitive to low values of migration frequency. However, as such values are increased, the behavior of the models is more stable. The logistic models obtain worse predictions as we enlarge the migration frequency value, and for lager values, they are no longer useful. HYPx and SUMx are quite stable (with small oscillations) and very accurate for lager values of migration. Figure 9 (right) shows the MSE of the takeover time prediction. Specially interesting is the case of LOG2 model, that obtains a very accurate fitting of the growth curves but is quite inaccurate to predict the takeover time. With the exception of the LOG2 model, there is not any significant difference in the MSE value among the models, although the first variants of the models (LOG1, HYP1, and SUM1) obtain a slightly better takeover time value than its corresponding second one.
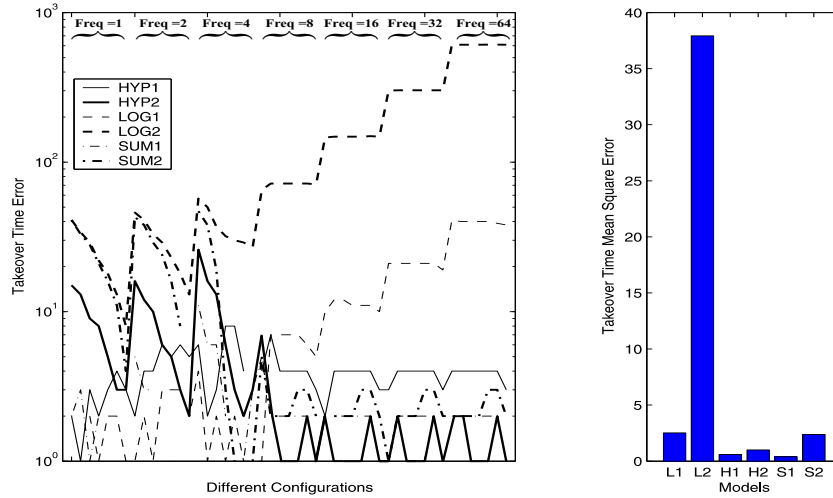
Fig. 9. Error between actual and predicted values of takeover time

We conclude this section by showing a closed equation for the takeover time calculation for the new models presented in this paper: SUMx models. This formula (8) is derived from the growth curve equation (7) of these models:

$$Takeovertime = freq \cdot (N - 1) - \frac{1}{b} \cdot Ln\left(\frac{\varepsilon \cdot N}{a \cdot (1 - \varepsilon \cdot (N - 1))}\right) \quad . \quad (8)$$

where $freq$ is the migration frequency, N is the number of islands and ε is the the expected level of accuracy (a small value near zero).

6 Conclusions

In this paper we have performed an analysis of the growth curves and takeover regimes of distributed evolutionary algorithms. We compared the well-known logistic model, a hypergraph model and a newly proposed model consisting in a sum of logistic definition of the component takeover regimes. A second variant of each model has been also proposed for the shake of accuracy. In this article we have shown how the models appropriately captured the effects of both migration frequency and migration rate.

Although every model has its own advantages, either simplicity (LOGx), extensibility (HYPx), or accuracy (SUMx), SUM2 is the model that better fitting obtained, while its predicted takeover time was of a similar accuracy with respect to the rest. However, much needs still to be said on them, since subtle factors could provoke deviations from the predicted behavior, like it may occur for the takeover regime near to the moment in which one subpopulation is completely filled with the optimum solution.

As a future work we plan to check the results presented in this paper on additional topologies and selection methods. Dealing with bounding cases of

widely spread algorithms is hopefully the right way to improve on our knowledge of the algorithms that most researchers are using in practice.

Acknowledgments

References

1. Alba, E., Tomassini, M.: Parallelism and Evolutionary Algorithms. IEEE Transactions on Evolutionary Computation **6** (2002) 443–462
2. Gordon, V.S., Whitley, D.: Serial and Parallel Genetic Algorithms as Function Optimizers. In Forrest, S., ed.: Proceedings of the Fifth International Conference on Genetic Algorithms, Morgan Kaufmann (1993) 177–183
3. Alba, E., Troya, J.M.: Influence of the Migration Policy in Parallel dGAs with Structured and Panmictic Populations. Applied Intelligence **12** (2000) 163–181
4. Sarma, J., De Jong, K.: An Analysis of the Effects of Neighborhood Size and Shape on Local Selection Algorithms. In Voigt, H.M., Ebeling, W., Rechenberg, I., Schwefel, H.P., eds.: PPSN IV. Volume 1141 of LNCS., Springer (1996) 236–244
5. Sarma, J., De Jong, K.: An Analysis of Local Selection Algorithms in a Spatially Structured Evolutionary Algorithm. In Bäck, T., ed.: Proceedings of the 7th International Conference on Genetic Algorithms, Morgan Kaufmann (1997) 181–186
6. Gorges-Schleuter, M.: An Analysis of Local Selection in Evolution Strategies. In Banzhaf, W., Daida, J., Eiben, A.E., Garzon, M.H., Honavar, V., Jakiela, M., Smith, R.E., eds.: Proceedings of the Genetic and Evolutionary Computation Conference. Volume 1., Orlando, Florida, USA, Morgan Kaufmann (1999) 847–854
7. Rudolph, G.: Takeover Times in Spatially Structured Populations: Array and Ring. In Lai, K.K., Katai, O., Gen, M., Lin, B., eds.: 2nd Asia-Pacific Conference on Genetic Algorithms and Applications, Global-Link Publishing (2000) 144–151
8. Giacobini, M., Tettamanzi, A., Tomassini, M.: Modelling Selection Intensity for Linear Cellular Evolutionary Algorithms. In Liardet, P., et al., eds.: Artificial Evolution, Sixth International Conference, Springer Verlag (2003)
9. Giacobini, M., Alba, E., Tomassini, M.: Selection Intensity in Asynchronous Cellular Evolutionary Algorithms. In Cantú-Paz, E., ed.: Proceedings of the Genetic and Evolutionary Computation Conference, Chicago, USA (2003) 955–966
10. Sprave, J.: A Unified Model of Non-Panmictic Population Structures in Evolutionary Algorithms. In Angeline, P.J., Michalewicz, Z., Schoenauer, M., Yao, X., Zalzala, A., eds.: Proceedings of the Congress of Evolutionary Computation. Volume 2., Mayflower Hotel, Washington D.C., USA, IEEE Press (1999) 1384–1391
11. Goldberg, D.E., Deb, K.: A Comparative Analysis of Selection Schemes Used in Genetic Algorithms. In Rawlins, G.J., ed.: Foundations of Genetic Algorithms. Morgan Kaufmann, San Mateo, CA (1991) 69–93
12. Chakraborty, U.K., Deb, K., Chakraborty, M.: Analysis of Selection Algorithms: A Markov Chain Approach. Evolutionary Computation **4** (1997) 133–167
13. Cantú-Paz, E.: 7. Migration, Selection Pressure, and Superlinear Speedups. In: Efficient and Accurate Parallel Genetic Algorithms. Kluwer (2000) 97–120