TAKEOVER TIME IN PARALLEL POPULATIONS WITH MIGRATION

Günter Rudolph
Universität Dortmund
Fachbereich Informatik
Lehrstuhl für Algorithm Engineering
44221 Dortmund / Germany
Guenter.Rudolph@uni-dortmund.de

Abstract The term *takeover time* regarding selection methods used in evolutionary algorithms denotes the (expected) number of iterations of the selection method until the entire population consists of copies of the best individual, provided that the initial population consists of a single copy of the best individual whereas the remaining individuals are worse. Here, this notion is extended to parallel subpopulations that exchange individuals according to some migration paths modelled by a directed graph. We develop upper bounds for migrations on uni- and bidirectial rings as well as arbitrary connected graphs where each vertex is reachable from every other vertex.

Keywords: Takeover time, spatially structured population, migration model

1. Introduction

The term *takeover time* regarding selection methods used in evolutionary algorithms (EAs) was introduced by Goldberg and Deb [7]. Suppose that a finite population of size n consists of a single best individual and $n - 1$ worse individuals. The takeover time of some selection method is the (expected) number of iterations of the selection method until the entire population consists of copies of the best individual.

The calculations in [7] for spatially unstructured (i.e., panmictic) populations implictly assume that at least one copy of the best individual is kept in the population although some selection method may erase all best copies by chance. If a selection method is *elitist*, i.e., the best individual survives selection with probability 1, this kind of extinction is precluded. At a first glance it is surprising that most results on the takeover time are approximations (without bounds) [7] or obtained numerically by an underlying Markov chain model [2, 11].

Apparently, selection in panmictic populations is the most difficult case for deriving rigorous results on the takeover time. If only a single individual is generated in each generation (steady-state EA) the Markov model looses some of its complexity as has been shown by Smith and Vavak [11]. Mathematically rigorous results have been provided by Rudolph [9, 10] for some of these non-generational selection methods. In case of populations with a spatial structure (at the level of individuals) the notion of the takeover time must be extended appropriately. This has been done by Rudolph [8] who developed bounds on the takeover time for arbitrary connected population structures and even an exact expression for a structure like a ring. These results have been extended by Giacobini et al. [4, 5, 6].

Recently, Alba and Luque [1] have considered spatially structured populations that are structured at the level of subpopulations (in contrast to individuals). In this population model the subpopulations are panmictic and from time to time some individuals migrate between the subpopulations according to some connectivity graph: The vertices of the graph are the subpopulations whereas the directed edges are the migration paths. In [1] the authors develop a plausible approximation (without bounds) for some special cases.

This was the starting point of this work: We show how to derive rigorous bounds for the takeover time for parallel populations with migration. For this purpose some mathematical facts are introduced in section 2 before the analysis is presented in section 3.

2. Mathematical Preliminaries

In the course of the analysis given in section 3 we need bounds on Harmonic numbers:

Definition 1
The symbol H_n denotes *nth Harmonic number* for some $n \in \mathbb{N}$ where

$$H_n = \sum_{i=1}^{n} \frac{1}{i} \,.$$

Likewise, the *nth Harmonic number of 2nd order* $H_n^{(2)}$ is given by

$$H_n^{(2)} = \sum_{i=1}^{n} \frac{1}{i^2}$$

for $n \geq 1$. □

Notice that

$$\log(n) \le H_n \le \log(n) + 1$$

for $n \ge 2$ and

$$1 \le H_n^{(2)} \le \frac{\pi^2}{6}$$

for $n \ge 1$.

Definition 2
A random variable G is *geometrically distributed* with support $\mathbb{N}$ if $\mathsf{P}\{\, G = k \,\} = p\,(1-p)^{k-1}$ for some $p \in (0,1) \subset \mathbb{R}$. □

The expectation and variance of G are

$$\mathsf{E}[\, G \,] \;=\; \frac{1}{p} \;\text{ resp. }\; \mathsf{V}[\, G \,] \;=\; \frac{p}{(1-p)^2}. \tag{1}$$

Definition 3
Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed (i.i.d.) random variables. Then $X_{1:n}$ denotes the minimum and $X_{n:n}$ the maximum of these random variables. □

Let $\mathsf{D}[\, X \,] = \sqrt{\mathsf{V}[\, X \,]}$ denote the standard deviation of some random variable X. There exists a general result regarding bounds on the expectation of the minimum and maximum:

Theorem 1 (David 1980, p. 59 and 63)
Let $X_1, X_2, \ldots, X_n$ be an i.i.d. sequence of random variables. The bounds

$$\begin{aligned}
\mathsf{E}[\, X_{1:n} \,] \;&\ge\; \mathsf{E}[\, X_1 \,] - \frac{n-1}{\sqrt{2\,n-1}}\,\mathsf{D}[\, X_1 \,] \\
\mathsf{E}[\, X_{n:n} \,] \;&\le\; \mathsf{E}[\, X_1 \,] + \frac{n-1}{\sqrt{2\,n-1}}\,\mathsf{D}[\, X_1 \,]
\end{aligned}$$

are valid regardless of the distribution of the X_i. □

3. Analysis

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote a directed graph where each vertex $v \in V$ represents a subpopulation and each directed edge $e = (v, v') \in E$ a migration path from subpopulation v to subpopulation v'. Random variable $X_v^{(t)}$ specifies the number of individuals with best fitness at iteration $t \ge 0$ of subpopulation $v \in V$ with $X_k^{(0)} = 1$ for a single subpopulation k and

$X_v^{(0)} = 0$ for $v \neq k$. The number of individuals s in each subpopulation is constant over time, identical for all subpopulations, and finite. Moreover, we make the following *general assumptions:*

(A1) Selection in subpopulations is elitist.

(A2) Migration takes place every mth generation with finite $m \in \mathbb{N}$.

(A3) Emmigration policy: a copy of the best individual travels along each migration path.

(A4) Immigration policy: replace the worst individual of the subpopulation with the immigrant (if it is better than the worst one).

Let $T_v = \min\{t \geq 0 : X_v^{(t)} = s\}$ be the random takeover time of subpopulation $v \in V$ and A_v the random arrival time, i.e., the number of iterations until the first individual with best fitness arrives at subpopulation $v \in V$. In general, the arrival times are not identically distributed. Their distributions depend on the connectivity or migration graph and in which subpopulation the initial best individual has emerged. If the migration path is vertex-symmetric (like Cayley graphs) the latter dependency vanishes. Here, we shall assume that the initial best individual emerges at vertex $v = 0$ and we rename the other vertices accordingly. Then

$$T = \max\{T_0, A_1 + T_1, A_2 + T_2, \ldots, A_n + T_n\} \tag{2}$$

is the *takeover time* of the migration model with $n + 1$ subpopulations considered here. Notice that random variables T_v are i.i.d. for $v \geq 1$ whereas the distribution of T_0 is different: Once a best copy has arrived at subpopulation $v \geq 1$, every mth generation at least one another best copy immigrates to this subpopulation regardless of the selection process within the subpopulation. Therefore it takes at most $m\,s$ iterations until all individuals in some subpopulation $v \geq 1$ are copies of the best individual regardless of the selection process. Thus,

$$T_v \leq m\,s \tag{3}$$

with probability 1 (w.p. 1) for $v \geq 1$. If m is large the bound above becomes useless since it is likely that the takeover event happens before the first migration interval is over. Therefore we define random variable T'_v which is the takeover time of subpopulation v if no further migration takes place once a best copy has arrived. As a consequence, we have

$$T_v \leq T'_v \tag{4}$$

w.p. 1 for all $v \geq 0$. Notice that $T'_0, T'_1, \ldots, T'_n$ are i.i.d. random variables.

3.1 Uni-directional Ring Topology

Suppose that the subpopulations are placed at the vertices of a uni-directional ring. Then the takeover time in eqn. (2) specializes to

$$T = \max\{T_0, m + T_1, 2\,m + T_2, \ldots, n\,m + T_n\} \tag{5}$$

for a finite migration interval $m \in \mathbb{N}$. Once a best individual has emerged at vertex 0 it takes m generations until this best individual migrates to vertex 1. Now it takes again m iterations until a best copy migrates to vertex 2 and so forth. As soon as a best copy has arrived at some vertex v it takes T_v iterations at vertex v until all individuals are copies of the best individual. Evidently, T can be bracketed as follows:

$$n\,m + \min\{T_0, \ldots, T_n\} \le T \le n\,m + \max\{T_0, \ldots, T_n\}\,. \tag{6}$$

Using (4) in the right hand side (r.h.s.) of inequality (6) we obtain the bound

$$T \le n\,m + \max\{T'_0, \ldots, T'_n\}$$

for the takeover time T and hence the bound

$$\mathsf{E}[\,T\,] \le n\,m + \mathsf{E}[\,T'_{n+1:n+1}\,] \tag{7}$$

for the expected takeover time. Usage of (3) in the r.h.s. of inequality (6) yields $\mathsf{E}[\,T\,] \le n\,m + m\,s$ which leads to the bound

$$\mathsf{E}[\,T\,] \le n\,m + \min\{m\,s, \mathsf{E}[\,T'_{n+1:n+1}\,]\} \tag{8}$$

in consideration of (7). Owing to Theorem 1 the bound in (7) can be expressed in terms of the expectation $\mathsf{E}[\,T'_0\,]$ and standard deviation $\mathsf{D}[\,T'_0\,]$ of T'_0. We obtain

$$\mathsf{E}[\,T\,] \le n\,m + \mathsf{E}[\,T'_0\,] + \frac{n\,\mathsf{D}[\,T'_0\,]}{\sqrt{2\,n+1}}\,. \tag{9}$$

But as long as nothing is known about the selection operation within the subpopulations the distribution and therefore the moments of T'_0 remain unknown. Therefore we assume that each subpopulation runs a steady-state EA with a selection method that does not erase any copy of the best individual contained in the current population. In this case expectation and variance can be calculated as follows [9]: If i denotes the number of best copies of the current population then the value of i is a nondecreasing sequence. Let $p_{i,i+1}$ be the probability that the next population will contain $i+1$ best copies and $p_{i,i} = 1 - p_{i,i+1}$ the probability that the number of best copies will not change, provided the current number of best copies is i. Then the random number G_i of generations until i changes to $i+1$ is geometrically distributed with expectation and variance

$$\mathsf{E}[\,G_i\,] \;=\; \frac{1}{p_{i,i+1}} \quad \text{resp.} \quad \mathsf{V}[\,G_i\,] \;=\; \frac{1-p_{i,i+1}}{p_{i,i+1}^2}$$

for $i = 1, \ldots, s-1$. Since $G_1, \ldots, G_{s-1}$ are mutually independent we obtain

$$\mathsf{E}[\,T_0'\,] \;=\; \sum_{i=1}^{s-1} \mathsf{E}[\,G_i\,] \;=\; \sum_{i=1}^{s-1} \frac{1}{p_{i,i+1}} \tag{10}$$

$$\mathsf{V}[\,T_0'\,] \;=\; \sum_{i=1}^{s-1} \mathsf{V}[\,G_i\,] \;=\; \sum_{i=1}^{s-1} \frac{1-p_{i,i+1}}{p_{i,i+1}^2} \tag{11}$$

for the takeover time T_0'. Next, we choose a specific selection method to exemplify our approach developed so far. The method called 'Replace Worst'-selection first draws two individuals at random with uniform probability. Subsequently the better one of the pair replaces the worst individual of the entire population. Therefore, i is incremented if at least one copy of the best individual is drawn. We obtain

$$p_{i,i+1} \;=\; 1 - \left(1 - \frac{i}{s}\right)^2 \;=\; \frac{i\,(2\,s - i)}{s^2}$$

and finally owing to (10)

$$\mathsf{E}[\,T_0'\,] \;=\; \frac{1}{2}\,(s\,H_{2\,s-1} - 1). \tag{12}$$

The result for the expectation above can be found in [9] already. Here, we also need a result for the variance. According to (11) we obtain

$$\begin{aligned}
\mathsf{V}[\,T_0'\,] \;&=\; \sum_{i=1}^{s-1} \frac{1-p_{i,i+1}}{p_{i,i+1}^2} \;=\; \sum_{i=1}^{s-1} \frac{(s-i)^2}{i^2} \cdot \frac{s^2}{(2\,s-i)^2} \\
&\le\; \sum_{i=1}^{s-1} \left(\frac{s}{i} - 1\right)^2 \quad \text{since} \quad \frac{s}{2\,s-i} \le 1 \\
&=\; \sum_{i=1}^{s-1} \left(\frac{s^2}{i^2} - \frac{2\,s}{i} + 1\right) \\
&=\; s^2\,H_{s-1}^{(2)} - 2\,s\,H_{s-1} + s - 1 \\
&\le\; s^2\,\frac{\pi^2}{6} - 2\,s\log(s-1) + s - 1 \quad \text{if} \quad s \ge 3
\end{aligned}$$

and since $s/(2\,s-i) \geq 1/2$

$$\begin{aligned}
\mathsf{V}[\,T_0'\,] &= \sum_{i=1}^{s-1} \frac{(s-i)^2}{i^2} \cdot \frac{s^2}{(2\,s-i)^2} \geq \frac{1}{4} \sum_{i=1}^{s-1} \left(\frac{s}{i} - 1\right)^2 \\
&= \frac{1}{4}\,(s^2\,H_{s-1}^{(2)} - 2\,s\,H_{s-1} + s - 1) \\
&\geq \frac{1}{4}\,(s^2 - 2\,s\log(s-1) - s - 1\,)
\end{aligned}$$

revealing that $\mathsf{V}[\,T_0'\,] = \Theta(s^2)$ or $\mathsf{D}[\,T_0'\,] = \Theta(s)$. Insertion in (9) yields the bound

$$\begin{aligned}
\mathsf{E}[\,T\,] &\leq n\,m + \frac{s\,H_{2\,s-1} - 1}{2} + \frac{n}{\sqrt{2\,n+1}} \sqrt{\frac{s^2\,\pi^2}{6} - 2\,s\,\log(s-1) + s - 1} \\
&\leq n\,m + \frac{s\,\log(2s)}{2} + \sqrt{\frac{n}{2}} \cdot \sqrt{\frac{s^2\pi^2}{6} + 1} \quad (\text{for } s \geq 2) \\
&= n\,m + \frac{s\,\log(2s)}{2} + s\,\pi \sqrt{\frac{n}{12}} \cdot \sqrt{1 + \frac{6}{s^2\pi^2}} \\
&\leq n\,m + \frac{s\,\log(2s)}{2} + s\,\pi \sqrt{\frac{n}{6}} \\
&= \mathcal{O}(n\,m + s\log s + s\sqrt{n})
\end{aligned}$$

and taking into account the bound given in (8) we obtain

$$\mathsf{E}[\,T\,] \leq n\,m + s\,\min\left\{\,m, \frac{\log(2s)}{2} + \pi\sqrt{\frac{n}{6}}\,\right\}. \tag{13}$$

A closer inspection of the upper bound (13) reveals that the bound could be strengthened with respect to the additive part $\pi\sqrt{n/6}$ which stems from the generality of Theorem 1. If the distribution of the random variables are taken into account then the bound for the maximum will become more accurate. We have made 30 independent experiments for each combination of $(n+1) \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$, $s \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 1000\}$, and $m \in \{1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 100\}$. Here, we only present the 10 worst results with regard to absolute (see table 1) and relative deviation (see table 2) between the bound in (13) and the observed mean.

Finally, we sketch a potential avenue to strengthen the result; its elaboration remains for future work. Recall from the discussion leading to (10) that the G_i are geometricly distributed random variables with parameter $p_{i,i+1}$ and that T_0' is just the sum of the G_i for $i = 1, \ldots, s-1$. Thus, the maximum of $n+1$ samples of T_0' is the maximum of $n+1$

n+1	s	m	min	max	mean	bound	abs. Δ	Δ%
1000	100	50	50149	50253	50196.8	54118.8	3922.0	7.81
1000	100	100	100107	100236	100166.6	104068.8	3902.2	3.90
1000	100	40	40149	40282	40207.3	43960.0	3752.7	9.33
1000	90	50	50111	50257	50178.7	53699.9	3521.2	7.02
1000	90	100	100078	100243	100157.2	103649.9	3492.7	3.49
1000	90	40	40111	40240	40169.7	43560.0	3390.3	8.44
1000	80	50	50103	50198	50140.1	53281.2	3141.1	6.26
1000	80	100	100072	100178	100114.4	103231.2	3116.8	3.11
1000	80	40	40095	40189	40143.5	43160.0	3016.5	7.51
1000	100	30	30153	30258	30202.7	32970.0	2767.3	9.16

Table 1. Results of experiments with the ten worst absolute deviations (abs. Δ) between bound and observed mean.

n+1	s	m	min	max	mean	bound	abs. Δ	Δ%
10	100	5	170	196	184.5	544.8	360.3	195.29
10	90	5	157	187	171.9	492.8	320.9	186.66
10	80	5	144	172	157.2	441.0	283.8	180.52
10	70	5	132	161	144.6	389.4	244.8	169.33
10	100	4	146	175	164.7	436.0	271.3	164.72
10	60	5	115	145	128.7	338.2	209.5	162.81
10	90	4	141	160	150.7	396.0	245.3	162.77
10	70	4	116	135	124.1	316.0	191.9	154.63
10	80	4	127	153	140.3	356.0	215.7	153.74
20	100	5	215	247	236.8	595.0	358.2	151.27

Table 2. Results of experiments with the ten worst relative deviations (Δ%) between bound and observed mean.

sums of geometric random variables. Since $\max\{a_1 + b_1, a_2 + b_2\} \le \max\{a_1, a_2\} + \max\{b_1, b_2\}$ we obtain an upper bound by the sum over the maxima of $s-1$ i.i.d. (!) geometric random variables. Unfortunately, the expectation of the maximum of geometric random variables cannot be determined exactly, in contrast to its minimum. But we can use the asymptotic theory of extreme value distributions [3] for getting some evidence that the maximum increases by order $\log(n)\,\mathsf{D}[\,T_0'\,]$ rather than order $\sqrt{n}\,\mathsf{D}[\,T_0'\,]$. Thus, we **conjecture** that

$$\mathsf{E}[\,T\,] = \mathcal{O}(n\,m + s\,\min\{m, \log s + \log n\})\,.$$

3.2 Bi-directional Ring Topology

The modifications of the results required in case of subpopulations at the vertices of a ring with bi-directional migration paths are straightforward: It takes $(n+1)\,m/2$ generations until an individual from each of the two possible migration paths arrive at the last vertex if n is odd (i.e., if the number of subpopulations is even). Therefore the upper bounds are

$$\mathsf{E}[\,T\,] \le \frac{(n+1)\,m}{2} + \max\{T'_0, T'_1, \ldots, T'_n\}$$

and

$$\mathsf{E}[\,T\,] \le \frac{(n+1)\,m}{2} + m\,s\,.$$

In the following we can use the same arguments and bounds as those from the preceding subsection.

3.3 Almost Arbitrary Connected Topology

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote the directed graph describing the migration paths between subpopulations. Needless to say, we assume that the graph is connected and that each vertex can be reached from any other vertex of the graph. As the preceding two examples have shown, the takeover time can be bounded by the time to reach each vertex in the graph (which is bounded by the *diameter* of the graph) plus the time required for takeover in $n+1$ parallel subpopulations. Consequently, the expected takeover time of (almost) arbitrary graphs can be bounded by the two bounds

$$\mathsf{E}[\,T\,] \le \operatorname{diam}(\mathcal{G})\,m + \max\{T'_0, T'_1, \ldots, T'_n\}$$

and

$$\mathsf{E}[\,T\,] \le \operatorname{diam}(\mathcal{G})\,m + m\,s\,.$$

Of course, these bounds can be improved if more information about a graph is known. For example, if we have a d-regular bi-directional graph then at least one best copy enters the population initially, d best copies will leave at the next migration event, and from now on d copies of the best individual will enter the subpopulation at each migration event.

4. Conclusions

It has been shown that the takeover time in parallel populations with migration is bounded by the diameter of the migration graph plus the time until takeover in parallel population occurs. These takeover times are dependent on the selection operation deployed in each subpopulation. Here, we have developed bounds for a particular non-generational selection method. It is conjectured that the bounds can be improved considerably as soon as a sufficiently tight bound for $\mathsf{E}[\max\{T'_0, T'_1, \ldots, T'_n\}\,]$ has been developed. In case of non-generational selection methods an appropriate bound for the maximum of geometrically distributed random variables is required. These tasks and the development of tight lower bounds will be part of future work.

References

[1] E. Alba and A. Luque. Growth curves and takeover time in distributed evolutionary algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2004)*, pages 864–876. Springer, Berlin, 2004.

[2] U. Chakraborty, K. Deb, and M. Chakraborty. Analysis of selection algorithms: A Markov chain approach. *Evolutionary Computation*, 4(2):133–167, 1996.

[3] H. A. David. *Order Statistics*. Wiley, New York, 2nd edition, 1981.

[4] M. Giacobini, M. Tomassini, and A. Tettamanzi. Modeling selection intensity for linear cellular evolutionary algorithms. In P. Liardet et al., editors, *Sixth International Conference on Artificial Evolution (EA 2003)*, Springer, Berlin, 2003.

[5] M. Giacobini, E. Alba, M. Tomassini, and A. Tettamanzi. Modeling selection intensity for toroidal cellular evolutionary algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2004)*, pages 1138–1149, Springer, Berlin, 2004.

[6] M. Giacobini, M. Tomassini, and A. Tettamanzi. Takeover time curves in random and small-world structured populations. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2005)*, pages 1333–1340, Springer, Berlin, 2005.

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

[8] G. Rudolph. On takeover times in spatially structured populations: Array and ring. In K. K. Lai et al., editors, *Proceedings of the 2nd Asia-Pacific Conference on Genetic Algorithms and Applications*, pages 144–151, Global-Link Publishing Company, Hong Kong, 2000.

[9] G. Rudolph. Takeover times and probabilities of non-generational selection rules. In D. Whitley et al., editors, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2000)*, pages 903–910, Morgan Kaufmann, San Fransisco (CA), 2000.

[10] G. Rudolph. Takeover times of noisy non-generational selection rules that undo extinction. In V. Kurkova et al., editors, *Proceedings of the 5th International Conference on Artificial Neural Nets and Genetic Algorithms (ICANNGA 2001)*, pages 268–271, Springer, Vienna, 2001.

[11] J. Smith and F. Vavak. Replacement strategies in steady state genetic algorithms: Static environments. In W. Banzhaf and C. Reeves, editors, *Foundations of Genetic Algorithms 5*, pages 219–233, Morgan Kaufmann, San Francisco (CA), 1999.