

Quality-Time Analysis of Multi-Objective Evolutionary Algorithms

Jian-Hung Chen
Institute of Information
Science
Academia Sinica
Taipei 115, Taiwan
jh.chen@ieee.org

Shinn-Ying Ho
Institute of Bioinformatics
National Chiao Tung University
HsinChu 300, Taiwan
syho@mail.nctu.edu.tw

David E. Goldberg
Department of General
Engineering
University of Illinois at
Urbana-Champaign
Urbana, IL 61801, USA
deg@uiuc.edu

ABSTRACT

A quality-time analysis of multi-objective evolutionary algorithms (MOEAs) based on schema theorem and building blocks hypothesis is developed. A bicriteria OneMax problem, a hypothesis of niche and species, and a definition of dissimilar schemata are introduced for the analysis. In this paper, the convergence time, the first and last hitting time models are constructed for analyzing the performance of MOEAs. Population sizing model is constructed for determining appropriate population sizes. The models are verified using the bicriteria OneMax problem. The theoretical results indicate how the convergence time and population size of a MOEA scale up with the problem size, the dissimilarity of Pareto-optimal solutions, and the number of Pareto-optimal solutions of a multi-objective optimization problem.

Categories and Subject Descriptors

Theory of Computation [Analysis of Algorithms and Problem Complexity]: General

General Terms

Algorithms

Keywords

Convergence, Dissimilar schemata, Multi-objective evolutionary algorithms, Population sizing

1. INTRODUCTION

The great success for evolutionary algorithms (EAs) techniques [1, 6, 14], including evolutionary programming (EP), evolutionary strategy (ES), genetic algorithm (GA), came in the 1980s when extremely complex optimization problems from various disciplines were solved, thus facilitating

the undeniable breakthrough of evolutionary computation as a problem-solving methodology. Recently, multi-objective evolutionary algorithms (MOEAs) have been recognized to be well-suited for solving multi-objective optimization problems (MOOPs) because their abilities to exploit and explore multiple solutions in parallel and to find a widespread set of non-dominated solutions in a single run.

Although MOEAs have been shown to be effective for solving many real-world applications and exploring complex non-linear search spaces as efficient optimizers, MOOPs present three types of difficulty for practitioners of EAs. The first difficulty is that, search spaces of problems depend on a large number of parameters (decision variables), known as large parameter optimization problems (LPOPs) [13, 16]. LPOPs pose a great challenge to decision makers due to the large parametric space to choose from and the possibility of large infeasible and non-uniform areas. Second, MOOPs are usually multi-modal problems, Pareto-optimal solutions spreads on search spaces and are composed of different decision variables. A decision variable may be optimal for a Pareto-optimal solution but bad for the other one. A Pareto-optimal solution may be similar to the other solutions in some variables, but indifferent in the rest of the variables. This causes difficulty in locating the multiple peaks of solutions, and hence diversity preservation becomes an important issue in designing MOEAs [3, 4]. The last difficulty is that, the number of Pareto-optimal solutions in a MOOP usually grows as the number of objectives of the MOOP increases [3, 4]. In continuous parameters problems, the number of Pareto optimal solutions may be infinite. This gives another challenge in finding all the Pareto-optimal solutions. There have been many efforts to design EAs faster, but only a few theoretical work have been done in analyzing how these three difficulties affect the convergence of MOEAs [17].

The objectives of this paper is to conduct quality-time analysis of multi-objective evolutionary algorithms (MOEAs) based on design decomposition theory of genetic algorithms [8, 10], in order to solve multi-objective optimization problems, quickly, reliably, and accurately. The issues focus on the following three topics:

1. The convergence time of MOEAs. Since MOEAs process building blocks (BBs), it is important to understand the growth of BBs in solving MOOPs under the existence of multiple objective functions. In particular, we are interested in identifying what kind of characteristics of MOOPs will affect the convergence time of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO'05, June 25–29, 2005, Washington, DC, USA.
Copyright 2005 ACM 1-59593-010-8/05/0006 ...\$5.00.

MOEAs. For example, when the number of Pareto-optimal solutions in a MOOP grows or the problem size of a MOOP increases, how does the convergence time of MOEAs change?

2. The first and last hitting time. The first and last hitting time stand for the time for MOEAs to obtain the first and last Pareto-optimal solutions. Identifying these time may provide hints for designing efficient MOEAs.
3. The population sizing model for MOEAs. Because GAs work on the fitness of an individual containing multiple building blocks (locally superior partial solutions), they face a statistical decision problem when deciding among competing BBs. For example, a bad BB might be selected over its competitors because it comes with other good BB in the same individual. Given a particular BB, fitness contribution of the BBs from other partitions in the individual can be considered as noise to the fitness. To solve MOOPs using MOEAs, due to a MOOP consists of multiple Pareto-optimal solutions, good BBs of a Pareto-optimal solution may not also be good BBs of the other Pareto-optimal solutions. This produces difficulties for deciding BB well. As a result, it is essential to develop a population sizing model for MOEAs to make statistically correct decisions among competing BBs.

In this paper, a bicriteria OneMax problem is extended from OneMax problem for the analysis. A hypothesis of niche and species and a definition of dissimilar schemata are introduced for the above-mentioned topics. By integrating the selection intensity [21], the noise model [18], the hypothesis of niche and species, and the definition of dissimilar schemata, the convergence time model of MOEAs for the bicriteria OneMax problem is constructed. By making use of the takeover time of GAs [9], the first and last hitting time of Pareto-optimal solutions are derived. Based on the Gambler's Ruin model [11], the noise model, the hypothesis of niche and species, and the definition of dissimilar schemata, a population sizing model is derived to give an approximated bound. The theoretical results indicates that convergence time and population sizing of MOEAs is affected by the problem size, the dissimilarity of Pareto-optimal solutions, and the number of Pareto-optimal solutions, in a MOOP.

This paper is organized as follows. Section 2 introduces OneMax and bicriteria OneMax problem and the convergence time model of the OneMax problem. Section 3 derives convergence-time of MOEAs by making uses of the hypothesis of niche and species, the noise model, and the definition of dissimilar schemata. The first and last hitting time models of Pareto-optimal solutions is given in Section 4. Section 5 briefly described the gambler's ruin model and derived the population sizing model for MOEAs. Section 6 presented the experimental verification. Section 7 concludes this paper.

2. BACKGROUND

2.1 OneMax and Bicriteria OneMax Problems

Before conducting the quality-time analysis of MOEAs, a brief outline of OneMax problem and its extension - bicriteria OneMax problem is presented. The bicriteria OneMax

problem is used as the verification of the quality-time analysis of MOEAs.

The OneMax problem is well-known and well-studied in the context of GAs. The OneMax problem is a bit-counting problem where the fitness value of each binary string is equal to the number of 1s in it. The simplicity of the OneMax problem makes it a prime candidate to study the performance of GAs. OneMax problem is defined as follows:

$$f = \sum_{i=1}^{\ell} x_i, \quad (1)$$

where x_i is the value of the i th allele. The global optimum of the OneMax problem is a string with all 1s and its fitness value is equal to the string length ℓ . OneMax problem has the following properties:

1. BBs are of equal salience: Every BB has the same contribution to the fitness value.
2. Alleles converge uniformly.
3. The fitness is distributed binomially.

Therefore, the mean of fitness values can be written as $\mu_t = \ell p_t$, and the variance of fitness values is $\sigma_t^2 = \ell p_t(1 - p_t)$, where p_t is the proportion of correct BBs in generation t .

A bicriteria OneMax problem is developed for analyzing MOEAs. The bicriteria OneMax problem is defined by

$$\text{Maximize} \begin{cases} f_1 = \ell - d(A, A_1) \\ f_2 = \ell - d(A, A_2) \end{cases}, \quad (2)$$

where A is the string to be evaluated, A_1 , and A_2 are two fixed strings with length ℓ , and $d(A, A_i)$ is Hamming distance function to calculate the string A and A_i . If the fixed string A is an all-1s string, then the corresponding objective function f_1 will be that of the OneMax problem. The number of Pareto-optimal solutions, Q , in the bicriteria OneMax problem can be calculated by

$$Q = 2^{d(A_1, A_2)}. \quad (3)$$

In this paper, unless otherwise mentioned, A_1 is an all-1s string, and A_2 is an all-1s string except the last b bits of A_2 are 0s. Therefore, $d(A_1, A_2) = b = 4$ and $Q = 16$.

2.2 Convergence Time Model of the OneMax Problem

For the OneMax problem, the convergence model is derived by using the response to selection equation [21]:

$$\Delta\mu = \mu_{t+1} - \mu_t = I\sigma_t. \quad (4)$$

This equation was derived by calculating the difference between mean fitness of two populations at times t and $t + 1$ using the *selection intensity* I , and the population's fitness variance σ_t^2 at time t . Equation 4 depicted as the expected increase in the average fitness of a population after selection is performed upon a population whose fitness is distributed according to a unit normal distribution. The selection intensity for different selection schemes is tabulated elsewhere [20]. For tournament selection, I is a constant depending on the tournament size s , approximated in terms of s by the relation [2]:

$$I = \sqrt{2(\log(s) - \log(\sqrt{4.14 \log(s)}))} \quad (5)$$

Using Equation 5 and substituting the mean and variance values in terms of proportion of correct BBs, p_t :

$$p_{t+1} - p_t = \frac{I}{\sqrt{\ell}} \sqrt{p_t(1-p_t)}, \quad (6)$$

where ℓ is the string length of the OneMax problem. Approximating the above difference equation with a differential equation yields

$$\frac{dp_t}{dt} = \frac{I}{\sqrt{\ell}} \sqrt{p_t(1-p_t)}. \quad (7)$$

Integrating the above equation with an initial proportion of correct BBs as p_0 , we obtain

$$p_t = \frac{1}{2} \left[1 - \cos\left(\frac{It}{\sqrt{\ell}} + 2 \sin^{-1} \sqrt{p_0}\right) \right]. \quad (8)$$

Assuming that the initial population is randomly generated, therefore $p_0=0.5$. Equation 8 can be rewritten as a function of time:

$$p_t = \frac{1}{2} \left[1 + \sin\left(\frac{It}{\sqrt{\ell}}\right) \right]. \quad (9)$$

The convergence time of the OneMax problem, t_{conv} , can be determined by substituting $p_t = 1$:

$$t_{conv} = \frac{\pi\sqrt{\ell}}{2I} \quad (10)$$

3. CONVERGENCE TIME MODEL OF THE BICRITERIA ONEMAX PROBLEM

In this section, convergence-time model of MOEAs for bicriteria OneMax problem is derived. First, a hypothesis of niche and species is introduced. Then, the convergence time model with noises [18] is described for tackling the disrupt noises caused by multiple species. A definition of dissimilar schemata extended from the schema theory is proposed to calculate disrupt noises.

3.1 Hypothesis of Niche and Species

By extending Equation 10, we can proceed to derive the convergence time model for the bicriteria OneMax problem. Given a MOOP with Q Pareto-optimal solutions, Q Pareto-optimal solutions can be regarded as Q niches based on the theory of niche and schema theory [7, 14]. In the worst case, to ensure MOEAs is capable of searching Q Pareto-optimal solutions, it is assumed that the population was divided into Q species (subpopulations), $\aleph_q, q = 1, 2, \dots, Q$, and each species optimizes its own niche (Pareto-optimal solution), as shown in Figure 1. Therefore, the optimal schemata of a species is its Pareto-optimal solution. Let the schemata of species be H_q , where the fixed positions are the maximum common string of all individuals in its species and the others are "don't care". Since species are in the same population, a schemata of a species may be disrupted by schemata of the other species due to genetic operators. The disruption between species can be further classified into the following two types:

1. **Species disrupt noise:** The fixed 0/1 schemata of H_{origin} are altered to "don't care" schemata by the corresponding positions of other schemata H_{other} . Thus, a species requires more time for fixing its "don't care" schemata.

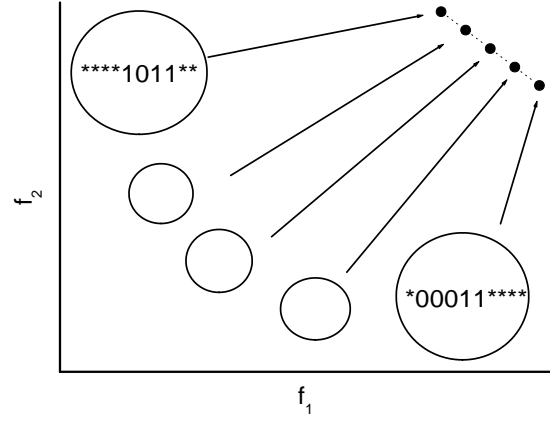


Figure 1: The population was divided into several species, and each species optimizes its own niche (Pareto-optimal solution).

2. **Species hitchhiking effect:** The "don't care" schemata of H_{origin} are altered to fixed 0/1 schemata by the corresponding positions of the other schemata H_{other} . If the altered schemata are located in the similarity regions of their optimal schemata, the change is good for the schemata H_{origin} . On the contrary, the change is bad for the schemata H_{origin} .

However, because multiple species exist in the population, thereby the chance of altering a 0/1 schema to a "don't care" schema is bigger than that of altering a "don't care" schema to a 0/1 schema. Therefore, in this paper, only the species disrupt noise is considered in the convergence model. The species hitchhiking effect can be modeled using the same way.

3.2 Modeling Noise in the Convergence Time Model

To model the species disrupt noise into the convergence model of MOEAs, we can use the convergence time model with noisy fitness functions [18] to predict convergence time in the presence of external noise caused by other species.

The noisy fitness function F_1 can be written as

$$F_1 = F + N_1, \quad (11)$$

where F is the actual fitness distribution $F \sim \mathcal{N}(\mu_{\aleph_q,t}, \sigma_{\aleph_q,t}^2)$, and N_1 is the external noise, $N_1 \sim \mathcal{N}(0, \sigma_{N_1}^2)$. Since both the actual fitness and noise are normally distributed, the noisy fitness function is also normally distributed:

$$F_1 \sim \mathcal{N}(\mu_{\aleph_q,t}, \sigma_{\aleph_q,t}^2 + \sigma_{N_1}^2). \quad (12)$$

Using a bivariate normal distribution, the expected true fitness value of an individual is:

$$E(F | f_1) = E(F) + \frac{\text{cov}(F_1, F)}{\sigma_{F_1}^2} (f_1 - \mu_{F_1}), \quad (13)$$

where $E(F)$ is the expected value of F , and $\text{cov}(F_1, F)$ is the covariance between F_1 and F . From Equations 12 and 13, we know $E(F) = \mu_{\aleph_q,t}$, $\sigma_{F_1}^2 = \sigma_{\aleph_q,t}^2 + \sigma_{N_1}^2$, and $\mu_{F_1} = \mu_{\aleph_q,t}$.

The covariance between F_1 and F is given by

$$\begin{aligned} \text{cov}(F_1, F) &= E(F_1, F) - E(F_1)E(F) \\ &= E((F + N_1)F) - \mu_{\aleph_q, t}^2 \\ &= E(F^2) + E(N_1F) - \mu_{\aleph_q, t}^2. \end{aligned} \quad (14)$$

Since the actual fitness and the noise are independent of each other, $E(N_1F) = E(N_1)E(F) = 0$. Therefore, the above equation can be written as

$$\begin{aligned} \text{cov}(F_1, F) &= E(F^2) - \mu_{\aleph_q, t}^2 \\ &= \sigma_{\aleph_q, t}^2 \end{aligned} \quad (15)$$

Equation 13 can be written as

$$E(F | f_1) = \mu_{\aleph_q, t} + \frac{\sigma_{\aleph_q, t}^2}{\sigma_{\aleph_q, t}^2 + \sigma_{N_1}^2} (f_1 - \mu_{\aleph_q, t}), \quad (16)$$

The selection intensity equation (Equation 4) can be utilized to obtain the expected noisy mean fitness after selection.

$$\begin{aligned} \mu_{F_1, t+1} &= \mu_{F_1, t} + I\sigma_{F_1} \\ &= \mu_{\aleph_q, t} + I\sqrt{\sigma_{\aleph_q, t}^2 + \sigma_{N_1}^2}. \end{aligned} \quad (17)$$

Using the aforementioned equation and Equation 16, we can compute the expected true fitness value.

$$\begin{aligned} E(F | \mu_{F_1, t+1}) &= \mu_{\aleph_q, t} + \frac{\sigma_{\aleph_q, t}^2}{\sigma_{\aleph_q, t}^2 + \sigma_{N_1}^2} \\ &\quad (\mu_{\aleph_q, t} + I\sqrt{\sigma_{\aleph_q, t}^2 + \sigma_{N_1}^2} - \mu_{\aleph_q, t}). \\ \mu_{\aleph_q, t+1} &= \mu_{\aleph_q, t} + \frac{I\sigma_{\aleph_q, t}^2}{\sqrt{\sigma_{\aleph_q, t}^2 + \sigma_{N_1}^2}}. \end{aligned} \quad (18)$$

Assumed a species \aleph_q receives M external disrupt noises from other species $\aleph_i, i \neq q$. The convergence time of a species \aleph_q can be extended from Equation 18 with M external disrupt noises, expressed as

$$\Delta\mu = \mu_{\aleph_q, t+1} - \mu_{\aleph_q, t} = I \frac{\sigma_{\aleph_q, t}^2}{\sqrt{\sigma_{\aleph_q, t}^2 + \sum \rho_{q,i}^2}}. \quad (19)$$

3.3 Dissimilar Schemata in Species

To derive the species disrupt noise $\rho_{q,i}^2$ of the species \aleph_q from the other species \aleph_i , let two species \aleph_q and \aleph_i are defined as a number of individuals containing two locally superior schemata H_q, H_i at generation t . H_q is likely to be disrupt by H_i on their *dissimilar schemata*. The number of dissimilar schemata $r_{q,i}$ among two schemata H_q and H_i can be calculated using a humming-distance-like equation, defined as:

$$\begin{aligned} r_{q,i} &= \sum u(h_a, h_b), \quad h_a \in H_q, h_b \in H_i, \quad a = b. \\ u(h_a, h_b) &= \begin{cases} 1, & \text{if } h_a \neq h_b, \\ 0, & \text{if } h_a = * \parallel h_b = h_b. \end{cases} \end{aligned} \quad (20)$$

The species disrupt noise $\rho_{q,i}^2$ can be derived as:

$$\rho_{q,i}^2 = r_{q,i} \sigma_{\aleph_q, bb}^2. \quad (21)$$

For example, let $H_1 = (****1011**)$ and $H_2 = (*00011****)$ be the schemata of two species \aleph_1 and \aleph_2 . The dissimilar schemata of H_1 among H_1 and H_2 are located at the positions 6, 7 and 8. Thereby, the number of the dissimilar

schemata $r_{1,2} = 3$. The dissimilar schemata of H_2 among H_1 and H_2 are located at the positions 2, 3, 4 and 6. The number of the dissimilar schemata $r_{2,1} = 4$.

3.4 The Convergence Time Model

Plugging Equation 21 into Equation 19, the convergence time of a species \aleph_q can be written as:

$$\Delta\mu = \mu_{\aleph_q, t+1} - \mu_{\aleph_q, t} = I \frac{\sigma_{\aleph_q, t}^2}{\sqrt{\sigma_{\aleph_q, t}^2 + \sum r_{q,i} \cdot \sigma_{\aleph_q, bb}^2}}, \quad (22)$$

where $\rho_{q,i}^2$ is the species disrupt noise of the species \aleph_q from the other species \aleph_i .

Equation 24 can be approximated by substituting constants. For the bicriteria OneMax domain, assumed that each species has same proportion of correct BBs, let p_t be the proportion of correct BBs in the species at generation t , therefore the mean fitness μ_{\aleph_q} at generation t equals ℓp_t , the average building-block variance $\sigma_{\aleph_q, bb}^2$ is approximated by $p_t(1-p_t)$, and the variance of fitness at generate t equals $\ell p_t(1-p_t)$. Let the dissimilar schemata of two niches in every generation be a constant $r_{q,i} = r$, then the total disrupt noise from M species can be simplified as:

$$\sum r_{q,i} \sigma_{\aleph_q, bb}^2 = rM p_t(1-p_t). \quad (23)$$

The population is converged to optimal when $p_t = 1$. Equation 19 now yields

$$p_{t+1} - p_t = \frac{I}{\sqrt{\ell}} \frac{\sqrt{p_t(1-p_t)}}{\sqrt{1 + \frac{rM}{\ell}}}. \quad (24)$$

Therefore, the convergence time of a species \aleph_q can be derived as:

$$t_{conv, \aleph_q} = \frac{\pi}{2I} \sqrt{\ell + rM}. \quad (25)$$

Finally, the convergence time of the whole population is written as:

$$t_{conv} = \max\{t_{conv, \aleph_q}\}. \quad (26)$$

If the number of external disrupt noises M is taken as 0 in Equation 25, then the above relation Equation 26 reduces to

$$t_{conv} = \frac{\pi\sqrt{\ell}}{2I}, \quad (27)$$

which agrees with existing convergence-time models for the OneMax problem.

From the approximated convergence time model in Equation 25, it is shown that the convergence time of MOEAs go as $O(\sqrt{\ell + rM})$. These factors are discussed as follows.

1. String length ℓ : String length stands for the problem size of the bicriteria OneMax problem. The convergence time model suggest that the number of generations go as $O(\sqrt{\ell})$ in the case of a uniformly scaled problem.
2. Dissimilar schemata in species r : Dissimilar schemata stand for the dissimilarity of Pareto-optimal solutions. Let's recall that the optimal schemata of a species is a Pareto-optimal solution, the minimum dissimilar schemata of two species can be obtained using their Pareto-optimal solutions. However, in reality, the dissimilar schemata of two species $r_{q,i}$ changes over time.

In order to approximate the convergence time model, $r_{q,i}$ can be set to a constant, such as a proportion of string length ℓ . From the minimum dissimilar schemata, it indicates that the convergence time increases as the degree of dissimilar schemata (the dissimilarity of Pareto-optimal solutions) increases. On the contrary, because of species hitchhiking noise, the convergence time decreases as the degree of similar schemata (the similarity of Pareto-optimal solutions) increases.

3. Number of species M : The number of species stands for the number of Pareto-optimal solutions. If the population is assumed to be divided into Q species for Q Pareto-optimal solutions, then M can be set to the number of Pareto-optimal solutions minus one ($M = Q - 1$) in Equation 25. However, in a real run of a MOEA with a finite population size, a species may only receive a limited number of external noises because the number of individuals in a species is smaller than the number of noises. In this case, M can be set to the number of individuals in the species or a proportion of Q .

4. THE FIRST AND LAST HITTING TIME

The first and last hitting time stand for the time for MOEAs to obtain the first and last Pareto-optimal solutions. Because the convergence time model is constructed, by making use of the takeover time of a Pareto-optimal solution in a species, a backward method can be used to derive the first and last hitting time.

Takeover time analysis considers the time from an initial proportion of Pareto-optimal solution until the species is converged. Following standard practice [9], assuming an initial proportion $P_0 = 1/n_{\aleph_q}$ and a final proportion $P_{final} = (n_{\aleph_q} - 1)/n_{\aleph_q}$, where n_{\aleph_q} is the species size unless stated otherwise. The difference equation for tournament selection is

$$P_{t+1} = 1 - (1 - P_t)^s, \quad (28)$$

where s is the tournament size. This equation may be solved if the complementary proportion $Q_t = 1 - P_t$ is used. Substituting into the equation yields the following:

$$Q_{t+1} = Q_t^s, \quad (29)$$

from which the solution may immediately be derived:

$$Q_t = Q_0^{s^t}. \quad (30)$$

Assuming that a species start with a single Pareto-optimal solution ($Q_0 = (n_{\aleph_q} - 1)/n_{\aleph_q}$) and end with all but one Pareto-optimal solution converged ($Q_{final} = 1/n_{\aleph_q}$) permits the takeover time to be approximated directly. Taking the natural log of the expression twice and recognizing that $\ln(1 - x) \approx x$ yields the takeover time:

$$t_s = \frac{\ln n_{\aleph_q} + \ln \ln n_{\aleph_q}}{\ln s}. \quad (31)$$

Asymptotically, if \aleph_q is small, $\ln \ln n_{\aleph_q}$ may be ignored.

From Equation 19, it is shown that a species can converge to a niche more quickly if and only if this species receives no disrupt noise from other species. Thus, the convergence time is equal to Equation 27. Thereby, the first hitting time can be written as:

$$t_{first} = \min\{t_{conv, \aleph_q}\} - t_s. \quad (32)$$

The last hitting time can be written as:

$$t_{last} = \max\{t_{conv, \aleph_q}\} - t_s. \quad (33)$$

For the bicriteria OneMax problem, the first hitting time of the bicriteria OneMax problem is:

$$t_{first} = \frac{\pi\sqrt{\ell}}{2I} - \frac{\ln n_{\aleph_q} + \ln \ln n_{\aleph_q}}{\ln s}. \quad (34)$$

The last hitting time of the bicriteria OneMax problem is:

$$t_{last} = \frac{\pi}{2I} \sqrt{\ell + \sum r_{q,i}} - \frac{\ln n_{\aleph_q} + \ln \ln n_{\aleph_q}}{\ln s}. \quad (35)$$

5. POPULATION SIZING MODEL OF THE BICRITERIA ONEMAX PROBLEM

The population sizing model of the bicriteria OneMax problem is extended from the gambler's ruin model of GAs [11]. It shows that the population sizing model for a binary encoded GAs is:

$$n = -2^{k-1} \log(\alpha) \frac{\sqrt{\pi(m-1)\sigma_{bb}^2}}{d}, \quad (36)$$

where n is the population size, k is the building block (BB) length, α is the failure rate, and σ_{bb}^2 is the variance of the BBs. For an OneMax with string length $m = 100$, $k = 1$, $\sigma_{bb}^2 = 0.25$.

Similar to Section 3, by extending the population sizing model for noisy environment [19], the population size n_{\aleph_q} for a species \aleph_q with M external noises can be rewritten as follows:

$$n_{\aleph_q} = -\frac{1}{2} \chi^k \log(\alpha) \frac{\sqrt{\pi}}{d} \sqrt{(m-1)\sigma_{\aleph_q, bb}^2 + \sum r_{q,i} \sigma_{\aleph_i, bb}^2}, \quad (37)$$

where $\sigma_{\aleph_i, bb}$ is the average building-block variance of the species \aleph_i . Assumed $\sigma_{\aleph_i, bb} = \sigma_{\aleph_q, bb}$, and let $r_{q,i} = r$, Equation 39 can be simplified as:

$$n_{\aleph_q} = -\frac{1}{2} \chi^k \log(\alpha) \sigma_{\aleph_q, bb} \frac{\sqrt{\pi}}{d} \sqrt{(m-1) + rM}. \quad (38)$$

For bicriteria OneMax problems with string length ℓ and the size of BBs $k = 1$, then $m = \ell$.

Afterward, the population sizing model of MOEAs can be written as:

$$n = \sum_{t=1}^T n_{\aleph_t}, \quad (39)$$

where T is the number of species in population. It is not easy to determine T , because the schemata of Pareto-optimal solutions are often overlapped, so that the number of species required is smaller than the number of Pareto-optimal solutions. For the bicriteria OneMax problem in Section 2.1, the minimal number of species required is two ($T = 2$), because the schemata of the Pareto-optimal solutions are overlapped. All the Pareto-optimal solutions could be generated by uniform crossover within one generation using two optimal schemata A_1 and A_2 .

6. EXPERIMENTAL VERIFICATION

To verify the convergence time, the first hitting time and the last hitting models, and the population sizing model of MOEAs, the bicriteria OneMax problem with $b = 4$ is used in the paper. A selectorecombinative MOGA with a

binary tournament selection without replacement, uniform crossover ($p_c = 1$) and fitness sharing is used for the analysis. An external archive is used only for storing the non-dominated solutions found so far. The guideline of fitness sharing given by Deb and Goldberg [5] is used to calculate the niche radius for fitness sharing. The used fitness assignment strategy is the generalized Pareto-based scale-independent fitness function [12, 13]. The used GPSIFF is briefly described below. Let the fitness value of an individual Y be a tournament-like score obtained from all participant individuals by the following function:

$$F(Y) = p - q + c, \quad (40)$$

where p is the number of individuals which can be dominated by the individual Y , and q is the number of individuals which can dominate the individual Y in the objective space, c is a constant to make fitness values positive.

All the experimental results are obtained from 100 independent runs of the selectorecombinative MOGA, per test problem. Box plots is used to depicted some empirical results. A box plot provides an excellent visual result of a distribution. The box stretches from the lower hinge (defined as the 25th percentile) to the upper hinge (the 75th percentile) and therefore contains the middle half of the scores in the distribution. The median is shown as a line across the box.

6.1 Verification of the Convergence Time Model

The convergence-time model predicted in Equation 25 is verified with empirical results for the bicriteria OneMax problem with string lengths, $\ell = 50, 100, 150, 200, 250, 300, 350, 400, 450,$ and 500 , as shown in Figure 2. The bold line in this figure is the convergence time of MOEAs for solving the bicriteria OneMax problem with $b = 4$, which is obtained using Equation 25. The approximated convergence time is obtained by the following manner. Because the disrupt noises may act on the same schemata, so that the minimum disrupt noise from the other species from the other species can be derived using the most dissimilar Pareto-optimal solutions of the bicriteria OneMax problem, $\rho_{q,i}^2 = b = 4$. Thereby, the theoretical convergence time model in Equation 25 is approximated using the minimum disrupt noise from the other species plus a proportion of ℓM , where $rM = 4 + 0.02\ell M$ and $M = 15$. Figure 2 clearly indicates the the model predict the convergence time well. The results indicate that the convergence time of MOEAs is affected by the problem size, the dissimilarity of Pareto-optimal solutions of a MOOP and the number of Pareto-optimal solutions of a MOOP.

6.2 Verification of the Hitting Time

The first and last hitting time predicted in Equations 34 and 35 are verified with empirical results for the bicriteria OneMax problem with string length $\ell = 100$ and $b = 4$ in Figures 3 and 4. The dotted line in Figure 3 is the convergence time of the OneMax problem with string length $\ell = 100$. The dotted line in Figure 4 is the convergence time of the bicriteria OneMax problem with string length $\ell = 100$. The bold lines are the predicted first and last hitting time. From the figures, it is shown that the experimental results agree with the predicted first and last hitting.

6.3 Verification of the Population Sizing Model

The population sizing model predicted in Equations 39

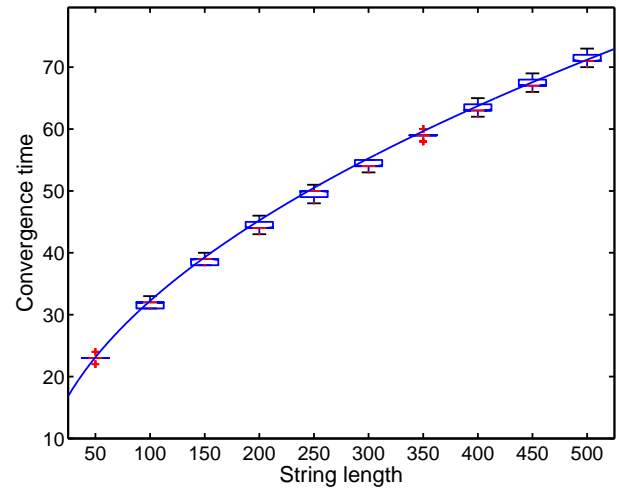


Figure 2: Verification of the convergence time model for different problem sizes.

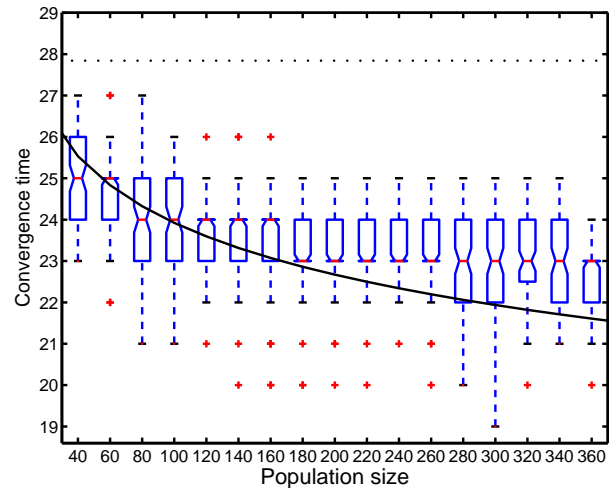


Figure 3: Verification of the first hitting time of the bicriteria OneMax problem for different population sizes. The bold line is the predicted first hitting time.

is verified with empirical results for the bicriteria OneMax problem with string length, $\ell = 50, 100, 150, 200, 250, 300, 350, 400, 450,$ and 500 , and $b = 4$ in Figure 5. All the experimental results are obtained from 100 independent runs of the selectorecombinative MOGA with fitness sharing, per test problem. Because the minimum dissimilar schemata of A_1 and A_2 is $b = 4$, the term rM in Equation 38 is approximated using $rM = 4 + 0.02\ell M$. Since the number of species M is varied during the run time of a MOGA, M is approximated using the number of Pareto-optimal solutions minus one, $M = 2^b - 1 = 15$. Finally, T is approximated using the minimum number of species required to generate all the Pareto-optimal solutions, $T = 2$.

The verification of the population sizing model for different problem sizes is reported in Figure 5. The results indicate that the population sizing model is affected by the problem size, the dissimilarity of Pareto-optimal solutions,

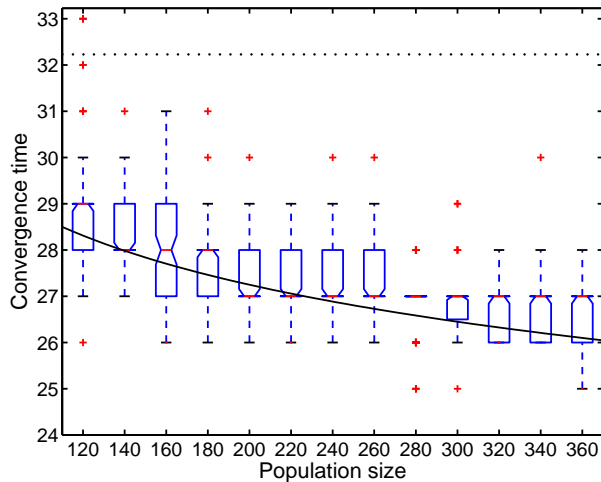


Figure 4: Verification of the last hitting time of the biCriteria OneMax problem for different population sizes. The bold line is the predicted last hitting time.

and the number of Pareto-optimal solutions of a MOOP.

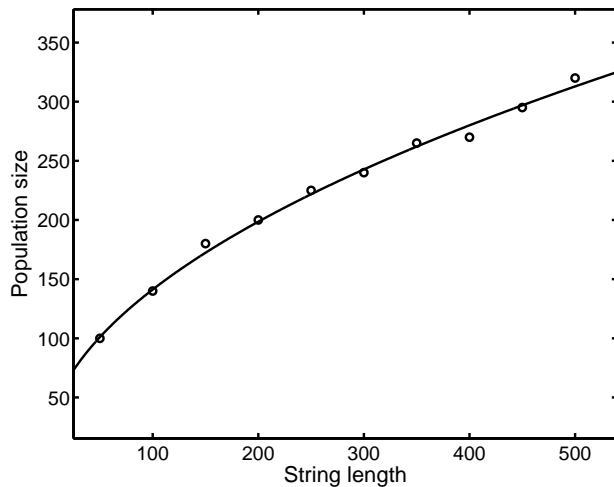


Figure 5: Verification of the population sizing model for different problem sizes.

7. CONCLUSIONS

In this paper, a quality-time analysis of MOEAs based on the schema theorem and BBs hypothesis is presented. By making use of simple and dimensional models, better understanding of the effects of problem characteristics on MOEAs are gained in theory. The convergence time model of MOEAs, the first and the last hitting time of Pareto-optimal solutions are constructed by integrating the selection intensity, the noise model, the hypothesis of niche and species, and takeover time. It indicates that convergence time and population sizing of MOEAs are affected by the problem size, the dissimilarity of Pareto-optimal solutions, and the number of Pareto-optimal solutions, in a MOOP. The developed models should be helpful to MOEA practi-

tioners in understanding the behavior of MOEAs in solving MOOPs.

The following concluding remarks are made according the developed models:

1. Maintaining sufficient species during the run of a MOEA. If the number of species M is smaller than the number of Pareto-optimal solutions, the MOEA may fail in obtaining all the Pareto-optimal solutions. If the number of species M is large, the MOEA will suffer slow convergence. Niching techniques, such as fitness sharing, are suitable methods for maintaining species.
2. Minimizing disrupt noises in species. If the good BBs of a species is often disrupted, this species will suffer slow convergence. Mate selection schemes, such as similarity-based mating scheme [15], is beneficial in reducing disrupt noises.
3. Maximizing hitchhiking effects in species. If the desired schemata can be hitchhiked, then the convergence speed of the species can be speed up.

8. REFERENCES

- [1] T. Bäck, D. B. Fogel, and Z. Michalewics. *Handbook of evolutionary computation*. Institute of Physics Publishing, 1998.
- [2] T. Blickle and L. Thiele. A mathematical analysis of tournament selection. *Proceedings of the Six International Conference on Genetic Algorithms*, pages 9–16, 1995.
- [3] C. A. Coello Coello, D. A. Van Veldhuizen, and G. B. Lamont. *Evolutionary algorithms for solving multi-objective problems*. Genetic algorithms and evolutionary computation ; 5. Kluwer Academic, New York, 2002.
- [4] K. Deb. *Multi-objective optimization using evolutionary algorithms*. Wiley-Interscience series in systems and optimization. John Wiley & Sons, 2001.
- [5] K. Deb and D. E. Goldberg. An investigation of niche and species formation in genetic function optimization. *Proceedings of the Third International Conference on Genetic Algorithms*, pages 42–50, 1989.
- [6] D. E. Goldberg. *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley Pub. Co., 1989.
- [7] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Publishing Co., Reading, MA, January 1989. ISBN: 0-201-15767-5.
- [8] D. E. Goldberg. *The Design of Innovation: Lessons from and for Competent Genetic Algorithms*, volume 7 of *Genetic Algorithms and Evolutionary Computation*. Kluwer Academic Publishers, June 2002. ISBN: 1-4020-7098-5.
- [9] D. E. Goldberg and K. Deb. A comparative analysis of selection schemes used in genetic algorithms. *Foundations of Genetic Algorithms*, 1:69–93, 1991.
- [10] D. E. Goldberg and G. Liepens. Theory tutorial, 1991. (Tutorial presented at the 1991 International Conference on Genetic Algorithms, La Jolla, CA).
- [11] G. Harik, E. Cantú-Paz, D. E. Goldberg, and B. L. Miller. The gambler’s ruin problem, genetic

- algorithms, and the sizing of populations. *Proceedings of the 1997 IEEE International Conference on Evolutionary Computation*, pages 7–12, 1997.
- [12] S.-Y. Ho and X.-I. Chang. An efficient generalized multiobjective evolutionary algorithm. In *Proceedings of the Genetic and Evolutionary Computation Conference 1999: Volume 1*, pages 871–878. Morgan Kaufmann Publishers, 1999.
- [13] S.-Y. Ho, L.-S. Shu, and J.-H. Chen. Intelligent evolutionary algorithms for large parameter optimization problems. *IEEE Transaction on Evolutionary Computation*, 8(6):522–541, December 2004.
- [14] J. H. Holland. *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor, MI, 1975. ISBN: 0-262-58111-6.
- [15] H. Ishibuchi and Y. Shibata. Mating scheme for controlling the diversity-convergence balance for multiobjective optimization. In *Proceeding of Genetic and Evolutionary Computation - GECCO 2004, Part I*, volume 3102 of *Lecture Notes in Computer Science*, pages 1259–1271. Springer, 2004.
- [16] K. KrishnaKumar, S. Narayanaswamy, and S. Garg. Solving large parameter optimization problems using a genetic algorithm with stochastic coding. *Genetic Algorithms in Engineering and Computer Science*, 1995. G. Winter, J. Periaux, M. Galán, P. Cuesta (Eds), John Wiley & Sons.
- [17] M. Laumanns, L. Thiele, and E. Zitzler. Running time analysis of multiobjective evolutionary algorithms on pseudo-boolean functions. *IEEE Transactions on Evolutionary Computation*, 8(2):170–182, April 2004.
- [18] B. L. Miller. *Noise, sampling, and efficient genetic algorithms*. doctoral dissertation, University of Illinois at Urbana-Champaign, Urbana, 1997.
- [19] B. L. Miller. *Noise, Sampling, and Efficient Genetic Algorithms*. PhD thesis, University of Illinois at Urbana-Champaign, Urbana, IL, May 1997.
- [20] B. L. Miller and D. E. Goldberg. Genetic algorithms, selection schemes, and the varying effects of noise. *Evolutionary Computation*, 4(2):113–131, 1996.
- [21] H. Mühlenbein and D. Schlierkamp-Voosen. Predictive models for the breeder genetic algorithm: I. Continuous parameter optimization. *Evolutionary Computation*, 1(1):25–49, 1993.