Structural Engineering and Mechanics, Vol. 26, No. 2 (2007) 191-210
DOI: http://dx.doi.org/10.12989/sem.2007.26.2.191

Discrete approaches in evolution strategies based optimum design of steel frames

O. Hasançebi†

Department of Civil Engineering, Middle East Technical University, 06531 Ankara, Turkey

(Received July 14, 2006, Accepted November 24, 2006)

Abstract. The three different approaches (reformulations) of evolution strategies (ESs) have been proposed in the literature as extensions of the technique for solving discrete problems. This study implements an extensive research on application, evaluation and comparison of them in discrete optimum design of steel frames. A unified formulation is first developed to explain these approaches, so that differences and similarities between their inherent search mechanisms can clearly be identified. Two examples from practical design of steel frames are studied next to measure their performances in locating the optimum. Extensive numerical experimentations are performed in both examples to facilitate a statistical analysis of their convergence characteristics. The results obtained are presented in the histograms demonstrating the distribution of the best designs located by each approach. In addition, an average improvement of the best design during the course of evolution is plotted in each case to compare their relative convergence rates.

Keywords: structural optimization; evolution strategies; discrete optimization; steel frames.

1. Introduction

Starting from the early 1990's, the research in structural optimization has focused on the utilization of new optimization techniques in an effort to catch up with increasing demands of the problems encountered in the field (Hajela 1999). In this context, genetic algorithms (GAs) have probably received the uppermost interest and credit due to their robustness and generality. Today, a vast amount of publications have been accumulated in the literature on the versatile use and successful applications of GAs. It is a well-known fact that GAs belong to a family of heuristic search and optimization techniques known as evolutionary algorithms (EAs). The explored potentials and promises of GAs, not only brought the technique a worldwide popularity in structural optimization, but also led to opening of research on other EA techniques, including evolution strategies (ESs).

The fundamentals of ESs were originally laid in the pioneering studies of Rechenberg (1965, 1973) and Schwefel (1965) at the Technical University of Berlin. They developed the first (simplest) variant of ESs, which implements on the basis of two designs; a parent and an offspring individual. Today, the modern variants of ESs are accepted as $(\mu + \lambda) - \text{ES}$ and $(\mu, \lambda) - \text{ES}$, which were again developed by Schwefel (1977, 1981). Both variants employ design populations

† Assistant Professor, E-mail: oguzhan@metu.edu.tr

consisting of μ parent and λ offspring individuals, and are intended to carry out a self-adaptive search in continuous design spaces. The extensions of these variants to solve discrete optimization problems were put forward in the following three studies in the literature: Cai and Thierauf (1993), Bäck and Schütz (1995), and Rudolph (1994). Amongst them, the one proposed by Cai and Thierauf (1993) refers to a non-adaptive reformulation of the technique and has probably found the most applications in discrete structural optimization, which were reported in Cai and Thierauf (1996), Papadrakakis and Lagaros (1998), Lagaros *et al.* (2002), and Rajasekaran *et al.* (2004). The approach proposed by Bäck and Schütz (1995) corresponds to an adaptive reformulation of the technique, which incorporates a self-adaptive strategy parameter called mutation probability. A literature survey turns up a few recent publications reporting a successful use of this approach in discrete optimum design of structural systems (Papadrakakis *et al.* 2003, Ebenau *et al.* 2005). Another adaptive reformulation of ESs is presented by Rudolph (1994) for general non-linear mathematical optimization problems. It is to the best knowledge of the author that this approach has not been used in any application of engineering optimization yet. The relative effectiveness of discrete approaches of ES in structural optimization still poses a question due to a lack of research on this particular topic. Besides, additional research is necessary to uncover potentials of Rudolph's approach in the field.

In the present study, the discrete reformulations of ESs are thoroughly investigated in the context of optimum design of steel frames. First, the fundamental principles of ESs are briefly overviewed to provide an adequate insight into their robust search mechanism. Next, the three abovementioned discrete approaches of the technique are explained based on a unified formulation. A refinement of the Rudolph's approach is also accomplished here, resulting in an increased performance of the approach for the problems of interest. The success and effectiveness of the approaches in optimum design of steel frames are experimented and quantified using two numerical examples from professional practice. For both examples, a certain number of independent runs are conducted with each approach to measure its stochastic performance in locating the optimum. The measured performances of the approaches are then processed to establish their relative convergence rates and reliabilities. The results are thoroughly discussed and certain conclusions are drawn regarding their more effective use in structural optimization applications.

2. Evolution strategies

The modern variants of ESs, i.e., $(\mu + \lambda) - \text{ES}$ and $(\mu, \lambda) - \text{ES}$, employ an optimization procedure featuring a generation based iteration of the technique for a population of individuals (potential solutions). Accordingly, the first step in this procedure is to establish an initial population consisting of μ number of parent individuals. Typically, this is performed through a random initialization. Next, the initial population is measured (evaluated), and each parent individual is given a fitness score. The fitness score of an individual quantifies the merit of its solution, and is distributed according to how well the individual satisfies the objective function and constraints. Once the initial population is evaluated, it goes through recombination and mutation evolutionary operators to yield the offspring population. Recombination is applied first, facilitating a trade of design information between the μ parents to generate λ new (offspring) individuals. A variety of different recombination operators exist to carry out this task and their relative effectiveness is an ongoing research subject (Bäck 1996). Mutation is applied to each offspring individual thereafter, resulting in a new set of design components for the individual. It is essential to emphasize that mutation is conceived of as the main operator of ESs owing to its dominating role in an effective search of the design space. Evaluation of the offspring population is then fulfilled the same way as the parents. Next, selection operator is implemented to determine the surviving individuals out of parent and offspring populations. The manner the selection is carried out identifies the only difference between $(\mu + \lambda)$ and (μ, λ) variants of ESs. In (μ, λ) variant, the parents are all left to die out, and the best μ individuals are chosen deterministically out of λ offspring in reference to their fitness scores. In $(\mu + \lambda)$ variant, however, the parents are also involved in this mechanism, and the best μ individuals are chosen from μ parents and λ offspring. It follows that the lifespan of an individual in $(\mu, \lambda) - \text{ES}$ is strictly defined as one generation, whereas the individual is permitted to remain alive until it is overwhelmed by subsequent individuals in $(\mu + \lambda)$ variant. At a first glance, $(\mu + \lambda) - \text{ES}$ seems to be more superior in the sense that it increases the selection pressure and comes up with a promise of guaranteed evolution (Bäck 1996). However, according to Schwefel (1981), this opinion may be deceptive. He hypothesized that $(\mu + \lambda)$ variant is less likely to escape from mis-adapted strategy parameters and local optima, as compared to $(\mu, \lambda) - \text{ES}$. In fact, a recent study by the author (2006) produced results justifying this hypothesis. The selected (surviving) individuals make up the parent population of the next generation. The aforementioned process is iterated in the same way over a certain number of generations, producing a new offspring population from the parent one at each generation. Further details and implementation specifics of the ES optimization procedure can be found in Bäck (1996), Bäck and Schwefel (1993, 1995), Beyer and Schwefel (2002), and others.

3. Discrete approaches

As stated previously, the three extensions of ESs to solve discrete optimization problems were proposed by Cai and Thierauf (1993), Bäck and Schütz (1995), and Rudolph (1994). They all employ the general optimization routine of $(\mu + \lambda)$ and (μ, λ) variants of ESs discussed above, and differ from each other in terms of the application of mutation only. In the following, a general formulation of the mutation in ESs is presented first. Next, implementation of the mutation in these approaches is discussed with respect to a unified formulation to achieve comparability amongst their search strategies.

3.1 Mutation

In a discrete reformulation of ESs, an individual $(\vec{a})$ consists of two sets of components, which are defined as follows

$$\vec{a} = \vec{a}(\vec{x}, \vec{s}) \tag{1}$$

In Eq. (1), $\vec{x} = [x_1 \ldots x_i \ldots x_n]$ stands for the design vector, corresponding to that component of the individual where the information related to n number of independent design variables is stored. The second component $\vec{s}$ represents the set of strategy parameters employed by the individual for establishing an automated problem-specific search mechanism in exploring the design space. An online self-update of these parameters is performed during the search, so that varying characteristics of the design space are instantly considered, and the search strategies are refined accordingly. This capability of ESs is characterized by the term "self-adaptation" by Schwefel (1981), and plays an important role in the success of the technique.

Every offspring individual is subjected to mutation, resulting in a new set of values for the design variables ($\vec{x}'$) and strategy parameters ($\vec{s}'$) of the individual, Eq. (2). This implies that not only the design information, but also the search strategy of the individual is altered during this process.

$$mut(\vec{a}(\vec{x}, \vec{s})) = \vec{a}'(\vec{x}', \vec{s}') \tag{2}$$

As a general procedure, mutation of the strategy parameters is performed first. The mutated values of the strategy parameters are then used to mutate the design vector. Mutation of the design vector causes the individual to move to a new point within the design space, and can be formulated as follows

$$\vec{x}' = \vec{x} + \vec{z} \tag{3}$$

where $\vec{z} = [z_1 \ldots z_i \ldots z_n]$ refers to an n-dimensional random vector. The mutated design vector $\vec{x}' = [x_1' \ldots x_i' \ldots x_n']$ is simply obtained by adding this random vector to the unmutated design vector $\vec{x}$.

As far as discrete size optimum design of structures is concerned, design variables correspond to cross-sectional areas of the structural members, which are chosen from ready sections in a given profile list. To identify different sections in a profile list, each section is indicated with a separate index number between 1 and n_s, where n_s denotes the total number of ready sections in the profile list. It is essential to highlight that the application of mutation for these problems is actually performed using these indexes. That is to say, a design variable initially corresponding to x_i-th ready section of the profile list is assigned to $x_i + z_i$-th section, after Eq. (3) is performed.

3.2 The approach proposed by Bäck and Schütz

In the reformulation of technique proposed by Bäck and Schütz (1995), an individual is defined as follows

$$\vec{a} = \vec{a}(\vec{x}, \vec{s}(\vec{p})) \tag{4}$$

where $\vec{p} = [p_1 \ldots p_i \ldots p_n]$ is referred to as the vector of mutation probability, and represents the set of adaptive strategy parameters. They are used to control (adjust) probabilities of the design variables to undergo mutation. In its most general formulation, each design variable (x_i) is coupled with a separate mutation probability (p_i), yielding n mutation probabilities in all. Nevertheless, it has been experimented that the general form suffers from a poor convergence behaviour, and on the contrary the algorithm exhibits a satisfactory performance when a single mutation probability (p) is used for all the design variables of an individual (Bäck and Schütz 1995). Consequently, the number of mutation probabilities (strategy parameters) employed per individual is set to one, i.e., $\vec{a} = \vec{a}(\vec{x}, p)$.

Mutation is performed such that the strategy parameter p is mutated first using a logistic normal distribution (Eq. 5), which assures that the mutated value of p always remains within a range (0,1).

$$p' = \left(1 + \frac{1-p}{p} \cdot e^{-\gamma \cdot N(0,1)}\right)^{-1} \tag{5}$$

In Eq. (5), p' stands for the mutated value of p, and $N(0,1)$ represents a normally distributed random variable with expectation 0 and standard deviation 1. The factor γ here refers to the learning rate of p, and is set to the following recommended value: $\gamma = 1/\sqrt{2\sqrt{n}}$. Once p' is obtained from Eq. (5), the design vector $(\vec{x})$ of the individual is mutated next as in Eq. (6).

$$z_i = \begin{cases} 0 & , \text{ if } r_i > p' \in [0, 1] \\ u_i \in \{-x_i + 1, \ldots, n_s - x_i\}, & \text{if } r_i \le p' \in [0, 1] \end{cases} \tag{6}$$

In this process, for each design variable x_i a random number r_i is generated anew in a real interval [0,1]. If $r_i > p'$, the variable is not mutated, that is $z_i = 0$ and $x_i' = x_i$. Otherwise $(r_i \le p')$, it is mutated according to a uniform distribution based variation, in which a uniformly distributed integer random number (u_i) sampled between $-x_i + 1$ and $n_s - x_i$ is assigned to z_i. In this way, mutated value of the design variable $(x_i' = x_i + z_i)$ is enforced to remain within 1 and n_s with all sections having an equal probability of being selected.

3.3 The approach proposed by Cai and Thierauf

In the discrete reformulation by Cai and Thierauf (1993), an individual is described with a null set of adaptive strategy parameters $\vec{s}(\phi)$, as follows

$$\vec{a} = \vec{a}(\vec{x}, \vec{s}(\phi)) \tag{7}$$

Mutation probability (p) is also employed here. Unlike the former approach, however, it is set to an appropriate static value between 0.1 and 0.4 throughout the optimization process (Cai and Thierauf 1996). This implies that every time a predefined percentage of design variables is probabilistically mutated for all the individuals, as in Eq. (8).

$$z_i = \begin{cases} 0 & , \text{ if } r_i > p \in [0, 1] \\ \pm(\kappa_i + 1), & \text{if } r_i \le p \in [0, 1] \end{cases} \tag{8}$$

Again here, for each design variable x_i a random number r_i is generated anew in a real interval [0,1], and is compared with the constant mutation probability. In case of $r_i \le p$, the variable is mutated according to a Poisson distribution based variation. For this, a Poisson distributed integer random number (κ_i) is sampled first, and either a positive or negative value of $\kappa_i + 1$ is then assigned to z_i under equal probability. In Statistics, the Poisson distribution is described by the following probability function

$$P(\kappa) = \frac{(c)^{\kappa}}{\kappa!} e^{-c}, \quad \kappa \in \{0, 1, 2, \ldots, +\infty\} \tag{9}$$

where the parameter c corresponds to both the mean and variance of the distribution. For some selected values of this parameter ($c = 3, 5, 8, 10$ and 15), a graphical representation of the probability distribution is plotted in Fig. 1.

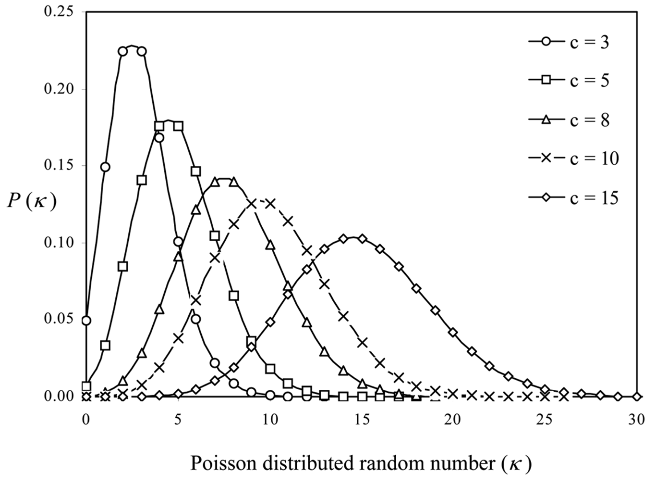


Fig. 1 Poisson distribution for some selected values of c

3.4 The approach proposed by Rudolph

Another adaptive reformulation of ESs is developed by Rudolph (1994) for solving general non-linear mathematical optimization problems with unbounded integer design spaces. In this approach, mutation of a design variable is performed based on a geometric distribution in the form of

$$P(g) = \frac{1}{\psi + 1}\left(1 - \frac{1}{\psi + 1}\right)^{g}, \quad g \in \{0, 1, 2, \ldots, +\infty\} \tag{10}$$

where g represents a geometrically distributed integer random number, and ψ corresponds to the mean (expectation) of this particular distribution. For some selected values of this parameter (ψ= 5, 10, 20 and 40), the variation in probability distribution pattern is shown in Fig. 2.

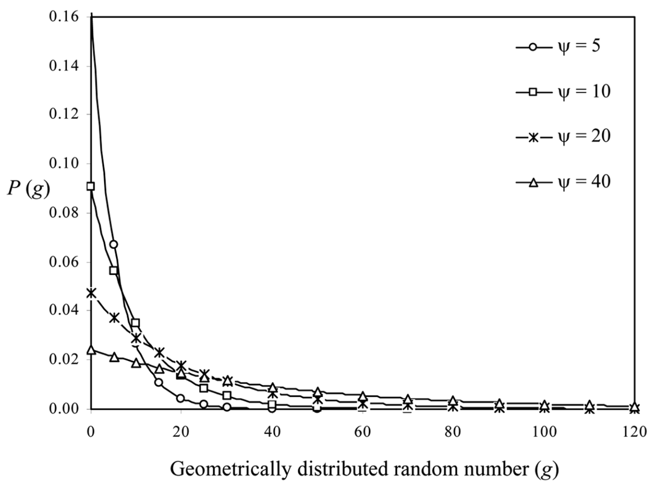


Fig. 2 Geometric distribution for some selected values of ψ

Rudolph's approach basically rests on a variable-wise and adaptive implementation of the parameter ψ throughout the search. The idea here is to let each variable develop a useful probability distribution pattern of its own (by adjusting ψ) for successful applications of mutation. Consequently, each design variable x_i of an individual is coupled with a different ψ_i, $i \in \{1, 2, \ldots, n\}$ parameter, and the individual is described as follows

$$\vec{a} = \vec{a}(\vec{x}, \vec{s}(\vec{\psi})) \tag{11}$$

According to Rudolph's approach, all design variables of an individual are subjected to mutation. When interpreted in view of discrete function optimization in mathematics, this strategy is plausible, as it causes an n-dimensional mutation of the individual to a next grid point in the vicinity of the former. However, structural optimization problems are such that the overall behaviour of a structural system might be very sensitive to changes in a few design variables owing to significant variations in the properties of ready sections. For a successful operation of mutation for these problems, it is essential to limit the number of design variables mutated at a time in an individual, as practiced by the former approaches. To this end, a refinement of Rudolph's approach is accomplished here, where the parameter p is incorporated and coupled with the original set of strategy parameters $\vec{\psi}$ for a harmonized implementation of the mutation operator. Accordingly, in the refined form of the Rudolph's approach, an individual is described as follows

$$\vec{a} = \vec{a}(\vec{x}, \vec{s}(p, \vec{\psi})) \tag{12}$$

In this framework, the parameter p is mutated first via Eq. (5). Analogous to former approaches, a random number $r_i \in [0,1]$ is then generated anew for each design variable x_i and its associated strategy parameter ψ_i. If $r_i > p'$, neither x_i nor ψ_i is mutated, i.e., $\psi_i' = \psi_i$ and $z_i = 0$. If not, ψ_i is mutated first according to a lognormal distribution based variation (Eq. 13), and is enforced to remain greater than 1.0 to preserve effectiveness of the mutation operator.

$$\psi_i' = \begin{cases} \psi_i & , \text{ if } r_i > p' \in [0, 1] \\ \psi_i \cdot e^{\tau \cdot N_i(0,1)} \geq 1.0 & , \text{ if } r_i \leq p' \in [0, 1] \end{cases} \tag{13}$$

In Eq. (13), ψ_i' stands for the mutated value of ψ_i. The factor τ here refers to the learning rate of this parameter, and is set to a recommended value of $1/\sqrt{n}$ for all individuals (Rudolph 1994). Then, two geometrically distributed integer random numbers $(g_{i,1}, g_{i,2})$ are sampled using the value of ψ_i', and x_i is mutated by the difference of these two numbers, Eq. (14).

$$z_i = \begin{cases} 0 & , \text{ if } r_i > p' \in [0, 1] \\ g_{i,1} - g_{i,2} & , \text{ if } r_i \leq p' \in [0, 1] \end{cases} \tag{14}$$

As a final point, it is worthwhile to mention that most programming language libraries fall short of providing a function to sample the geometrically distributed numbers $g_{i,1}, g_{i,2}$. However, one can easily generate them using Eq. (15).

$$g_{i,1}, g_{i,2} = \left\lfloor \frac{\log(1 - r_i)}{\log(1 - 1/(1 + \psi_i'))} \right\rfloor \tag{15}$$

4. Numerical examples

In this section, two examples will be studied thoroughly to investigate numerical success and effectiveness of the abovementioned discrete approaches of ESs in sizing of steel frames for minimum weight. A complete mathematical formulation of the optimum design problem is posed in Saka (1991), Huang and Arora (1997), etc., and thus will not be repeated here. The computational specifics of the algorithm developed to solve the numerical examples can be found in Hasançebi and Ulusoy (2005). Furthermore, a stepwise implementation of the algorithm is illustrated in Ulusoy *et al.* (2004) for a simple frame example. In all the tests performed in the following, a (μ, λ) variant of ESs is employed with the following population parameters: $\mu = 15, \lambda = 100$.

4.1 Industrial building

The first example is an industrial building (Fig. 3), which is taken from the existing literature (Grierson and Lee 1984, Erbatur *et al.* 2000). It consists of two side frames and a gable roof in between, and has 27 members in all. The structure is subjected to three loading conditions considering various combinations of lateral wind load and vertical loads, such as roof, floor and crane loads. The symmetry of the structure about centerline plus the continuity requirement are employed to group 27 members into 12 independent size variables, as shown in Fig. 3. In addition to the beam and column members of the side frames (groups 1 through 6), the top and bottom chord members of the gable roof (groups 7 and 8) are designed as axial-flexural members. On the other hand, the vertical and inclined web members (groups 9 through 12) of the gable roof are exposed to axial forces only, and thus they are designed as axial members. All members are to be sized using standard sections in AISC-ASD (1989) specification, as follows: the beam and column members are to be selected from 295 wide-flange sections (W); the chord members are to be selected from 267 tee sections cut from wide-flange (WT); and the web members are to be selected

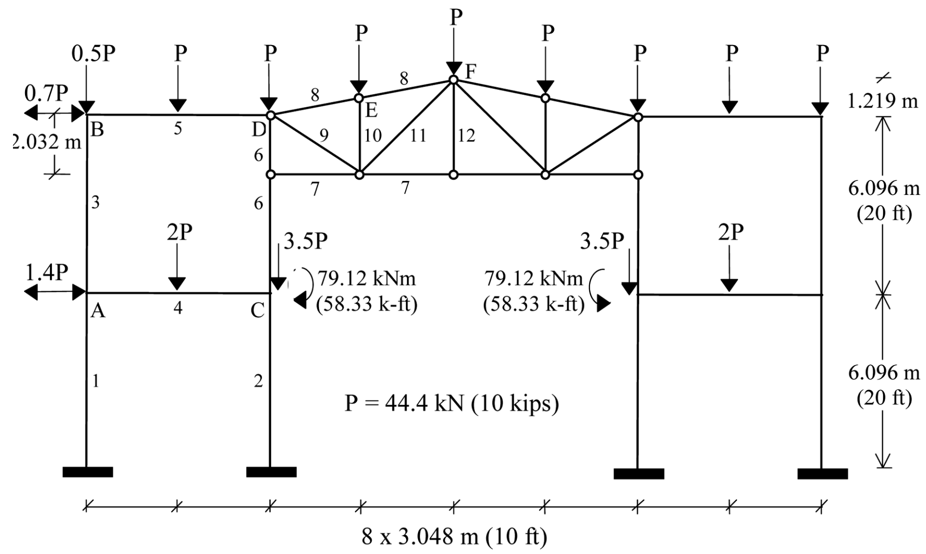


Fig. 3 Industrial building

Table 1 The design data for industrial building

Constraint Data		
Displacement constraints:	Lateral displacements of joints A, B, C, D, E, F for load cases 2 and 3 ≤ 3.81 cm (1.5 in)	
Stress constraints:	Allowable stress in tension ≤ 165.48 MPa (24 ksi) Allowable stress in comp. ≤ 110.32 MPa (16 ksi)	
Loading Data		
Case number	Load combinations, P = 44.4 kN (10 kips)	
1	roof + floor + crane loads	
2	case I + wind load from left	
3	case I + wind load from right	
Profile Lists for Structural Members		
Group number	Behaviour	Profile List (AISC)
1, 2, 3, 4, 5, 6	flexural members	wide flange (W)
7, 8	flexural members	tee cut from wide-flange (WT)
9, 10, 11, 12	axial members	double angle (DL)
Material Properties		
Modulus of elasticity	206850 MPa (30×10^3 ksi)	
Density of material	7.85×10^{-6} kg/mm^3 (100.283 lb/in^3)	
Yield stress	248 MPa (36 ksi)	

from 91 double angle sections (DL). The constraints are imposed on combined member stresses in tension and compression, as well as on horizontal displacements of some selected nodes. The complete and detailed list of design data is given in Table 1.

The three discrete approaches of ESs are applied to solve the problem under a total of six test cases. In this context, Bäck and Schütz approach is implemented under two test cases. In the first test case (BS_1), a standard Bäck and Schütz algorithm is applied with an initial mutation probability of $p_{ini} = 0.25$ for all individuals. The second test case (BS_2) is identical to the first one, except that p is enforced to remain above an imposed minimum mutation probability of $p_{min} = 1/n_d = 0.08$ throughout the search. Three test cases are considered with Cai and Thierauf approach by choosing the mean (variance) of the Poisson distribution as $c = 3$, 5, and 8 under a constant mutation probability of $p = 0.20$, referred to as CT_1, CT_2, and CT_3 in the following, respectively. Finally, a single test case is considered with refined Rudolph (RR) approach based on the choice of the following values: $p_{ini} = 0.25$, $\psi_{ini} = 10$, and $p_{min} = 0.08$.

Considering the probabilistic nature of ESs, for each test case twenty independent runs are performed to obtain statistically significant data for measuring the stochastic performance of the approaches in locating the optimum. The generation number in these runs is adjusted to 400 in order to limit computational requirements to a justifiable amount. Hence, a total of $400 \times 100 = 40000$ function evaluations are carried out in each run, which takes 4 min of computing time on a PC with Pentium IV 2.4 GHz processor. Amongst a total of 120 independent runs performed in this way, a solution with the minimum design weight of 3466.7 kg has been located twice by CT_2 and three times by both CT_3 and RR. This design (shown in Table 2) is considered to be the optimum solution of the industrial building problem reached in the present study, and it is better than the

Table 2 The optimum design for the industrial building

Members	Sizing Group	Ready section	Area, cm^2 (in^2)
Beam and Column Members	1	W6 × 9	17.29 (2.68)
	2	W21 × 68	129.03 (20.00)
	3	W12 × 14	26.84 (4.16)
	4	W21 × 44	83.87 (13.00)
	5	W12 × 16	30.39 (4.71)
	6	W14 × 22	41.87 (6.49)
Chord Members	7	WT3 × 4.25	8.06 (1.25)
	8	WT3 × 6	11.48 (1.78)
Web Members	9	DL2 × 2 × 3/16	9.23 (1.43)
	10	DL2 × 2 × 1/8	6.19 (0.96)
	11	DL2 × 2 × 1/8	6.19 (0.96)
	12	DL2 × 2 × 1/8	6.19 (0.96)
Weight (kg)	3466.71 kg (7642.65 lb)		

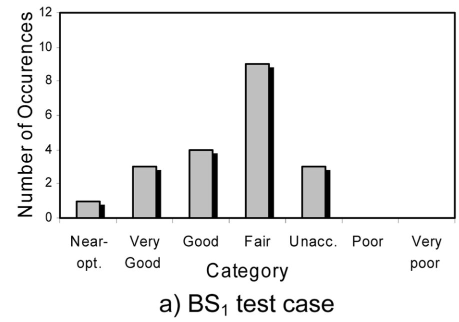
Table 3 The mean best designs and their standard deviations obtained in the six test cases of the industrial building

Discrete Approach	Test Case	Parameter Sets	Mean Best Design (kg)	Standard Deviation
Bäck and Schütz	BS_1	$p_{ini} = 0.25$	3590.86	59.57
	BS_2	$p_{ini} = 0.25$, $p_{min} = 0.08$	3538.05	45.56
Cai and Thierauf	CT_1	$p = 0.20$, $c = 3$	3662.23	218.75
	CT_2	$p = 0.20$, $c = 5$	3548.89	91.56
	CT_3	$p = 0.20$, $c = 8$	3505.43	34.02
Refined Rudolph	RR	$p_{ini} = 0.25$, $\psi_{ini} = 10$, $p_{min} = 0.08$	3494.90	16.98

existing solutions in the literature. The optimum weight was reported to be 4078.6 kg and 3943.2 kg in Grierson and Lee (1984) and Erbatur *et al.* (2000), respectively. In fact, a comprehensive study by Hasançebi and Ulusoy (2005) evaluating the performance of ESs in some benchmark problems led to the conclusion that they are very powerful search tools.

For each test case, the best designs obtained in twenty runs are averaged to find the mean best design and its corresponding standard deviation (Table 3). In addition, the distributions of the best designs within the six test cases are plotted in the histograms shown in Fig. 4. For this, the best designs are classified into the seven categories such as, near-optimum, good, fair, poor, etc., reflecting various levels of their quality when compared to the optimum. For example, a design which is 0-1% higher than the optimum (i.e., between 3466.7-3501.4 kg) is conceived as a near-optimum solution, or a design is deemed as poor if it is 7-10% higher than the optimum, etc. The abscissa in a histogram shows these categories, whereas the numbers of occurrences of runs falling into these categories are shown in the ordinate.

A comparison of Figs. 4(a) and 4(b) clearly documents the significance of p_{min} parameter in the BS approach. It has been observed that as the search process carries on, the parameter p usually



Category	% Range	Weight Range (kg)
Near-optimum	% 0 – 1	3466.7 - 3501.4
Very good	% 1 – 2	3501.5 - 3536.0
Good	% 2 – 3	3536.1 - 3570.7
Fair	% 3 – 5	3570.8 - 3640.0
Unacceptable	% 5 – 7	3640.1 - 3709.4
Poor	% 7 – 10	3709.5 - 3813.4
Very Poor	> % 10	>3813.5

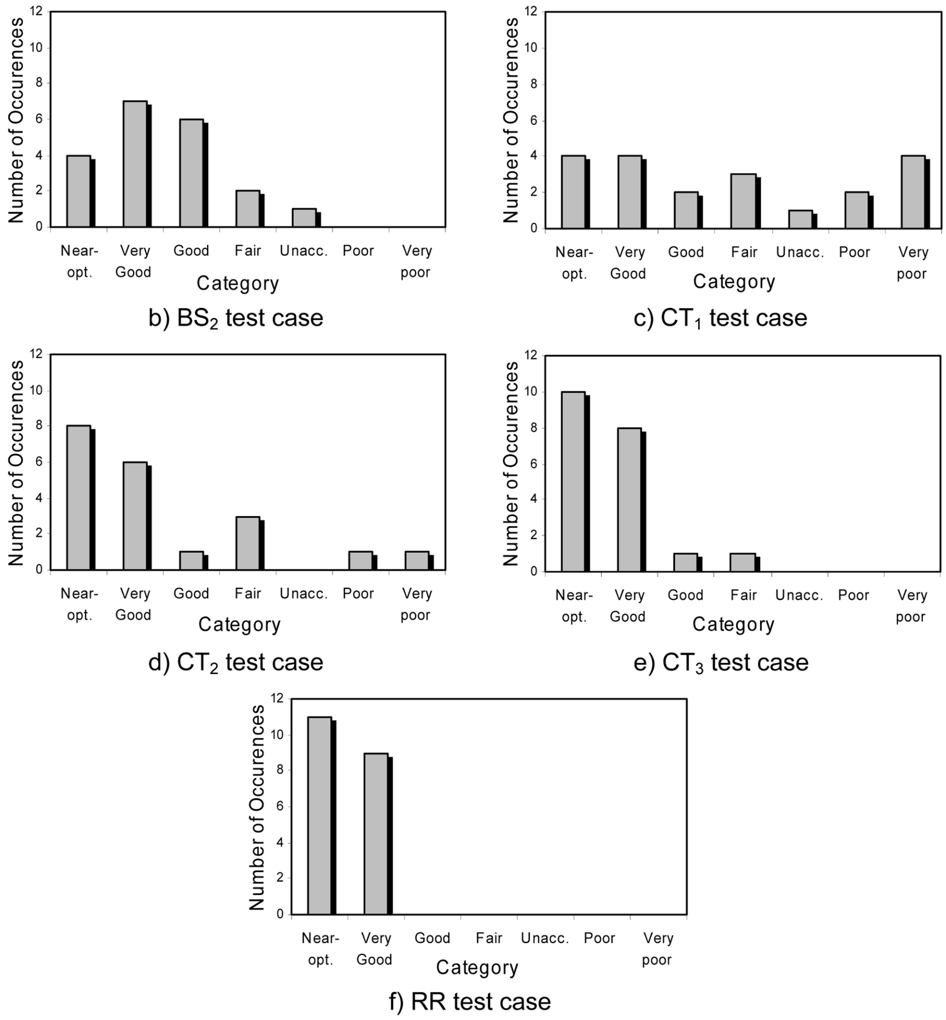


Fig. 4 The distribution of the best designs (histograms) in the six test cases of industrial building

tends to approach values around 0, resulting in a completely lost search ability. The utilization of p_{min} then guarantees preservation of a minimum search ability throughout the process, and thus an improved distribution is produced in BS_2 as compared to BS_1. Also, it is remarkable to note that despite the fact that BS_1 and BS_2 located near-optimum solutions only a few times, in none of the runs performed they got stuck in poor or very poor local optima. This is due to a uniform distribution based search strategy, such that mutation performed in wide design regions reduces the probability of entrapment in poor local optima, yet hampers an exploitative search to reach better design points in favourable regions.

A comparison of the histograms in Figs. 4(c), 4(d) and 4(e) reveals that the parameter c has a strong influence on the distribution pattern and performance of the CT approach. In contrast to a relatively flat distribution of CT_1 indicating a large diversity of local optima located, CT_3 identified a high number of successful runs yielding best designs in near-optimum and very good categories. Apparently, the use of wider design transitions for mutation in CT_3 let the algorithm escape from local optima more easily, in comparison to other two test cases. It is worthwhile to mention that a further increase of c beyond eight did not seem to cause an additional improvement in the performance of the algorithm. Since small mutations occur seldom in this case due to the nature of Poisson distribution, the advantage gained with wide design transition is offset by a reduction in the exploitative search ability.

As seen from Fig. 4(f), the best performance amongst the test cases is exhibited by RR with eleven near-optimum and nine very good best designs. The mean best design in this test case turns out be 3494.9 kg with a small standard deviation of 16.98, reflecting its high convergence reliability. Apart from a self-adaptive manipulation of step size parameters (ψ), RR approach takes advantage of utilizing a geometric distribution based variation for the application of mutation. In fact, this particular distribution has two very useful features that contribute to an effective search of the design space. Firstly, for all values of ψ, small mutations are always favoured in comparison to large ones. Secondly, even under small values of ψ, it yields a probability distribution facilitating an occasional occurrence of large mutations. As a result, an exploitative search is mainly conducted by the algorithm, however, design transitions in wide regions are permitted from time to time to escape from local optima.

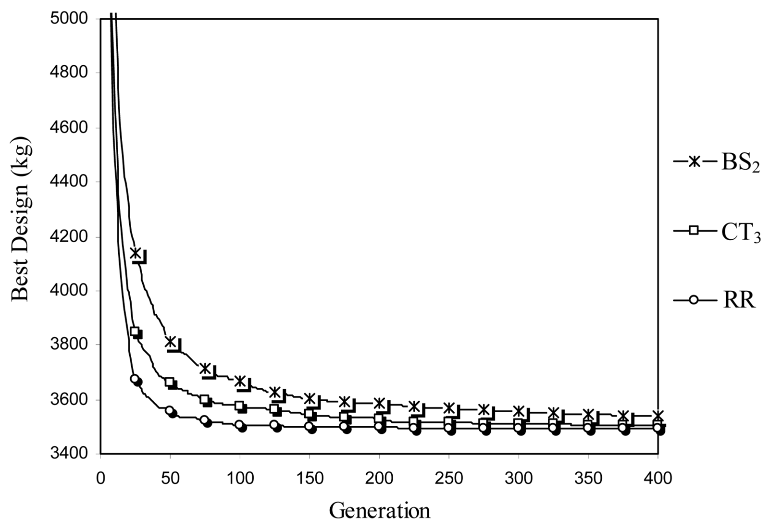
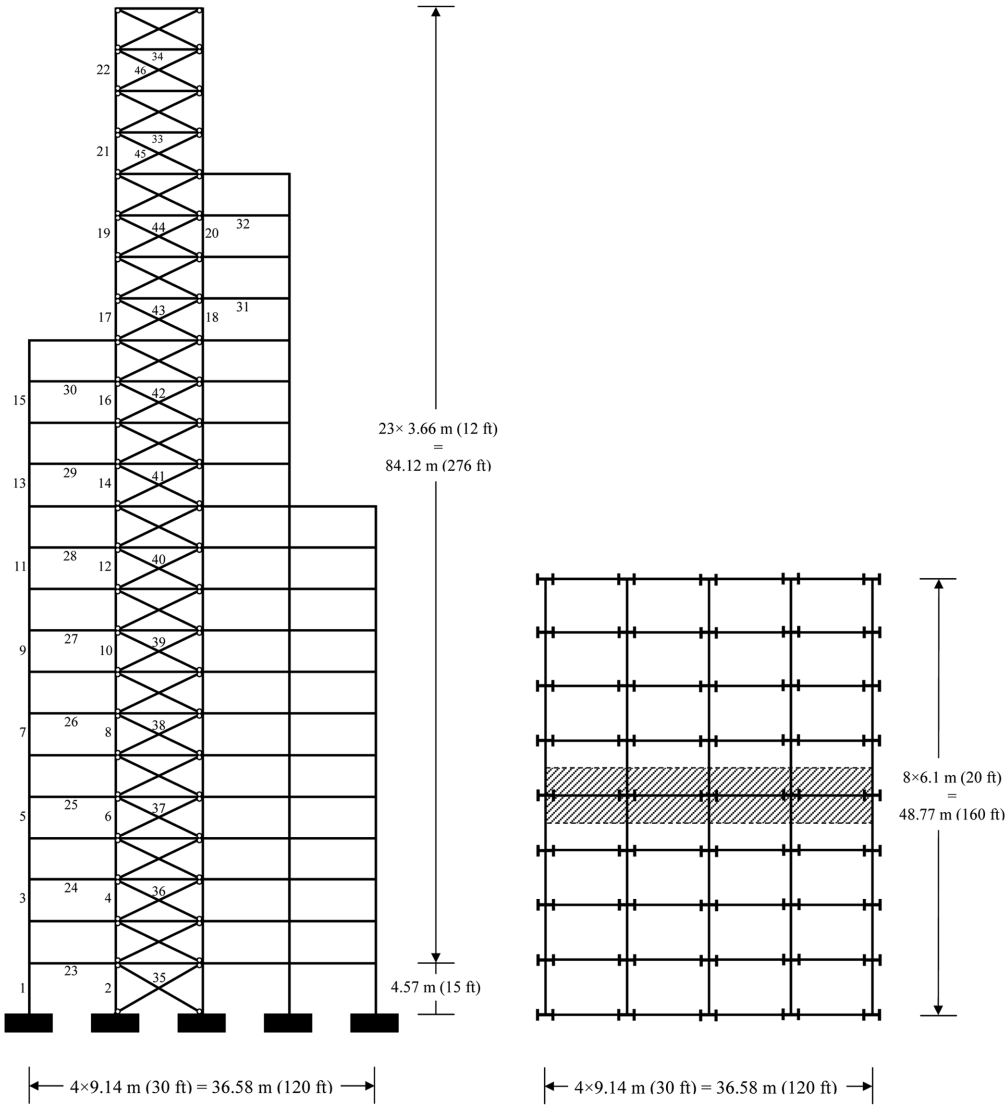


Fig. 5 The average variation of the best design during the course of evolution in BS_2, CT_3 and RR test cases of the industrial building

In order to assess convergence rates of the approaches, variation of the best design during the course of evolution in BS_2, CT_3 and RR test cases is plotted in Fig. 5, after averaging the results of twenty runs for each. It is observed that all the approaches exhibit a linear convergence behavior up to around 25 generations, and the rate of convergence decreases continually thereafter. It appears that RR is the fastest approach for this problem, while BS_2 is the slowest one. Assuming that an acceptable design is higher than the optimum by 5% or lesser, such a design is first produced at 112th generation with BS_2, at 56th generation with CT_3 and at 29th generation with RR approach on average.



(a) Side view (b) Plan view

Fig. 6 216-member irregular frame

4.2 216-Member irregular frame

The second problem refers to the optimum design of a 4-bay, 216-member irregular braced frame, which serves to measure performances of the discrete approaches in a larger design space. Fig. 6 shows the plan and elevation view of this structure, which actually represents one of the interior frameworks of a steel building along the short side. All the beams and columns of the frame are rigidly connected, while the diagonals of X-braced truss are pin connected. The 216 members of the frame are grouped into 46 independent size variables to satisfy practical fabrication requirements. That is, exterior columns are grouped together as having the same section over two adjacent stories, as are interior columns, beams and diagonals, as indicated in Fig. 6. The wide-flange (W) profile list consisting of 295 ready sections is used to size all the beam and column members, while the diagonals are to be selected from 267 tee sections cut from wide-flange (WT).

For design purpose, the frame is subjected to various types of vertical and lateral loads based on the provisions of ASCE 7-98 (1998). These loads cover dead (D), live (L), snow (S) and wind (W) loads, which are all combined under a single loading condition of D + L + S + W. Dead, live and snow loads are all applied as uniformly distributed loads on the beams using the tributary area shown in Fig. 6. All floors, except the roof are subjected to a design dead load of 2.88 kN/m^2 (60.13 lb/ft^2) and a design live load of 2.39 kN/m^2 (50 lb/ft^2). The beams of the roof level are subjected to the design dead load plus snow load. The design snow load (in kN/m^2) is computed as per ASCE standard using the following equation: $p_s = 0.7C_sC_eC_tIp_g$, where C_s is the roof slope factor, C_e is the exposure factor, C_t is the temperature factor, I is the importance factor, and p_g is the ground snow load. For a heated residential building having a flat and fully exposed roof, these factors are chosen as follows: $C_s = 1.0, C_e = 0.9, C_t = 1.0, I = 1.0$ and $p_g = 1.20$ kN/m^2 (25 lb/ft^2), resulting in a design snow load of 0.75 kN/m^2 (15.75 lb/ft^2). Wind loads are applied to the frameworks as lateral point loads at every floor level, and are computed according to ASCE standard using the following equation: $p_w = (0.613K_zK_{zt}K_dV^2I)(GC_p)$, where p_w is the design wind pressure in kN/m^2, K_z is the velocity exposure coefficient, K_{zt} is the topographic factor, K_d is the wind direction factor, V is the basic wind speed, G is the gust factor, and C_p is the external pressure coefficient. Assuming that the building is located in a flat terrain with a basic wind speed of $V = 65$ m/s (105 mph) and exposure category B, the following values are used for these parameters: $K_{zt} = 1.0, K_d = 0.85, I = 1.0, G = 0.85$, and $C_p = 0.8$ for windward face and −0.5 for leeward face.

The strength/stability requirements of all members are prescribed in accordance with the provisions of AISC-ASD (1989) specification. For displacement constraints, the top story drift and inter-story drifts are limited to a maximum value of H/400 and h/400, respectively, where H is the total building height, and h is the story height.

Each of the three discrete approaches of ESs is applied to solve the problem under a separate test case. In the test case with Bäck and Schütz (BS) approach, mutation probability parameter is initially set to $p_{ini} = 0.25$ for all individuals, and a minimum mutation probability of $p_{min} = 1/n_d = 0.02$ is imposed to assure at least mutation of one design variable per individual on average. The Cai and Thierauf (CT) approach is implemented by setting c and p parameters to constant values of 10 and 0.05, respectively. Finally, for the application of refined Rudolph (RR) approach, p_{ini} is set to 0.25 for all individuals while initial values of ψ are chosen as 20.

For each test case considered, twenty independent runs are conducted by executing the respective algorithms of the discrete approaches over 500 generations. This implies that a total of 500 × 100 =

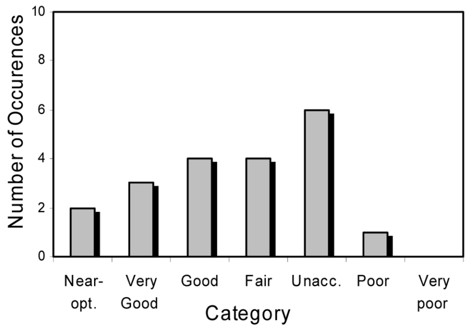
Table 4 The optimum design for the 216-member irregular frame

Members	Sizing Group	Ready section	Area, cm^2 (in^2)	Members	Sizing Group	Ready section	Area, cm^2 (in^2)
Columns	1	W21 × 111	83.12 (32.7)	Diagonals	35	WT12 × 31	23.28 (9.16)
	2	W33 × 387	287.25 (113)		36	WT10.5 × 25	18.71 (7.36)
	3	W24 × 103	77.02 (30.3)		37	WT10.5 × 25	18.71 (7.36)
	4	W40 × 268	200.31 (78.8)		38	WT9 × 20	14.95 (5.88)
	5	W21 × 101	75.75 (29.8)		39	WT9 × 20	14.95 (5.88)
	6	W40 × 244	182.26 (71.7)		40	WT10.5 × 22	16.50 (6.49)
	7	W27 × 84	63.04 (24.8)		41	WT7 × 13	9.79 (3.85)
	8	W27 × 217	162.18 (63.8)		42	WT7 × 11	8.26 (3.25)
	9	W21 × 62	46.52 (18.3)		43	WT7 × 11	8.26 (3.25)
	10	W27 × 194	144.89 (57.0)		44	WT7 × 11	8.26 (3.25)
	11	W24 × 55	41.18 (16.2)		45	WT7 × 15	11.24 (4.42)
	12	W27 × 146	109.05 (42.9)		46	WT7 × 11	8.26 (3.25)
	13	W30 × 116	86.94 (34.2)				
	14	W33 × 130	97.36 (38.3)				
	15	W27 × 102	76.26 (30.0)				
	16	W24 × 104	77.79 (30.6)				
	17	W30 × 99	73.97 (29.1)				
	18	W18 × 97	72.45 (28.5)				
	19	W30 × 90	67.11 (26.4)				
	20	W24 × 84	62.79 (24.7)				
	21	W8 × 48	35.84 (14.1)				
	22	W18 × 40	30.00 (11.8)				
Beams	23	W16 × 77	57.45 (22.6)				
	24	W24 × 94	70.41(27.7)				
	25	W21 × 93	69.40 (27.3)				
	26	W18 × 86	64.31 (25.3)				
	27	W16 × 77	57.45 (22.6)				
	28	W18 × 86	64.31 (25.3)				
	29	W24 × 94	70.41 (27.7)				
	30	W30 × 90	67.11 (26.4)				
	31	W33 × 118	88.21 (34.7)				
	32	W27 × 94	70.41 (27.7)				
	33	W40 × 149	111.34 (43.8)				
	34	W14 × 61	45.50 (17.9)				
Weight (kg)		179294.04 kg (395269.05 lb)					

50000 function evaluations are carried out in each run, which takes 268 min of computing time on a PC with Pentium IV 2.4 GHz processor. Amongst 60 independent runs performed in this way, a minimum design weight of 179294.04 kg is produced by RR. This design is tabulated in Table 4 and is considered to be the optimum solution of the problem reached in the present study.

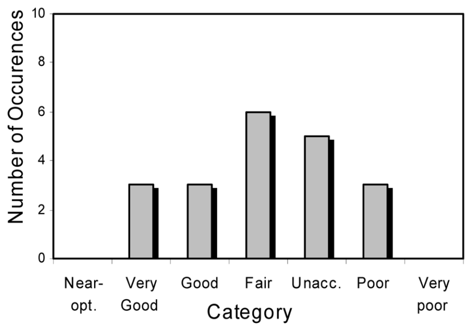
Table 5 The mean best designs and their standard deviations obtained in the three test cases of 216-member irregular frame

Discrete Approach	Test Case	Parameter Sets	Mean Best Design (kg)	Standard Deviation
Bäck and Schütz	BS	$p_{ini} = 0.25$, $p_{min} = 0.02$	186078.75	3788.32
Cai and Thierauf	CT	$p = 0.05$, $c = 10$	187107.26	3805.40
Refined Rudolph	RR	$p_{ini} = 0.25$, $\psi_{ini} = 20$, $p_{min} = 0.02$	183187.77	3242.43

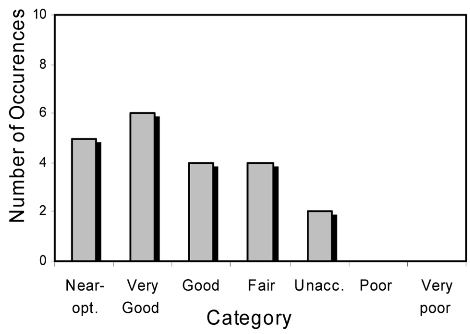


a) BS test case

Category	% Range	Weight Range (kg)
Near-optimum	% 0 – 1	179294.0 - 181087.0
Very good	% 1 – 2	181087.1 - 182879.9
Good	% 2 – 3	182880.0 - 184672.9
Fair	% 3 – 5	184673.0 - 188258.7
Unacceptable	% 5 – 7	188258.8 – 191844.6
Poor	% 7 – 10	191844.7 – 197223.4
Very Poor	> % 10	>197223.5



b) CT test case



c) RR test case

Fig. 7 The distribution of the best designs (histograms) in the three test cases of 216-member irregular frame

The mean best design and its corresponding standard deviation measured in each test case are tabulated in Table 5, and the distributions of the best designs within these test cases are presented in the histograms shown in Fig. 7, using the same quality scale introduced formerly. In comparison to the previous example, somewhat reduced performances of all the three approaches are observed in Fig. 7. The reduction in their performances can be attributed to the fact that the design space of the current problem is not only expanded due to increased number of design variables, but also is strongly controlled by additional constraints stipulated by AISC-ASD (1989) specification.

When compared in terms of relative performances, BS is observed to yield a more favourable histogram than CT for this example. While near-optimum solutions were located in two runs by the former, none of the runs performed with the latter got to produce a design in this category. Likewise, BS got stuck in a poor local optimum only once, whereas this happened three times with CT. These results manifest an increased usefulness of wide design transition based mutation in an

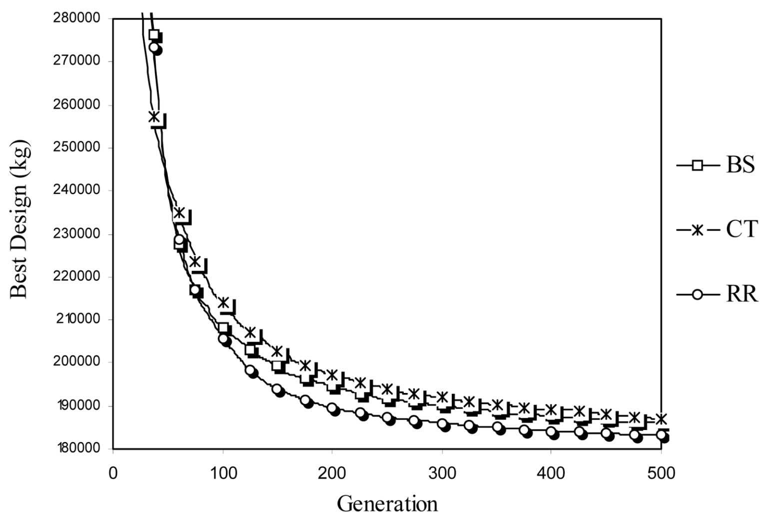


Fig. 8 The average variation of the best design during the course of evolution in the three test cases of 216-member irregular frame

efficient exploration of large and complicated design spaces. However, still the best performance is exhibited by RR approach with five near-optimum and six very good best designs.

It is worth mentioning that extensive numerical experiments need to be performed to determine the useful settings of the CT approach, because its performance is strongly dependent on the choice of its parameter set. For instance, a number of experiments performed with the optimal parameter setting of the industrial building problem (i.e., $p = 0.20$ and $c = 8$) led to all very poor solutions in this example. In the light of the results obtained, it can be stated that an appropriate value of p parameter usually lies in the range of $[1/n_d, 4/n_d]$, changing inversely with the number of design variables used. Also, it is recommended that the parameter c be employed with values between 8 and 12 for sizing problems utilizing ASCE ready sections.

Shown in Fig. 8 is the average variation of the best design during the course of evolution for the three approaches. Interestingly, despite its localized search characteristics, CT exhibits a higher convergence rate than do BS and RR in the first 50 generations. This is due to adaptive parameters incorporated in the latter such that these parameters first need to be adjusted to suitable values required for an efficient search. When a long-term execution of the algorithms is considered, however, RR appears to be the fastest approach again. The first acceptable solution is located at 258th generation with RR, at 362nd generation with BS, and at 441st generation with CT approach on average.

5. Conclusions

In the literature, three different reformulations of ESs have been proposed by Cai and Thierauf (1993), Bäck and Schütz (1995) and Rudolph (1994) as adaptation and extension of the technique to deal with discrete optimization problems. This study has concentrated on application, evaluation and comparison of these approaches in discrete optimum design of steel frames, where members are sized using standard sections.

The principle differences in the approaches extend to the implementation means of adaptation during search process along with the type of distribution used to mutate design variables. Adaptation is a significant feature since it enables an algorithm to establish and upgrade search strategies according to varying regional characteristics of the design space. In Bäck and Schütz (BS) approach, this feature is incorporated into the algorithm through the mutation probability parameter (p), by which the degree of mutability of each individual is tuned online during the search. High values of this parameter usually lead to design transitions to remote regions, and thereby preventing stagnation in poor local optima, whereas its low values are useful for local search. The algorithm's best effort to perform a local search is, however, undermined somewhat by uniform distribution, which completely disregards the concept of locality. In this particular distribution, a design variable is mutated into any value in the discrete set under equal probability.

Cai and Thierauf (CT) approach refers to non-adaptive reformulation of ESs, where a constant mutation probability is implemented in conjunction with Poisson's distribution based mutation. The dominancy between global and local search characteristics of the approach is controlled by the parameter c, the mean and variance of this particular distribution. By increasing c to some level, the global search can be promoted, resulting in increased search efficiency. However, the performance of the approach does not seem to improve more or may even worsen with further increase of this parameter due to degenerated local search ability.

In Rudolph's approach, adaptation is executed in a rather different manner. In this approach, for each design variable of an individual an adaptive parameter ψ is employed to adjust shape and flatness of a geometric distribution used to mutate the variable. In this way, the range of probabilistic variation is tuned for each design variable specifically in order to produce successful moves in the design space. In addition to adaptation strategy of this kind, the enhanced performance of Rudolph's approach can be attributed to the usefulness and suitability of geometric distribution for mutation operation. Through this distribution, small mutations are considered more preferable, yet a non-trivial probability is allocated to large mutations even under low values of ψ to avoid entrapment in local optima. It has been noted, however, that the use of Rudolph's approach in its original formulation may lead to a poor algorithm for sizing problems owing to the application of mutation on every design variable. Hence, a refinement of this approach has been accomplished, referred to as refined Rudolph's (RR) approach, where the parameter p is incorporated and coupled with ψ in order to limit the number of design variables mutated in an individual.

Two numerical examples (industrial building and 216-member irregular frame) have been studied in depth to measure stochastic performances of the three discrete approaches in locating the optimum. The distributions of the best designs located under different test cases are presented in the form of histograms in Figs. 4 and 7 to identify convergence reliabilities of the approaches. Furthermore, the average variations of the best design in these test cases are monitored and plotted in Figs. 5 and 8 to compare their relative convergence rates. In general, the following conclusions are drawn based on the results obtained from numerical experiments. The relative performance of the BS approach enhances with the increasing size of the design space, yet to some extent it suffers from its missing emphasis on an exploitative search. The success of the CT approach is strongly reliant on a suitable choice of its parameter setting, which is problem-specific in most cases. An optimal parameterization of this approach may lead to an effective algorithm for small to moderate size design spaces. Finally, the RR approach optimally combines the two requirements of an efficient search. It primarily performs an exploitative search, yet to a lesser degree wide design transitions are encouraged to escape from local optima.

References

AISC-ASD (American Institute of Steel Construction-Allowable Stress Design). (1989), *Manual of Steel Construction-Allowable Stress Design.*

ASCE 7-98 (American Society of Civil Engineers). (1998), *Minimum Design Loads for Buildings and Other Structures.*

Bäck, T. (1996), *Evolutionary Algorithms in Theory and Practice*, Oxford University Press, New York.

Bäck, T. and Schütz, M. (1995), "Evolutionary strategies for mixed-integer optimization of optical multilayer systems", *Proc. of the 4th Annual Conf. on Evolutionary Programming*, McDonnel, J.R., Reynolds, R.G. and Fogel, D.B. (eds.), MIT Press, Cambridge, MA, 33-51.

Bäck, T. and Schwefel, H.-P. (1993), "An overview of evolutionary algorithms for parameter optimization", *Evolution. Computation*, **1**, 1-23.

Bäck, T. and Schwefel, H.-P. (1995), "Evolution strategies I: Variants and their computational implementation", *In Genetic Algorithms in Engineering and Computer Science*, Winter, G., Perieaux, J., Gala, M. and Cuesta, P. (eds.), Wiley, Chichester, 111-126.

Beyer, H.-G. and Schwefel, H.-P. (2002), "Evolution strategies: A comprehensive introduction", *Natural Computing*, **1**, 3-52.

Cai, J. and Thierauf, G. (1993), "Discrete structural optimization using evolution strategies", *In Neural Networks and Combinatorial Optimization in Civil and Structural Engineering*, Topping, B.H.V. and Khan, A.I. (eds.), Civil-Comp Limited, Edinburgh.

Cai, J. and Thierauf, G. (1996), "Evolution strategies for solving discrete optimization problems", *Adv. Eng. Softw.*, **2**, 177-183.

Ebenau, C., Rottschäfer, J. and Thierauf, G. (2005), "An advanced evolutionary strategy with an adaptive penalty function for mixed-discrete structural optimization", *Adv. Eng. Softw.*, **36**, 29-38.

Erbatur, F., Hasançebi, O., Tütüncü, İ. and Kılıç, H. (2000), "Optimal design of planar and space structures with genetic algorithms", *Comput. Struct*., **75**, 209-224.

Grierson, D.E. and Lee, W.H. (1984), "Optimum synthesis of steel frameworks using standard sections", *J. Struct. Mech*., **12**, 335-370.

Hajela, P. (1999), "Nongradient methods in multidisciplinary design optimization - status and potential", *J. Aircraft*, **36**, 255-265.

Hasançebi, O. (2006), "Optimization of truss bridges within a specified design domain using evolution strategies", *Eng. Optimiz.*, in press.

Hasançebi, O. and Ulusoy, A.F. (2005), "Discrete and continuous structural optimization using evolution strategies", *Proc. of the 8th Int. Conf. on the Application of Artificial Intelligence to Civil,* Structural and Environmental Engineering, Topping, B.H.V. (ed.), Civil-Comp Press, Rome.

Huang, M.-W. and Arora, J.S. (1997), "Optimal design of steel structures using standard sections", *Struct. Multidiscip. O.*, **14**, 24-35.

Lagaros, N.D., Papadrakakis, M. and Kokossalakis, G. (2002), "Structural optimization using evolutionary algorithms", *Comput. Struct*., **80**, 571-589.

Papadrakakis, M. and Lagaros, N.D. (1998), "Advanced solution methods in structural optimization based on evolution strategies", *Eng. Comput.*, **15**, 12-34.

Papadrakakis, M., Lagaros, N.D. and Fragakis, Y. (2003), "Parallel computational strategies for structural optimization", *Int. J. Numer. Meth. Eng*., **58**, 1347-1380.

Rajasekaran, S., Mohan, V.S. and Khamis, O. (2004), "The optimization of space structures using evolution strategies with functional networks", *Eng. Comput.*, **20**, 75-87.

Rechenberg, I. (1965), "Cybernetic solution path of an experimental problem", *Royal Aircraft Establishment*, Library translation No. 1122, Farnborough, Hants., UK.

Rechenberg, I. (1973), *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, Frommann-Holzboog, Stuttgart.

Rudolph, G. (1994), "An evolutionary algorithm for integer programming", In *Parallel Problem Solving from Nature – PPSN III International Conference on Evolutionary Computation*, Davidor, Y., Schwefel, H.-P. and Männer, R. (eds.), Volume 866 of *Lecture Notes in Computer Science*, Springer, Berlin, 139-148.

Saka, M.P. (1991), “Optimum design of steel frames with stability constraints”, *Comput. Struct.*, **41**, 1365-1377.
Schwefel, H.-P. (1965), *Kybernetische Evolution als Strategie der experimentellen Forschung in der Strömungstechnik*, Diploma thesis, Technical University of Berlin.
Schwefel, H.-P. (1977), “Numerische optimierung von computer-modellen mittels der evolutionsstrategie”, Volume 26 of *Interdisciplinary Systems Research*, Basel, Birkhäuser.
Schwefel, H.-P. (1981), *Numerical Optimization of Computer Models*, Wiley, Chichester.
Ulusoy, A.F., Erbatur, F. and Hasançebi, O. (2004), “Discrete optimization of rigid frames via evolution strategies”, *Proc. of the 3rd Int. Conf. on Advances in Structural Engineering and Mechanics (ASEM’04)*, Choi, C.-K., Kim, S.-H. and Kwak, H.G. (eds.), Seoul, Korea.