A Lévy Flight-Inspired Random Walk Algorithm for Continuous Fitness Landscape Analysis

Yi Wang, Guangdong University of Science and Technology, China*
Kangshun Li, South China Agricultural University, China

ABSTRACT

Heuristic algorithms are effective methods for solving complex optimization problems. The optimal algorithm selection for a specific optimization problem is a challenging task. Fitness landscape analysis (FLA) is used to understand the optimization problem's characteristics and help select the optimal algorithm. A random walk algorithm is an essential technique for FLA in continuous search space. However, most currently proposed random walk algorithms suffer from unbalanced sampling points. This article proposes a Lévy flight-based random walk (LRW) algorithm to address this problem. The Lévy flight is used to generate the proposed random walk algorithm's variable step size and direction. Some tests show that the proposed LRW algorithm performs better in the uniformity of sampling points. Besides, the authors analyze the fitness landscape of the CEC2017 benchmark functions using the proposed LRW algorithm. The experimental results indicate that the proposed LRW algorithm can better obtain the structural features of the landscape and has better stability than several other RW algorithms.

KEYWORDS

Algorithm selection, Fitness landscape analysis, Lévy flight, Random walk algorithm

1. INTRODUCTION

Evolutionary algorithms guided by heuristic principles have demonstrated their prowess in effectively addressing a multitude of real-world challenges (Zolpakar et al., 2021). In this dynamic landscape, the performance of diverse algorithms across distinct optimization problems exhibits notable variation. It is evident that a universal algorithmic solution capable of efficiently tackling all optimization problems remains an elusive aspiration (Singh et al., 2021). Regrettably, the present landscape of evolutionary algorithm research often overlooks the vital interplay between the specific optimization problem and the algorithmic approach adopted. Nevertheless, a promising avenue emerges through the analysis of an optimization problem's fitness landscape, serving as a valuable compass in the design and tailoring of evolutionary algorithms (Lu et al., 2019).

DOI: 10.4018/IJCINI.330535 *Corresponding Author

A fitness landscape takes form by compiling fitness values within the solution space, as highlighted by (Fragata et al. 2019). The tool of choice for unearthing the intricacies of optimization problems, Fitness Landscape Analysis (FLA), discerns essential attributes like entropy, neutrality, and smoothness. Scholars have been keenly focused on integrating FLA methodologies into evolutionary algorithms (Li et al., 2020). The exploration of landscape topologies has been pursued through various metrics, as evidenced by Malan et al.'s three-pronged approach (Malan et al., 2013). Similarly, Sallam et al. delved into landscape insights concerning diverse Differential Evolution (DE) operator selections (Sallam et al., 2017). Li et al. introduced dynamic fitness landscape analysis techniques to understand optimization problems comprehensively. Additionally, landscape characteristics are useful in assessing DE performance (Li et al., 2019). Among notable contributions, Huang et al. conceived a multi-objective differential evolution, titled Landscape Ruggedness Multi-Objective Differential Evolution (LRMODE), wherein information entropy informed the landscape's structural examination, steering DE's strategy selection (Huang et al., 2020). Tan et al.'s Fitness Landscape-Based Differential Evolution (FLDE) stands out, leveraging fitness landscape features to train the K-nearest neighbors (KNN) algorithm. This approach aptly guides mutation strategy selection based on distinct optimization problems (Tan et al., 2021). Within this literature, the exploration of landscape topologies chiefly centers on dissecting the sample points gleaned from a random walk. Methods such as auto-correlation coefficients, entropic measures, and dispersion metrics have been explored in the context of simple random sampling, extracting pivotal landscape attributes tied to optimization problems (Lang et al., 2019).

A random walk (RW) is a technique employed in various fields, beginning from an initial point in the defined space, creating neighboring points through a mutation operator based on the initial point, selecting the next point randomly from these neighbors, and subsequently generating a new set of neighborhood points based on the chosen point. This process continues iteratively. An illustrative example includes the work of Flyvbjerg et al. (1992), who introduced a straightforward random walk (SRW) algorithm for landscape analysis. Expanding on this concept, Malan et al. (2014) introduced the progressive random walk (PRW) algorithm, which involves multiple walks to sample neighborhood structures in continuous space. Another approach Jana et al. (2018) proposed is the chaos-based random walk (CRW) algorithm, which leverages chaotic mappings to generate random numbers and gain insights into local landscape features. However, it's important to note that the uniformity of sampling within the search space of a random walk algorithm significantly relies on the chosen sampling points. Consequently, the significance of RW algorithms or sampling methods is closely tied to landscape analysis (FLA), which offers valuable insights into the structural characteristics of the search space (Lang et al. 2020).

This paper introduces an innovative Lévy flight-based random walk (LRW) algorithm designed to extract essential fitness landscape features within a continuous search space. In the LRW approach, the Lévy flight mechanism governs both the step size and the directional changes of the random walk. The subsequent position of the walk is determined by the product of a random number and the adjustable step size. The LRW algorithm is compared against three other RW algorithms through a series of experiments to validate its effectiveness. Additionally, the proposed LRW algorithm is applied to analyze the landscape characteristics of the CEC2017 function set. The comprehensive experimental results unequivocally demonstrate the superior performance of the LRW algorithm in comparison to the other three RW algorithms.

The subsequent content of this paper is outlined as follows: In Section 2, we explore the simple random walk algorithm. Section 3 delves into an insightful discussion of several techniques tailored to analyze fitness landscapes. Transitioning to Section 4, we unveil the meticulously crafted framework underpinning our novel algorithm. The empirical findings and ensuing discussions are revealed in Section 5, shedding light on the experimental results. Summarizing the culmination of our study, Section 6 encapsulates the key takeaways and presents a prospective outlook on potential avenues for future research.

2. SIMPLE RANDOM WALK ALGORITHM

A random walk embodies a stochastic process that involves sampling within a designated search space at specific step intervals. This approach finds application across various disciplines, serving as a fundamental model to elucidate the nature of stochastic activities and operational behaviors. Diverse random walk algorithms have been proposed to dissect combinatorial optimization problems, and research findings underscore the pivotal role of step size selection in these algorithms. Many studies have advocated employing a probabilistic distribution-based strategy to determine the random step distribution (Ochoa et al. 2019). The computation of a comprehensive global fitness landscape poses significant challenges. Typically, fitness landscapes are computed through sampling or by traversing the domain space through random walks. **Algorithm 1** elucidates the workings of a specific instance, the random increasing walk algorithm.

A procedural outline for the random walk method unfolds: an initial landscape point is randomly generated within the domain space, and a specific walking strategy generates neighboring points, subsequently assessing whether these points fall within the defined problem bounds. The resulting sequence of the random walk is then established.

3. FITNESS LANDSCAPE ANALYSIS THEORY

The architecture of a fitness landscape encompasses an array of fitness values intricately woven within the solution space. These landscape characteristics serve as beacons illuminating the optimization problem's intricacy, directing evolutionary algorithms in their quest for optimal solutions. A diverse array of features captures the essence of fitness landscape structures, including ruggedness, neutrality, deception, and evolvability. The realm of landscape analysis has birthed several metrics designed to scrutinize the intricacies of optimization problems. In the ensuing sections, we explore two widely employed methodologies for dissecting fitness landscapes, shedding light on their fundamental significance.

3.1 Fitness Distance Correlation

Jones introduced the pioneering concept of fitness distance correlation (FDC) to gauge optimization problems' intricacies. Jones' innovative approach posits that problem complexity can be quantified by analyzing the interplay between fitness values and the spatial separation of solutions from the optimal point (Jones et al. 1995). This dynamic relationship is succinctly encapsulated by the FDC

Algorithm 1. Random increasing walk algorithm

Input: the problem's dimensions (D) and domain (*domain*),
the walk steps (*steps*), and step size (*size*).
1. Initial an empty sequence $walk = \phi$;
2. Initialize $count = 1$;
3. Generate a random number for a $walk[1]$;
4. while $count < steps$ **do**
5. for i=1,2,..,D **do**
6. Generate a random number ($step$) within [0, $size_i$];
7. Set $walk[count+1]_i = walk[\text{count}]_i + step$;
8. if $walk[count+1]_i > \max(domain)$, **then**
9. Set $walk[\text{count}+1]_i = walk[\text{count}]_i - domain$;
10. end if
11. end for
12. $count = count+1$;
13. end while
Output: the random walk sequence *walk*.

coefficient, which is a correlation measure that reveals the nuanced connection between fitness values and the proximity of solutions to the global optimum across the entire search space. The FDC coefficient is calculated as the correlation coefficient between fitness values and the distances to the nearest global optimum for all solutions within the search space. The fitness values in the model are determined by a sample of m solutions and the associated fitness value for each solution $F=\{f_1,f_2,...,f_m\}$. The Euclidean distance $D=\{d_1,d_2,...,d_m\}$ is computed from each solution in the sample to the optimal solution in the model. The optimal solution is an ideal solution to a known problem. Thus, the FDC is defined as (Tan et al. 2022):

$$FDC = \frac{C_{FD}}{\delta_F \delta_D} \tag{1}$$

where C_{FD} is the co-variance of F and D, which can be calculated by:

$$C_{FD} = \frac{1}{m}\sum_{i=1}^{m}\left(f_i - \bar{f}\right)\left(d_i - \bar{d}\right) \tag{2}$$

where δ_F and δ_D are the variances of F and D, respectively, and $\bar{f}$ and $\bar{d}$ are the mean values of F and D, respectively.

3.2 Ruggedness of Information Entropy

Information entropy is related to the number and distribution of locally optimal solutions. Malan et al. proposed a fitness landscape measurement method based on an information entropy theory (Malan et al. 2009). A random walk strategy is used to explore the character of a fitness landscape. Assume that a random walk is performed to obtain a series of fitness values $\{f_1,f_2,...,f_n\}$. A string $S(\varepsilon)=s_1,s_2,s_3,...,s_n$ of symbols $S_i \in \{-1, 0,1\}$ can be calculated by the following principle (Vassilev et al. 1997):

$$S_i\left(\varepsilon\right) = \begin{cases} -1, & if\ f_i - f_{i-1} < -\varepsilon \\ 0, & if\ \left|f_i - f_{i-1}\right| \le \varepsilon \\ 1, & if\ f_i - f_{i-1} > \varepsilon \end{cases} \tag{3}$$

the accuracy of the string $S(\varepsilon)$ calculation depends on the parameter ε.

According to the definition of the string $S(\varepsilon)$, the calculation of entropy $H(\varepsilon)$ is computed as follows:

$$H\left(\varepsilon\right) = -\sum_{p \ne q} P_{[pq]} \log_6 P_{[pq]} \tag{4}$$

where p and q are elements from the set {-1, 0, 1}, and $p \neq q$, the total number of different combinations $[pq]$ is equal to 6. The probability $P_{[pq]}$ can be calculated by:

$$P_{[pq]} = \frac{n_{[pq]}}{n} \tag{5}$$

where $n_{[pq]}$ is the number of the different combinations $[pq]$ in the string $S(\varepsilon)$.

The string $S(\varepsilon)$ is sensitive to the size of parameter ε. By increasing the value of ε, the entropy $H(\varepsilon)$ will become more prominent, which indicates the landscape is more straightforward. When the value of ε takes the maximum, the string $S(\varepsilon)$ is the collection of 0, which is called the information stability and is also represented by ε^*. Hence, the value of ε^* would be equal to the most significant difference in fitness values. The ε^* can be calculated by:

$$\varepsilon^* = \max\left\{f_i - f_{i-1}\right\}_{i=2}^{n} \tag{6}$$

To depict the ruggedness of a fitness landscape more intuitively, a single value R_f is used to measure the ruggedness of information entropy:

$$R_f = \max_{\forall\varepsilon\in[0,\,\varepsilon^*]} \left\{H\left(\varepsilon\right)\right\} \tag{7}$$

All ε in the set

$$\left\{\varepsilon \mid \varepsilon = 0, \frac{\varepsilon^*}{128}, \frac{\varepsilon^*}{64}, \frac{\varepsilon^*}{32}, \frac{\varepsilon^*}{16}, \frac{\varepsilon^*}{8}, \frac{\varepsilon^*}{4}, \frac{\varepsilon^*}{2}, \varepsilon^*\right\} \tag{8}$$

Therefore, the higher the R_f value, the more rugged landscape, while the value of R_f for a smooth landscape is small. The value of R_f captures the ruggedness, smoothness, and neutrality of landscapes. The ruggedness of information entropy is an effective measure index to estimate the complexity of a problem.

4. THE PROPOSED ALGORITHM

Fitness landscape analysis (FLA) serves as a critical tool for unraveling the inherent structural complexities of optimization problems. This intricate landscape structure is meticulously fashioned by deploying a Random Walk (RW) algorithm, which diligently generates an array of neighboring points. Thus, the efficacy of FLA is inherently intertwined with the characteristics of the chosen RW algorithm. The overarching objective lies in the development of an RW algorithm that comprehensively traverses the entire search space, thus enabling the algorithm to discern the defining attributes of the optimization problem. This endeavor is necessitated by the realization that the existing sampling points within the search space often fall short, undermining the thoroughness of landscape analysis. To address this limitation, we introduce the Lévy flight-based random walk algorithm. This innovative method harnesses the power of Lévy flights to furnish both the direction and the scaling factor for the step size of the walk. In the subsequent sections, we delve into an exposition of the Lévy flight concept and unveil the proposed algorithm's intricacies.

4.1 Lévy Flight

Lévy flight is a distinctive flight pattern observed in numerous flying creatures in the natural world. This pattern involves an intriguing interplay of both lengthy and brief steps, adhering to the principles of the Lévy distribution (Kaidi et al. 2022). Nonetheless, the practical implementation of the Lévy distribution function presents certain complexities. To address this, Mantegna proposed a methodology for simulating the Lévy distribution, enabling the computation of Lévy flight steps through the following formula (Jensi et al. 2016):

$$L\left(s\right) \sim \frac{\lambda\Gamma\left(\lambda\right)\sin\left(\frac{\pi\lambda}{2}\right)}{\pi}\frac{1}{s^{1+\lambda}}, \left|s\right| \to \infty \tag{9}$$

where, $\Gamma(\lambda)$ is the standard gamma function with an index λ, and significant steps ($s>s_0>0$) which is generated by the following transformation formula:

$$s = \frac{u}{\left|v\right|^{\lambda^{-1}}} \tag{10}$$

where u and v are two random samples drawn from a Gaussian normal distribution with a mean equal to zero, and standard deviations δ_u and δ_v can be expressed as:

$$u \sim N\left(0, \delta_u^2\right),\ v \sim N\left(0, \delta_v^2\right) \tag{11}$$

$$\delta_u = \left[\frac{\Gamma\left(1+\lambda\right) \times \sin\left(\pi \times \lambda/2\right)}{\lambda \times \Gamma\left(1+\lambda/2\right) \times 2^{\left(\lambda-1\right)/2}}\right]^{1/\lambda}, \delta_v = 1 \tag{12}$$

The parameter λ is set to 1.5. In fact, a Lévy flight diagram with 500 steps is shown in Fig.1.

4.2 LRW Algorithm

A D-dimensional random sampling point in the continuous search space can be denoted as $X = \{x_1, x_2,\dots,x_i, x_D\}$, and the domain of the search space for each sampling point is limited in $\left\{x_1^{\min}, x_1^{\max}\right\}$,..., $\left\{x_i^{\min}, x_i^{\max}\right\}, \left\{x_D^{\min}, x_D^{\max}\right\}$ where $x_i^{\min}$ and x_i^{max} indicate the lower and upper bounds of the individual x_i. The initial end of the walk in the search space is generated by:

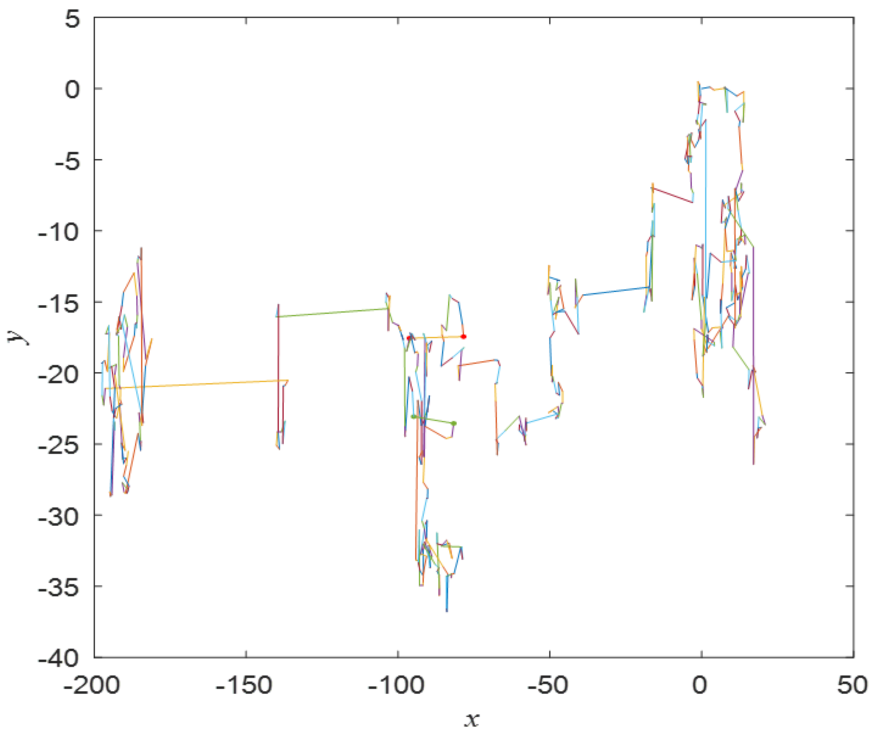
$$x_{1,i} = x_{1,i}^{\min} + \left(x_{1,i}^{\max} - x_{1,i}^{\min}\right) * rand\left(0, 1\right) \tag{13}$$

where $rand(0,1)$ is a uniformly distributed random number in [0,1]. First, generating a random point as a current point, the next point is obtained by disturbing each dimension of the current point, which uses a Lévy step size and the random number to generate the perturbing steps. The next point in the walk is generated recursively as:

$$x_{j,i} = x_{j-1,i} + \delta * s \tag{14}$$

where $j=1,2,\dots,N$. $\delta \in \left(-1, 1\right)$, s is the Lévy flight step size. Each random point in the walking process needs to be restricted to a search space boundary. The general framework of the proposed LRW algorithm is presented in **Algorithm 2**.

Figure 1. A Lévy flight diagram with 500 steps



Algorithm 2. Lévy random walk algorithm

Input: the problem's dimensions (*D*) and domain (*domain*), the walk steps (*steps*), and step size (*size*).
1. Initial an empty sequence *walk* $=\phi$;
2. Initialize *count* = 1;
3. Generate a random number for a *walk*[1];
4. while *count* < *steps* **do**
5. for i=1,2,..,*D* **do**
6. Generate a random number (*step*) within [0, $size_i$];
7. Set $walk[count+1]_i = walk[\text{count}]_i + step_* \delta$;
8. if $walk[count+1]_i$ >domain **then**
9. Set $walk[\text{count}+1]_i$ =max(domain);
10. else
11. Set $walk[\text{count}+1]_i$ =min(domain);
12. end if
13. end for
14. *count* =*count*+1;
15. end while
Output: the random walk sequence *walk*.

In essence, the effectiveness of an RW algorithm is intricately tied to the judicious selection of its step size. Employing larger step sizes often results in seemingly random leaps across the search space, resembling abrupt shifts to arbitrary locations. Consequently, utilizing such leaps to estimate landscape metrics like ruggedness tends to lack meaningful interpretation. Conversely, adopting exceedingly small step sizes confines the walk to a minute region within the search space, rendering a comprehensive understanding of the entire fitness landscape structure elusive. Hence, the decision regarding the RW algorithm's step size must strike a balance, ensuring a reasonable proximity between successive points in the walk that, in turn, facilitates enhanced coverage of the search space. To this end, we consider four distinct RW algorithms including SRW, PRW, CRW, and LRW. Each comprising 200 steps characterized by various step sizes, namely 5, 10, and 15. These configurations are visually depicted in Fig. 2, Fig.3, Fig.4, and Fig.5, respectively.

When contrasted with smaller step sizes, employing a step size of 15 yields a more uniformly distributed set of sampling points. Furthermore, Fig. 2 showcases how SRW tends to cluster

Figure 2. One independent sample walk of 200 steps with variable step size by SRW; (a) *s* = 5, (b) *s* = 10, (c) *s* =15

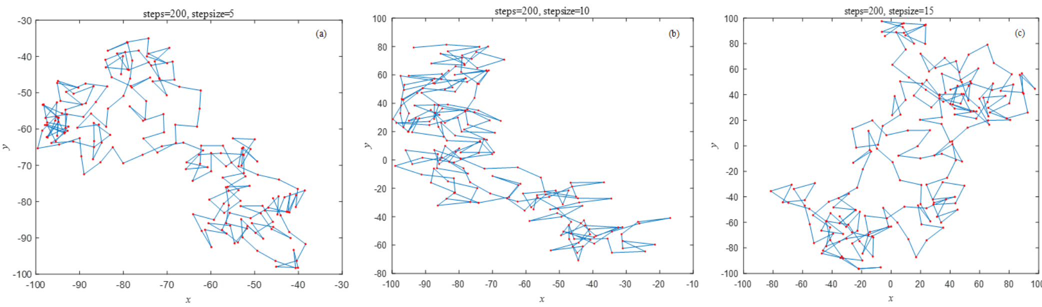


Figure 3. One independent sample walk of 200 steps with variable step size by PRW; (a) *s* = 5, (b) *s* = 10, (c) *s* =15

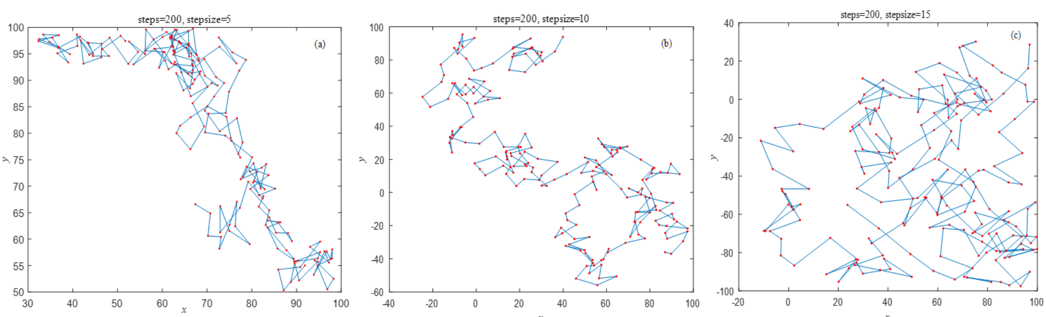


Figure 5. One independent sample walk of 200 steps with variable step size by LRW; (a) *s* = 5, (b) *s* = 10, (c) *s* =15

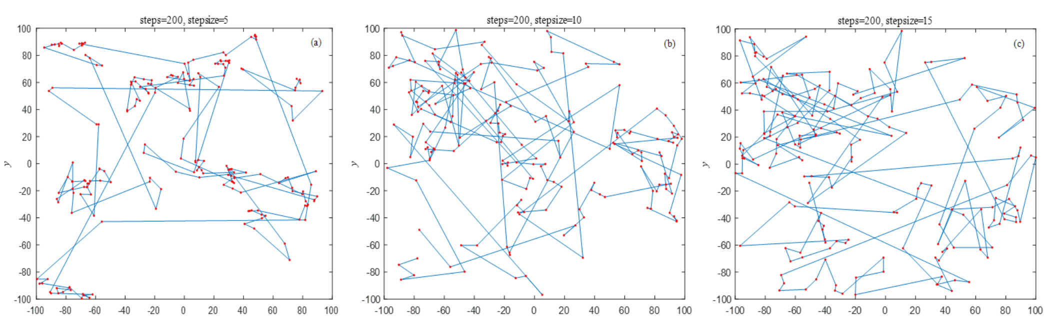
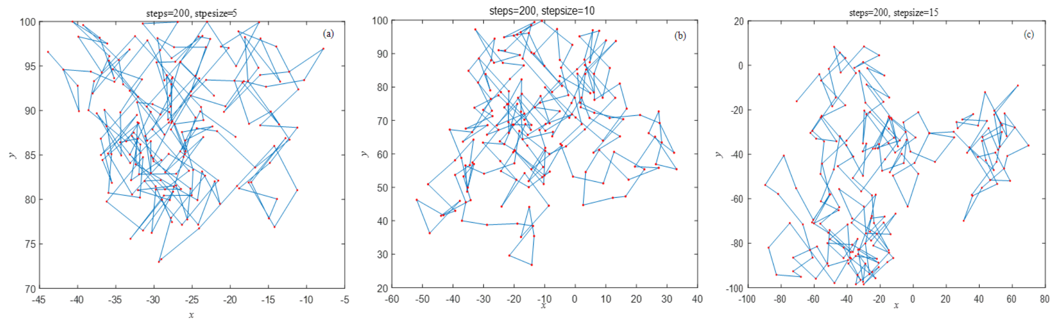


Figure 4. One independent sample walk of 200 steps with variable step size by CRW; (a) s = 5, (b) s = 10, (c) s =15



around two distinct angular regions for varying step sizes. Meanwhile, the step-increasing strategy employed by PRW leads to a concentration of random points along a single line, evident in Fig. 3. The landscape coverage depicted in Fig. 4 reveals that CRW outperforms SRW and PRW in terms of overall coverage, albeit with an area surrounding the target space remaining unexplored. However, it is the LRW algorithm, as depicted in the visual representation (Fig. 5), that shines by offering a significantly broader coverage area. This observation leads us to conclude that the Lévy variable-step random walk strategy proves to be an effective means of comprehensive search space sampling. Empirical experiments corroborate these findings, revealing that a step size of 15 yields satisfactory results in terms of search space coverage when employing the LRW algorithm for landscape analysis.

4.3 Coverage Testing Walks

The assessment of the coverage area within the continuous search space hinges on quantifying the deviation from a uniform sampling distribution. To accomplish this, a histogram serves as a tool to scrutinize the distribution of a selected set of points, facilitating the estimation of the probability distribution across the continuous search space. In our experimental context, Fig. 6 visually portrays the distributions of four walks comprising 10,000 points each within a two-dimensional space confined within the interval of [-100, 100]. This histogram encompasses 100 bins of equal dimensions, resulting in an average of 100 points allocated per bin.

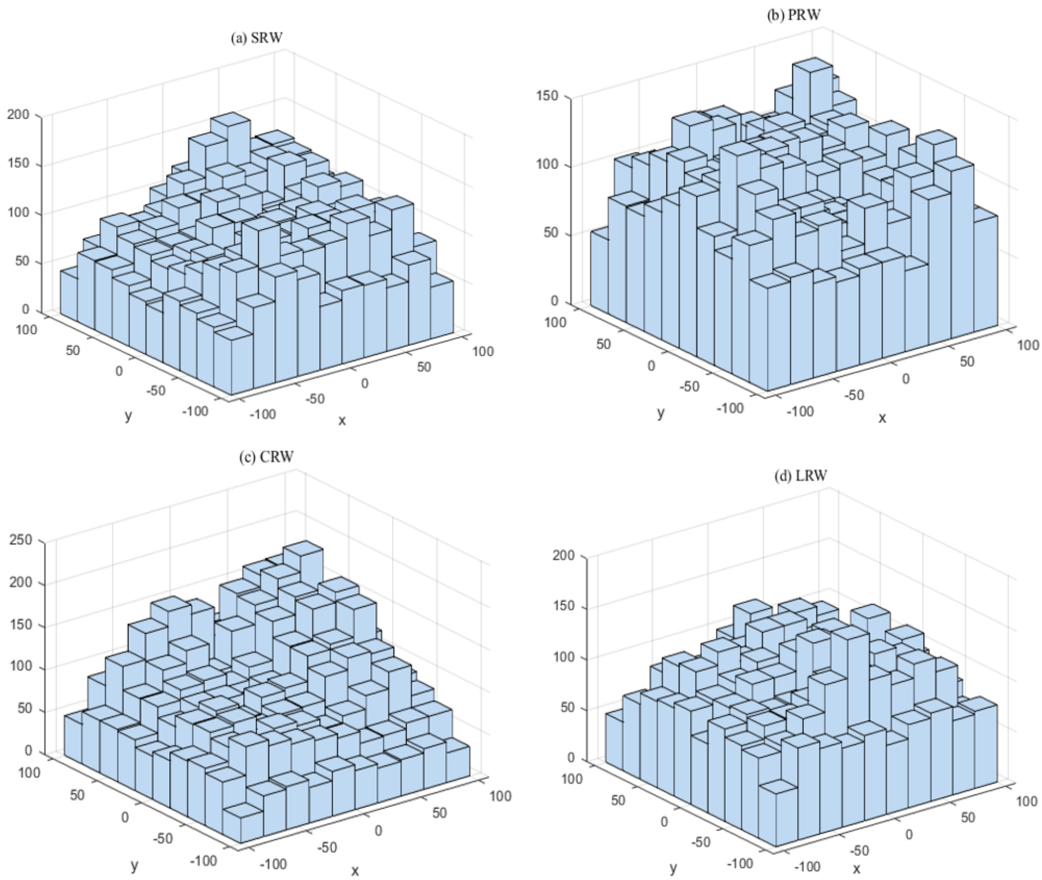
An insightful observation drawn from Fig. 6 is the discernible deviation of frequencies obtained from SRW, PRW, CRW, and LRW, extending beyond the mean value of 100. Particularly striking is the distribution of sample points attributed to LRW, which notably achieves a more accurate and comprehensive search space coverage when juxtaposed with the other three random walk methods. The issue of clustering within the search space is vividly elucidated by the histogram representation of these walks.

5. EXPERIMENTAL RESULTS AND ANALYSIS

5.1 Benchmark Functions

To evaluate the efficacy of the proposed LRW algorithm, we employ both information entropy and fitness distance correlation (FDC) as analytical tools to scrutinize the landscapes of the CEC2017 benchmark functions, as outlined by Mohamed et al. (2017). This benchmark compilation encompasses a total of 30 functions, each characterized by distinct attributes. These problems are broadly categorized into four types: unimodal functions (f_1-f_3), basic multimodal functions (f_4-f_{10}), hybrid functions (f_{11}–f_{20}), and composite functions (f_{21}–f_{30}). The search space

Figure 6. Histogram of four random walk algorithms for a sample of 10000 points



for all these functions is defined within the domain of [-100, 100]^D, following the specifications put forth by Wu et al. (2016).

During the experimental phase, a total of thirty independent walks were executed on each benchmark function, encompassing a total of $10^4{*}D$ steps for each walk. The step size (s) was consistently set to 15 throughout the entirety of the experiment. The dimensions of the functions were considered as D=30 and 50, reflecting the varying complexities of the optimization landscapes. Moreover, for the purpose of comparison, the SRW, PRW, and CRW were pitted against the LRW. The yardstick for evaluating the performance of these algorithms in landscape analysis is the accuracy they offer. Specifically, the entropic measure (H) was computed based on the maximum value of ε for each individual walk.

5.2 The Results of FDC

The fitness landscape analysis results of FDC for 30D, and 50D are shown in Table 1 and Table 2. Three well-known random walk algorithms, including SRW, PRW, and CRW, are compared with the LRW algorithm. It can be observed from Table 1 and Table 2 that the FDC values are different for each benchmark function, and the mean values obtained by all algorithms are similar.

Table 1. The results of FDC for 30*D*

Functions	SRW		PRW		CRW		LRW	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
f_1	0.849	9.254E-02	0.829	4.875E-02	0.814	8.206E-02	0.827	3.439E-02
f_2	0.146	1.608E-01	0.087	5.117E-02	0.106	1.001E-01	0.069	4.146E-02
f_3	0.074	2.395E-01	0.042	1.242E-01	0.030	1.813E-01	0.052	8.461E-02
f_4	0.592	2.704E-01	0.664	7.553E-02	0.626	1.648E-01	0.668	4.253E-02
f_5	0.776	1.496E-01	0.818	3.885E-02	0.791	1.074E-01	0.813	3.017E-02
f_6	0.648	9.961E-02	0.723	4.158E-02	0.691	8.446E-02	0.734	2.845E-02
f_7	0.876	8.415E-02	0.872	3.145E-02	0.882	5.547E-02	0.870	2.266E-02
f_8	0.756	1.115E-01	0.800	3.468E-02	0.795	7.726E-02	0.782	4.438E-02
f_9	0.558	1.198E-01	0.646	4.098E-02	0.594	9.844E-02	0.656	3.638E-02
f_{10}	0.060	6.437E-02	0.082	2.346E-02	0.081	7.732E-02	0.086	3.052E-02
f_{11}	0.008	2.575E-01	0.139	1.143E-01	0.104	2.105E-01	0.114	8.496E-02
f_{12}	0.640	2.162E-01	0.653	6.660E-02	0.625	1.608E-01	0.634	6.464E-02
f_{13}	0.447	2.564E-01	0.475	1.224E-01	0.493	2.154E-01	0.490	7.744E-02
f_{14}	0.116	2.574E-01	0.146	1.052E-01	0.041	2.523E-01	0.160	1.001E-01
f_{15}	0.234	3.003E-01	0.236	9.472E-02	0.269	2.409E-01	0.253	8.650E-02
f_{16}	0.307	3.577E-01	0.361	1.273E-01	0.399	1.792E-01	0.375	7.777E-02
f_{17}	0.345	1.769E-01	0.232	8.007E-02	0.257	1.955E-01	0.284	7.402E-02
f_{18}	0.107	2.783E-01	0.193	1.342E-01	0.233	2.464E-01	0.190	7.998E-02
f_{19}	0.265	3.002E-01	0.238	1.303E-01	0.258	2.797E-01	0.244	8.798E-02
f_{20}	0.158	8.698E-02	0.160	4.318E-02	0.148	9.205E-02	0.179	3.459E-02
f_{21}	0.186	2.675E-01	0.064	1.269E-01	0.039	2.379E-01	0.085	1.022E-01
f_{22}	0.019	1.251E-01	-0.015	6.807E-02	-0.007	7.529E-02	-0.008	5.286E-02
f_{23}	-0.125	3.063E-01	-0.158	1.600E-01	-0.261	2.349E-01	-0.162	1.228E-01
f_{24}	0.107	3.472E-01	0.127	1.558E-01	0.119	3.369E-01	0.122	1.171E-01
f_{25}	0.404	2.332E-01	0.345	1.035E-01	0.350	2.174E-01	0.324	7.520E-02
f_{26}	0.110	3.312E-01	0.119	1.328E-01	0.123	2.436E-01	0.124	8.516E-02
f_{27}	0.042	3.319E-01	0.087	1.338E-01	0.134	2.910E-01	0.028	8.915E-02
f_{28}	0.049	2.861E-01	0.012	1.149E-01	0.026	2.630E-01	0.006	1.239E-01
f_{29}	0.053	2.480E-01	0.050	1.262E-01	0.107	2.007E-01	0.083	9.399E-02
f_{30}	-0.038	3.558E-01	-0.054	1.688E-01	-0.062	3.046E-01	0.000	1.024E-01

When the problem's dimensionality changes, the magnitude of the difference in the FDC value of each benchmark function is not very significant. As the number of problem dimensions increases, all algorithms' stability worsens. For 30D optimization, LRW is significantly better than the other three algorithms, CRW demonstrates the worst performance, and PRW is slightly inferior to LRW. Fig. 7 visually represents the mean and standard deviation of the FDC for each benchmark function.

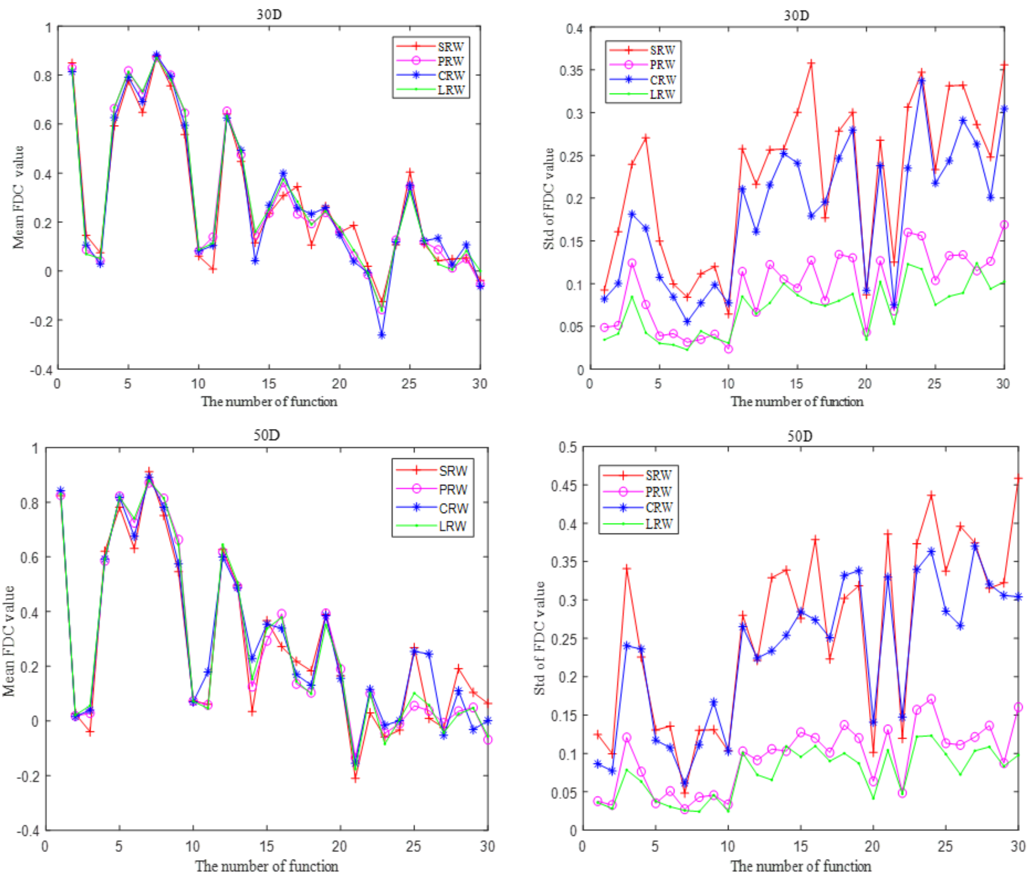
Table 2. The results of FDC for 50*D*

Functions	SRW		PRW		CRW		LRW	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
f_1	0.824	1.246E-01	0.826	3.783E-02	0.842	8.639E-02	0.830	3.608E-02
f_2	0.026	9.964E-02	0.018	3.279E-02	0.016	7.712E-02	0.021	2.825E-02
f_3	-0.040	3.407E-01	0.029	1.208E-01	0.040	2.402E-01	0.056	7.850E-02
f_4	0.621	2.256E-01	0.586	7.622E-02	0.592	2.360E-01	0.593	6.343E-02
f_5	0.782	1.305E-01	0.821	3.512E-02	0.817	1.172E-01	0.818	3.775E-02
f_6	0.630	1.355E-01	0.725	5.099E-02	0.675	1.075E-01	0.741	3.053E-02
f_7	0.912	4.827E-02	0.872	2.730E-02	0.891	6.132E-02	0.879	2.576E-02
f_8	0.751	1.303E-01	0.813	4.313E-02	0.781	1.114E-01	0.815	2.426E-02
f_9	0.545	1.309E-01	0.663	4.568E-02	0.574	1.668E-01	0.645	4.561E-02
f_{10}	0.073	1.033E-01	0.073	3.354E-02	0.070	1.030E-01	0.072	2.458E-02
f_{11}	0.061	2.801E-01	0.058	1.027E-01	0.179	2.650E-01	0.046	1.006E-01
f_{12}	0.627	2.212E-01	0.618	9.129E-02	0.599	2.241E-01	0.645	7.213E-02
f_{13}	0.496	3.290E-01	0.492	1.057E-01	0.488	2.337E-01	0.507	6.549E-02
f_{14}	0.034	3.388E-01	0.125	1.031E-01	0.229	2.539E-01	0.153	1.091E-01
f_{15}	0.367	2.761E-01	0.293	1.276E-01	0.354	2.847E-01	0.338	9.557E-02
f_{16}	0.272	3.786E-01	0.392	1.202E-01	0.339	2.739E-01	0.379	1.096E-01
f_{17}	0.218	2.231E-01	0.135	1.012E-01	0.170	2.507E-01	0.145	9.004E-02
f_{18}	0.183	3.019E-01	0.103	1.371E-01	0.131	3.320E-01	0.100	1.001E-01
f_{19}	0.389	3.186E-01	0.394	1.201E-01	0.385	3.381E-01	0.351	8.705E-02
f_{20}	0.165	1.014E-01	0.190	6.371E-02	0.155	1.407E-01	0.202	4.134E-02
f_{21}	-0.210	3.858E-01	-0.135	1.312E-01	-0.155	3.296E-01	-0.176	1.042E-01
f_{22}	0.030	1.195E-01	0.101	4.847E-02	0.116	1.473E-01	0.101	4.734E-02
f_{23}	-0.059	3.732E-01	-0.045	1.569E-01	-0.016	3.397E-01	-0.084	1.219E-01
f_{24}	-0.034	4.364E-01	-0.010	1.709E-01	0.001	3.631E-01	0.005	1.232E-01
f_{25}	0.268	3.377E-01	0.055	1.133E-01	0.254	2.853E-01	0.101	9.899E-02
f_{26}	0.009	3.957E-01	0.037	1.112E-01	0.245	2.663E-01	0.058	7.258E-02
f_{27}	-0.025	3.744E-01	-0.007	1.215E-01	-0.052	3.701E-01	-0.043	1.035E-01
f_{28}	0.191	3.153E-01	0.036	1.363E-01	0.110	3.206E-01	0.025	1.086E-01
f_{29}	0.105	3.223E-01	0.049	8.742E-02	-0.032	3.058E-01	0.047	8.332E-02
f_{30}	0.064	4.586E-01	-0.068	1.604E-01	0.001	3.042E-01	-0.056	9.783E-02

5.3 The Results of Entropy

The fitness landscape analysis results of information entropy for different dimensions are shown in Table 3 and Table 4. In the experiments, we only consider the maximum value of information entropy. Similarly, we can find from the Tables that the H values are different for each benchmark function. The information entropy value does not change much for optimization problems with different dimensions. The standard deviations of the values obtained by the four algorithms are all within 0.1

Figure 7. Mean and standard deviation of FDC value for all random walk algorithms with different dimensions



with relatively minor variations, except the CRW algorithm deviates more than 0.1 on some functions. The overall performance of the LRW algorithm seems to have a clear advantage from the calculation results. The version of SRW and CRW is lower than PRW. The Fig. 8 shows the variation curve of each algorithm on different functions.

5. CONCLUSION

A random walk algorithm is an effective strategy for dissecting the landscape structure inherent to optimization problems. In this research endeavor, we introduce a Lévy flight-based random walk (LRW) algorithm tailored to characterize optimization problem attributes, including information entropy and fitness distance correlation (FDC). This novel algorithm capitalizes on the principles of Lévy flight, thereby enhancing its performance. The proposed LRW algorithm's efficacy is meticulously assessed using the comprehensive CEC 2017 test suite. The empirical findings, acquired through experimentation on standard benchmark functions, unequivocally demonstrate the LRW algorithm's proficiency in discerning essential landscape features like ruggedness and deception. The performance of the LRW algorithm notably surpasses its counterparts in terms of stability and accuracy. Future research endeavors will explore additional features, such as fitness cloud and auto-correlation, employing the proposed random walk algorithms as a basis for investigation. This

Table 3. The results of entropy for 30*D*

Functions	SRW		PRW		CRW		LRW	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
f_1	0.814	7.403E-03	0.823	7.472E-03	0.811	9.591E-03	0.826	5.054E-03
f_2	0.415	2.908E-02	0.396	4.099E-03	0.452	3.025E-02	0.396	4.727E-03
f_3	0.573	6.330E-02	0.540	4.564E-02	0.527	7.700E-02	0.554	3.378E-02
f_4	0.799	1.428E-02	0.801	8.094E-03	0.805	9.343E-03	0.803	1.089E-02
f_5	0.855	6.074E-03	0.858	7.192E-03	0.844	7.229E-03	0.854	6.002E-03
f_6	0.863	6.698E-03	0.865	8.985E-03	0.842	8.844E-03	0.866	6.859E-03
f_7	0.841	6.193E-03	0.833	5.973E-03	0.831	1.117E-02	0.835	6.066E-03
f_8	0.855	7.364E-03	0.860	9.434E-03	0.842	9.585E-03	0.856	5.139E-03
f_9	0.861	1.070E-02	0.858	1.061E-02	0.857	1.112E-02	0.852	1.039E-02
f_{10}	0.861	7.389E-03	0.870	6.880E-03	0.843	7.533E-03	0.867	6.878E-03
f_{11}	0.586	7.231E-02	0.571	3.368E-02	0.540	7.372E-02	0.570	2.457E-02
f_{12}	0.818	9.355E-03	0.822	8.651E-03	0.810	7.620E-03	0.826	6.491E-03
f_{13}	0.811	9.773E-03	0.818	9.669E-03	0.808	8.258E-03	0.822	8.903E-03
f_{14}	0.750	1.949E-02	0.737	1.219E-02	0.762	1.824E-02	0.744	1.077E-02
f_{15}	0.803	1.351E-02	0.803	1.080E-02	0.806	1.025E-02	0.814	9.991E-03
f_{16}	0.779	2.094E-02	0.763	1.225E-02	0.781	1.718E-02	0.765	1.291E-02
f_{17}	0.625	8.849E-02	0.620	3.808E-02	0.615	7.898E-02	0.611	4.539E-02
f_{18}	0.780	1.684E-02	0.783	1.121E-02	0.789	1.589E-02	0.786	1.065E-02
f_{19}	0.802	9.725E-03	0.811	1.003E-02	0.811	9.591E-03	0.826	5.054E-03
f_{20}	0.853	3.970E-03	0.868	5.563E-03	0.452	3.025E-02	0.396	4.727E-03
f_{21}	0.817	2.210E-02	0.823	7.472E-03	0.527	7.700E-02	0.554	3.378E-02
f_{22}	0.857	8.349E-03	0.396	4.099E-03	0.805	9.343E-03	0.803	1.089E-02
f_{23}	0.834	1.178E-02	0.540	4.564E-02	0.844	7.229E-03	0.854	6.002E-03
f_{24}	0.818	7.640E-03	0.801	8.094E-03	0.842	8.844E-03	0.866	6.859E-03
f_{25}	0.795	1.246E-02	0.858	7.192E-03	0.831	1.117E-02	0.835	6.066E-03
f_{26}	0.801	1.324E-02	0.865	8.985E-03	0.842	9.585E-03	0.856	5.139E-03
f_{27}	0.819	1.161E-02	0.833	5.973E-03	0.857	1.112E-02	0.852	1.039E-02
f_{28}	0.814	7.403E-03	0.860	9.434E-03	0.843	7.533E-03	0.867	6.878E-03
f_{29}	0.415	2.908E-02	0.858	1.061E-02	0.540	7.372E-02	0.570	2.457E-02
f_{30}	0.573	6.330E-02	0.870	6.880E-03	0.810	7.620E-03	0.826	6.491E-03

extension of our work promises to contribute to a more comprehensive understanding of optimization landscape characteristics.

ACKNOWLEDGEMENT

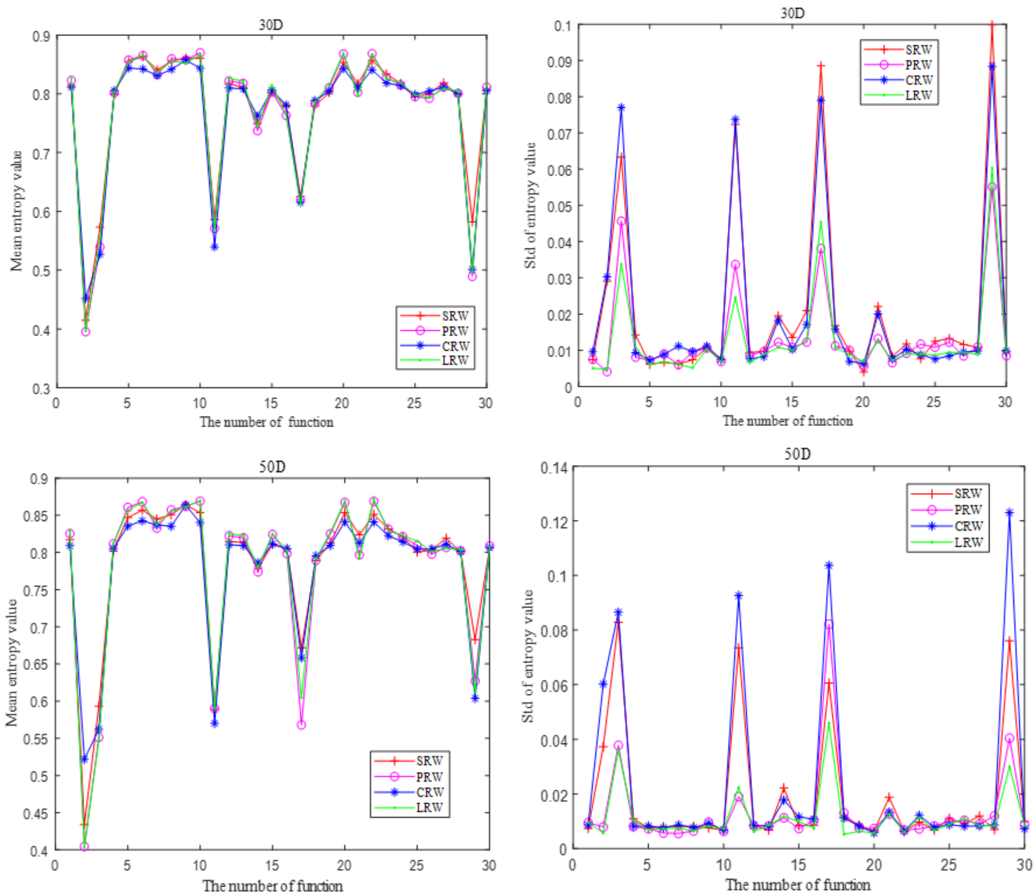
This work is supported by the Dongguan social development technology project (20231800910352), the Guangdong Youth Characteristic Innovation Project (2020KTSCX166,), and Project of

Table 4. The results of entropy for 50*D*

Functions	SRW		PRW		CRW		LRW	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
f_1	0.817	7.357E-03	0.825	9.451E-03	0.809	8.637E-03	0.829	9.529E-03
f_2	0.434	3.723E-02	0.405	7.924E-03	0.522	6.021E-02	0.404	5.655E-03
f_3	0.593	8.280E-02	0.551	3.776E-02	0.562	8.650E-02	0.548	3.603E-02
f_4	0.802	1.089E-02	0.811	8.017E-03	0.805	8.201E-03	0.812	1.028E-02
f_5	0.847	7.508E-03	0.861	7.245E-03	0.836	8.232E-03	0.857	6.906E-03
f_6	0.857	7.990E-03	0.868	5.704E-03	0.842	7.604E-03	0.867	7.272E-03
f_7	0.845	8.156E-03	0.833	5.536E-03	0.837	8.643E-03	0.838	7.445E-03
f_8	0.851	8.055E-03	0.857	6.420E-03	0.835	7.642E-03	0.829	9.529E-03
f_9	0.864	7.546E-03	0.862	9.614E-03	0.809	8.637E-03	0.404	5.655E-03
f_{10}	0.853	6.729E-03	0.869	6.332E-03	0.522	6.021E-02	0.548	3.603E-02
f_{11}	0.817	7.357E-03	0.590	1.899E-02	0.562	8.650E-02	0.812	1.028E-02
f_{12}	0.434	3.723E-02	0.822	8.416E-03	0.805	8.201E-03	0.857	6.906E-03
f_{13}	0.593	8.280E-02	0.820	8.211E-03	0.836	8.232E-03	0.867	7.272E-03
f_{14}	0.802	1.089E-02	0.825	9.451E-03	0.842	7.604E-03	0.838	7.445E-03
f_{15}	0.847	7.508E-03	0.405	7.924E-03	0.837	8.643E-03	0.855	6.443E-03
f_{16}	0.857	7.990E-03	0.551	3.776E-02	0.835	7.642E-03	0.861	8.204E-03
f_{17}	0.845	8.156E-03	0.811	8.017E-03	0.864	9.205E-03	0.870	7.256E-03
f_{18}	0.851	8.055E-03	0.861	7.245E-03	0.840	6.826E-03	0.590	2.234E-02
f_{19}	0.864	7.546E-03	0.868	5.704E-03	0.570	9.260E-02	0.825	6.161E-03
f_{20}	0.853	6.729E-03	0.833	5.536E-03	0.810	8.476E-03	0.829	9.529E-03
f_{21}	0.589	7.352E-02	0.857	6.420E-03	0.809	8.141E-03	0.404	5.655E-03
f_{22}	0.815	8.370E-03	0.862	9.614E-03	0.786	1.772E-02	0.548	3.603E-02
f_{23}	0.813	6.709E-03	0.869	6.332E-03	0.811	1.158E-02	0.812	1.028E-02
f_{24}	0.779	2.216E-02	0.590	1.899E-02	0.805	1.073E-02	0.857	6.906E-03
f_{25}	0.811	8.366E-03	0.822	8.416E-03	0.658	1.037E-01	0.867	7.272E-03
f_{26}	0.805	8.540E-03	0.820	8.211E-03	0.796	1.121E-02	0.838	7.445E-03
f_{27}	0.672	6.059E-02	0.774	1.123E-02	0.809	8.191E-03	0.855	6.443E-03
f_{28}	0.789	1.151E-02	0.824	7.365E-03	0.841	5.728E-03	0.861	8.204E-03
f_{29}	0.813	8.597E-03	0.798	1.022E-02	0.813	1.327E-02	0.870	7.256E-03
f_{30}	0.854	6.367E-03	0.568	8.213E-02	0.841	6.451E-03	0.590	2.234E-02

Guangdong Provincial Department of Education (2020KTSCX166), the Key research project of Guangdong University of Science and Technology (GKY-2021KYZDK-8), the Guangdong University of Science and Technology Key discipline research ability enhancement project (GKY-2022ZDXKTS), the Guangdong University youth innovative talent project (2021KQNCX120), and the Guangdong Institute of Science and Technology general project (GKY-2021KYYBK-20,GKY-2022KYYBK-25).

Figure 8. Mean and standard deviation of entropy value for all random walk algorithms with different dimensions



REFERENCES

Flyvbjerg, H., & Lautrup, B. (1992). Evolution in a rugged fitness landscape [J]. *Physical Review A*, *46*(10), 6714–6723. doi:10.1103/PhysRevA.46.6714 PMID:9907980

Fragata, I., Blanckaert, A., & Louro, M A D. (2019). Evolution in the light of fitness landscape theory [J]. *Trends in Ecology & Evolution*, *34*(1), 69–82. doi:10.1016/j.tree.2018.10.009 PMID:30583805

Huang, Y., Li, W., Tian, F., & Meng, X. (2020). A fitness landscape ruggedness multiobjective differential evolution algorithm with a reinforcement learning strategy *Applied Soft Computing*, *96*, 106693. doi:10.1016/j.asoc.2020.106693

Jana, N. D., Sil, J., & Das, S. (2018). Continuous fitness landscape analysis using a chaos-based random walk algorithm *Soft Computing*, *22*(3), 921–948. doi:10.1007/s00500-016-2397-2

Jensi, R., & Jiji, G. W. (2016). An enhanced particle swarm optimization with Lévy flight for global optimization *Applied Soft Computing*, *43*, 248–261. doi:10.1016/j.asoc.2016.02.018

Jones T., & Forrest S. (1995). Fitness distance correlation as a measure of problem difficulty for genetic algorithms[C]. */ICGA, 95*, 184-192.

Kaidi, W., Khishe, M., & Mohammadi, M. (2022). Dynamic Lévy flight chimp optimization *Knowledge-Based Systems*, *235*, 107625. doi:10.1016/j.knosys.2021.107625

Lang, R., & Engelbrecht, A. (2019). On the robustness of random walks for fitness landscape analysis. [C]//*2019 IEEE Symposium Series on Computational Intelligence (SSCI).* IEEE.

Lang, R. D., & Engelbrecht, A. P. (2020). Decision Space Coverage of Random Walks[C]//2020 *IEEE Congress on Evolutionary Computation (CEC).* IEEE.

Li, K., Liang, Z., Yang, S., Chen, Z., Wang, H., & Lin, Z. (2019). Performance analyses of differential evolution algorithm based on dynamic fitness landscape. *International Journal of Cognitive Informatics and Natural Intelligence*, *13*(1), 36–61. doi:10.4018/IJCINI.2019010104

Li, W., Tian, F., & Li, K. (2020). Fitness Distance Correlation Strategy for Solving the RGV Dynamic Scheduling Problem. *International Journal of Cognitive Informatics and Natural Intelligence*, *14*(3), 20–40. doi:10.4018/IJCINI.2020070102

Lu, H., Zhou, R., Fei, Z., & Guan, C. (2019). Spatial-domain fitness landscape analysis for combinatorial optimization *Information Sciences*, *472*, 126–144. doi:10.1016/j.ins.2018.09.019

Malan, K. M., & Engelbrecht, A. P. (2009). Quantifying ruggedness of continuous landscapes using entropy[C]//2009. *IEEE Congress on evolutionary computation*. IEEE.

Malan, K. M., & Engelbrecht, A. P. (2013). Ruggedness, funnels and gradients in fitness landscapes and the effect on PSO performance[C]//2013. *IEEE Congress on Evolutionary Computation*. IEEE.

Malan, K. M., & Engelbrecht, A. P. (2014). A progressive random walk algorithm for sampling continuous fitness landscapes[C]//2014. *IEEE Congress on Evolutionary Computation (CEC).* IEEE.

Mohamed, A. W., Hadi, A. A., & Fattouh, A. M. (2017). LSHADE with semi-parameter adaptation hybrid with CMA-ES for solving CEC 2017 benchmark problems[C]//2017. *IEEE Congress on evolutionary computation (CEC).* IEEE.

Ochoa, G., & Malan, K. (2019). Recent advances in fitness landscape analysis[C]. *Proceedings of the genetic and evolutionary computation conference companion*. IEEE.

Sallam, K. M., Elsayed, S. M., Sarker, R. A., & Essam, D. L. (2017). Landscape-based adaptive operator selection mechanism for differential evolution *Information Sciences*, *418*, 383–404. doi:10.1016/j.ins.2017.08.028

Singh, P. K., Singh, N., & Negi, R. (2021). Short-term wind power prediction using hybrid auto regressive integrated moving average model and dynamic particle swarm optimization. *International Journal of Cognitive Informatics and Natural Intelligence*, *15*(2), 111–138. doi:10.4018/IJCINI.20210401.oa9

Tan, Z., Li, K., & Wang, Y. (2021). Differential evolution with adaptive mutation strategy based on fitness landscape analysis *Information Sciences*, *549*, 142–163. doi:10.1016/j.ins.2020.11.023

Tan, Z., Tang, Y., Huang, H., & Luo, S. (2022). Dynamic fitness landscape-based adaptive mutation strategy selection mechanism for differential evolution *Information Sciences*, *607*, 44–61. doi:10.1016/j.ins.2022.05.115

Vassilev, V. (1997). Information analysis of fitness landscapes[C]//. *Proceedings of the Fourth European Conference on Artificial Life: ECAL97*. IEEE.

Wu ,G., Mallipeddi, R., & Suganthan, P. N. (2016). *Problem definitions and evaluation criteria for the CEC 2017 competition and special session on constrained single objective real-parameter optimization[J].* Nanyang Technol. Univ., Singapore, Tech. Rep.

Zolpakar, N. A., Yasak, M. F., & Pathak, S. (2021). A review: Use of evolutionary algorithm for optimization of machining parameters *International Journal of Advanced Manufacturing Technology*, *115*(1), 31–47. doi:10.1007/s00170-021-07155-7

Yi Wang received the B.S. degree in electronic engineering from Guangxi Normal University and the Ph.D. degree from the College of artificial Intelligence, South China Agricultural University. She has published over ten research articles. Her research interests include evolutionary optimization, fitness landscape, and large-scale optimization. KANGSHUN LI has been a Full Professor with South China Agricultural University, Guangzhou, China, since 2007. He has also been a Visiting Proffessor with The Chinese University of Hong Kong and the University of Calgary. He is currently a Professor of intelligent computation, image identification, data mining, evolutionary modeling, and neural network. He has published two books, 24 journal articles, and 34 conference proceedings articles. He is the Vice Chairman of the IEEE.