Applied Soft Computing 51 (2017) 192–208

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Measures in the time and frequency domains for fitness landscape analysis of dynamic optimization problems

Hui Lu[a,*], Jinhua Shi[a], Zongming Fei[b], Qianlin Zhou[a], Kefei Mao[a]

[a] School of Electronic and Information Engineering, Beihang University, Beijing 100191, PR China
[b] Department of Computer Science, University of Kentucky, Lexington, KY 40506-0495, USA

ARTICLE INFO

Article history:
Received 25 May 2016
Received in revised form
24 September 2016
Accepted 23 November 2016
Available online 29 November 2016

Keywords:
Dynamic fitness landscape
Landscape analysis
Evaluation criteria
Measures in time and frequency domains
Evolutionary optimization

ABSTRACT

Dynamic optimization problems (DOPs) have attracted increasing attention in recent years. Analyzing the fitness landscape is essential to understand the characteristics of DOPs and may provide guidance for the algorithm design. Existing measures for analyzing the dynamic fitness landscape, such as the dynamic fitness distance correlation and the severity of change, cannot give a comprehensive evaluation of the landscape and have many disadvantages. In this paper, we used Discrete-time Fourier transform (DTFT) and dynamic time warping (DTW) distance to acquire information of fitness landscape from frequency and time domains. Five measures are proposed, including the stationarity of amplitude change, the keenness, the periodicity, the change degree of average fitness and the similarity. They can reflect the features of fitness landscape from the aspects of outline, keenness, period, fitness value and similarity degree, respectively. These criteria can obtain essential information that cannot be acquired by existing criteria, and do not depend on the distribution of variables, the prior information of solutions and algorithms. To illustrate the performance of the five measures, experiments are conducted based on three types of standard DOPs with a two-peak function. In addition, we also apply these criteria on the test task scheduling problem for illustrating the fairness and adaptability. The experiment results show that these criteria can reflect the change characteristics of dynamic fitness landscape, and are consistent with the theoretical analysis.

1. Introduction

Dynamic optimization problems (DOPs) have been widely studied in recent years [1–3]. To deal with DOPs, it is essential to detect the environment change, such as the arrival of new tasks to the current schedule and the machine breakdown to the resource dispatch. It will make algorithm design much easier if we have prior knowledge of the characteristics of the solution space, especially for the change in dynamic process.

The analysis of fitness landscape has become an effective approach for understanding the characteristics of the solution space. For DOPs, the fitness landscape will change in terms of shape, height, and period during the optimization process. It makes the evaluation and selection of algorithms much more difficult. Through the analysis of fitness landscape, we can propose measures of landscape, like ruggedness, peak number, height, separation, and clustering in the solution space, to reflect the dynamic change of the landscape. Therefore, the analysis of dynamic fitness landscape is very important to help us understand the characteristics of DOPs. Using landscape properties to provide guidelines for algorithm design is a megatrend [4].

Recent researches proposed a series of criteria to extract the features of dynamic landscape. Some popular criteria are the modality, the dynamic fitness distance correlation, the mean severity of change, the mean optimal fitness difference, the mean distance between main metastable states and mean percentage of time to reach the main metastable state.

The modality [5] reflects the number of optima in a dynamic fitness landscape. It changes over time because the number of local optima varies in the dynamic environment. As a result, the computational process is clumsy. In addition, it has the limitation that it can only be easily calculated for constructed fitness landscapes for which an equation-like mathematical description is available [5]. However, this condition cannot always be fulfilled in the dynamic optimization problems.

Dynamic fitness distance correlation [6] measures the relationship between the distance to the nearest optimum and the fitness for a subset of solutions. This criterion can reflect the average

* Corresponding author.
E-mail address: mluhui@vip.163.com (H. Lu).

http://dx.doi.org/10.1016/j.asoc.2016.11.041
1568-4946/

hardness of the dynamic problem. However, it needs the prior knowledge of the optimal solution, and the variables must have a bivariate normal distribution. In addition, it only considers the difference between the landscape after each change and the initial landscape, but does not reflect the difference between two consecutive changes.

Mean severity of change and mean optimal fitness difference are two measures to evaluate dynamic landscape. Both of them concentrate on the variation of the optimal solution. Mean severity of change is proposed by Branke [7] to measure the average variation of the optimal solution in Euclidean distance. The mean optimal fitness difference [7] reflects the average difference between the fitness values of the optimal solutions in constant change cycles. These criteria can reflect the variation of spatial location and fitness value of the optimal solution after a series of changes, but they also need the information of the optimal solution. In addition, there is no connection with the difficulty of the problem. The mean distance between main metastable states and the mean percentage of time to reach the main metastable state [6] are also used to evaluate the dynamic landscape. However, they rely on specific algorithms, and different algorithms have different results.

To sum up, recent evaluation criteria for dynamic fitness landscape have many defects. Firstly, some of them need the information of optimal solutions, such as dynamic fitness distance correlation, mean severity of change and mean optimal fitness difference. Secondly, they have special requirement for the encoding of solutions. For example, the dynamic fitness distance correlation and mean severity of change are only useful for discrete representation. Thirdly, some of them have specific requirement for the distribution of variables, such as the fitness distance correlation. It works best when the variables follow a bivariate normal distribution, there is no guarantee that this will be the case if we have a random sample of fitnesses [8]. Fourthly, the adaptability to solve different problems is not obvious. We cannot obtain all the solutions for real-world problems. Therefore, we cannot use the criteria like mean distance between main metastable states and mean percentage of time to reach the main metastable state. Fifthly, some criteria depend on specific algorithms. Finally, existing criteria cannot reflect the periodicity, and the similarity in landscape.

In this paper, we focus on the dynamic fitness landscape from frequency domain and time domain. The stationarity of amplitude change, the keenness, the periodicity and the change degree of average fitness are proposed to characterize the fitness landscape by analyzing the frequency spectrum. The stationary of amplitude change can be used to describe the outline of fitness landscape. The keenness measures the acute degree of the landscapes. The periodicity can reflect the length of the repeating component in each landscape. The change degree of average fitness indicates the change of the whole fitness. In addition, the similarity of fitness landscape is proposed from the time domain. It uses dynamic time warping (DTW) distance to compare the similarity of different fitness landscapes.

In order to verify these criteria, we choose a DOP benchmark generator with a two-peak function for generating standard dynamic optimization problems. It can produce benchmark DOPs from binary static optimization, and make it possible to explore the properties of dynamic fitness landscape. The DOP generator used in this paper is proposed by Tinós and Yang [6], which is realized by three types of transformations occurring in the fitness landscapes. We can obtain three types of DOPs by simulation, and the change rules of them are different. Based on these fitness landscapes, we obtain statistical results through multiple runs and perform quantitative analysis. The experiment results can reflect the characteristics of three types of DOPs and are consistent with the theoretical analysis.

In addition, we apply these criteria to analyze the characteristics of the test task scheduling problem (TTSP) in automatic test system to illustrate the adaptability. We discuss the dynamic change from the aspect of the task number and instrument allocation. The experimental results can reflect the potential information of fitness landscape, and indicate that the number of task and the order of task sequence play a main role in terms of fitness landscape, while the instrument scheme only change the neighborhood information. The experiment also illustrates the effect and adaptability of the proposed criteria.

Our proposed measures are essential criteria for dynamic fitness landscape because they have no special requirement for the encoding of solutions, the optimal solutions, and the distribution of variables. In addition, they only reflect the characteristics of fitness landscape itself and do not depend on a specific algorithm. In other words, they are more general and can be widely used for analyzing various DOPs. Furthermore, they describe the fitness landscape from a new perspective which explores the information in terms of period, similarity and the degree of acute, which cannot be reflected by existing criteria.

The rest of the paper is organized as follows. Section 2 summarizes related work on characterizing optimization problems. In Section 3, we propose five metrics to describe dynamic fitness landscape. Experimental results of benchmark DOPs and TTSP are presented in Section 4 . Finally, Section 5 concludes the paper.

2. Related work

According to the No Free Lunch (NFL) theory [9], it is true that no one optimization algorithm is superior to the other at all times. In general, the design of the algorithm for a specific optimization problem lacks the adaptability analysis because the prior knowledge of the problem is deficient. In fact, the selection of an algorithm for the corresponding problem is implicit. Nowadays, analyzing the characteristics of the problems and providing guidance for the selection of algorithms has become popular. Malan and Engelbrecht [10] pointed out that many attempts at characterizing optimization problems have focused on finding a measure that could divide problems into those that are easy and those that are hard to solve [11–13]. These attempts have not been very successful. Therefore, instead of trying to find one measure of hardness, a more realistic approach could be to determine the characteristics of a problem. Analyzing the fitness landscape can explore the features of the solution space from different aspects, like ruggedness and modality. It has attracted much attention and become an important strategy for a better understanding of the problems in the hope that researchers will have better guidance in selecting appropriate algorithms. Study on the fitness landscape can be divided into three basic stages.

The first stage is from the late 1980s to the mid-1990s. Researchers focused on the analysis of the ruggedness, which relates to the number and distribution of local optima. If neighborhood solutions have very different fitness values, the result is a rugged landscape. The opposite of a very rugged landscape is a landscape with a single large basin of attractions or a flat landscape with no features. Techniques for measuring ruggedness are diverse. The autocorrelation function is one way proposed by Weinberger [14], which obtains a sequence of fitness values from a random walk in landscape, and calculates the correlation with the same sequence of values a small distance away. Weinberger also proposed the concept of correlation length [9], which is based on the results of autocorrelation function, and is extended in references [15–17]. Lipsitch put forward another computation method of correlation length one year later [18]. The principle is calculating the standard correlation coefficient between the fitness of points and the fitness of each of 30 i-mutant neighbors of the points. These techniques