Generating Customized Landscapes in Permutation-Based Combinatorial Optimization Problems

Leticia Hernando, Alexander Mendiburu, and Jose A. Lozano(✉)

Intelligent Systems Group,
Department of Computer Science and Artificial Intelligence,
University of the Basque Country UPV/EHU, San Sebastián, Spain
{leticia.hernando,alexander.mendiburu,ja.lozano}@ehu.es

Abstract. Designing customized optimization problem instances is a key issue in optimization. They can be used to tune and evaluate new algorithms, to compare several optimization algorithms, or to evaluate techniques that estimate the number of local optima of an instance. Given this relevance, several methods have been proposed to design customized optimization problems in the field of evolutionary computation for continuous as well as binary domains. However, these proposals have not been extended to permutation spaces. In this paper we provide a method to generate customized landscapes in permutation-based combinatorial optimization problems. Based on a probabilistic model for permutations, called the Mallows model, we generate instances with specific characteristics regarding the number of local optima or the sizes of the attraction basins.

Keywords: Combinatorial optimization problems · Landscape generator · Mallows model · Permutation space · Local optima

1 Introduction

Generating instances of combinatorial optimization problems (COPs) is an essential factor when comparing and analyzing different metaheuristic algorithms, and when evaluating algorithms that estimate the number of local optima of an instance. The design of a tunable generator of instances is of high relevance as it allows to control the properties of the instances by changing the values of the parameters.

Given the significance of this topic, several proposals have been presented in the literature. For example, a generator for binary spaces is proposed in [1], or a more recent work [2] shows a software framework that generates multimodal test functions for optimization in continuous domains. Particularly, the study developed in [3] has high relevance with our paper. The authors proposed a continuous space generator based on a mixture of Gaussians, which is tunable by

G. Nicosia and P. Pardalos (Eds.): LION 7, LNCS 7997, pp. 299–303, 2013.
DOI: 10.1007/978-3-642-44973-4_33,

a small number of user parameters. Based on that work we propose a generator of permutation-based COPs instances based on a mixture of a probabilistic model for permutations called the Mallows model.

The rest of the paper is organized as follows. The Mallows model is explained in Sect. 2. In Sect. 3 we present our generator of instances of COPs based on permutations. Finally, future work is given in Sect. 4.

2 Mallows Model

The Mallows model [4] is an exponential probability model for permutations based on a distance. This distribution is defined by two parameters: the central permutation σ_0, and the spread parameter θ. If Ω is the set of all permutations of size n, for each $\sigma \in \Omega$ the Mallows distribution is defined as:

$$p(\sigma) = \frac{1}{Z(\theta)} e^{-\theta d(\sigma_0, \sigma)}$$

where $Z(\theta) = \sum_{\sigma' \in \Omega} e^{-\theta d(\sigma_0, \sigma')}$ is a normalization term and $d(\sigma_0, \sigma)$ is the distance between the central permutation σ_0 and σ. The most commonly used distance is the Kendall tau. Given two permutations σ_1 and σ_2, it counts the minimum number of adjacent swaps needed to convert σ_1 into σ_2. Under this metric the normalization term $Z(\theta)$ has closed form and does not depend on σ_0:

$$Z(\theta) = \prod_{j=1}^{n-1} \frac{1 - e^{-(n-j+1)\theta}}{1 - e^{-\theta}}.$$

Notice that if $\theta > 0$, then σ_0 is the permutation with the highest probability. The rest of permutations $\sigma' \in \Omega - \{\sigma_0\}$ have probability inversely exponentially proportional to θ and their distance to σ_0. So, the Mallows distribution can be considered analogous to the Gaussian distribution on the space of permutations.

3 Instance Generator

In this section we show a generator of instances of COPs where the solutions are in the space of permutations. Our generator defines an optimization function based on a mixture of Mallows models.

The generator proposed in this paper uses $3m$ parameters: m central permutations $\{\sigma_1, ..., \sigma_m\}$, m spread parameters $\{\theta_1, ..., \theta_m\}$ and m weights $\{w_1, ..., w_m\}$. We generate m Mallows models $p_i(\sigma|\sigma_i, \theta_i)$, one for each σ_i and θ_i, $\forall i \in \{1, ..., m\}$. The objective function value for each permutation $\sigma \in \Omega$ is defined as follows:

$$f(\sigma) = \max_{1 \leq i \leq m} \{w_i p_i(\sigma|\sigma_i, \theta_i)\}.$$

Landscapes with different properties, and hence different levels of complexity, are obtained by properly tuning these parameters.

Some of these interesting properties are analyzed here. The first relevant factor we consider is that all central permutations σ_i's were local optima. Clearly, in order to be local optima, $\{\sigma_1, ..., \sigma_m\}$ have to fulfill that $d(\sigma_i, \sigma_j) \geq 2$, $\forall i \neq j$. A second constraint is that the objective function value of σ_i has to be reached in the ith Mallows model, i.e.:

$$f(\sigma_i) = \max_{1 \leq k \leq m} \{w_k p_k(\sigma_i|\sigma_k, \theta_k)\} = w_i p_i(\sigma_i|\sigma_i, \theta_i) = w_i \frac{e^{-\theta_i d(\sigma_i, \sigma_i)}}{Z(\theta_i)} = \frac{w_i}{Z(\theta_i)} \quad (1)$$

Moreover, in order to be σ_i a local optimum the following constraint has to be fulfilled:

$$f(\sigma_i) > f(\sigma), \quad \forall \sigma \;\; s.t. \;\; d(\sigma_i, \sigma) = 1. \quad (2)$$

To satisfy (2), and taking into account the constraint (1), we need to comply with:

$$\forall j = 1, ..., m, \quad \frac{w_i}{Z(\theta_i)} > w_j p_j(\sigma), \quad \forall \sigma \; s.t. \; d(\sigma_i, \sigma) = 1.$$

However, taking into account that if $\sigma \in \Omega$ is s.t. $d(\sigma_i, \sigma) = 1$, then $d(\sigma_j, \sigma) = d(\sigma_j, \sigma_i) - 1$ or $d(\sigma_j, \sigma) = d(\sigma_j, \sigma_i) + 1$, Eq. (2) can be stated as:

$$\frac{w_i}{Z_i(\theta_i)} > \frac{w_j}{Z_j(\theta_j)} e^{-\theta_j (d(\sigma_i, \sigma_j) - 1)} \quad , \qquad \forall j \in \{1, 2, ..., m\}, i \neq j. \quad (3)$$

Notice that once the parameters θ_i's have been fixed, the previous inequalities are linear in w_i's. So the values of w_i's could be obtained as the solution of just a linear programming problem. However, we have not defined any objective function to be optimized in our linear programing problem. This function can be chosen taking into account the different desired characteristics for the instance.

For example, one could think about a landscape with similar sizes of attraction basins. In this case, and without loss of generality, we consider that σ_1 is the global optimum and that σ_m is the local optimum with the lowest objective function value. Our objective function tries to minimize the difference between the objective function values of these two permutations (and implicitly minimize the difference of the objective function values of all the local optima). In addition we have to include new constraints to comply with these properties in the objective function values. This landscape can be generated as follows:

1. Choose uniformly at random m permutations in Ω: $\sigma_1, \sigma_2, ..., \sigma_m$, such that $d(\sigma_i, \sigma_j) \geq 2, \;\; \forall i, j \in \{1, ..., m\}, \;\; i \neq j$.
2. Choose uniformly at random in the interval $[a, b]$ (with $b > a > 0$) m spread parameters : $\theta_1, \theta_2, ..., \theta_m$.
3. Solve the linear programming problem in the weights w_i's:

$$\min \left\{ \frac{w_1}{Z(\theta_1)} - \frac{w_m}{Z(\theta_m)} \right\}$$

$$\frac{w_i}{Z(\theta_i)} > \frac{w_{i+1}}{Z(\theta_{i+1})} \qquad (\forall i \in \{1, 2, ..., m-1\})$$

$$w_i/Z(\theta_i) > w_j \frac{e^{-\theta_j(d(\sigma_i,\sigma_j)-1)}}{Z(\theta_j)} \quad (\forall i, j \in \{1, 2, ..., m\},\ ?i > j)$$

4. Assign to each $\sigma \in \Omega$ the objective function value:

$$f(\sigma) = \max_i \{w_i \frac{e^{-\theta_i d(\sigma_i,\sigma)}}{Z(\theta_i)}\}$$

4 Conclusions and Future Work

In this paper we introduce a framework to generate instances of COPs with permutation search spaces that is based on [3]. We create the landscapes based on a Mallows mixture. The aim is to obtain different kinds of instances depending on the central permutations $\sigma_1, ...\sigma_m$, the values of the spread parameters $\theta_1, ..., \theta_m$ and the values of the weights $w_1, ...w_m$.

Once the values of θ_i's are fixed, and the σ_i's are chosen, some linear constraints in w_i's have to be fulfilled in order to be all σ_i's local optima. These constraints can be accompanied by a function to be optimized, and therefore w_i's can be obtained as solutions of a linear programming problem. This optimization function is a key element when creating the instances under desired characteristics. One function is explained in Sect. 4, but obviously one could think of many other functions. For example, if we want to create an instance with a big size of attraction basin of the global optimum σ_i, our intuition leads us to think that we have to maximize the difference between the objective function value of σ_i and the other local optima, where σ_i is the local optimum that is further to the rest of local optima. However, if we want a global optimum σ_i with a small size of attraction basin, we could think about minimizing the difference between the objective function value of σ_i and the value of its neighbors, where σ_i has to be the local optimum that is nearer on average to the other local optima.

A remarkable point is that in the example we have taken the local optima uniformly at random. However, they can be chosen taking into account different criteria, such as the distance between them. For example, we can choose all the local optima as close as possible, or choose them maintaining the same distance, while the global optimum is far from them.

A more tunable model, and therefore more interesting when trying to create instances with different levels of complexity, can be obtained using the Generalized Mallows model [5]. This model uses a decomposition of the Kendall-tau distance and different spread parameters are assigned to each of the index $i \in \{1, 2, ..., n\}$, where n is the size of the permutations. So the parameters of the model ascend to $2m + n * m$. Apart from that, the Mallows model can be used with other distances such as the Hamming distance, the Cayley distance, etc.

We believe that by controlling the parameters we would we able to create instances with similar characteristics to those existing for famous COPs, such as the Traveling Salesman Problem, the Flowshop Scheduling Problem, the Linear Ordering Problem, etc. Moreover, we think that the model could be flexible enough to represent the complexity of real-world problems.

Acknowledgements. This work has been partially supported by the Saiotek, Etortek and Research Groups 2007-2012 (IT- 242-07) programs (Basque Government), TIN2010-14931 (Spanish Ministry of Science and Innovation) and COMBIOMED network in computational biomedicine (Carlos III Health Institute). Leticia Hernando holds a grant from the Basque Government.

References

1. De Jong, K.A., Potter, M.A., Spears, W.M.: Using problem generators to explore the effects of epistasis. Seventh International Conference on Genetic Algorithms, pp. 338–345. Morgan Kaufmann, San Francisco (1997)
2. Rönkkönen, J., Li, X., Kyrki, V., Lampinen, J.: A framework for generating tunable test functions for multimodal optimization. Soft Comput., 1–18 (2010)
3. Gallagher, M., Yuan, B.: A general-purpose tunable landscape generator. IEEE Trans. Evol. Comput. **10**(5), 590–603 (2006)
4. Mallows, C.L.: Non-null ranking models. Biometrika **44**(1–2), 114–130 (1957)
5. Fligner, M.A., Verducci, J.S.: Distance based ranking models. J. Roy. Stat. Soc. **48**(3), 359–369 (1986)